



Methods uncovering usability issues in medication-related alerting functions: results from a systematic review

Romarc Marcilly, Francis Vasseur, Elske Ammenwerth, Marie- Catherine Beuscart-Zephir

► To cite this version:

Romarc Marcilly, Francis Vasseur, Elske Ammenwerth, Marie- Catherine Beuscart-Zephir. Methods uncovering usability issues in medication-related alerting functions: results from a systematic review. Studies in Health Technology and Informatics, 2014, 10.3233/978-1-61499-432-9-885 . hal-04160481

HAL Id: hal-04160481

<https://hal.science/hal-04160481>

Submitted on 12 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Methods uncovering usability issues in medication-related alerting functions: results from a systematic review

Romarc MARCILLY^{a1}, Francis VASSEUR^b, Elske AMMENWERTH^c and Marie-Catherine BEUSCART-ZEPHIR^a

^a INSERM CIC-IT, Lille ; Univ Lille Nord de France ; CHU Lille ; UDSL EA 2694 ; F-59000 Lille, France

^b Université Lille Nord de France ; Pôle de Santé Publique, CHRU Lille ; UDSL EA2694 ; F-59000 Lille, France

^c Institute of Health Informatics, UMIT, 6060 Hall in Tyrol, Austria

Abstract. This paper aims at listing the methods used to evaluate the usability of medication-related alerting functions and at knowing what type of usability issues those methods allow to detect. A sub-analysis of data from this systematic review has been performed. Methods applied in the included papers were collected. Then, included papers were sorted in four types of evaluation: “expert evaluation”, “user-testing/simulation”, “on site observation” and “impact studies”. The types of usability issues (usability flaws, usage problems and negative outcomes) uncovered by those evaluations were analyzed. Results show that a large set of methods are used. The largest proportion of papers uses “on site observation” evaluation. This is the only evaluation type for which every kind of usability flaws, usage problems and outcomes are detected. It is somehow surprising that, in a usability systematic review, most of the papers included use a method that is not often presented as a usability method. Results are discussed about the opportunity to provide usability information collected after the implementation of the technology during their design process, *i.e.* before their implementation.

Keywords. Human engineering; usability; systematic review; decision support, clinical; method

Introduction

Usability studies are now mandatory to ensure an optimal and secure use of health technologies [1]. Usability is “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specific context of use” [2]. It deals with the features of the graphical user interface along with the fitting between the system behavior and its users’ needs. Numerous usability methods are listed in books, standards and papers (e.g. [2;3]). Yet, each method applies at a different step of the product lifecycle [2] and they do not allow uncover the same kinds of usability issues [4]. Moreover, as far as we know, for health

¹ Corresponding author : Romarc Marcilly, CIC-IT Biocapteurs & e-santé, Innovation & usages, Maison Régionale de la Recherche Clinique, 6 rue du professeur Laguesse, 59037 Lille. E-mail : romarc.marcilly@univ-lille2.fr / romarc.marcilly@gmail.com

technologies, there is no precise inventory of the methods used and of the usability issues they allow to retrieve.

Medication-related alerting functions are health technologies for which usability issues are recurrently highlighted (e.g. [5]). We performed a systematic review identifying clear facts about the types of usability flaws reported for this technology [6;7]. Usability flaws are concrete descriptions of the characteristics of the system that do not match usability design principles. When the system is put in use, usability flaws may cause usage problems, *i.e.* bothering of and even impairing user's experience with the technology. Finally, they may also generate negative outcomes in the work system, including patient safety issues. Results from this systematic review are presented in [6]. The present paper deals with original data and results not published elsewhere. It reports a sub-analysis aiming at answering the following questions: "what are the methods used that detect facts on usability flaws in medication-related alerting systems? What type of usability issues those methods allow to detect?"

1. Method

This review targeted original evaluation studies of medication-related alerting functions supporting the management of e-prescription by physicians, pharmacists and nurses and used in general hospital or in primary care general practice. Papers were searched in references databases (PubMed, Scopus and Ergonomics Abstract; last update on June 2013) and this search was completed by searching references lists of included papers. To be included, papers must report facts (not hypotheses) on usability flaws identified voluntarily or incidentally and the method used must be sufficiently well-described. Papers reporting evaluation of the perceived usability (e.g. through questionnaires) were excluded. For each paper included in the analysis process, two authors (RM, MCBZ) extracted together the used methods and the uncovered usability flaws, usage problems and negative outcomes. Usability flaws were categorized by both reviewers using first an existing usability heuristics [8] and then developing sub-categories through an inductive process.

