Université Lille 2
Droit et Santé

Université Lille Nord de France
Pôle de Recherche et d'Enseignement Supérieur

Année 2014

THÈSE DE DOCTORAT

présentée par

Romaric MARCILLY

en vue de l'obtention du grade de
Docteur de l'Université de Lille 2
Discipline : Informatique Médicale

TOWARDS A USABILITY KNOWLEDGE BASE TO SUPPORT HEALTH INFORMATION TECHNOLOGY DESIGN AND EVALUATION: APPLICATION TO MEDICATION-RELATED ALERTING SYSTEMS

Présentée et soutenue publiquement le 15 octobre 2014

Directeurs de thèse :	Francis VASSEUR	Docteur (HDR) *Université de Lille 2*
	Elske AMMENWERTH	Professeur des Universités *UMIT*
Rapporteurs :	Alain VENOT	Professeur des Universités *Université de Paris 13*
	JM Christian BASTIEN	Professeur des Universités *Université de Lorraine*
Président du jury :	Régis BEUSCART	Professeur des Universités *Université de Lille 2*
Examinateurs :	Christian LOVIS	Professeur des Universités *Université de Genève*
	Christian NØHR	Professeur des Universités *Aalborg University*
	Paul TURNER	Professeur des Universités *University of Tasmania*

Acknowledgments

I would like to thank the members of the dissertation committee, and especially the reviewers, for having accepted to evaluate this work. I am very grateful to them for that.

I also wish to express my gratitude to my PhD directors, Pr. Elske Ammenwerth and Dr. Francis Vasseur for having accepted to supervise my work and for their innumerable precious advices. I would like also to express my sincere thanks to Dr. Marie-Catherine Beuscart-Zéphir for her frequent supervision on my work: even if our agendas were difficult to combine; her advices, her support for the extraction, the analysis and the interpretation of the data was more than precious. Without her, this PhD could not have come to an end.

I would also like to express my sincere thanks to my employer, Lille' academic hospital for having funded my work. I would like to extend my thanks to my lab (Lille' CIC-IT) manager, Pr. Regis Logier and especially to the manager of the Human Factors team, Dr. Sylvia Pelayo, for proposing me this PhD and for allowing me to join in all the conferences I needed to be part of. Thank you Sylvia for all your support.

The Research Host Team 2694, coordinated by Pr. Alain Duhamel, provided me an essential financial, technical, administrative and human support to successfully complete this work. Together, Lille' CIC IT, INSERM, Lille' academic hospital and research host team 2694 provided me with ideal conditions to carry out this work.

I would like also to express my appreciation to the Doctoral School "Biologie et Santé" and especially Laurence Fofana for committing herself to providing all the required assistance to ensure everything was going as well as possible during the PhD.

Practically, this work could not have been achieved without the support of Lille 2 Medicine faculty's library. Their very efficient work in retrieving the required papers for my review was critical in the success of this work. I would also like to thank Dr. Benoit Thirion and Pr. Stéfan Darmoni for the great support in the definition of the requests.

I would like to take this opportunity to thank all my colleagues in Lille' CIC-IT, especially the Human Factors team, who have been involved in this PhD either by informal exchanges or more directly by retrieving and/or analyzing data. I thank also the members of the CERIM lab for their precious help which made my tasks easier.

I would like to address my warm thanks to Dr. Julie Niès for her great moral support and for her precious reviews and comments on this manuscript. I also wish to express my grateful thanks to Dr. Erin Roehrer for her valuable feedbacks on this document. A special thank to Dr. Linda Peute for having encouraged me.

At several occasions during congresses and seminars I met people who influenced my work by exchanging with me. I would like to thank them. I would like to extend my thanks to my former

supervisors who trained me in experimental, ecological and cognitive psychology, Dr. Marion Luyat and Dr. Françoise Anceaux. Also, I would like to express my sincere thanks to Pr. Francis Six who trained me in ergonomics. Many thanks to Dr. Justine Forrière for our helpful exchanges.

A special thank to my colleagues PhD students in Lille' CIC-IT: Sabrina Bras Da Costa, Antoine Lamer, Maxime Baas, Clément Wawrzyniak. You're next in line.

A warm thank to all the debriefers and to my sport and music coaches.

Finally I would like to thank my beloved parents, my brother and my sister, along with my grand-parents, uncles and aunts, my cousins, my brother-in-law and my friends for having trusted and encouraged me all along this work. You are my fuel.

To Emmanuel. This PhD is yours.

VERS UNE BASE DE CONNAISSANCE EN UTILISABILITE POUR AIDER LA CONCEPTION ET L'EVALUATION DE TECHNOLOGIES DE L'INFORMATION EN SANTE : APPLICATION AUX SYSTEMES D'ALERTE MEDICAMENTEUX

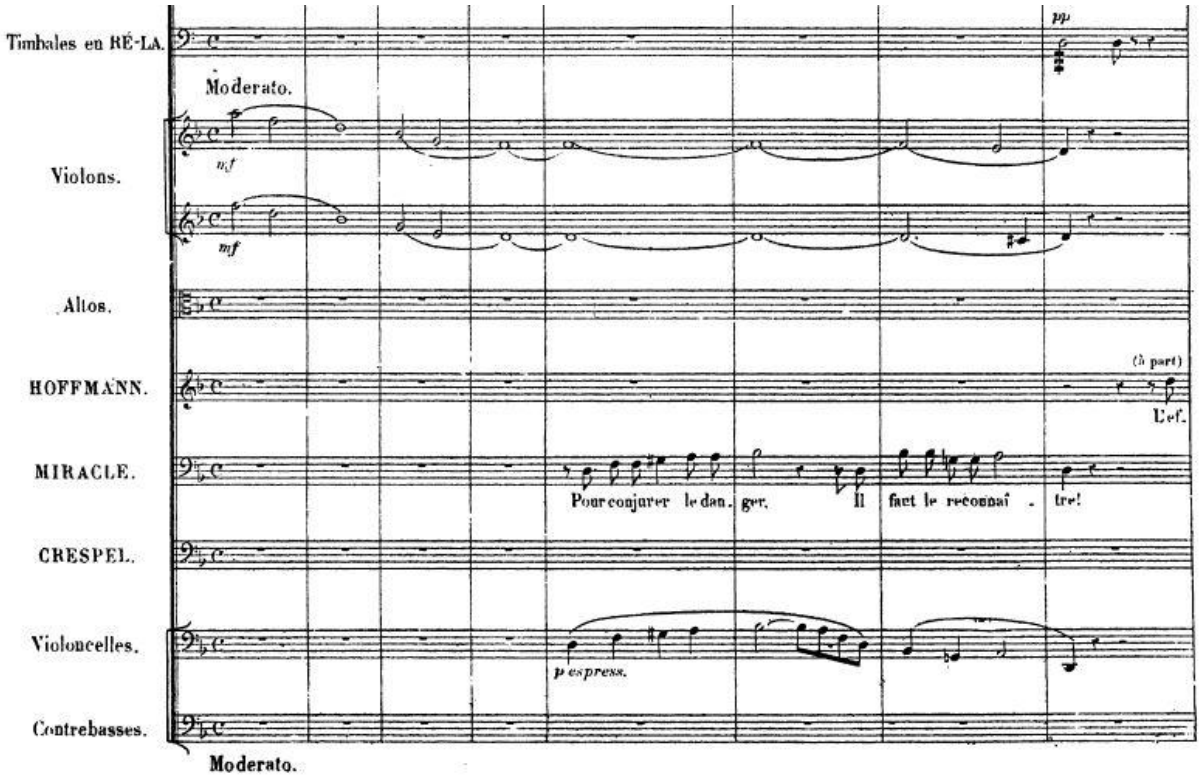
Introduction. Les Technologies de l'Information en Santé (TIS) sont de plus en plus utilisées pour améliorer la qualité des soins et la sécurité du patient. Cependant, certains problèmes d'utilisabilité peuvent amenuiser leur impact et peuvent même induire de nouveaux problèmes dont la mise en danger du patient. Pour éviter ces effets négatifs, il est notamment nécessaire d'améliorer l'utilisabilité des TIS ce qui requiert l'application de connaissances d'utilisabilité éprouvées. Les connaissances en utilisabilité appliquée aux TIS sont rares, éparpillées à travers diverses supports et peu utilisables. Par ailleurs, leur couverture en termes de problèmes d'utilisabilité est peu connue. Ce travail a deux objectifs: (i) participer à l'amélioration de l'accumulation de la connaissance en utilisabilité pour les TIS, (ii) fournir une connaissance structurée sur l'utilisabilité des TIS et dont la couverture est établie. Le domaine d'application est celui des systèmes d'alerte médicamenteux.

Méthode. Deux analyses indépendantes de la littérature ont été menées : d'un côté, identifier et organiser les problèmes d'utilisabilité des systèmes d'alerte médicamenteux ainsi que leurs conséquences ; de l'autre, identifier et synthétiser les principes d'utilisabilité spécifiques à ces systèmes. Les résultats de ces analyses ont été croisés afin de connaitre la couverture desdits principes en termes de problèmes d'utilisabilité.

Résultats. La revue systématique a identifié 13 types de problèmes d'utilisabilité dans les systèmes d'alerte médicamenteux. Les conséquences de ces problèmes sur le clinicien et son système de travail sont variées et ont un grand pouvoir de nuisance (*e.g.*, fatigue, erreur d'interprétation). Au total, 63 principes d'utilisabilité permettent de rendre compte de tous les problèmes d'utilisabilité identifiés. Ils sont organisés en 6 thèmes : améliorer le ratio signal-bruit, être en adéquation avec l'activité des cliniciens, supporter le travail collaboratif, afficher les informations pertinentes, rendre le système transparent et fournir des outils utiles. Le croisement des deux ensembles de données révèle une bonne correspondance entre les principes d'utilisabilité énoncés et les problèmes d'utilisabilité réellement observés.

Discussion. Une liste structurée des principes d'utilisabilité illustrés par des exemples réels de leur violation a été développée à partir de ce travail. Cette liste peut aider les concepteurs et les experts en Facteurs Humains à comprendre et à appliquer les principes d'utilisabilité durant la conception et l'évaluation de systèmes d'alerte médicamenteux. L'utilisabilité appliquée aux TIS est une discipline relativement récente qui souffre d'un déficit de structuration et de capitalisation de ses connaissances. Ce travail montre qu'il est possible d'accumuler et de structurer les données d'utilisabilité des TIS. Ce travail pourrait être poursuivi en développant une base de connaissance en utilisabilité appliquée aux TIS afin de tendre vers une « utilisabilité fondée sur les preuves ».

Table of contents



Dr Miracle: "Pour conjurer le danger, il faut le reconnaître" /

"To ward off the danger, it must be recognized".

Excerpt from the second (sometimes third) act of "The Tales of Hoffmann", opéra fantastique by Jacques Offenbach (libretto by Jules Barbier) based on the short stories by E. T. A. Hoffman.

General Introduction

1. Introduction

Health Information Technology (HIT) is increasingly disseminated and implemented to improve patient safety and healthcare quality and performance. Nonetheless, HIT applications face acceptance issues [1] may be abandoned [2] and/or induce medical errors [3] that may ultimately lead to patient harm or death [4-7]. One major cause of those problems is of ergonomics or Human Factors (HF) nature[1]. Besides organizational ergonomics problems pertaining to wrong implementation choices, the ergonomic quality of HIT, *i.e.*, its (poor) usability is often pointed out [5-7]. Usability is the "extend 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" [8]. When considering a particular category of technology or tool, usability refers to those characteristics of the product that make it easy to use and easy to learn how to use by their intended users. Such characteristics result from the proper implementation of usability design principles during the design and development lifecycle of the product.

To prevent problems originating in usability issues (*e.g.,* use errors), usability must be considered all along the design and evaluation process of HIT. This need is ever more recognized and is now part of the essential requirements that govern the European Conformity (CE) marking of medical devices, that also applies to certain categories of HIT [9], *e.g.,* typically Decision Support Systems(DSS).

Optimizing HIT usability is a core objective of research and practice of the HF team of the Lille Clinical Investigation Center for Innovative Technology[2] (CIC-IT). For this purpose, the User-Centered Design (UCD) process [10] and corresponding usability methods [11;12] are regularly applied to design and to evaluate various HIT systems or applications [13-19]. However, be it in our lab or in other research centers, applying processes and methods is necessary but not sufficient to design usable HIT: there is a need of HF knowledge to design efficiently an HIT [20;21]. In all (re-)design or evaluation project, two types of knowledge are necessary:

- Knowledge of the UCD process and corresponding usability methods;
- Knowledge of usability design principles that apply to the category of product / technology under consideration.

The knowledge of UCD process and usability methods is well established and described in many handbooks, manuals and standards, including a harmonized standard adapting the usability engineering process to prevent the risk of use errors with medical devices [22].

[1] "Ergonomics" and "Human Factors" are synonyms.
[2] http://www.cic-it-lille.com/

The knowledge of usability design principles is unfortunately much less stable, structured, and available. There are several reasons for that situation:

- The technology (including HIT) evolves rapidly; it is difficult to find specific and established usability design principles for truly innovative products.
- The usability design principles need a certain level of generality to cover significant categories of tools. One such category is interactive computer systems, for which user-interface design guidelines (also called ergonomics heuristics or criteria) have been patiently accumulated over the last 40 years.
- Usability design principles are essentially based on the application of fundamental knowledge from (cognitive) psychology to the specificities of the type of tool under consideration *e.g.,* specified users, specified goals, and modalities of human-machine interaction. The principles identify human capabilities and limitations that must be considered when designing this type of tool. But the set of usability design principles relevant for a type of product is also regularly extended or filtered. Those changes are based on observations and reports of usability defects identified in similar products undergoing formal usability evaluations or already put into use (*e.g.,* reports of incidents related to usability).

