Long-term follow-up of childhood cancer survivors: clinical decision support and research participation

Kilsdonk, E.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
From an expert-driven paper guideline to a user-centred decision support system: a usability comparison study

Ellen Kilsdonk, Linda W.P. Peute, Rinke J. Riezebos, Leontien C.M. Kremer, Monique W.M. Jaspers

Abstract

Objective: To assess whether a user-centred prototype clinical decision support system (CDSS) providing patient-specific advice better supports healthcare practitioners in terms of (a) types of usability problems detected and (b) effective and efficient retrieval of childhood cancer survivor’s follow-up screening procedures compared to an expert-driven paper-based guideline

Materials and methods: A user-centred design (UCD) process was employed to design a prototype CDSS. Usability problems in information retrieval with the paper-based guideline were assessed by think-aloud analysis with 13 participants. Both simple and more complex tasks were applied. The analysis provided input for the UCD process of the prototype. The usability of the prototype CDSS was subsequently evaluated by think-aloud analysis with the same participants. Usability problems of the paper-based guideline and the prototype CDSS were compared by using the classification of usability problems scheme. In addition, efficiency (time to complete task) and effectiveness (completeness of retrieved screening procedures) of information retrieval of participants in the expert-driven paper-based guideline and the user-centred prototype CDSS were compared.

Results: Usability problems in both the paper-based guideline and the CDSS prototype were mainly classified as ‘incongruent with participants’ mental model’. The prototype CDSS reduced this type of problem from 17 to 6 problems. The time to perform simple information retrieval tasks increased by 58 s when using the prototype CDSS, however, it resulted in a 58% improvement in task completeness compared to the paper-based guideline. The time to perform complex scenarios decreased by 3:50 min with the prototype CDSS, with 17% higher completeness compared to the paper-based guideline.

Conclusions: Analysis showed that usability problems experienced by healthcare practitioners when using a paper-based guideline could be overcome by implementing the guideline in a user-centred CDSS design. Although different types of usability problems were experienced with the prototype CDSS, they did not inhibit effective and efficient performance of tasks in the system. The usability problem analysis of the paper-based guideline effectively supported comparison of usability problems found in the two information retrieval systems and it supported the UCD of the CDSS.
1. Introduction

In clinical practice, evidence-based guidelines are developed and implemented to guide and support healthcare practitioners’ performance in making more effective and efficient clinical decisions in patient care [1]. However, although the underlying principles of the guidelines are generally acknowledged, uptake of and compliance with paper-based guidelines in clinical practice is generally low [2]. Research into the causes of this low uptake shows that next to management and workload related issues, paper-based guidelines appear not to support healthcare practitioners’ practical information needs in patient care. This might be caused by the expert-driven design process used to develop paper-based guidelines [3].

An expert-driven design process is characterised by a focus on scientific validation of a guideline. In the design process of guidelines, experts give higher priority to reaching consensus than to content-related issues. A lack of focus on the presentation of this content in the guideline leads to complex and intricate content structures. As a result, clinicians experience difficulties in using the resulting paper-based guideline in daily practice [3], leading to a lack of familiarity with the guideline and low guideline adherence. Consequently, evidence-based clinical practice based on paper-based guidelines is hampered [4].

A possible solution for improving the use of guidelines into clinical practice is to communicate them through clinical decision support systems (CDSSs), preferably linked to an electronic medical record [5]. CDSSs may overcome problems in the use of paper-based guidelines in clinical practice by offering health care practitioners patient-specific advice based on the guideline recommendations. CDSSs have indeed been shown to improve guideline adherence by increasing healthcare providers’ knowledge of preferred practice, by reducing inertia to previous practice, and by reducing guideline complexity [6]. However, it is important to realise that the usability of a CDSS – the extent to which it accommodates clinicians’ cognitive workflow – is critical for the effective and efficient use of that CDSS [7]. A computer system’s usability is strongly tied to the extent to which a user’s mental model, a set of beliefs of how the system works, matches and predicts the actual actions of the system [8]. If a system design does not match the user’s mental model of the system then users will have a hard time learning how to use the system. This is a problem commonly
found with CDSS automation [9]. Designing a CDSS that optimally supports healthcare practitioners in clinical practice is thus challenging. One approach used to achieve these challenging goals is referred to as user-centred design (UCD) [10]. UCD is a system-development methodology that explicitly focusses on analysing end users needs, mental processes, limitations and preferences in order to design a system that meets end users requirements. Usability testing is a fundamental part of UCD for evaluating whether end users are supported in achieving their tasks in the system in an effective and efficient manner [10]. Since research has shown that many CDSSs are not usable [11], UCD has been strongly recommended for CDSS development [11,12]. Insight into the context of use and usability problems that healthcare practitioners encounter while processing information using a paper-based guideline can be useful in revealing the requirements for the design of a CDSS. Such a UCD approach may enhance CDSS acceptance among healthcare practitioners because it may help design a system that supports the end users’ needs that were not met or provided for in the paper-based guideline [3].