The same experts used an extraction grid to list all methods used in the included papers. Then, based on the methods applied and on their aim, papers were sorted by both reviewers independently in four categories:

- "Expert evaluations": usability audits performed in-lab by several experts and involving any actual user
- "User-testing/simulation": observation of intended users using the technology following usage scenarios and while thinking-aloud in lab
- "On site observation": observation of the actual usage of the technology by actual users in the work system
- "Impact analyses": retrospective analyses of the results of the activities with the system or of the experience of actual users

While "expert evaluation" and "user-testing/simulation" are used during the design process of the product, "on-site observations" and "impact studies" are per se related to the evaluation of the product once implemented in a work system, e.g. in "post market surveillance".

Then, for each kind of evaluation, the number of papers reporting usability flaws, usage problems and negative outcomes was counted. For usability flaws, results were sorted according to their category.

2. Results

Out of the 6380 papers identified, a total of 26 papers matched the inclusion criteria and were analyzed in detail. Included papers propose a great variety of types of methods (cf. Table 1). The agreement score between both experts on the categorization of the papers according to the type of evaluation is almost perfect good (Cohen's $\kappa = 0.94$). The only disagreement was solved during a meeting. "Expert evaluation" is reported in 4 papers, "user-testing/simulation" in 6, "on site observation" in 11, and "impact studies" in 7. One paper reports both "expert evaluation" and "user-testing/simulation" and another "user-testing/simulation" and "on-site observation".

Table 1. Number of papers for the four types of evaluation. Sum is over 26 because two papers use two types of evaluation (one both "expert evaluation" and "user-testing/simulation" evaluations and the other both "on site observation" and "user-testing/simulation").

Type of evaluation	Instance of typical methods applied	Number of papers
"Expert evaluation"	Cognitive walkthrough, heuristic inspection	4
"User-testing/simulation"	User-testing, simulation	6
"On site observation"	Observation of the actual use (shadowing)	11
"Impact studies"	Focus groups/interviews on the usage of implemented systems, retrospective analysis with expert review	7

Both papers that used two types of evaluations did not allow distinguishing unambiguously the type of usability issues uncovered by each type of evaluation. Therefore, they were not included in the subsequent analysis. Table 2 presents the number of papers reporting usability flaws, usage problems and negative outcomes according to the type of evaluation used.

Table 2. Number of papers reporting usability flaws, usage problems and negative outcomes according to the type of evaluation. Additionally, usability flaws are sorted according to their category.

Usability issues			“Expert evaluations”	“User-testing”	“On site observations”	“Impact studies”
Usability flaws	General	Guidance	1	3	1	4
		Workload	1	3	6	3
		Significance of codes	0	2	1	2
		Consistency	1	0	1	0
		Explicit control	0	0	2	0
		Adaptability	0	0	1	0
	Specific	Error management	0	0	1	0
		Low signal-to-noise ratio	0	2	9	5
		Alert content	1	2	4	1
		Transparency	2	0	5	2
		Alert appearance	1	0	5	2
		Tasks and control distribution	0	1	3	2
		Alert features	1	1	3	0
Usage problems		0	4	10	5	
Negative outcomes		0	3	7	4	

Only papers using "on site observations" report every category of usability flaws along with usage problems and outcomes. This exhaustive coverage is mainly due to one paper [9] that used on-line appendices to provide the complete list of uncovered usability flaws. Without this paper, adaptability and error management categories of flaws would not have been represented. "Impact studies" report 3 out of 7 general

categories of flaws and all but one category of specific flaws. They report also both usage problems and outcomes. As for “user-testing/simulations”, they report all categories of specific flaws along with usage problems and outcomes but only 3 out of 7 categories of general flaws. Finally, “expert evaluations” report 3 out of 7 categories of general flaws and 4 out of 6 categories of specific flaws; this type of evaluation does not report any usage problem or outcome.

3. Discussion

This paper aimed at answering two questions: “what are the methods used that detect facts on usability flaws in medication-related alerting systems? What type of usability issues those methods allow to detect?” Results show that a wide range of methods is applied (from heuristics inspection to shadowing). Once the methods categorized in four types of evaluations, results confirmed that those types do not report similar usability flaws, as it has already been noticed in the literature for other types of technologies (e.g. [4]), and that not all of them report usage problems and outcomes.