Accumulation of empirical evidence regarding usability flaws of a given type of product is therefore critical to improve the accuracy and efficiency of sound usability design principles.

As regards HIT, the process of accumulating empirical data on usability flaws and establishing accurate and efficient sets of usability design principles per category of applications or systems is still in its infancy. A few papers are published that provide usability recommendations to design various kinds of HIT. However, recommendations are scattered across various documents (sometimes hard to access) and they are not structured and/or worded so as to make them easy to understand and apply by designers and even by HF experts. Furthermore, the coverage of each set of recommendations is rarely described. Nonetheless, there are also some promising initiatives to provide lists of HIT usability principles for specific HIT and medical devices with a precisely defined coverage (*e.g.,* FDA's [23] and NHS's guidelines [24]). However, those initiatives are essentially based on experts' consensus without any mention of their empirical support in terms of actual usability problems.

A significant source of empirical usability data rests on the results of usability studies. Those data are still seldom shared across the HF community. When HF researchers publish their results, usability data lack of comprehensiveness and precision [25]. Descriptions of usability issues are often entangled with other related issues. For instance, in usability studies' publications, usability problems noticed in the technology are often described together with their consequences for the user and / or the work system without clear distinction.

Altogether, those barriers prevent HF experts from using optimally existing HIT usability knowledge during HIT usability projects: concretely, for each project, HF experts waste time searching and understanding this knowledge. At a larger scale, those reasons prevent the HF community from accumulating HIT usability knowledge and they slow down the progress of usability researches on HIT.

The aim of the present PhD is two-fold:

- To participate in improving the accumulation of usability knowledge for HIT;
- To provide synthetic structured easy-to-use HIT usability knowledge with a clear coverage.

Those two topics are closely intertwined and will be addressed together. Addressing both topics implies questioning methods that are currently used to cumulate usability data along with methods used to report them.

Accumulating empirical usability knowledge requires being able to target only pure usability data. This implies to know precisely what HIT usability refers to. To define it precisely we developed a "usability framework" that structures concepts around usability. It is presented in details in chapter 1. In this framework, the usability cause (what is directly related to the usability of the technology) is clearly distinguished from their consequences. It allows considering the usability of a technology from various perspectives: the technology itself, its user and the work system it is implemented in. Four usability concepts are structured together that allow considering usability from various perspectives (Figure 1):

- Usability design principles, *i.e.*, recommendations in terms of usability for the design of the technology.
- Usability flaws, *i.e.*, violations of usability design principles; they are descriptions at the level of the technology.
- Usage problems, *i.e.*, how the user is experiencing the usability flaws while interacting with the technology.
- Negative outcomes, *i.e.*, impact of the usability flaws and corresponding usage problems on the work system and its performance, including the patient, the workflow, the technology etc.

Looking for the consequences of the violations of usability design principles to develop usability knowledge may be seen as similar to looking for evidence-based usability. The concept of evidence in medicine comes from Sackett's works [26]. This author defines evidence-based medicine as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of the individual patients. It means integrating individual clinical expertise with the best available external clinical evidence from systematic research." This concept has been extended to software design [27] and to health informatics [28] amongst other fields. By analogy, evidence-based usability could be defined as "the conscientious, explicit and judicious use of current best evidence to support decision with regard to the design, development and evaluation of technologies in terms of usability". Usability data come mainly

from observational studies (mostly case reports): those methods are considered as providing a lower quality of evidence compared to systematic review of randomized clinical trials (RCT) [29]. Even if usability methods do not provide the "current best evidence" possible, it is the best evidence currently available in usability field because it is not possible to perform RCT to evaluate the usability of HIT. Besides, another difference between "evidence-based medicine" and "evidence-based usability" rests on the available corpus of data. While studies of medication or technology evaluation may report just as well results in favor of or unfavorable towards using the medicament or technology, usability studies mainly report negative usability aspects of technology.

In order to develop the usability knowledge, we propose to gather usability data reported in the literature by performing a systematic review of usability observational studies. This searching process focuses on the elements of the "usability framework" that are directly related to usability and that are less likely to be influenced by theoretical considerations: the usability flaws. Usability design principles are also directly related to usability but there wording may easily be impacted by their authors' theoretical background. Then, the negative consequences of the usability flaws for the user of the technology and for the work system have to be searched. For this purpose, a secondary analysis of the papers included in the systematic review process has to be performed. Finally, once the associations between usability flaws and their consequences for the user and the work system are identified, the related usability design principles must be listed and be matched with the usability flaws.

The outcome of this process will be a list of usability mistakes not to make completed with their known consequences and the related usability design principles to fix them.

The proposed approach has been applied to a specific type of HIT. We focus on decision support systems (DSS). However, DSS embrace a very large range of systems (*e.g.,* alerting systems, pathway, order sets) [30]. For consistency sake, we focus on a specific type of DSS, namely medication alerting systems. Those systems display in real-time an appropriate clinical or pharmaceutical knowledge at the point of decision to help clinicians make informed decision. Those systems are supposed to "achieve large gains in performance, [to] narrow gaps between knowledge and practice, and [to] improve safety" [31]. Indeed, alerting systems help improve providers' performance with drug ordering [32]. There is also evidence that CPOE augmented with such DSS enhance healthcare quality and safety [33], even more so when advanced decision support functions are available [34]. However, other studies reveal that their intended positive impact is not always achieved [35; 36]. It is also noticed that they face acceptance and usage problems [37-39]. A poor usability is a well-known cause of those issues [31;40]. There are long debates about the usefulness and acceptance of those systems that materialize in a large amount of publications ensuring material for the systematic review.

In order to ensure that the data considered in the systematic review are of good validity, the methods that have been used to uncover them are investigated. One study is dedicated to exploring whether the method used has an impact on the usability issues that are reported. Moreover, a commonly reported issue during the collection of usability data is the lack of comprehensiveness and organization of the usability

evaluations reports. In order to improve the accumulation of usability knowledge for HIT, a usability evaluation reporting form based on the "usability framework" was developed and a preliminary evaluation of its perceived usefulness has been performed.

2. Research questions and outlines of this thesis

This PhD work addresses 7 questions:

- What are the usability flaws of medication alerting functions identified in published studies? (addressed in chapter 2)
- What types of usage problems and negative outcomes originating in identified usability flaws are reported in usability studies of medication alerting functions? (addressed in chapter 3)
- What are the cause-consequence links between usability flaws, usage problems and negative outcomes reported in medication alerting functions? (addressed in chapter 3)
- What are the usability design principles specific to medication alerting systems reported in literature? (addressed in chapter 4)
- How well usability flaws reported in the literature are matched with the usability design principles? (addressed in chapter 4)
- What are the methods used that detect facts on usability in medication alerting systems? What type of usability issues those methods allow to report? (addressed in chapter 5)
- How useful is a usability reporting form that distinguishes clearly usability design principles, usability flaws, usage problems and negative outcomes? (addressed in chapter 6)

Those questions are explored from chapter 2. Chapters are organized according to the "usability framework" described in details in chapter 1. Chapter 2 to 4 constitutes the core section of the PhD work: gathering and structuring the usability knowledge about medication alerting systems. Figure 1 represents schematically the structure of those chapters. Chapter 5 and 6 address the question of the methods usually applied to uncover usability issues in medication alerting systems along with the question of the usefulness of the "usability framework" to report usability data in a structured way. Since those questions do not represent the core of the PhD work, they are not presented in Figure 1.

Chapter 1 introduces and explores the "usability framework" that is proposed to gather various types of usability data. The consequences of this framework in terms of methods for data collection and analysis are discussed. First considerations related to a systematic review on the usability flaws in medication alerting functions are presented.

Chapter 2 presents the first application of the "usability framework" in a systematic review on the type of usability flaws reported in medication alerting systems. Identified instances of usability flaws are

categorized (a) first through general usability flaws categories and (b) then through sub-categories of usability flaws specific to alerting systems.

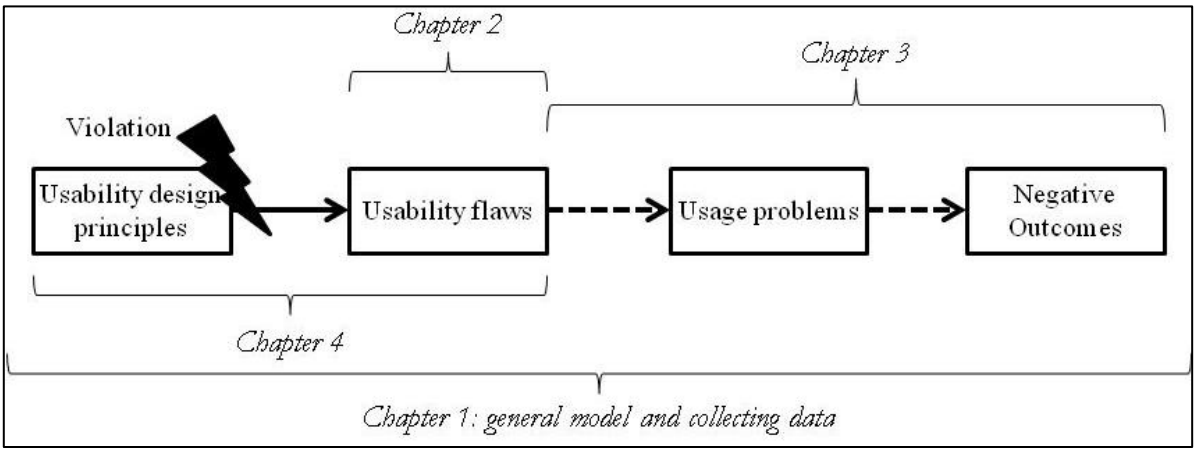


Figure 1. Schematic representation of the core chapters of the thesis according to the "usability framework" that described the origin of usability flaws in the violation of usability design principles and the propagation of their consequences to the user ("usage problem") and to the work system ("negative outcomes"), adapted from [41]. The dotted lines represent the potential propagation of the usability flaws.

Chapter 3 presents how the consequences of usability flaws identified in chapter 2 were searched. Both usage problems and negative outcomes were sought along with their associations with categories of usability flaws. The categorization highlights that those consequences deal with different areas of user's experience and of the work system.

Chapter 4 presents how usability desgn principles dedicated to medication alerting systems reported in the literature were identified and synthesized and how the usability flaws identified in chapter 2 were matched with those principles.

Chapter 5 provides feedbacks on the application of usability methods to uncover usability flaws in medication alerting systems. Attention is paid to the types of usability issues uncovered.

Chapter 6 presents the preliminary evaluation of the reporting form adapted from the "usability framework". This form aims (a) at helping manufacturers understand the "usability-induced use-errors" concept and (b) at allowing an easy collection of inputted data to analyze the consequences of usability design principles violations.

References

[1] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004 Mar;11(2):104-12.

[2] Kaplan B, Harris-Salamone KD. Health IT success and failure: recommendations from literature and an AMIA workshop. J Am Med Inform Assoc 2009 May;16(3):291-9.

[3] Kushniruk A, Triola M, Stein B, Borycki E, Kannry J. The relationship of usability to medical error: an evaluation of errors associated with usability problems in the use of a handheld application for prescribing medications. Stud Health Technol Inform 2004;107(Pt 2):1073-6.

[4] Han YY, Carcillo JA, Venkataraman ST, Clark RS, Watson RS, Nguyen TC, et al. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. Pediatrics 2005 Dec;116(6):1506-12.

[5] Magrabi F, Ong MS, Runciman W, Coiera E. An analysis of computer-related patient safety incidents to inform the development of a classification. J Am Med Inform Assoc 2010 Nov;17(6):663-70.

[6] Magrabi F, Ong MS, Runciman W, Coiera E. Using FDA reports to inform a classification for health information technology safety problems. J Am Med Inform Assoc 2012 Jan;19(1):45-53.

[7] Samaranayake NR, Cheung ST, Chui WC, Cheung BM. Technology-related medication errors in a tertiary hospital: a 5-year analysis of reported medication incidents. Int J Med Inform 2012 Dec;81(12):828-33.

[8] International Standardization Organization. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability (Rep N° 9241-11). Geneva: International Standardization Organization; 1998.

[9] Council Directive 2007/47/EC, 2007:247:0021:0055, European Parliament Council, (2007).

[10] International Standardization Organization. Ergonomics of human system interaction - Part 210: Human centered design for interactive systems (Rep N°9241-210). Geneva: International Standardization Organization; 2010.

[11] Beuscart-Zephir MC, Elkin P, Pelayo S, Beuscart R. The human factors engineering approach to biomedical informatics projects: state of the art, results, benefits and challenges. Yearb Med Inform 2007;109-27.

[12] International Standardization Organization. Ergonomics of human-system interaction - Usability methods supporting human-centred design (Rep N° 16982). Geneva: International Standardization Organization; 2002.

[13] Beuscart-Zephir MC, Pelayo S, Degoulet P, Anceaux F, Guerlinger S, Meaux JJ. A usability study of CPOE's medication administration functions: impact on physician-nurse cooperation. Stud Health Technol Inform 2004;107(Pt 2):1018-22.

[14] Beuscart-Zephir MC, Anceaux F, Menu H, Guerlinger S, Watbled L, Evrard F. User-centred, multidimensional assessment method of Clinical Information Systems: a case-study in anaesthesiology. Int J Med Inform 2005 Mar;74(2-4):179-89.

[15] Beuscart-Zephir MC, Pelayo S, Bernonville S. Example of a Human Factors Engineering approach to a medication administration work system: potential impact on patient safety. Int J Med Inform 2010 Apr;79(4):e43-e57.

[16] Hackl WO, Ammenwerth E, Marcilly R, Chazard E, Luyckx M, Leurs P, et al. Clinical evaluation of the ADE scorecards as a decision support tool for adverse drug event analysis and medication safety management. Br J Clin Pharmacol 2013 Sep;76 Suppl 1:78-90.