We followed a UCD approach in the development of a CDSS prototype implementing an expert-driven paper-based guideline. The guideline is used in daily practice by healthcare practitioners to define follow-up screening procedures for childhood cancer survivors. The UCD approach focussed on the usability problems healthcare practitioners encountered with the paper-based guideline as input into a CDSS design that implemented the guideline. This study explores the type of usability problems experienced by clinicians when using the expert-driven paper-based guideline and compared them with the usability problems when using a user-centred CDSS prototype; it further assesses the effectiveness and efficiency of clinician information retrieval of the two systems.

2. Methods

2.1. Research context

Over the last decades, advances in diagnosis and treatment of childhood cancer have dramatically increased long-term survival [13]. As a result, the number of childhood cancer survivors is growing rapidly and it has become increasingly clear that the former disease and the received treatments can significantly impair long-term health and quality of life by causing late effects. For this reason, the
long-term effects after childhood cancer task force (LATER) was started in 2004 by the Dutch childhood oncology group (DCOG LATER). LATER identified a nationwide cohort of over 6000 survivors and patient data including general patient information, medical history, tumour diagnosis and detailed treatment information, collected through paper-based and electronic medical records. In a collaborative effort of DCOG LATER, an evidence-based guideline was developed for screening childhood cancer survivors for the treatment-related late effects. The goal of the DCOG LATER guideline is to promote uniform and high-quality follow-up care across the seven university medical centres in The Netherlands involved in follow-up care of childhood cancer survivors [14].

The DCOG LATER guideline was developed by 16 multidisciplinary teams, each covering a specific clinical domain (e.g. cardiology, and nephrology). The teams were instructed to follow a uniform procedure in developing the guideline. The expert-based approach of designing the guideline resulted in a paper-based guideline document of 78 pages structured according to the clinical domains covered by the multidisciplinary teams. Subdivision of the 16 clinical domains resulted in a total of 24 domains in the paper-based guideline. Each domain, for instance cardiology, first describes domain-specific inclusion criteria (e.g. survivors that received radiation to the thorax). It then continues to describe the potential resulting long-term effects (e.g. arrhythmia) related to the inclusion criteria. Next, it describes the screening procedures to detect these long-term effects (e.g. an echocardiogram) and the frequency with which to perform these screening procedures. Lastly, it describes the advice that needs to be given to the patient and specific points of interest for the anamnesis and physical examination.

The DCOG LATER guideline is currently communicated through a paper-based format and provides information on screening protocols based on the treatment that a cancer survivor received in childhood. To prepare a survivor’s follow-up screening visit, the healthcare practitioner retrieves a patient’s historical diagnosis and treatment data from the paper-based and electronic medical records at their site. Based on the childhood cancer history, they determine the patient-specific screening procedures, points of interest for the anamnesis and physical examination, and the advice to be given to the patient by reviewing the contents of the paper-based guideline. Because of multi-modal cancer treatments and the
large variety of related possible late effects, healthcare practitioners experience difficulties in defining the overall follow-up screening procedures for individual patients. They have to go through more than 700 clinical rules to define the screening procedure for each individual cancer survivor. Consequently, defining the overall follow-up screening procedures for individual childhood cancer survivors by using the paper-based guideline, proves challenging for healthcare practitioners in daily clinical practice. The healthcare practitioners involved in follow-up care of childhood cancer survivors expressed a need for a CDSS providing them with patient-tailored advice to better support them in defining childhood cancer survivor’s follow-up screening procedures than the paper-based guideline.

2.2. Research design
This study consisted of two phases: a usability evaluation of the paper-based guideline and a formative usability study of the prototype CDSS. Thirteen healthcare practitioners from the follow-up clinics participated in both study phases. The usability problems that practitioners encountered in defining the follow-up screening procedures for individual patients by use of the paper-based guideline were analysed by use of the think-aloud method [15]. The results from this usability study provided input into the requirements and design of a working prototype CDSS. A formative usability study was performed on the CDSS design to assess if healthcare practitioners were more effective and efficient in defining the overall follow-up screening procedures for individual patients by using the prototype CDSS than by using the paper-based guideline. The paper-based guideline evaluation was conducted in April 2011; the formative usability study of the prototype CDSS in June 2011.