Since this systematic review focused on usability, it was somehow logical to think that included papers would be usability studies. However, we took as main inclusion criteria for the papers the report of facts on usability flaws and not a specific type of method. Results surprisingly show that usual usability methods such as “expert evaluations” and “user-testing/simulation” do not constitute the larger proportion of the papers included in the review; those evaluations do not provide the best coverage of types of flaws, usage problems and negative outcomes. As expected, “impact studies” provide insight on usage problems and outcomes but do not cover all types of usability flaws. On the contrary, “on site observations”, that are not promoted as usual usability evaluation methods, are the most numerous studies included and provide a complete coverage of flaws, usage and outcomes. This type of evaluation allows an evaluation in the context of use after the technology is implemented. The flaws it highlights emerge from the integration into the clinical context; they are not specific to the characteristics of a given context. In sum, the evaluation of the technology while interacting with the work organization, different users’ profiles and other technologies uncovers flaws that could have not been detected out of the work system (*i.e.* in-lab).

Even if there may be reporting and publication biases, it is not deniable that “on site observations” provide a valuable insight on the question of the usability of medication-related alerting functions and more generally on health technologies. This result questions the coverage power of usual usability methods used during the design process (*i.e.* “expert evaluation” and “user-testing/simulation”). More specifically, just like other papers do (e.g. [4]) they question the power of detection of “user-testing/simulation” and their coverage in term usability flaws. Currently, the question of the power of “user-testing/simulation” is handled through calculating the optimal size of the sample of participants to observe in order to uncover all significant usability flaws (e.g. [10]). Results from the present study show that the coverage is not only a matter of sampling but also a matter of learning from the actual use of the technology.

Yet, for medical devices such as medication-related alerting functions, it is not allowed to implement a technology that has not been formerly CE marked [1]. Therefore, “on site observations” of the usage of such technologies are not possible.

An approach to bypass this issue is to capitalize the usability facts that have already been reported in the literature to elaborate a kind of “usability checklist”. This

checklist should be supported by empiric evidence: it must rest on reliable and exhaustive databases of usability facts reported (usability flaws and related usage problems and negative outcomes). For this purpose, it is necessary that all usability facts be precisely reported in on-line appendices of published papers [11]. Ultimately, such a grid specific to a given technology (e.g. alerting functions or infusion pumps) could be used during the design process of this technology, before it be CE marked and used in a work system. It would allow providing Human Factors experts and designers during the design process with the usability knowledge coming from the actual usage of similar systems. Therefore, it would be possible to anticipate issues that are uncovered generally only after the implementation of the system. This type of tool would increase the power of pre-implementation usability evaluations.

Acknowledgment

This work has been supported by Lille Academic Hospital and by Lille 2 University.

References

- [1] European Parliament and Council. Council Directive 2007/47/EC (2007), Official Journal L247/21.
- [2] International Standardization Organization. Ergonomics of human-system interaction -- Usability methods supporting human-centred design (ISO/TR 16982:2002). Geneva, International Standardization Organization, 2002.
- [3] MC Beuscart-Zephir, P Elkin, S Pelayo, R Beuscart. The human factors engineering approach to biomedical informatics projects: state of the art, results, benefits and challenges. *Yearb Med Inform* (2007): 109-27.
- [4] MW Jaspers. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *Int J Med Inform* 78 (2009): 340-53.
- [5] JS Ash, DF Sittig, EM Campbell, KP Guappone, RH Dykstra. Some unintended consequences of clinical decision support systems. *AMIA Annu Symp Proc* (2007): 26-30.
- [6] R Marcilly, E Ammenwerth, F Vasseur, E Roehrer, MC Beuscart-Zephir. Usability flaws of medication-related alerting functions: a systematic review. *Int J Med Inform*. Submitted.
- [7] R Marcilly, MC Beuscart-Zephir, E Ammenwerth, S Pelayo. Seeking evidence to support usability principles for medication-related clinical decision support (CDS) functions. *Stud Health Technol Inform* 192 (2013), 427-31.
- [8] DL Scapin, JMC Bastien. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behaviour and Information Technology* 6 (1997), 220-31.
- [9] AL Russ, AJ Zillich, MS McManus, BN Doebbeling, JJ Saleem. Prescribers' interactions with medication alerts at the point of prescribing: A multi-method, in situ investigation of the human-computer interaction. *Int J Med Inform* 81 (2012), 232-43.
- [10] M Schmettow, W Vos, JM Schraagen. With how many users should you test a medical infusion pump? Sampling strategies for usability tests on high-risk systems. *Journal of biomedical informatics* 46 (2013), 626-41.
- [11] LW Peute, KF Driest, R Marcilly, S Bras Da Costa, MC Beuscart-Zephir, MW Jaspers. A framework for reporting on human factor/usability studies of health information technologies. *Stud Health Technol Inform* 194 (2013), 54-60.