[17] Marcilly R, Leroy N, Luyckx M, Pelayo S, Riccioli C, Beuscart-Zephir MC. Medication related computerized decision support system (CDSS): make it a clinicians' partner! Stud Health Technol Inform 2011;166:84-94.

[18] Marcilly R, Bernonville S, Riccioli C, Beuscart-Zephir MC. Patient safety-oriented usability testing: a pilot study. Stud Health Technol Inform 2012;180:368-72.

[19] Pelayo S, Bras Da CS, Leroy N, Loiseau S, MacKeon D, Trancard D, et al. Application of the medical device directive to software: methodological challenges. Stud Health Technol Inform 2013;192:437-41.

[20] Russ AL, Zillich AJ, Melton BL, Russell SA, Chen S, Spina JR, et al. Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation. J Am Med Inform Assoc 2014 Mar 25.

[21] Tsopra R, Jais JP, Venot A, Duclos C. Comparison of two kinds of interface, based on guided navigation or usability principles, for improving the adoption of computerized decision support systems: application to the prescription of antibiotics. J Am Med Inform Assoc 2014 Feb;21(e1):e107-e116.

[22] International Electrotechnical Commision. Medical devices - Application of usability engineering to medical devices (Rep N° 62366). Geneva: International Standardization Organization; 2007.

[23] http://www fda gov/MedicalDevices/DeviceRegulationandGuidance/HumanFactors/ucm124829 htm 2014 June 28

[24] http://www mscui net/ 2014 June 28

[25] Peute LW, Spithoven R, Bakker PJ, Jaspers MW. Usability studies on interactive health information systems; where do we stand? Stud Health Technol Inform 2008;136:327-32.

[26] Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996 Jan 13;312(7023):71-2.

[27] Kitchenham BA, Dybå T, Jørgensen M. Evidence-based Software Engineering. ICSE '04 Proceedings of the 26th International Conference on Software Engineering 2004;273-81.

[28] Ammenwerth E. Evidence based health informatics. Stud Health Technol Inform 2010;151:427-34.

[29] Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. JAMA 1995 Dec 13;274(22):1800-4.

[30] Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. J Biomed Inform 2012 Dec;45(6):1202-16.

[31] Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, et al. Ten commendments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003 Nov;10(6):523-30.

[32] Jaspers MW, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. J Am Med Inform Assoc 2011 May 1;18(3):327-34.

[33] Gandhi TK, Weingart SN, Seger AC, Borus J, Burdick E, Poon EG, et al. Outpatient prescribing errors and the impact of computerized prescribing. J Gen Intern Med 2005 Sep;20(9):837-41.

[34] Ammenwerth E, Schnell-Inderst P, Machan C, Siebert U. The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. J Am Med Inform Assoc 2008 Sep;15(5):585-600.

[35] Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. JAMA 1998 Oct 21;280(15):1339-46.

[36] Ranji SR, Rennke S, Wachter RM. Computerised provider order entry combined with clinical decision support systems to improve medication safety: a narrative review. BMJ Qual Saf 2014 Apr 12.

[37] Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. AMIA Annu Symp Proc 2007;26-30.

[38] Kuperman GJ, Bobb A, Payne TH, Avery AJ, Gandhi TK, Burns G, et al. Medication-related clinical decision support in computerized provider order entry systems: a review. J Am Med Inform Assoc 2007 Jan;14(1):29-40.

[39] van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006 Mar;13(2):138-47.

[40] Seidling HM, Phansalkar S, Seger DL, Paterno MD, Shaykevich S, Haefeli WE, et al. Factors influencing alert acceptance: a novel approach for predicting the success of clinical decision support. J Am Med Inform Assoc 2011 Jul;18(4):479-84.

[41] Marcilly R, Beuscart-Zephir MC, Ammenwerth E, Pelayo S. Seeking Evidence to Support Usability Principles for Medication-Related Clinical Decision Support (CDS) Functions. Stud Health Technol Inform 2013;192:427-31.

Chapter 1. Seeking evidence to support usability principles for medication-related CDS functions

Romaric Marcilly, Marie-Catherine Beuscart-Zéphir, Elske Ammenwerth & Sylvia Pelayo

Published in MedInfo 2013 conference proceedings

Nominated for distinguished paper

Abstract

There is a need for evidence-based usability principles to support the design of usable medication-related computerized Clinical Decision Support (CDS) functions and systems. Such evidence requires establishing scientific relationships between usability principles, their violation in terms of usability flaws, issuing usage problems and their consequences or outcomes in the clinical work and patient care. This kind of evidence is not currently directly available in scientific evaluation studies of medication CDS functions. A possible proxy to seek evidence is systematic review of existing scientific evaluation reports. We rely on a four-stage framework describing the chain of consequences and inferences linking usability principles to clinical outcomes to design the systematic review methodology and interpretation principles. This paper describes the four-stage framework and the resulting consequences for the systematic review design.

Keywords

Usability; Human Engineering; Decision support systems; Evaluation

1.2. Introduction

Medication computerized Clinical Decision Support (CDS) functions have been shown to have a positive impact on patient safety by improving prescribing practices [1] and reducing Adverse Drug Events [2]. However, they remain difficult to implement and face acceptance problems [3]. Moreover, they may also generate technology induced errors, *i.e.,* latent types of errors related to the usage of this technology [4]. The root causes of such problems are usually of Human Factors (HF) nature and more specifically related to usability features. However, the evidence of the relationship between the observed/reported outcomes, the usage problems experienced by the users, the usability flaws involved in the usage problems and the usability principle whose violation led to the usability flaws remains unsubstantiated most of the time.

This paper presents a four-stage framework describing the chain of consequences and inferences linking usability principles to clinical outcomes. The consequences of the framework in terms of method and interpretive expertise are discussed and applied to the design of a systematic review of usability evaluation studies of medication related CDS functions.

1.3. Background

Figure 2 describes the proposed four-stage framework putting together usability principles, usability flaws, usage problems and outcomes in the work system. It can be applied to any kind of Healthcare Information Technology (HIT) system. Two processes connect the four stages. The top-down process describes the propagation of usability flaws until they finally impact the healthcare system's performance through deteriorated outcomes, *e.g.,* medical errors. The bottom-up process describes the research, evaluation and expert consensus process which allows (i) identifying and characterizing actual usage problems and (ii) infer from them elements of evidence to support usability principles for corresponding HIT systems.

Usability flaws of technical systems such as Computerized Prescriber Order Entry (CPOE), CDS functions or medical devices result from violations of usability principles in the design of those systems. Most of those usability flaws create usage problems when the system is put into use. The importance of the usage problems experienced by the users depends on several variables (*e.g.,* the nature of the usability feature violated or the type of task supported by the faulty function). Ultimately, these usage problems may actually negatively impact the healthcare system performance, *e.g.,* by slowing down the clinical workflows or generating medical errors characterized as technology-induced errors [4]. Again, the scope and importance of the negative outcomes depend on several variables, *e.g.,* the severity of the usage problems but also characteristics of the context of use and of the socio-technical organization in which the system has been implemented along with the capacity of adaptation ingrained in the work system. The bottom-up process aims at linking the outcomes identified by evaluation and impact studies back to usage problems and their usability root causes (usability flaws) and ultimately to the corresponding usability

principles that have not been complied with. Given the multiplicity of other HF or technical factors that are intertwined with usability variables in the described top-down process, this inference work is far from trivial and requires a sound expertise in HF and usability of HIT applications. Moreover, most of impact studies such as clinical trials of CPOE or CDS functions lack qualitative analysis that would allow identifying the usage patterns of the HIT system acting as intermediate variables explaining the observed outcomes. Qualitative HIT evaluation studies may prove more informative but still require HF expertise in the analysis of work systems to differentiate organizational vs. usability issues, given that the report of the evaluation study provides enough details to make this inference. Usability studies aiming specifically at identifying usage problems often fail linking the observed usage problems with causing usability flaws [5]. Such studies require usability expertise in HIT systems to properly infer the usability flaws. Finally, usability studies listing usability flaws of evaluated systems do not always provide the necessary level of details (*e.g.,* screenshots) to establish clear links with violated usability principles and descriptions of potential related usage problems.

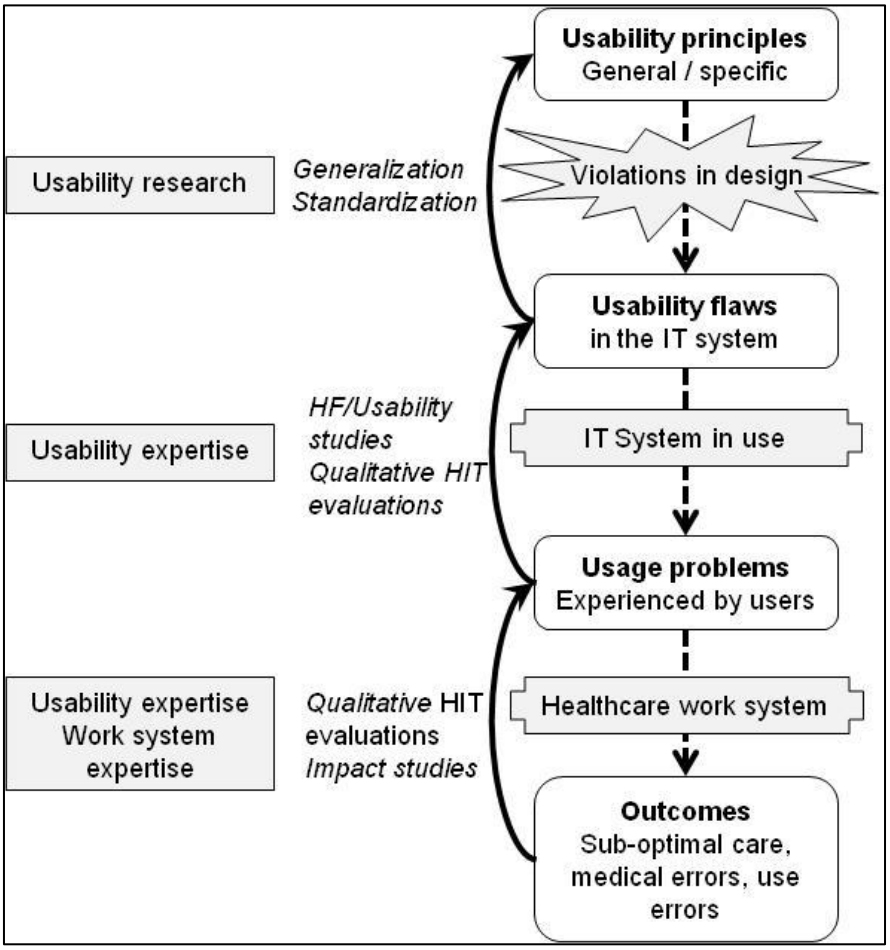


Figure 2. Emergence and potential propagation (dotted lines) of usability flaws to healthcare outcomes and seeking evidence for usability principles.

Usability principles are the core part of the framework. They can be divided into two categories: (1) methodological principles to apply the user centered design or the usability engineering process [6, 7] and (2) usability principles and features of the targeted products [8, 9]. This paper deals only with the latter category. The most important part of those principles is reported in standards elaborated on the basis of international expert consensus. They may be relatively general (*e.g.,* usability principles for the design of Graphical User Interface (GUI) [10] that would apply to all HIT applications) or they may be more specific to a category of product (*e.g.,* medical devices of a certain kind). Unfortunately, most standards face several problems that prevent usability principles from an easy and unambiguous interpretation by non-experts [11].

During the last decade, there have been several initiatives to identify the most important usability principles for medication CDS systems. These attempts are mostly based on the experience of the authors in the domain and on lessons learned from medication CDS functions design and implementation projects in which they participated [12-15], or derived from a specific theoretical approach of the cognitive processes involved in the interaction of the users with medication CDS functions [16].

1.4. Rationale

The great variety of usability principles, whether recommended by standards or scientific publications, and their lack of comprehensible organization prevent developers, and even HF experts, from identifying those they should apply, and therefore from applying them completely and correctly. Moreover, the lack of evidence to support usability principles may also lead systems' developers to question the legitimacy of the stated principles. We aim at seeking evidence supporting the organization and prioritizing of those usability principles.

The systematic review method is a technique allowing the emergence of evidences from HIT evaluations' published reports. As far as we know, it has been used only once in the field of medical management systems to find evidence of usability flaws in CPOE systems [17]. In order to seek evidence to support usability principles for medication CDS functions, we designed a systematic review based on the four-stage framework. This review proposes to answer two main questions:

- What features are characterizing medication CDSS usability?
- Do those features generate usage problems of the CDS function and ultimately outcomes in the work system?

Applying this framework allows identifying precisely the inferences necessary to jump from one stage of the framework to another. It also supports the search query through the delimitation of the scope of relevant evaluation studies, the definition of inclusion/exclusion criteria and the design of the interpretive grid for the analysis of final set of papers.

1.5. Systematic review design and process

The design of the systematic review follows as far as possible good practice recommendations [18-20]. The key concepts involved in the review, "medication CDS functions" and "usability", have been defined in the light of the framework. This supports the latter definition of inclusion or exclusion criteria of the papers.

1.5.1. Concepts definition

1.5.1.1. Usability

Usability is defined as "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" [21]. Usability goes way beyond the features of the GUI (*e.g.,* legibility of the texts, layout and prompting of information and tools), and deals more generally with the fitting between the system behavior and its users' needs. Considering medication CDS functions, one of the most important usability features which violations may engender important usage problems and negative outcomes is the compatibility of the system with clinicians' activities, essentially of cognitive nature [16]. For the review (mostly the papers analysis phase) we consider four dimensions in the usability concepts:

- HCI characteristics;
- How the system responds to users' actions;
- Organization or accuracy of the knowledge incorporated;
- Availability of functions required to support users' tasks, especially of cognitive nature.

Therefore, the analysis of retrieved papers, both for decision upon inclusion or exclusion and for final systematic analysis, requires a deep HF knowledge of intended users' needs, activities and working procedures.