In the paper-based guideline evaluation, healthcare practitioners were asked to prepare two patient visits to the follow-up clinic on the basis of two fictitious patient scenarios, one simple and one complex. For each scenario, healthcare practitioners were asked to define the relevant follow-up screening procedures by using the paper-based guideline. The patient scenarios were constructed from real-life patient cases. An information sheet described the patient’s general information (age, sex, etc.), information about his/her cancer history, and information about the treatment the survivor had received in childhood. An
expert physician on late effects of treatments for childhood cancer reviewed and validated the scenario descriptions. For each scenario, a list of all relevant screening procedures was constructed by reviewing the paper-based guideline contents. These lists were verified by the expert physician. Participants were randomly assigned to either start with the simple scenario or to start with the complex scenario. They received instructions about the think-aloud method and a training task to practice the verbalisation of thoughts prior to starting their first scenario. During task performance, a microphone recorded the participants’ verbal utterances and a video camera recorded the participants’ activities with the paper-based guideline.

In the formative CDSS usability study, the healthcare practitioners were again asked to prepare two clinic visits on the basis of fictitious patient scenarios and to define the overall follow-up screening procedures by using the prototype CDSS. Two new patient scenarios, one simple and one complex, were developed which were similar to the ones used in the first study phase in terms of the number of screening procedures that applied to the fictitious patient. The expert physician once again validated the scenario descriptions and verified the list of relevant screening procedures constructed for each scenario.

A microphone recorded the participants’ verbal utterances. Screen and mouse clicks, face expressions, and voice and time records of the participants working with the prototype CDSS were captured with Morae software (TechSmith Corporation, Okemos, MI, USA).

2.3. Study population
The study participants were recruited through the coordinators of the long-term follow-up clinics by email. Participation was on a voluntary basis but participants were only included if they bear responsibility for preparing follow-up visits and defining the overall follow-up screening procedures for individual survivors by using the paper-based guideline.

2.4. Usability analysis
For both the paper-based guideline evaluation and formative CDSS usability study, all participants’ verbal utterances were transcribed to verbal protocols.
The verbal protocols were divided into segments, each segment representing a single comment from the participant. To detect what types of usability problems occurred in each user session, each protocol segment was linked to the video recordings (revealing the activities of healthcare providers on page sections of the paper-based guideline or CDSS screens at certain timestamps in the sessions). This analysis provided detailed insight into the origin of each usability problem encountered in the user sessions of the paper-based guideline and prototype CDSS.

Each time a unique usability problem was revealed it was added to the list of usability problems. When a specific usability problem had already been revealed and listed (because the session with another participant had already revealed the same usability problem), the frequency of that particular usability problem problem was increased. Two lists of usability problems were constructed; one concerning usability problems revealed with the paper-based guideline and one with the prototype CDSS.

2.5. Classification of usability problems
Each usability problem was mapped to the classification of usability problems (CUP) scheme, developed by Hvannberg and Law [16], by two independent researchers (EK, RR). The classification of usability problems according to the CUP scheme of the two researchers was compared and differences in classification were resolved through discussion. The CUP scheme focusses on usability problems concerning the structure and information content of an interactive tool. Because of this focus, the CUP scheme was considered suitable for classifying both paper-based guideline and prototype CDSS-related usability problems.

The CUP scheme contains attributes for describing a usability problem in detail. For the comparison of usability problems revealed with the paper-based guideline and the usability problems with the prototype CDSS, the CUP attributes severity and failure qualifier were considered relevant in the context of this study. Cohen’s kappa coefficient was calculated to assess inter-rater reliability for CUP-classification of failure quality for detected usability problems. Table 1 provides a description for the severity rating and the failure qualifier as defined by the CUP scheme.
2.6. The user-centred design of the prototype CDSS

Figure 1 provides an overview of the four stages within the UCD process followed to develop the prototype CDSS: (1) assessment of user needs and specification of the context of use, (2) specification of CDSS requirements and user needs, (3) design of the CDSS prototype, and (4) evaluation of the CDSS prototype.