1.5.1.2. Medication-related CDS functions

CDS functions refer to a very wide range of tools: documentation forms-templates, relevant data presentation, order-prescription creation facilitators (*e.g.,* order sets), protocol or pathway support, reference information and guidance, alerts and reminders (pushed or pulled) [16].

As the review focuses on HIT tools, medical devices in which medication CDS functions have been integrated are excluded from the analysis (*e.g.,* auto-injectors pens, pumps). Only medication CDS software used in hospital or general practice in the internal medicine field is considered. To increase results' homogeneity, we focus on alerting systems (alerts and reminders). As a result, the review includes studies of software supporting the management of e-prescriptions by physicians, pharmacists and nurses. Thus,

CDS functions integrated in Bar Coded Medication Administration and e-Medication Administration Records are excluded from the review.

1.5.2. Consequences of the framework on the systematic review

1.5.2.1. Queries definition

The literature was searched from 1980 to 2012 using PubMed, Scopus and Ergonomics Abstracts databases. Two semantic groups of key terms were constructed with the support of medical terminologies experts: terms related to CDS functions, alerts and CPOE and terms related to HF (cf. Table 1). As far as possible, MeSH terms were chosen for PubMed database thanks to the Health Multi-Terminology Portal [22]. The key terms have been slightly adapted for Scopus and Ergonomics Abstracts databases. Key terms in each group were combined with the operator "OR". Then both groups were combined using the operator "AND". Queries have been performed in March 2012 and updated on the 26th October 2012. The search retrieved a total of 5862 items.

Table 1 - Terms included in the queries according to their semantic category and to the queried database.

Databases	Terms related to CDSS, alerts & CPOE	Terms related to Human Factors
PubMed	"Medical order entry systems" ; "Medication alert system"; "Computerized physician order entry system"; "CPOE"; "Decision Support Systems, Clinical"; "Clinical decision support systems"; "CDSS"	"User-computer interface"; "Human engineering"; "Risk factors"; "Usability"; "Humans"
Scopus & Ergonomics Abstracts	"Medical order entry"; "Medication alert"; "Computerized physician order entry"; "CPOE"; "Clinical decision support"; "CDSS"	"User-computer interface"; "Human engineering"; "Risk factor"; "Human factor"; "Usability"; "Human-computer interaction"

1.5.2.2. Inclusion and exclusion criteria

Only studies published in English/French peer-reviewed journals and conference proceedings were considered. To be included the studies must report:

- The application of usability methods or of other qualitative methods aiming at evaluating CDS function(s) to report facts (not opinions) on usability flaws and usage problems;
- To specific standalone or integrated medication CDS functions. Functions that are not specifically dedicated to medications (*e.g.,* care protocols/pathways) were included if at least one medication related feature was available.

Papers reporting evaluation of several systems without distinguishing results according to the systems were excluded.

1.5.2.3. Selection and analysis of studies according to the evaluation methods applied

According to the four-stage framework, three kinds of data may be used to formalize usability principles: usability flaws, usage problems and their outcomes. These three kinds of data are retrievable through different sorts of methods.

Questionnaire and interview or focus group methods are usability methods [23]. However, they are often used to gather users' opinions about the system (perceived usability). Usually, such methods retrieve mostly feelings about a system, rarely usability facts. Therefore, studies based on those methods may be included only if they question explicitly specific usability features of the system to detect usability flaws or usage problems.

Usability flaws may be detected by classical usability evaluation methods resting on the standardized analysis of the system. During such evaluations, experts identify usability flaws by reference to a heuristic or to their knowledge about optimal human-machine interactions and to their knowledge of intended users work systems and procedures. Most known methods are expert evaluation and cognitive walkthrough. In those methods, usability flaws detection rests on hypothesizes about problems specific users may encounter.

Usage problems may be discovered only by making intended users actually use the system under evaluation. The most used methods are user testing with thinking-aloud and post-implementation surveillance. The former method aims at observing representative end-users using the test product in a simulated environment to identify usability flaws and rooms for improvement. The goal of the latter one is the same but either the users are observed during their actual use of the sys-tem once implemented in their work system, or users report by themselves usage problems they encounter.

Finally, the detection of the outcomes rests mainly on the socio-technical approach that proceeds by observation of the actual use of the system in the work system and by interviews of actual users of the system. To a lesser extent some outcomes can be detected in the results of impact evaluation studies including qualitative description of the system. Outcomes extracted by both methods may be reported with the usage problems that contribute to their appearance. There is often a need of usability expert inference to link them up together.

The identification of the methods resulting from the framework supports the process of papers' inclusion. Since the aim of this systematic review is to link up detected usability flaws, usage problems and their potential outcomes in the work system, papers relating the application of the aforementioned methods should be included. Nonetheless, data that can be extracted from each kind of papers are not the same and the analyses performed on them do not require the same type of skills and contextual information nor the same analyses to perform (Table 2). For instance, to analyze data issued from a socio-technical evaluation, the description of the work system in which the system is implemented must be retrieved while this information is not necessary to analyze data issued from an expert evaluation.

Therefore, the framework has also consequences on the construction of the grid through which each included paper is analyzed.

1.6. Systematic review process and analysis

Publications' relevance was screened by one author (RM) to exclude doubloons, posters and non-peer-reviewed papers. After a training session on 77 papers, two authors (RM & MCBZ) reviewed independently 471 papers' titles (Cohen's $\varkappa = 0.66$); remaining papers were shared-out for screening. Decision of inclusion or exclusion was based on the aforementioned criteria. If in doubt, the paper was included in the next step. The same process was applied for the screening at the abstracts (training on 44 papers, independent review on 73 papers, Cohen's $\varkappa = 0.69$; sharing-out remaining papers) and at the whole papers (training on 20 papers, independent review on 20 papers, Fleiss' $\varkappa = 0.95$; sharing-out remaining papers). The latter step involved three authors (RM, MCBZ & SP). 27 papers were eventually included in the analysis process.

In each paper, descriptions of usability flaws, usage problems and outcomes were extracted and categorized by two authors (RM & MCBZ). A content analysis is used to identify classes of usability flaws related to medication CDS and their reported links with usage problems and outcomes in the work system.

Table 2 - Methods applied, data to retrieve, analyses to perform on the data and requisites for the analyst and the analyzed papers according to the three stages of the framework that support the inference of elements of evidence to support usability principles.

	Usability flaws	**Usage problems**	**Outcomes**
Methods	Expert evaluation & cognitive walkthrough	User testing & post implementation surveillance	Socio-technical & impact evaluation methods whose results may be reinterpreted in terms of usability
Retrievable data	Description of usability flaws; hypotheses about their consequences	Description of usage problems; facts/hypotheses about their causes and their consequences in the work system	Description of facts consecutive to the implementation of the system; hypotheses about their causes
Analysis to perform	Organization of the usability flaws through a common frame	Organization of the usage problems and inference on their linkage to usability problems and outcomes	Organization of the outcomes and inference on their linkage to usage problems
Requisites for the analyst	Usability expertise	Usability expertise; Expertise of the work process	Usability expertise; Expertise of the work process
Requisites for the analyzed paper	Description of the system under evaluation	Description of the system under evaluation and of the work system in which the system will be/is implemented	Description of the system under evaluation and of the work system in which the system is implemented

1.7. Discussion

This paper describes a general framework linking up usability principles, usability flaws, usage problems and their potential outcomes in the work system. It presents also the application of this framework to the design of a systematic review on the usability features of medication CDS functions. The selection process of the systematic review is finished and will be reported along with the results in another paper.

First comments about the advantages of the application of the framework can be drawn. The framework has consequences on the systematic review design at two levels. First, it supports the selection process by facilitating the identification of the kinds of methods the selected papers must apply. Second, it allows developing the final analysis grid through with each included paper is analyzed. As compared to previous similar work [17], the framework's added-value is that it allows establishing relations between usability flaws, usage problems and outcomes.

The analysis process is under progress. For now, only few papers have been identified that links up usability flaws, usage problems and their outcomes. Therefore, inferences have to be drawn to link up the retrieved usability flaws, usage problems, and their outcomes. By providing an architecture that articulates the different kinds of data, this framework allows reviewers be aware of those inferences. In sum, the review supported by the framework is informed and requires usability and work system expertise to draw inferences.

As it is noticed in Table 2, linking up usability, usage and outcomes results requires contextual information (*e.g.,* description of the work system and of the system under evaluation). Yet, in most of published papers this information is missing or ambiguous: most often the CDS functions is not clearly described, and the work system in which it is implemented is not described at all.

During the data analysis, an unexpected difficulty arose related to the level of description of the usability/usage/outcome results. Indeed, this level greatly varies from one paper to another: papers report either raw descriptions of unique usability flaws, or categories of similar usability flaws. This lack of homogeneity makes it difficult (1) to compare findings from a paper to another and (2) to distinguish usability flaws from usage problems. Moreover, in most publications, place is limited forcing authors to report only a subset of issues that they found in their study. This makes quantitative comparison between studies or groups of studies impossible; only a qualitative synthesis can be performed.

To support more complete reports of usability or qualitative impact evaluation, a consideration on the elements to report for HF studies (especially the complete report of usability flaws or usage problems and system description in on-line appendices) and on the way to report them has been engaged [24] as it has been done for general HIT evaluations [25].

1.8. Conclusion

The proposed framework supports performing an informed systematic review in which drawn inferences and evidences are highlighted. It is used to find evidence to organize medication CDS functions' usability principles but it could be used to organize other usability principles requiring evidence.

References

[1] Kawamoto K, Houlihan CA, Balas EA, & Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005: 330:765.

[2] Ammenwerth E, Schnell-Inderst P, Machan C, & Siebert U. The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. J Am Med Inform Assoc 2008: 15(5):585-600.

[3] Ash JS, Berg M, & Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004: 11(2):104-12.

[4] Kushniruk AW, Triola MM, Borycki EM, Stein B, & Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. Int J Med Inform 2005: 74(7-8):519-26.

[5] Beuscart-Zéphir MC, Pelayo S, Borycki E, & Kushniruk A. Human Factors Considerations in Health IT Design and Development. 2d ed, Carayon P ed. CRC Press, 2011: 649–670

[6] Association for the Advancement of Medical Instrumentation. Human factors design process for medical devices (ANSI/AAMI HE74). Arlington, VA: AAMI, 2001.

[7] International Standardization Organization. Ergonomics of human system interaction - Part 210: Human centered design for interactive systems (Rep N°9241-210). Geneva, International Standardization Organization, 2010.

[8] International Standardization Organization. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 12: Presentation of information (ISO 9241-12). Geneva, International Standardization Organization, 1998.

[9] Association for the Advancement of Medical Instrumentation. Human factors engineering-design of medical devices (ANSI/AAMI HE75). Arlington, VA: AAMI, 2009.

[10] Nielsen J. Usability engineering. London, Academic Press, 1993.

[11] Bras da Costa S, Pelayo S, Bastien JMC, & Beuscart-Zéphir MC. Issues in the implementation of the ISO 62366:2007 standard for medical devices' usability and safety: a case study. Proceedings of the International Conference of Healthcare Systems Ergonomics and Patient Safety. Spain, CRC Press, 2011.

[12] Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, Spurr Cynthia, Khorasani R, Tanasijevic M, & Middleton B. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003:10 (6):523-30.

[13] Phansalkar S, Edworthy J, Hellier E, Seger DL, Schedlbauer A, Avery AJ, & Bates D. A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. J Am Med Inform Assoc 2010: 17 (5): 493-501.

[14] Sittig DF, Wright A, Osheroff JA, Middleton B, Teich JM, Ash JS, Campbell E, & Bates DW. Grand challenges in clinical decision support. J Biomed Inform 2008: 41(2): 387-92.

[15] Pelayo S, Marcilly R, Bernonville S, Leroy N, & Beuscart-Zéphir MC. Human factors based recommendations for the design of medication related clinical decision support systems (CDSS). Stud Health Technol Inform 2011: 169: 412-4066.

[16] Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, & Middleton B. Interface design principles for usable decision support: A targeted review of best practices for clinical prescribing interventions. J Biomed Inform 2012: 45 (6): 1202-1216.

[17] Khajouei R, & Jaspers MW. The impact of CPOE medication systems' design aspects on usability, workflow and medication orders: a systematic review. Methods Inf Med 2010: 49(1):3-19.

[18] Cochrane handbook for systematic reviews for interventions. Higgins JPT, Green S, eds. England, Wiley-Blackwell, Cochrane Book Series, 2008.

[19] Systematic reviews: CRD's guidance for undertaking re-views in health care. University of York, Center of reviews and Dissemination, 2008.

[20] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, & Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339-b2700.

[21] International Standardization Organization. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability (ISO 9241-11). Geneva, International Standardization Organization. 1998.

[22] CISMeF. Health Multi-Terminology Portal. Available from URL: http://pts chu-rouen fr/index html?lang=en [cited 2012 Oct 26]

[23] International Standardization Organization. Ergonomics of human-system interaction -- Usability methods supporting human-centered design (ISO/TR 16982:2002). Geneva, International Standardization Organization. 2002

[24] Peute LW, Driest KF, Marcilly R, Bras Da Costa S, Beuscart-Zéphir MC & Jaspers MWM. A framework for reporting on Human Factor/Usability studies of Health In-formation Technologies. Context Sensitive Health Informatics conference. Copenhagen, 2013 (accepted).

[25] Talmon J, Ammenwerth -E, Brender J, de Keizer N, Nykänen P, Rigby M. STARE-HI--Statement on reporting of evaluation studies in Health Informatics. Int J Med In-form 2009: 78(1): 1-9

Chapter 2. Usability flaws of medication-related alerting functions: a systematic review

Romaric Marcilly, Elske Ammenwerth, Francis Vasseur, Erin Roehrer & Marie-Catherine Beuscart-Zéphir

Abstract

Introduction. Medication-related alerting functions may include usability flaws that limit their optimal use. A first step on the way to preventing usability flaws is to understand the characteristics of these usability flaws. This systematic review aims to analyze the type of usability flaws found in medication-related alerting functions.