The in-depth analysis of the protocol segments and video recordings of the user sessions with the paper-based guideline provided insight into the usability problems participants encountered and their unmet information needs in defining the screening procedures for the fictitious patient scenarios. Based on this analysis, the functional and non-functional requirements for the CDSS prototype’s graphical user interface (GUI) and information contents were specified and used as input for the design of a working prototype CDSS. Based on the usability analysis of the paper-based guideline, the organisational structure according to clinical domains was completely abandoned in the design of the prototype CDSS. Instead, the GUI of the prototype CDSS was designed

Table 1. Description of CUP attributes applied, e.g. severity and failure qualifier

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity</td>
<td>A problem is classified as severe, moderate or minor.</td>
</tr>
<tr>
<td>Failure qualifier</td>
<td>Describes how the participant experienced a usability problem. Contains the following categories: missing, incongruent mental model, irrelevant, wrong, better way and overlooked.</td>
</tr>
</tbody>
</table>

  - Severe: prevents the participant from completing a task
  - Moderate: hinders the completion of the task but workarounds are possible
  - Minor: does not hinder a task but still annoys participants
  - Incongruent: something in the tool is unclear, because it does not match with the test participant’s mental model or previous experience
  - Missing: participant fails to find something in the tool that was expected to be present
  - Better way: participant suggests that something could have been done differently in the tool
  - Wrong: participant can notice something went wrong with the tool, e.g. an apparent programming bug or typing error
  - Irrelevant: tool contains information that is not needed
  - Overlooked: participant is given a task but overlooks an entity, i.e. the participant fails to see existing entity or fails to realize to interact with it
as a tab-based structure with six tabs, to support ease of use in navigating and retrieving guideline information. A knowledge base was developed containing all the clinical rules that were extracted from the paper-based guideline. The clinical rules were extracted from the guideline and formalised to a computer-interpretable format by using the logical elements rule method (LERM) [17]. A decision engine applies the clinical rules by using individual patient data whenever a patient record is retrieved.

The prototype CDSS contains a home screen explaining its functionality. From the home screen menu, users can search for a specific patient by a unique identification number, a patient’s date of birth or surname. After the CDSS has retrieved the patient record, the tab-based structure containing six tabs appears. The first tab shows patient identification data and information about the childhood cancer. The second tab contains information on the cancer treatment the patient received. The third tab contains a list describing the screening procedures to be performed for the patient. The fourth tab contains points of interests for the anamnesis and physical examination specific to that patient. The fifth tab shows the specific advices that the healthcare practitioner should give the patient. The sixth and final tab contains the possible late effects that a patient could develop.

Figure 1. UCD process of the CDSS
as a result of the earlier cancer treatment. The prototype CDSS offers users the option to add, edit and delete content on the tabs for screening procedures, points of interest for anamnesis and physical examination and advice to accommodate for patient-specific situations. For instance, the tab screening procedures lists all the screening procedures that need to be performed on a patient. Healthcare practitioners are allowed to adapt this list when a patient’s condition does not allow the performance of a certain screening procedure. Fig. 2 shows a screenshot with the screening procedures tab opened. To support trust in the prototype CDSS, healthcare practitioners can also look up the evidence (decision rules) available for each screening procedure by clicking on the information buttons (i-buttons) displayed next to each screening procedure.

2.7. Outcome measures

Effectiveness was measured in terms of completeness with which the participants were able to retrieve relevant screening procedures for the patient scenarios by using the paper-based guideline and the CDSS prototype. After each session with the paper-based guideline and the CDSS prototype, the participants were asked to write down all patient-specific screening procedures for each scenario. For each patient scenario, the participants’ lists of screening procedures were compared to the list of all relevant screening procedures as defined by the guideline for that patient scenario (verified by an expert physician). For each participant, the number of correctly identified screening procedures per patient scenario was counted for using the paper-based guideline and the CDSS prototype. Per patient scenario and per participant, the percentage of correctly identified screening procedures with the paper-based guideline and with the prototype CDSS was calculated.

Efficiency was measured in terms of the time (minutes and seconds) the participants needed to define the overall follow-up screening procedures for each patient scenario by using the paper-based guideline or the prototype CDSS. The Wilcoxon signed-rank test was used to test for statistical significance of differences in effectiveness and efficiency between the paper-based guideline and the prototype CDSS.
**Figure 2.** Screenshot of prototype system, screening procedures tab

<table>
<thead>
<tr>
<th>Diagnostiek</th>
<th>Frequentie</th>
<th>Opmerking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Echocardiografie</td>
<td>2x per 5 jaar of 1x per 3 jaar</td>
<td></td>
</tr>
<tr>
<td>ECG</td>
<td>Eenzelf 5 jaar na diagnose</td>
<td></td>
</tr>
<tr>
<td>Glomerulaire functie en tubulaire functie</td>
<td>1x per 5 jaar</td>
<td>- bloed: creatinine, K, Mg, P, bicarbonaat, urine; albumine, creatinine, al-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>microglobuline, K, Mg, P</td>
</tr>
<tr>
<td>HBV</td>
<td>bij 1e LATER bezoek</td>
<td></td>
</tr>
<tr>
<td>HCV</td>
<td>bij 1e LATER bezoek</td>
<td></td>
</tr>
<tr>
<td>Bespalen Howell-Jolly bodys</td>
<td>eenmalig na einde therapie</td>
<td></td>
</tr>
<tr>
<td>Tympanometrie van 0.5 t/m 12.5 kHz en tympanometrie</td>
<td>1x per 5 jaar</td>
<td></td>
</tr>
<tr>
<td>VVV-vragenlijst</td>
<td>1x per 5 jaar</td>
<td></td>
</tr>
<tr>
<td>Gerichte tandartecontele op ontwikkelingsstoornissen, mondhygiene, speekselproductie</td>
<td>2x per 1 jaar / eenmalig 5 jaar na diagnose / rondom 10de-15de jaar</td>
<td></td>
</tr>
<tr>
<td>Voeg diagnostiek toe</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Results

Thirteen healthcare practitioners from six follow-up clinics responsible for preparing patient follow-up visits participated in this study: three male and three female expert physicians, and seven female expert physician assistants. The participants had a mean age of 44.9 years and a mean of 4.8 years experience in long-term follow-up of cancer survivors. On a scale of 1–5, the mean self-reported computer experience was 4.4. The participants’ demographic data are given in Table 2.

Table 2. Characteristics of healthcare practitioners participating in the paper-based guideline evaluation and formative usability study of the prototype CDSS.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>13</td>
<td>44.9</td>
<td>7.53</td>
<td>27-56</td>
</tr>
<tr>
<td>Years of experience</td>
<td>13</td>
<td>*</td>
<td>4.67</td>
<td>0.5-15.0</td>
</tr>
<tr>
<td>Self-reported computer experience</td>
<td>13</td>
<td>4.4</td>
<td>0.51</td>
<td>4-5</td>
</tr>
</tbody>
</table>

* regarding years of experience, mean is not measured.

3.1. Usability problems with the paper-based guideline and prototype CDSS

Table 3 shows the number of unique usability problems per failure qualifier and severity that participants experienced with the paper-based guideline and prototype CDSS. Inter-rater reliability of the classification of the paper-based guideline failure quality and CDSS were determined by $\kappa = 0.394$ ($P < .001$) and by $\kappa = 0.349$ ($P < .001$) respectively, both considered as fair agreement [18]. In using the paper-based guideline, the participants encountered 31 usability problems, 4 severe, 15 moderate and 12 minor. With the prototype CDSS, the participants encountered 22 usability problems, 5 severe, 10 moderate and 7 minor.

We classified 17 usability problems (1 minor, 13 moderate and 3 severe) with the paper-based guideline as incongruent, indicating that information presented in the guideline was unclear because it did not match with the participants’ mental model or previous experience within the LATER clinic. The majority of these problems related to the content being structured according to clinical domains. The participants’ initial search patterns through the paper-based guideline
Table 3. Number of unique usability problems encountered with paper-based guideline versus CDSS prototype, per CUP failure qualifier and severity (minor, moder = moderate, severe).

<table>
<thead>
<tr>
<th>Failure qualifier</th>
<th>Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>paper-based guidelines</td>
</tr>
<tr>
<td></td>
<td>minor</td>
</tr>
<tr>
<td>Incongruent</td>
<td>1</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
</tr>
<tr>
<td>Better way</td>
<td>4</td>
</tr>
<tr>
<td>Wrong</td>
<td>0</td>
</tr>
<tr>
<td>Irrelevant</td>
<td>7</td>
</tr>
<tr>
<td>Overlooked</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>12</strong></td>
</tr>
</tbody>
</table>

revealed that they tried to define individual screening procedures on the basis of the cancer treatment a patient had received. The participants gave up this strategy when realising that the inclusion criteria for certain screening procedures were structured according to organ system and not according to treatment. The remaining 14 usability problems were classified as missing (1 moderate), i.e. participants failed to find something that was expected to be present; better way (4 minor and 1 moderate), i.e. participant suggested that something could have been done differently; irrelevant (7 minor), i.e. the tool contains information that is not needed by users; and overlooked (1 severe), i.e. the participant fails to see an existing entity or fails to realise that he or she should interact with it. One participant mentioned that it was easy to miss screening procedures because of the guideline structure, which resulted in the severe usability problem of overlooked.