Method. Papers were searched via PubMed, Scopus and Ergonomics Abstracts databases, along with references lists. Paper selection, data extraction and data analysis was performed by two to three Human Factors experts. Meaningful semantic units representing usability flaws were the main data extracted. They were analyzed through qualitative methods: categorization through usability heuristics and through an inductive process for flaws specific to medication alerting functions.

Main results. From the 6,380 papers initially identified, 26 met all eligibility criteria. The analysis of the papers identified a total of 168 instances of usability flaws that could be classified into 13 categories of usability flaws representing either violations of general usability principles (i.e., they could be found in any system, e.g., guidance and workload issues) or infractions specific to medication-related alerting functions. The latter refer to issues of low signal-to-noise ratio, incomplete content of alerts, transparency, presentation mode and timing, missing alert features, tasks and control distribution.

Main conclusion. The list of 168 instances of usability flaws of medication-related alerting functions provides a source of knowledge for checking the usability of medication-related alerting functions during their design and evaluation process and ultimately constructs evidence-based usability design principles for these functions.

Keywords

User-computer interface; Human engineering; Decision support systems, clinical; Review, systematic; Usability; Alerting functions

2.1. Introduction

Computerized Clinical Decision Support (computerized CDS) functions may have a noteworthy impact on medication management safety [1]. Several studies have shown that they help to improve antibiotic use [2], drug dosing [3; 4], clinical practice [5; 6] and patient outcomes [7]. However, implemented systems may face acceptance problems [8; 9] that partly originate from poor usability. Poor usability may lead users to reject CDS functions or to adopt workarounds even if the CDS functions are of benefit.

Usability is "the extent to which a product can be used by specified users to achieve specified goals effectively, efficiently and satisfactorily within a specific context of use" [10]. Usability goes beyond the features of the Graphical User Interface (GUI; *e.g.,* legibility of texts), and deals with the functionalities of the product and with the fit between the system behavior and the needs of the users' [11].

Therefore, along with the study of the GUI characteristics, usability includes the analysis of the way in which the system responds to users' actions, of the organization and accuracy of the knowledge incorporated, and of the availability of functions supporting users' tasks. Poor usability of systems arises from the existence of usability flaws. Flaws are violations of usability design principles, and are additionally known as usability heuristics or usability criteria [12-15]. They may have an impact on users' experience with the system (usage problems) and generate negative outcomes in the work system (*e.g.,* performance/patient safety issues) [16]. The present review focuses on usability flaws.

Improving the usability of CDS functions is a necessity [17]. In the broad sense, according to [18], computerized CDS interventions refer to a wide range of tools: forms and templates (*e.g.,* to support proper drug order documentation), relevant data presentation (*e.g.,* to support optimal decision making), proactive drug order suggestions and order sets (*e.g.,* to ensure that a clinical situation is completely addressed), protocol supports/clinical pathways (*e.g.,* to avoid omissions in the care process), reference information/guidance (*e.g.,* to address known information needs) and alerts (*e.g.,* to prevent errors due to lack of knowledge) [18; 19]. These categories of tools are not exclusive, for instance, alerts may be integrated in order sets or in protocol supports. Within the whole range of available computerized medication CDS systems, alerting functions are known to face serious usability issues [17; 20].

One way to prevent such usability issues is to provide manufacturers and Human Factors experts with evidence-based usability design principles [16]. Currently, existing lists of usability design principles regarding medication alerting functions are not based on evidence but rather on expert consensus [17] or targeted review [19; 20]. This study is part of a project that aims at contributing to the emerging knowledge on usability design principles to complete the existing lists and identify the usability design principles that are supported by evidence in the literature. A first step in that direction is to systematically comprehend the usability characteristics of medication-related alerting functions.

The present systematic review focuses on medication-related alerting functions and addresses the following question: "What are the usability flaws of medication-related alerting functions identified in published studies?"

2.2. Method

This systematic review complies as far as possible with international methodological guidelines [21;22] as well as with reporting recommendations [23].

2.2.1. Eligibility criteria

This review considered only original studies reporting usability flaws and published after 1980 in peer-reviewed journals or conference proceedings. Only English and French speaking papers were included. Three eligibility criteria were defined:

- Only medication-related alerting functions supporting the prescribing of medications and used in general hospital or in primary care general practice were included. Surgery, dentistry, anesthesiology, emergency were excluded because the organization of the medication management of those wards is different from the general hospital's medication use process. Pathology or diagnosis management alerting functions were excluded when they did not include features to support medication decision-making. Alerting functions dedicated to the patients as primary end-users were also excluded.
- Usability studies as well as socio-technical studies and impact studies addressing (at least partially) usability issues were included. Only papers with a good quality reporting of the study were kept. Studies on more than one system were included if the results presented insights for each system separately.
- The review targeted studies that reported usability flaws in a descriptive and objective way. This excluded all studies reporting perceived usability assessment or feelings/opinions *e.g.,* collected through usability questionnaires.

2.2.2. Information sources and search

Information was searched for in on-line references databases. Themes of searched papers are at the intersection of two domains: "health technologies" and "ergonomics". Therefore three databases dealing with those themes were chosen: PubMed, Scopus and Ergonomics Abstract. This search was completed by searching references in the reviewed papers.

When possible, MeSH terms were chosen for PubMed. The terms were adapted for Scopus and Ergonomics Abstracts. Two sets of key terms were defined: on "alerting functions" and on "usability" (Table 3). In each set, terms were combined with the "OR" operator. Both sets were then combined with

the "AND" operator (cf. appendix 3 for the complete queries). As the Ergonomics Abstracts database is dedicated to Human Factor topics, only the first set of terms was searched. The language was restricted to English/French, publication date after 1980, and type of journals to medical journals. Searches were performed in April 2012 and updated on the 25th June 2013.

Table 3 - Key terms used in the queries according the database searched.

	PubMed	**Scopus**	**Ergonomics Abstracts**
Alerting functions terms	Medical order entry systems, Medication alert system, Computerized physician order entry system, CPOE, Decision Support Systems, Clinical, Clinical decision support systems, CDSS	Medical order entry, Medication alert, Computerized physician order entry, CPOE, Clinical decision support, CDSS	
Human Factors terms	User-computer interface, Human engineering, Risk factors, Humans, Usability	User-computer interface, Human engineering, Risk factor, Human factor, Usability, Human-computer interaction	Not applicable

2.2.3. Study selection process

The study selection was performed by usability experts with high expertise in Human Factors applied to health informatics and who had previous experience with medication management systems, CDS and alerting functions. At each step of the selection, the review process was over-inclusive; if in doubt, the item was included for an analysis at the next step.

One reviewer (RM) excluded duplicate publications, non-original studies and non-peer-reviewed papers. Then, two reviewers (RM & MCBZ) screened the title of the papers, after a joint training session on 77 papers that were chosen at random from amongst all the papers, the reviewers screened 471 randomly selected papers individually. The agreement score calculated on the review of the 471 papers was good (Cohen's $\varkappa = 0.67$) and discrepancies were solved through reconciliation meetings. After the review of the 471 papers all remaining papers were divided between both reviewers to be screened individually by title.

In the next step, the same two reviewers screened the abstracts of the selected papers. For papers without abstracts (n = 108), the full texts were screened directly. A joint training session on 44 randomly selected papers was performed and followed by parallel individual review on the abstract of 73 papers. The agreement score was again good (Cohen's $\varkappa = 0.69$) and thus the remaining papers were divided between both reviewers to complete the screening of the abstracts.

Once the screening of abstracts was completed, the full-texts of the selected papers were screened by three reviewers (MCBZ, SP and RM): after a training session on 20 randomly selected papers, reviewers individually analyzed 20 other papers. The agreement score was almost perfect (Fleiss' $\varkappa = .95$) and thus each reviewer screened a subset of the remaining papers. The excellent agreement score shows that the

eligibility criteria, including the definition of objective usability flaws, were sufficiently well-described to support non ambiguous decision-making. The motive for rejecting a paper was documented.

2.2.4. Data extraction and analysis

Data was collected for all included papers by two reviewers through several independent readings (MCBZ & RM). Where available, on-line appendices of the papers were also analyzed. During the data collection process any disagreement was solved by discussion. Authors of [24-28] were e-mailed to get more information.

A review reporting form was used for data extraction that included three sections:

- Description of alerting system:
  Alerting function may either be used standalone, or be integrated into a larger information system such as a Computerized Physician Order Entry (CPOE) or an Electronic Medical Record (EMR).
  Medication-related alerting functions may target several kinds of clinical information. Kuperman et al's [9] classes of medication-related decision support were used to categorize the functions: drug allergy checking, basic and advanced dosing guidance, formulary decision support, duplicate therapy checking, drug-drug interaction checking, advanced guidance for medication-associated laboratory testing, advanced checking for drug-disease interactions and contraindications, and finally advanced drug-pregnancy alerting. Those classes are non-exclusive: a single alerting function may include several classes of clinical information.
  Alerting functions may issue different modes of alerts [19] and were accordingly categorized as (a) interruptive (active, pushed alerts), *i.e.,* alerts designed as modal dialog boxes requiring an action to dismiss, (b) non-interruptive (passive, pulled alerts), *i.e.,* alerts displayed in a non-intrusive asynchronous presentation format or (c) mixed, *i.e.,* combining both interruptive and non-interruptive formats of alerts.
  Moreover, we considered the stage of development of the alerting function either during the design process ("under development") or when it is in use.
- Description of methods: The methods applied in the studies to collect usability data were extracted. This extraction was supported by a checklist of methods listed in usability standards [29] and in a focused review [30]. The checklist included the following items of observation, interviews, user-testing (including think-aloud), heuristic evaluation, focus groups, retrospective data analysis (expert review), cognitive walkthrough, questionnaires, telephone and e-survey, log files analysis, experimental design and performance measurement, critical incident analysis, creativity methods, contextual inquiry, collaborative evaluation, automated evaluation, brainstorming, document analysis, document-based methods, model-based methods, parallel design.

- Objective descriptions of usability flaws: Any violations of a usability principle described from the system's perspective. Usability flaws are descriptions of the characteristics of the systems that do not adhere to usability principles. They may concern the system's GUI, its behavior and the suitability of the knowledge implemented within it for users' needs and the availability of features needed to perform a task. Therefore the objective descriptions may describe all of the above mentioned dimensions. Items representing instances of usability flaws were searched for in the results and discussion sections of the included papers. Only items reported in a descriptive, objective and reproducible way were retained in the analysis process in order to get reliable data; hypotheses drawn by the authors of the included studies were not analyzed.

"Duplicate instances" of flaws, *i.e.,* descriptions of usability flaws detected in a given function and that were reported in several papers on the same study, as well as instances described several times in the same paper, were presented together in a single instance.

Once the data extraction was complete, each extracted usability flaw was categorized according to the usability design principle it violated. Design principles can be described in a variety of different ways [12-15] however it has been identified that differently named principles can reflect the same inherent usability concepts [31]. The main differences reside in their construction, in the precision level of the principles and in the instructions on how they should be applied. Scapin and Bastien's usability design principles [14] were chosen for this review as unlike other sets of design principles, they were based on a review of recommendations published in the literature and in standards and then reviewed by experts, not only from experience. Additionally, Scapin and Bastien's usability design principles [14] were considered by both reviewers as being most precise with easy to apply instructions.

Scapin and Bastien's [14] set of heuristics is composed of eight main usability design principles and 18 sub-principles (Table 4). The first seven principles (guidance, workload, explicit control, adaptability, error management, consistency and significance of codes) are applicable to any kind of computerized system. The eighth principle, "compatibility", considers how the characteristics of the system under design/evaluation fit:

- The characteristics of the tasks to be performed with or supported by the system;
- The characteristics of the typical end-user(s) (mental model, knowledge organization, cognitive tasks);
- The characteristics of the typical end-user(s) workflow.

The compatibility principle accounts for the specificity of the task to perform, *i.e.,* in the present context, the interaction of the clinicians with the alerting function. It is not a principle that has previously been divided into sub-categories. Therefore, to account for the different dimensions of the flaws specific to the medication alerting functions, sub-categories of "compatibility" flaws were developed by both reviewers (RM & MCBZ) in an interactive and inductive manner. During this process any difficulties and

disagreements were discussed between both reviewers to achieve clear, unambiguous, and mutually exclusive sub-categories with a high internal consistency and on which both reviewers were in complete agreement. At the end of the categorization process, each usability flaw was assigned to a unique category and sub-category.

Table 4 - Main Scapin and Bastien's usability principles (cf. [32] for sub-criteria's description)

Usability criteria	Definition
Guidance	Refers to the means available to advise, orient, inform, instruct, and guide the users throughout their interactions with a computer (messages, alarms, labels, etc.), including from a lexical point of view.
Workload	Concerns all interface elements that play a role in the reduction of the users' perceptual or cognitive load, and in the increase of the dialogue efficiency.
Explicit control	Refers to the system processing of explicit user actions, and to the control users have on the processing of their actions by the system.
Adaptability	Refers to system's capacity to behave contextually and according to the users' needs and preferences.
Error management	Refers to the means available to prevent or reduce errors and to recover from them when they occur. Errors are defined in this context as invalid data entry, invalid format for data entry, incorrect command syntax, etc.
Consistency	Refers to the way interface design choices (codes, naming, formats, procedures, etc.) are maintained in similar contexts, and are different when applied to different contexts.
Significance of codes	Qualifies the relationship between a term and/or a sign and its reference. Codes and names are significant to the users when there is a strong semantic relationship between such codes and the items or actions they refer to.
Compatibility	Refers to the match between users' characteristics (memory, perceptions, customs, skills, age, expectations, etc.) and task characteristics on the one hand, and the organization of the output, input, and dialogue for a given application, on the other hand.