Nine usability problems (1 minor, 5 moderate and 3 severe) with the prototype CDSS were classified as incongruent. Compared to the incongruent usability problems detected in the paper-based guideline, most of the incongruent usability problems with the prototype CDSS concerned the terminology used for labelling screening procedures. The participants were sometimes unfamiliar with the terminology used to label screening procedures because of a different labelling used within their institute or clinic. Four usability problems (1 minor, 2 moderate and 1 severe) were classified as missing. We also encountered usability
problems that we classified as wrong (1 severe), i.e. participant can notice that something went wrong, and overlooked (2 minor problems), i.e. participant fails to see an existing entity or fails to realise that he or she should interact with it. We found one missing usability problem when one participant expressed concerns about the CDSS prototype’s validity and indicated to miss a link in the prototype CDSS to scanned pages of the paper-based guideline to verify the CDSS advice. A similar problem was classified as overlooked, which related to an information button with which participants could look up evidence on why a screening procedure should be performed: one participant was looking for this information but failed to notice the information button. The wrong usability problem was classified as severe because it resulted in lower completeness of retrieved screening procedures. A participant accidentally deleted a screening procedure from the list and the prototype CDSS provided no option to undo the deletion. The participant could not recall which screening procedure had been deleted, which resulted in a lower completeness. The remaining 6 usability problems experienced with the prototype CDSS were classified as better way (3 minor and 3 moderate).

3.2. Effectiveness and efficiency of use of the paper-based guideline and the prototype CDSS

Effectiveness: Table 4 shows the mean percentage of completeness of relevant screening procedures retrieved while using the paper-based guideline and the prototype CDSS per scenario. When using the paper-based guideline, the participants retrieved on average 36% more relevant screening procedures for the complex scenario than for the simple scenario. With the prototype CDSS, the participants retrieved on average 5% more relevant screening procedures for the complex than for the simple scenario. For both the simple and complex scenarios, the Wilcoxon signed-rank tests showed a statistically significant improvement in completeness of retrieved screening procedures by using the CDSS prototype compared to using the paper-based guideline. On average, the participants showed an improvement of 58% in retrieving the relevant screening procedures for the simple scenario by using the prototype CDSS compared to using the paper-based guideline. For the complex scenario, the participants showed an improvement of 17% by using the prototype CDSS compared to using the paper-based guideline.
Table 4. Percentage total completeness of retrieving screening guidelines and time on task per condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Scenario</th>
<th>Completeness (%)</th>
<th>Time (mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper-based guideline</td>
<td>13</td>
<td>Simple</td>
<td>42%</td>
<td>7:21</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Complex</td>
<td>78%</td>
<td>12:59</td>
</tr>
<tr>
<td>Prototype CDSS</td>
<td>13</td>
<td>Simple</td>
<td>100%*</td>
<td>8:19</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Complex</td>
<td>95%*</td>
<td>9:09</td>
</tr>
</tbody>
</table>

* $p < .01$

Efficiency: Table 4 also shows the mean time the participants required to retrieve screening procedures with the paper-based guideline and the prototype CDSS per scenario. When using the paper-based guideline, the participants needed on average 5:38 min more to retrieve the screening procedures for the complex scenario than for the simple scenario. With the prototype CDSS, the participants needed on average 0:50 min more to retrieve the screening procedures for the complex scenario than for the simple scenario. Although it took the participants on average 0:58 min more to retrieve screening procedures for the simple scenario with the prototype CDSS compared to the paper-based guideline, use of the prototype CDSS resulted in 58% improvement in scenario completeness compared to use of the paper-based guideline. For the complex scenarios, the participants needed on average 3:50 min less to retrieve the relevant screening procedures, with 17% higher completeness when they used the prototype CDSS than when they used the paper-based guideline. For both the simple and complex scenarios, the Wilcoxon signed-rank tests showed no statistically significant differences between the times it took the participants to retrieve the screening guidelines with the CDSS prototype or with the paper-based guideline.

4. Discussion
4.1. Main findings

Paper-based formats of evidence-based clinical guidelines can limit their usability since the knowledge contained in the guidelines may not be easily accessible to healthcare practitioners [19]. Hence, extracting information from a paper-based guideline and determining the relevance for a specific patient can require additional effort from a healthcare practitioner. In this study, we performed a think-aloud analysis in order to gain insight into current usability problems
when using a paper-based guideline for follow-up screening of childhood cancer survivors. The results of this analysis provided input into a user-centred prototype CDSS to support healthcare providers in defining the overall follow-up screening procedures for childhood cancer survivors.