2.2.5. Bias assessment

To ensure a good validity of the eligible studies, two eligibility criteria concerning the report and the method of the studies were defined. Both criteria were selected because they were necessary if the study was to be reproduced. During the analysis process, two reviewers (RM & MCBZ) scored each paper individually on two 5-point Likert scales (from 1 = poor report or method to 5 = very good report or method) regarding the following criteria:

- Report: completeness and clarity in the description of the aim of evaluation study, the context of evaluation, the function under evaluation (including type of system, stage of development), the setting in which it is (to be) implemented, and in the results.
- Method: completeness and clarity of the applied methods. This included an assessment of the method(s) applied, a description of participants (number, profile and experience in the setting

and/or with the system), and examination of the study design (*i.e.,* single evaluation, comparison between systems and/or pre- post- re-engineering).

When any of the two scores was equal to one on the 5-point Likert scale, the paper was excluded. The usability flaws reported in the included studies may have differing description levels, creating a risk of bias across studies. By selecting studies reporting on usability flaws that were sufficiently self-explanatory (*i.e.,* usability flaws' descriptions that do not need supplementary information to be understood), the potential of bias was mitigated.

Moreover, publication and selective reporting biases had an impact on the review. Firstly, conference proceedings do not provide as much space for describing usability flaws as do journal papers with on-line appendices. Secondly, the focus of the study, *e.g.,* alerting function vs. EMR/CPOE that includes an alerting function or pure evaluation study vs. entire design cycle study, could also have an impact on the report of usability flaws. Since the aim of the review is to achieve a comprehensive description of all usability flaws reported in previous studies on medication-related alerting functions, those biases were handled by performing only qualitative analysis.

2.3. Results

2.3.1. Study selection

The study design is schematically described in Figure 3. The database searches and the searches within publication references identified 6,380 publications. After the removal of duplicate publications, non-original studies and non- peer-reviewed publications (n = 1109), screening of the titles and of the abstracts excluded 4,817 publications and rendered 454 papers eligible for further full-text review. Based on the full-text review, 419 papers were also excluded because they do not focus on the studied topic. Moreover, nine publications were excluded due to quality concerns in the reporting style, either the usability flaws and the method applied were not precisely described, or the discussion of the results confused flaws of different systems [33-41]. Finally, a total of 26 papers met our inclusion criteria and were used for detailed analysis. One [28] out of the two papers claiming to have on-line appendices did not provide them despite e-mailing the authors. On-line appendices of only one paper were analyzed [42].

2.3.2. Results of the bias assessment

Only studies with a good validity were included in the analysis: after rating the validity of the included papers, nine papers were excluded due to quality concerns in the reporting style. Overall, the report quality was average (mean score = 3.46, median = 3.5) and the method quality was relatively good (mean score = 3.88; median = 4) for the included publications.

For the differing levels of description in the reported usability flaws, only studies reporting self-explanatory usability flaws were selected. Amongst the 168 instances of usability flaws identified in the

studies, 155 are verbatim from the papers and only 13 (7.74%) needed rephrasing to make the usability flaws clearer. Rephrasing was based on complementary information provided in the papers, screenshots of the functions along with users' and designers' comments.

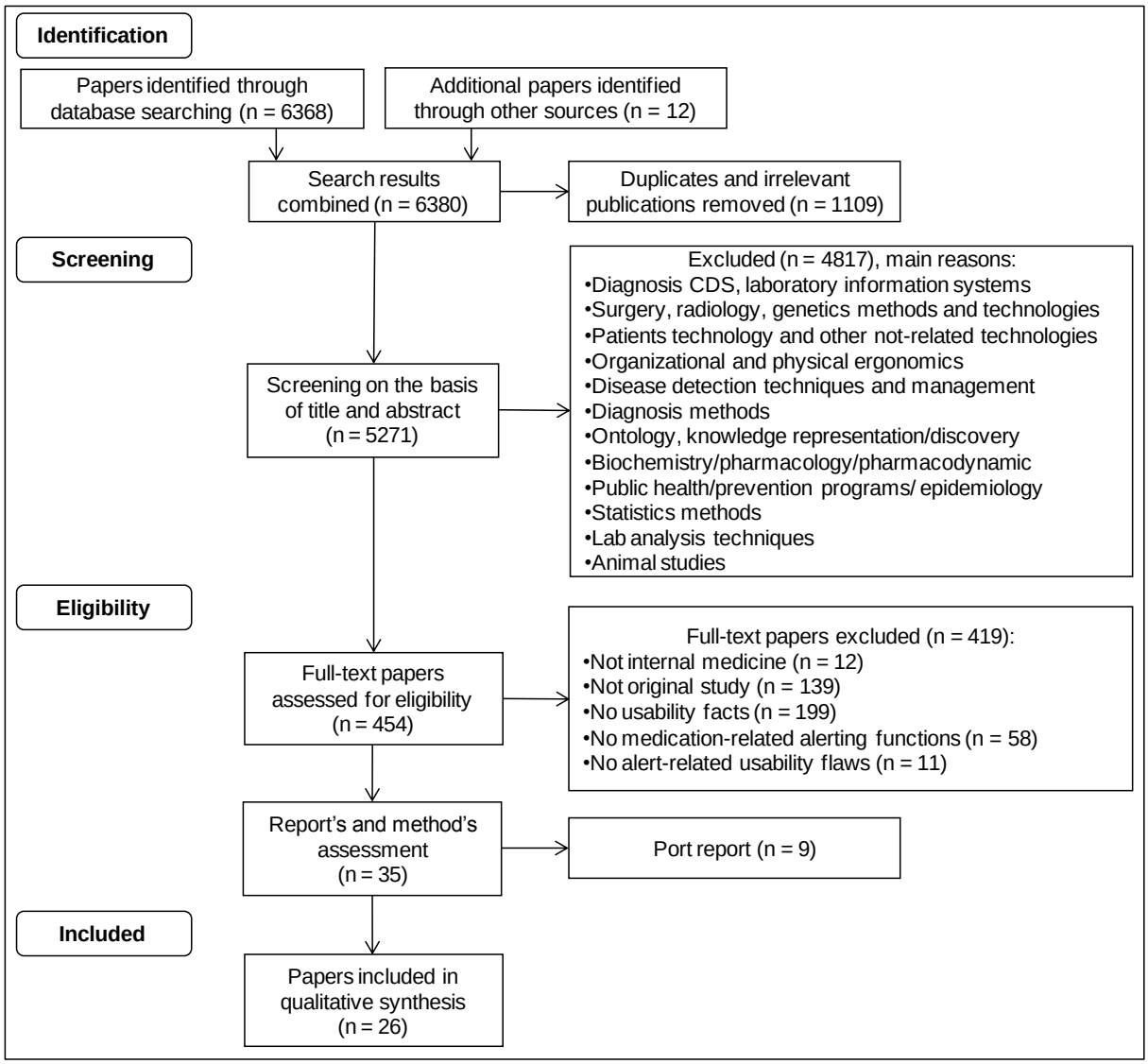


Figure 3. Study flow.

2.3.3. Characteristics of included papers

The set of 26 included papers comprises 10 conference proceedings and 16 journal papers. These papers report evaluations of 19 different systems integrating alerting. For two instances, 2 papers reported separate evaluations of the same system ([26; 27] & [43; 44]). Seven papers report evaluations of a CPOE that contains both non-interruptive and interruptive alerting functions. It was not identifiable whether both functions worked similarly or not therefore each of the 7 papers was analyzed separately. Amongst these papers, 3 report different evaluations of non-interruptive alerts [14; 45; 46] and 4 report a unique

study on interruptive alerts [42; 47-49]. Since the latter 4 papers provide the same results, "duplicate instances" of usability flaws are presented together.

The main characteristics of analyzed papers are summarized in Table 5. When the number of studies does not total 26 (or 19 functions), papers do not report on the issue.

There is no clear mention in any of the papers of the functions' design/implementation stage therefore deductions, based on the context of the studies, have been made to obtain this information. The large majority of the functions (16 in 23 papers) belong to system already in use, amongst them 1 paper reports an evaluation during a redesign process [50; 51] and 1, an evaluation without real patient data [52]. Two other systems were still under development at the time of evaluation [53] and 1 was acquired by the hospital but not yet implemented [28].

Two alerting functions are standalone software [53]. The other 17 functions (in 24 papers) are integrated either into EMR, CPOE or into another type of electronic patient records.

Classes of medical information targeted by the alerting functions [9] are not reported in a systematic manner. The most reported class is "drug-drug interaction checking" (for 10 alerting functions in 15 papers). There are also six alerting functions that include "drug-allergy checking", "duplicate therapeutic checking" or "basic dosing guidance" (respectively in 9, 11 and 8 papers). Evaluations of "advanced guidance for medication-associated laboratory testing" are reported in two alerting functions (in 6 papers). Evaluations of "advanced dosing guidance", "formulary decision support" and "advanced drug-pregnancy" alert are reported once.

Twelve alerting functions (17 papers) are interruptive, 3 are non-interruptive (5 papers) and 3 are mixed systems (3 papers).

Each paper included a detailed description of the methodology applied and the data collection methods used. The included papers proposed a great variety of methods. Eighteen papers combine at least two methods such as observations [24; 42; 45-49; 51; 54-56], interviews [24; 42-49; 51; 53-57], focus groups [55; 58-60], user testing [26; 52; 53], simulation [44; 50; 51], cognitive walkthrough [27; 52], heuristics evaluation [28; 61], questionnaire [50; 51; 53; 59], survey [51; 55], retrospective analysis [25; 43; 62] and log files analysis [51]. Eight papers apply one method amongst the above mentioned [25; 27; 28; 57; 58; 60-62].

2.3.4. Categories of identified usability flaws

Overall, 168 instances of usability flaws are reported and categorized. No inter-experts agreement score was calculated because both experts performed this categorization process, including the development of the sub-categories, together.

The subsequent sections of this review describe the categories and sub-categories of usability flaws. The ultimate aim of this research is to look for evidence for usability design principles dedicated to medication alerting functions, therefore a focus is drawn to the categories of flaws specific to those

functions. Table 6 proposes a synthesis of the categories of reported usability flaws. For more details, the reader may refer to appendix 3 that presents the complete list of instances of usability flaws.

Table 5 - Main characteristics of the selected papers.

System		Mode of alerting	Integration	Class[3]	Design stage	Methods
VA Computerized Patient Record System	[49]	Interruptive	In CPOE	DA, DT, DDI, AGML	In use	Observations, interviews
	[47]	Interruptive	In CPOE	DA, DT, DDI, AGML	In use	Observations, interviews
	[42]	Interruptive	In CPOE	DA, DT, DDI, AGML	In use	Observations, interviews
	[48]	Interruptive	In CPOE	DA, DT, DDI, AGML	In use	Observations, interviews
	[46]	Non-interruptive	In CPOE	No data	In use	Observations, interviews
	[50]	Non-interruptive	In CPOE	No data	Redesign of system in use	User testing, simulation, questionnaire
	[45]	Non-interruptive	In CPOE	No data	In use	Observations, interviews
Medicator ©	[26]	Interruptive	In CPOE	BDG, DT, DDI	In use	Heuristics evaluation, user testing, experimental design
	[27]	Interruptive	In CPOE	BDG, DT, DDI	In use	Cognitive walkthrough
Medicatie/ EVS ©	[43]	Interruptive	In CPOE	BDG, DT, DDI, AGML	In use	Retrospective analysis, interviews
	[44]	Interruptive	In CPOE	BDG, DT, DDI, AGML	In use	Simulation
Other	[24]	Non-interruptive	In patient record	BDG, DDI, ADG	In use	Observations, interviews
	[62]	Interruptive	In CPOE	DA, DT, DDI	In use	Retrospective analysis
	[58]	Interruptive	In CPOE	DA, DDI, ADP	In use	Focus group, Interviews
	[54]	Interruptive	In CPOE	No data	In use	Observations, interviews
	[60]	Mixed	In EMR	No data	In use	Focus group
	[52]	Interruptive	In CPOE	BDG	In use but evaluated with fake patients' data	Cognitive walkthrough, user testing
	[25]	No data	In CPOE	DA, DT	In use	Retrospective analysis
	[61]	Interruptive	In EMR	DDI	In use	Heuristics evaluation

[3] Acronyms for the classes [9]: DA, "drug allergy"; DT, "duplicate therapy"; DDI," drug-drug interaction"; BDG, "basic dosing guidance"; ADG, "advanced dosing guidance"; ADP, "advanced drug-pregnancy alert"; AGML, "advanced guidance for medication-associated laboratory testing"; FDS, "formulary decision support".

System		Mode of alerting	Integration	Class[3]	Design stage	Methods
	[51]	Non-interruptive	Standalone	No data	Under development	Observation, interviews, log files analysis, simulation, questionnaire, e-survey, telephone survey
	[57]	Mixed	In EMR /CPOE	DDI	In use	Interviews
	[55]	Interruptive	In CPOE	DA, BDG; FDS	In use	Observations, interviews, focus groups
	[53]	Interruptive	Standalone	DDI	Under development	Interviews, user testing, questionnaire
	[56]	Interruptive	In CPOE	DA, DT, ADP	In use	Observations, Interviews, e-survey, telephone survey
	[28]	Interruptive	In CPOE	BDG	Prior implementation	Heuristics evaluation
	[59]	Mixed	In patient record	DDI	In use	Focus groups, questionnaire

2.3.5. General usability flaws

Guidance infractions refer to issues related to prompting the user(s): important information is not highlighted, information is displayed in the visual periphery and instructions are unclear. There are also instances of the lack of distinction of alerts according to format: alerts with different severity levels are not visually distinguished as well as alerts of different types. There are also legibility issues along with lack of immediate feedback and heterogeneous presentation of alerts presenting the same severity level.

Workload-related usability flaws are mainly related to the excessive number of actions to be performed either to obtain information or to enter data. Other workload infractions refer to dense information and to non-concise information.

Violations of the significance of codes criterion are related to non-intuitive icons and wording. There are also issues with the consistency of the behavior of the system; it does not work the same way across use and according to the data it analyzes. Instances of explicit control issues are related to the fact that the system does not act as the user required and due to the lack of user control; there is no way to undo an action. An instance of adaptability flaw is also observed, as the system does not support all types of users.

Finally, an instance of error management flaw is reported; the message that is supposed to explain a problem related to the alerting function is not clear.

2.3.6. Medication-related alerting functions-specific usability flaws

Six categories of medication alerting functions-specific usability flaws are identified.

2.3.6.1. Low signal-to-noise ratio

The category of "low signal-to-noise ratio of alerts" deal with the failure to consider the context of use including the clinical context (patient clinical case), the setting context (user's expertise or ward habits, clinicians' priority, good practices and pharmacist knowledge), logistical context (care logic) or the care context (actions already taken by the clinicians). The category also deals with problems of reliability of the data triggering the alerts, there are several instances are reported of alerts that appear erroneously due to clinical data not being up-to-date. Seven issues of alert redundancy are observed. They include problems of re-appearance of the same alert throughout the same order entry, problems of system logic that does not consider the relevant solutions proposed by the users, and finally problems of the impossibility of de-activating a specific alert for a specific patient.

2.3.6.2. Alert content issue

This category refers to missing information in the alert and to the wrong content, rendering the alert useless. The missing information is related to one of three topics: the purpose of the alert (clinicians are facing alerts containing only the name of incriminated medications, without the reason why they are triggered, their severity and supporting scientific evidence), contextual information (patient's condition of importance regarding the alert, information necessary to interpret data within the alert such as lab results), and suggestions of actions to be taken to avoid (or to manage) the detected potential problem. Other instances are related to suggested actions that are "clinically erroneous" as stated in the studies analyzed.

Table 6 - General and medication-related alerting functions' specific categories of usability flaws and references of the papers from which they were retrieved.

General usability flaws	**Studies**
Guidance issues	
Prompting issues: unclear text, information highlight deficiency, alert/information far from the center of the screen, no detail.	[24; 42; 44; 47; 51; 53; 57; 59; 62]
No distinction by format (shape, color) of different severity alerts, types of alerts or types of message (system vs. medical alerts).	[42; 49; 58; 61]
Legibility issues: not sufficient inter-line space, font in capital letters, size of elements too small.	[42; 49; 53]
No feedback to inform the user that (s)he has just missed an alert.	[44]
Too much distinction by location: no grouping of same severity alerts.	[61]
Workload issues	
Minimal action: too many actions for entering information or obtaining information (*e.g.,* scrolling, tabs).	[42; 44; 46; 50; 53; 59; 61]
Information density: too much information of different kinds in the window, several alerts in the same window, alert content displayed in a one-paragraph format.	[24; 42; 47; 49; 51; 52]

Lack of concision.	[56-58]
Significance of codes issues	
Non-intuitive wording.	[26; 43;50; 51; 60]
Non-intuitive icons.	[45; 51]
Consistency issues	
Inconsistency of behavior of the system across use or according to data analyzed	[28; 47]
Explicit control issues	
Explicit user actions: system's action does not correspond to the action requested by the user.	[45]
User control: there is no way to undo an action.	[46]
Adaptability issues	
Lack of flexibility: the system does not support all user types.	[42]
Error management issues	
Quality of error message: problem messages are unclear.	[42]
Usability flaws specific to medication-related alerting functions	
Low signal to-noise-ratio	
Alerts are irrelevant regarding: expertise/ward habits, existing validated good practices, pharmaceutical knowledge, data considered, patient case, actions engaged, clinician's interest for at risk situations, care logic, no detail.	[24; 26; 42-44; 46-49; 58-60]
Low signal-to-noise ratio without specific description.	[42-45; 47; 49; 54; 56; 58; 60]
Alerts are redundant: alerts appear very frequently/ several times during the decision making, clinically relevant solutions from the clinicians are not accepted, no feature for turning-off a specific alert in a specific context.	[42 ;45; 46; 49; 56; 57; 60]
Alert content issues	
Information required to make a decision is missing: the actions that could be taken, patient data, the problem detected its evidence and its severity and information for interpreting data within the alert.	[42;43;47;49;51;53;61]
The alert's content proposes erroneous suggestions: the proposed action does not suit the clinical context, no detail.	[26;45]
Function is not transparent enough for the user	
The alerting function is not transparent about the way it works: no information about the alert severity scale, about the up-to-dateness of the alerts' rules or no detail.	[24;28;42;46;47;52;61]
The alerting function is not transparent about the data it uses: all available data is not used to trigger the alert or incomplete mapping.	[25;42;49;60]
Alert appearance issues: timing and mode	
Alert does not appear at the right moment to support the decision making process: before the decision process starts, at the wrong moment, after the decision is made.	[24;27;52;55-57;60]
Data processing is slow.	[42;45]
The alert's display mode does not suit the decision making process: not sufficiently intrusive, too intrusive.	[60]
Tasks and control distribution issues	
Alert not displayed to the right clinician or only to the pharmacist.	[44;55;59]
The alerting function allows users to enter comments that are displayed to no one.	[42;62]
Alert not transferable from one clinician to another.	[45]
Alert features issues	
Alert features are missing: no feature for reconsidering an alert later, no access to additional information from the alert, no action tool to solve the problem from the alert.	[42;44-46;51;61]

2.3.6.3. Function is not transparent enough for the user

Flaws in this category refer to the fact that the system does not provide sufficient information for the user to know how it works. Instances deal with a lack of transparency about the way the alerting function is operating: what knowledge is applied to trigger the alerts (and if it is up-to-date), how alerts are categorized by severity or what actions users can perform on the system. Some instances reveal a lack of information about the data that is used by the system to trigger alerts. Clinicians are not informed that some data is screened while other data is not, and that the medications' mapping implemented in the alerting function does not consider all possible usages of a medication.

2.3.6.4. Alert appearance issues : timing and mode

This category illustrates alert appearance timing issues: the alert appears after the decision is made or the alert appears before the decision-making process is started or just at the wrong moment. There are also instances of low processing, slowing down the appearance of the alert. Finally, there are issues with the mode of presentation of the alert, *i.e.,* their level of intrusiveness: alerts are either too intrusive, distracting the clinician from the decision-making process, or not sufficiently intrusive being unnoticed.

2.3.6.5. Tasks and control distribution issues

This category identifies that tasks and control distribution flaws are related to the behavior of the alerting functions that is unsuitable for the cooperative and distributed aspects of the medication use process. The alert is displayed to the clinicians who are not concerned by it (*e.g.,* physiotherapist or psychologist). It is displayed once to a single clinician, not to the whole team. The alert and users' comments on it may also not be transferrable between clinicians who take care of the patient.

2.3.6.6. Alert features issues

This category deals with the lack of alert features adapted to support decision-making or the instantiation of a decision. Some instances are related to the vlatility of alerts: the clinician cannot choose to reconsider an alert later in the decision making process. Moreover, the user cannot access additional information directly from the alert nor can they act to solve the problem highlighted by the alert through actions featured in the alert. The clinician must go back to the patient's record or ordering system to find information or to take action. Finally, there is an instance of a feature that does not suit clinicians' workflow and so the alert is recorded in the patient's progress note outside the template compelling clinicians to search for it in the entire note.

2.4. Discussion

This review fits into a project aiming at contributing to the emerging knowledge on usability design principles to complete the existing lists and identify those that are supported by evidence in the literature. The present systematic review aimed at answering the question "what are the usability flaws in medication-related alerting functions identified in published studies?" to provide comprehensive knowledge on the usability flaws already reported for those functions.

Two main kinds of usability flaws have been observed: general usability flaws and specific usability flaws. General usability flaws are related to infractions of good guidance practices, workload, significance of codes, explicit control, adaptability, error management and consistency (Table 4). General usability flaws can be observed whatever the type of computerized system. They are known for potentially making the use of the system harder [14].

The results also highlight types of usability flaws that are specific to medication-related alerting functions. These instances are distributed across six sub-categories providing an overview of the specific usability issues for those functions (Table 4): low signal-to-noise ratio, content issues, transparency issues, appearance (timing and mode) issues, tasks and control distribution issues and alert features issues. Those flaws are not dealing only with the GUI of the function but concern all components of the alerting functions including the knowledge implemented in it, its triggering model and the behavior of the function. The consequence of these flaws is that users may question the usefulness of the system when it does not provide relevant information to help clinicians make their decision. A severe consequence is the loss of impact of the alert, for instance, the low signal-to-noise ratio and the high intrusiveness of the alert have been identified as factors contributing to clinicians' "alert fatigue" that could ultimately lead to their rejection of the alerting function [5; 63].

This systematic review has limits and biases that must be considered. In the reviewed papers, the description of the functions under evaluation is not complete, preventing the analysis of whether usability flaws are related to a specific type of the alerting function. Another limitation is restricting the review process to peer-reviewed papers, potentially putting aside institutional or companies' reports. Relevant usability studies could have been missed. However, only a peer-reviewed publication process could guarantee that the methods applied in those studies are of sufficient quality. The quality of the review method reflects this, with only nine of the included papers excluded.

Publication and selective reporting could have potentially impacted the representativeness of the results. Moreover, some types of flaws are easier to picture in a few words or with a screenshot (*e.g.,* guidance) while others require long descriptions of a system's behavior to be understood (*e.g.,* error management). This may have impacted, at least partially, the representativeness of the results (*e.g.,* 1 instance for error management but 29 for guidance). This representativeness bias may have impacted the diversity of the usability flaws reported as compared to those actually existing in alerting functions.

However, all categories of flaws identified included instances that come from different papers: the overlapping of data from various sources ensures the reliability of these results.

The systematic review process highlighted papers that initially met the inclusion criteria but further examination revealed that significant data was missing from the methods and/or the results section of the papers. This problem has been identified by previous systematic reviews on usability characteristics of medical software (*e.g.,* [64] for CPOE). In this review, only papers with sufficient data in the study method have been analyzed; 9 papers were excluded after initial acceptance due to missing data. This represents a loss of about a fourth of the total number of papers that could have been analyzed. Moreover, even in the included papers, some information non-essential for the topic of the review (*e.g.,* the class of CDS, the context of use) was often briefly described. There is an actual need for reporting guidelines for usability-related papers [65] as has been done for HIT evaluations [66]. Moreover, this review has also illustrated the advantages of on-line appendices to describe usability flaws [42] exhaustively and in detail provides much useful data. The use of on-line appendices has to be encouraged for the reporting of the whole set of usability flaws uncovered by the study and the precise description of the system. This proposal is currently under examination in an international Delphi study [67]. A demand for high quality reports for Human Factor evaluation will enable a repository of high quality studies to be created. This repository is necessary to capitalize on usability data in order to ultimately look for evidence for usability principles.

The results of this review provide insight on the topic of usability flaws in alerting functions by precisely detailing the types of usability flaws reported for alerting functions. Moreover, in their current state they are directly useful, a list of 168 actual concrete instances of usability flaws that characterize those functions is now available. Even though it may not be representative of the entire set of usability flaws that could possibly be found in medication alerting functions, it is exhaustive considering what has been published on this topic.

As far as we know, it is the first time that such a list based on empirical illustrated knowledge has been proposed for medical software. This list could be used as an illustrated check-list for usability mistakes not to be made by Human Factor experts, designers and health informatics project managers to facilitate the identification and correction of potential usability flaws during the design, evaluation, procurement and implementation processes.

This work should be regularly updated to consider the evolution of the usability features of medication-related alerting functions for the present list represent the current state of knowledge on usability flaws problematic for those functions. Moreover, those data should also be completed by searching incident reports for medication-related alerting functions [68]. This will enable us to identify new kinds of flaws not yet reported in the literature and enhance the database developed by the current review.

This study reports the first mandatory step to seek evidence for usability design principles for medication-related alerting functions. The flaws identified may have an impact on users' experience with

the system and negatively affect the work system (*e.g.,* generating patient safety issues). This impact may have different severity levels considering its object from low to potentially harmful. Therefore, further studies must endeavor to identify the consequences of those flaws in terms of usage problems and usability-related outcomes in the work system. This will help us to weight the usability flaw categories in order to find out which ones are more dangerous and should be set as a priority and ultimately support the construction of evidence for related usability design principles that take clearly into account the consequences of the infractions of those principles [16]. The required next step will consist in matching categories of usability flaws in medication alerting functions with existing related usability design principles. The results of this operation will enable the identification of existing lists of usability design principles that are supported by evidence of the literature.

2.5. Conclusion

The present systematic review aimed at identifying the usability flaws that have been reported in previous studies on medication-related alerting functions. Results identified 168 instances of usability flaws that were categorized into eight categories of general flaws completed by six categories of flaws specific to medication-related alerting functions. The 168 instances represent 168 usability mistakes not to be made. This list can be used as a usability check-list during the design, the evaluation, the procurement and the implementation process of medication-related alerting functions.

Knowing those flaws is a first step to provide recommendations for improving the usability of medication-related alerting functions. Further studies are needed to identify the known potential consequences of those flaws in terms of usage problems and outcomes in the work system and to provide suitable and precise usability design principles.

References

[1] Jaspers MW, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. J Am Med Inform Assoc 2011 May 1;18(3):327-34.

[2] Thursky K. Use of computerized decision support systems to improve antibiotic prescribing. Expert Rev Anti Infect Ther 2006 Jun;4(3):491-507.

[3] Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. JAMA 1998 Oct 21;280(15):1339-46.

[4] Oppenheim MI, Vidal C, Velasco FT, Boyer AG, Cooper MR, Hayes JG, et al. Impact of a computerized alert during physician order entry on medication dosing in patients with renal impairment. Proc AMIA Symp 2002;577-81.