We categorised the types of usability problems healthcare practitioners encountered using the expert-driven paper-based guideline and the user-centred prototype CDSS by the CUP scheme. With the paper-based guideline, we found 17 usability problems that we classified as incongruent. With the prototype CDSS, the number of incongruent usability problems was reduced to 9, indicating a better fit with participants’ information processing strategies. We also found less irrelevant usability problems with the prototype CDSS (0 usability problems) compared to the paper-based guideline (7 usability problems), indicating that the prototype CDSS did not contain unnecessary information. Although the number of usability problems were reduced in the categories incongruent and irrelevant, we did find new problems with the prototype CDSS in other categories. There were more usability problems with the prototype CDSS compared to the paper-based guideline in the categories of missing (4 vs. 1), better way (6 vs. 5), wrong (1 vs. 0) and overlooked (2 vs. 0).

The degree to which the healthcare practitioners’ effectiveness and efficiency in retrieving screening procedures was impacted was also assessed. With the paper-based guideline, healthcare practitioners retrieved on average 36% more relevant screening procedures for the simple scenario compared to the complex scenario. The verbal protocols revealed that in performing the simple scenario, the participants often relied on their clinical knowledge to look up screening procedures specific to the treatment the patient had received. The study participants often skipped the screening procedures that should be performed for all childhood cancer survivors. The participants were more thorough in their search for information in the paper-based guideline for the complex scenario, resulting in a higher completeness. This is in line with the results of Patel et al. [20], who found that reliance on experience in solving clinical problems contributes to skipping steps in clinical guidelines. With the prototype CDSS, the completeness with which the study participants retrieved screening procedures improved 58% for the simple scenario and 17% for the complex scenario.
Implementations of CDSSs have been shown to improve healthcare practitioners’ performance, particularly concerning preventive care [5, 21]. The study of Verhoeven et al. [3] showed that a UCD process of a website yielded a more effective and efficient means of communicating infection control guidelines: healthcare practitioners’ performance in completing tasks as recommended by the guidelines improved, while it took them less time to complete tasks. Their results showed that a UCD process of a website supporting healthcare providers in evidence-based practice maximised the effectiveness and efficiency of healthcare practice. We likewise used a UCD approach for the development of a CDSS to support healthcare practitioners in retrieving follow-up screening procedures for childhood cancer survivors. The retrieval of screening procedures by healthcare practitioners significantly increased in terms of effectiveness (correctly identified screening procedures) when using the CDSS prototype compared to using the paper-based guideline. A CDSS design is more likely to lead to an improvement in effectiveness and efficiency measurements compared to the manual process employed in a paper-based guideline, provided that its design optimally supports healthcare practitioners [22]. Suboptimal design may lead to new and severe usability problems, inhibiting the potential improvement in healthcare efficiency and effectiveness [22]. The UCD process in this study provided input to a prototype CDSS design in which healthcare practitioners could review patient-specific screening procedures, points of interest for anamnesis, physical examination and advice, and possible late effects more effectively by use of the tab-based structure. The comparison of usability problems experienced in the paper-based guideline and the user-centred prototype CDSS indicated that the CDSS tab-based structure better mapped to the mental model of the participants in effectively retrieving screening procedures, but also introduced new usability problems. With regard to the UCD process of the CDSS, these new usability problems will have to be solved in a redesign of the CDSS.

Minimising the time spent on consulting clinical guidelines is crucial when attempting to improve the uptake of clinical guidelines in everyday practice [19]. The results of Verhoeven et al. [3] showed a decrease of approximately 4 min in the time required to execute tasks when a website that communicated infection control guidelines was used instead of paper-based guidelines. In our study, the average time healthcare practitioners required to retrieve screening guidelines
for the simple and complex scenarios did not significantly improve with the prototype CDSS, although the participants needed on average 3:50 min less to complete the complex scenario.

A trusted knowledge base and confidence in the peers developing a CDSS are considered important factors for effective implementation of a CDSS [23]. When using the paper-based guideline, participants did not express any concerns regarding the validity of the information content. However, some participants indicated that they were not sure about the validity of the guideline information provided by the CDSS prototype. They expressed a need for the paper-based guideline to validate the information provided by the CDSS prototype. The clinical rules in the knowledge base were constructed according to the LERM procedure [17] and verified by a group of experts on late effects of childhood cancer treatment. This approach guarantees the validity of the CDSS knowledge base. It has been found that links to reference material is highly valued by users of decision support systems [24]. The prototype CDSS provided reference material to the evidence supporting specific screening procedures. The material was provided through small information buttons listed next to the screening procedure. However, the information buttons for disclosing this evidence were often overlooked by the healthcare practitioners. Moreover, the CDSS prototype did not provide evidence-based reference information for the advice, anamnesis, physical examination and possible late effects. This information should be available in a redesigned CDSS prototype system.

4.2. Limitations
The healthcare practitioners that participated in this study were recruited through an email that was sent to the coordinators of the follow-up clinics. Participation in the study was voluntary, potentially leading to bias of the results. However, only a limited number of healthcare practitioners are specialised in late effects of treatment for childhood cancer (approximately four in each follow-up clinic). The number of participating healthcare practitioners is therefore representative of the entire population.

In classifying the usability problems, not all detected problems could be adequately mapped according to the CUP scheme. One specific problem was related to the
unavailability of resources in a specific follow-up clinic: requesting a particular blood test as specified in the guideline was not possible in that centre. During the usability session of the paper-based guideline, one participant mentioned the screening procedure but chose not to document it while stating that it could not be ordered anyway. This resulted in a lower effectiveness in task completion and lack of adherence to the guideline, respectively. In this study we mapped this problem into incongruent, although we underline the fact that the CUP scheme should be expanded to be able to specify and include workflow-related problems.

Furthermore, our observed agreement in the classification of failure quality for the usability problems was considered fair agreement. In the CUP scheme, usability problems are classified on a high aggregation level, which may negatively affect the reliability of the inter-rater agreement [23]. However, all disagreements in usability problem classifications were resolved after discussion, which contributed to the detailed analysis of the final classification of the usability problems.

The study participants did not receive any training or explanation prior to the usability analysis of the prototype CDSS. Training was not provided because we were specifically interested in observing first-time user experiences regarding the ease of learning to use the prototype system. We were specifically interested in learning whether the CDSS design would intuitively, and, consequently, effectively and efficiently support end users in determining follow-up procedures. However, the formative usability results of the CDSS prototype are potentially influenced by the participants’ exploratory behaviour on the CDSS interface and its functionalities. This may have resulted in more time needed to retrieve patient-specific screening procedures with the prototype CDSS than participants might have needed if they had not explored the CDSS. Research has shown that until software is learned, additional time is typically required to complete daily tasks [25]. The additional time that the participants may have needed to learn to use the CDSS prototype could also explain the slight decrease in efficiency for the simple scenario compared to the situation in which they used the paper-based guideline. Such a learning effect might have had a negative impact on the time participants needed to complete the complex scenario when using the CDSS. In addition, this could potentially have influenced the type of usability problems that were found. Usability problems that were found during the prototype CDSS
analysis are problems associated with first-time use. After prolonged use of the CDSS in practice, these usability problems might be overcome by end users.

4.3. Implications
The prototype CDSS increased the effectiveness with which healthcare practitioners retrieved patient-specific screening guidelines. Whether the system will also lead to better adherence of healthcare practitioners to the DCOG LATER guideline in clinical practice needs to be further investigated. However, this study shows that it is important to incorporate healthcare participants’ information-processing strategies when designing clinical guidelines to increase the effectiveness of a guideline containing a high number of clinical decision rules. Evidence from clinical studies keeps accumulating and, consequently, clinical guidelines often need to be revised [26]. The knowledge bases of CDSSs need to adapt to dynamics in clinical guideline revisions in order to sustain high system validity [27]. Some participants in our study stated that they had concerns with the validity of information provided by the prototype CDSS, although the clinical rules contained in the knowledge base had been validated by experts in late effects of childhood cancer treatments. Further research into contributing factors to distrust in guideline validity in CDSS development is necessary to minimise validity concerns.

4.4. Conclusion
This study provides insight into the use and application of usability evaluation of current practices in paper-based guidelines to provide input for a UCD of a prototype CDSS. Our research demonstrated that usability problems experienced by healthcare practitioners when using a paper-based clinical guideline could be overcome by implementing the guideline in a user-centred CDSS design. Other usability problems were found in the CDSS prototype and should be prevented in a redesign. Even though the CDSS was only a prototype, healthcare practitioners were more effective with the prototype than with the original paper-based guideline in defining the screening procedures for individual patient scenarios. Healthcare providers did not become more efficient in retrieving these procedures by use of the CDSS, but their exploratory behaviour and time needed to learn to operate the CDSS might explain these results.
References


From an expert-driven paper guideline to a user-centred decision support system: A usability comparison study