[5] Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005 Apr 2;330:765-73.

[6] Schedlbauer A, Prasad V, Mulvaney C, Phansalkar S, Stanton W, Bates DW, et al. What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior? J Am Med Inform Assoc 2009 Jul;16(4):531-8.

[7] Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. JAMA 2005 Mar 9;293(10):1223-38.

[8] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004 Mar;11(2):104-12.

[9] Kuperman GJ, Bobb A, Payne TH, Avery AJ, Gandhi TK, Burns G, et al. Medication-related clinical decision support in computerized provider order entry systems: a review. J Am Med Inform Assoc 2007 Jan;14(1):29-40.

[10] International Standardization Organization. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability (Rep N° 9241-11). Geneva: International Standardization Organization; 1998.

[11] Beuscart-Zephir MC, Borycki E, Carayon P, Jaspers MW, Pelayo S. Evolution of human factors research and studies of health information technologies: the role of patient safety. Yearb Med Inform 2013;8(1):67-77.

[12] Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. J Biomed Inform 2003 Feb;36(1-2):23-30.

[13] Schneiderman B, Plaisant C, Cohen M, Jacobs S. Designing the user interface: strategies for effective human-computer interaction. 5th edition ed. Boston: Addison Wesley Longman; 2008.

[14] Scapin D.L., Bastien J.M.C. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behaviour and Information Technology 1997;6(4-5):220-31.

[15] Nielsen J. Usability Engineering. Boston: Academic Press; 1993.

[16] Marcilly R, Beuscart-Zephir MC, Ammenwerth E, Pelayo S. Seeking Evidence to Support Usability Principles for Medication-Related Clinical Decision Support (CDS) Functions. Stud Health Technol Inform 2013;192:427-31.

[17] Sittig DF, Wright A, Osheroff JA, Middleton B, Teich JM, Ash JS, et al. Grand challenges in clinical decision support. J Biomed Inform 2008 Apr;41(2):387-92.

[18] Osheroff J, Pifer E, Sittig D, Jenders R, Teich J. Clinical decision support implementers' workbook. Chicago, Ohio: Healthcare Information Management and Systems Society; 2004.

[19] Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. J Biomed Inform 2012 Dec;45(6):1202-16.

[20] Phansalkar S, Edworthy J, Hellier E, Seger DL, Schedlbauer A, Avery AJ, et al. A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. J Am Med Inform Assoc 2010 Sep;17(5):493-501.

[21] Cochrane handbook for systematic reviews for interventions. England: Wiley-Blackwell; 2008.

[22] Center of reviews and dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. University of York: Center of reviews and dissemination; 2008.

[23] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100.

[24] Wipfli R, Betrancourt M, Guardia A, Lovis C. A qualitative analysis of prescription activity and alert usage in a computerized physician order entry system. Stud Health Technol Inform 2011;169:940-4.

[25] Hartmann Hamilton AR, Anhoj J, Hellebek A, Egebart J, Bjorn B, Lilja B. Computerised Physician Order Entry (CPOE). Stud Health Technol Inform 2009;148:159-62.

[26] Khajouei R, Peek N, Wierenga PC, Kersten MJ, Jaspers MW. Effect of predefined order sets and usability problems on efficiency of computerized medication ordering. Int J Med Inform 2010 Oct;79(10):690-8.

[27] Khajouei R, de Jongh D, Jaspers MW. Usability evaluation of a computerized physician order entry for medication ordering. Stud Health Technol Inform 2009;150:532-6.

[28] Chan J, Shojania KG, Easty AC, Etchells EE. Usability evaluation of order sets in a computerised provider order entry system. BMJ Qual Saf 2011 Nov;20(11):932-40.

[29] International Standardization Organization. Ergonomics of human-system interaction -- Usability methods supporting human-centred design (Rep N° 16982). Geneva: International Standardization Organization; 2002.

[30] Beuscart-Zephir MC, Elkin P, Pelayo S, Beuscart R. The human factors engineering approach to biomedical informatics projects: state of the art, results, benefits and challenges. Yearb Med Inform 2007;109-27.

[31] Nassar V. Common criteria for usability review. Work 2012;41 Suppl 1:1053-7.

[32] Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behaviour and Information Technology 1997;6(4-5):220-31.

[33] Ahearn MD, Kerr SJ. General practitioners' perceptions of the pharmaceutical decision-support tools in their prescribing software. Med J Aust 2003 Jul 7;179(1):34-7.

[34] Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. AMIA Annu Symp Proc 2007;26-30.

[35] Bouchand F, Thomas A, Zerhouni L, Dauphin A, Conort O. [Pharmacists' interventions before and after prescription computerization in an internal medicine department] . Presse Med 2007 Mar;36(3 Pt 1):410-8.

[36] Campbell EM, Sittig DF, Ash JS, Guappone KP, Dykstra RH. Types of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc 2006 Sep;13(5):547-56.

[37] Campbell EM, Guappone KP, Sittig DF, Dykstra RH, Ash JS. Computerized provider order entry adoption: implications for clinical workflow. J Gen Intern Med 2009 Jan;24(1):21-6.

[38] Jaderlund HL, Rudebeck CE, Petersson G. Usability of computerised physician order entry in primary care: assessing ePrescribing with a new evaluation model. Inform Prim Care 2011;19(3):161-8.

[39] Kanstrup AM, Christiansen MB, Nohr C. Four principles for user interface design of computerised clinical decision support systems. Stud Health Technol Inform 2011;166:65-73.

[40] Short D, Frischer M, Bashford J. Barriers to the adoption of computerised decision support systems in general practice consultations: a qualitative study of GPs' perspectives. Int J Med Inform 2004 May;73(4):357-62.

[41] Vaziri A, Connor E, Shepherd I, Jones RT, Chan T, de Lusignan S. Are we setting about improving the safety of computerised prescribing in the right way? A workshop report. Inform Prim Care 2009;17(3):175-82.

[42] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. 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 2012 Apr;81(4):232-43.

[43] van der Sijs H, Aarts J, van GT, Berg M, Vulto A. Turning off frequently overridden drug alerts: limited opportunities for doing it safely. J Am Med Inform Assoc 2008 Jul;15(4):439-48.

[44] van der Sijs H, van GT, Vulto A, Berg M, Aarts J. Understanding handling of drug safety alerts: a simulation study. Int J Med Inform 2010 May;79(5):361-9.

[45] Saleem JJ, Patterson ES, Militello L, Render ML, Orshansky G, Asch SM. Exploring barriers and facilitators to the use of computerized clinical reminders. J Am Med Inform Assoc 2005 Jul;12(4):438-47.

[46] Patterson ES, Nguyen AD, Halloran JP, Asch SM. Human factors barriers to the effective use of ten HIV clinical reminders. J Am Med Inform Assoc 2004 Jan;11(1):50-9.

[47] Russ A.L., Saleem J.J., McManus M.S., Zillich A.J., Doebbling B.N. Computerized medication alerts and prescriber mental models: observing routine patient care. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2009 p. 655-9.

[48] Russ A.L., Saleem J.J, McManus M.S., Frankel R.M., Zillich A.J. The Workflow of Computerized Medication Ordering in Primary Care is Not Prescriptive. 2010 p. 840-4.

[49] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. A human factors investigation of medication alerts: barriers to prescriber decision-making and clinical workflow. AMIA Annu Symp Proc 2009;548-52.

[50] Saleem JJ, Patterson ES, Militello L, Anders S, Falciglia M, Wissman JA, et al. Impact of clinical reminder redesign on learnability, efficiency, usability, and workload for ambulatory clinic nurses. J Am Med Inform Assoc 2007 Sep;14(5):632-40.

[51] Trafton J, Martins S, Michel M, Lewis E, Wang D, Combs A, et al. Evaluation of the acceptability and usability of a decision support system to encourage safe and effective use of opioid therapy for chronic, noncancer pain by primary care providers. Pain Med 2010 Apr;11(4):575-85.

[52] Horsky J, Kaufman DR, Patel VL. Computer-based drug ordering: evaluation of interaction with a decision-support system. Stud Health Technol Inform 2004;107(Pt 2):1063-7.

[53] Duke JD, Bolchini D. A successful model and visual design for creating context-aware drug-drug interaction alerts. AMIA Annu Symp Proc 2011;339-48.

[54] Ash JS, Sittig DF, Dykstra RH, Guappone K, Carpenter JD, Seshadri V. Categorizing the unintended sociotechnical consequences of computerized provider order entry. Int J Med Inform 2007 Jun;76 Suppl 1:S21-S27.

[55] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005 Mar 9;293(10):1197-203.

[56] Baysari MT, Westbrook JI, Richardson KL, Day RO. The influence of computerized decision support on prescribing during ward-rounds: are the decision-makers targeted? J Am Med Inform Assoc 2011 Nov;18(6):754-9.

[57] Feldstein A, Simon SR, Schneider J, Krall M, Laferriere D, Smith DH, et al. How to design computerized alerts to safe prescribing practices. Jt Comm J Qual Saf 2004 Nov;30(11):602-13.

[58] Weingart SN, Massagli M, Cyrulik A, Isaac T, Morway L, Sands DZ, et al. Assessing the value of electronic prescribing in ambulatory care: a focus group study. Int J Med Inform 2009 Sep;78(9):571-8.

[59] Kortteisto T, Komulainen J, Makela M, Kunnamo I, Kaila M. Clinical decision support must be useful, functional is not enough: a qualitative study of computer-based clinical decision support in primary care. BMC Health Serv Res 2012;12:349-57.

[60] Krall MA, Sittig DF. Clinician's assessments of outpatient electronic medical record alert and reminder usability and usefulness requirements. Proc AMIA Symp 2002;400-4.

[61] Zachariah M, Phansalkar S, Seidling HM, Neri PM, Cresswell KM, Duke J, et al. Development and preliminary evidence for the validity of an instrument assessing implementation of human-factors principles in medication-related decision-support systems--I-MeDeSA. J Am Med Inform Assoc 2011 Dec;18 Suppl 1:i62-i72.

[62] Chused AE, Kuperman GJ, Stetson PD. Alert override reasons: a failure to communicate. AMIA Annu Symp Proc 2008;111-5.

[63] van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006 Mar;13(2):138-47.

[64] Khajouei R, Jaspers MW. The impact of CPOE medication systems' design aspects on usability, workflow and medication orders: a systematic review. Methods Inf Med 2010;49(1):3-19.

[65] Peute LW, Spithoven R, Bakker PJ, Jaspers MW. Usability studies on interactive health information systems; where do we stand? Stud Health Technol Inform 2008;136:327-32.

[66] Talmon J, Ammenwerth E, Brender J, de Keizer N, Nykanen P, Rigby M. STARE-HI -statement on reporting of evaluation studies in health informatics. Yearb Med Inform 2009;23-31.

[67] Peute LW, Driest KF, Marcilly R, Bras Da Costa S, Beuscart-Zephir MC, Jaspers MW. A Framework for reporting on Human Factor/Usability studies of Health Information Technologies. Stud Health Technol Inform 2013;194:54-60.

[68] Magrabi F, Ong MS, Runciman W, Coiera E. Using FDA reports to inform a classification for health information technology safety problems. J Am Med Inform Assoc 2012 Jan;19(1):45-53.

Chapter 3. Impact of usability flaws in medication alerting systems on usage and work system

Romaric Marcilly, Elske Ammenwerth, Erin Roehrer, Sylvia Pelayo, Francis Vasseur & Marie-Catherine Beuscart-Zéphir

Abstract

Introduction. Previous research has shown that medication alerting systems face usability issues. Previously there have been no attempts of systematically exploring the consequences of usability flaws in those systems for users (i.e., usage problems) and the work system (i.e., negative outcomes) and at providing a comprehensive synthesis of those consequences. This paper aimed to explore and synthesizing consequences of usability flaws in terms of usage problems and negative outcomes in the work system.

Method. A secondary analysis of the 26 papers included in a systematic review on the usability flaws reported in medication alerting was performed. Usage problems and negative outcomes in the work system were extracted and sorted along with their links with usability flaws.

Results. Results show that bad usability causes a large variety of difficulties. It impacts the user from a cognitive, a behavioral, an emotional and an attitudinal perspective. Ultimately usability flaws have negative consequences for the workflow, for the effectiveness of technology, for the medication use process and more importantly in terms of patient safety. There are only few converging lines of ongoing influences from usability flaws leading to negatives outcomes derived from the associations reported in the literature.

Discussion. The presence of usability flaws in medication alerting systems impedes the users of the system and ultimately their work system, including patient safety issues. Therefore considering usability of those systems along with their technical characteristics may participate in improving their impact.

Keywords

Human engineering; Decision support system, clinical; Review, systematic; Usability; Usage; Patient safety; Alerting functions

3.1. Introduction

Health Information Technology (HIT) is a promising tool to improve the efficiency, the effectiveness and the safety of healthcare [1]. Nonetheless, around 40% of HIT fail or are rejected [2]. Identified factors of failure include the system not meeting users' needs and poor interface specifications [2;3]. These problems refer to usability issues.

Usability is the "extend 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" [4]. Usability goes beyond the features of the graphical user interface (*e.g.,* legibility of the texts, layout and prompting of information) and includes how the system responds to users' actions, the organization/accuracy of the knowledge incorporated in the system and the availability of features required to support users' (cognitive) tasks.

Usability is considered as a critical component of effective and safe use of HIT [5]. For instance, the analysis of the incident report system by Magrabi and colleagues shows that 45% of incidents affecting patient safety originate from problems of usability nature [6]. An analysis of an incident report database of a large tertiary hospital by Samaranayake found that 17% of incidents reported are related to the use of technology and that many of them come from a poor usability [7].

Usability flaws (also known as usability problems, infractions or defects) refer to "aspect[s] of the system and/ or a demand on the user which makes it unpleasant, inefficient, onerous or impossible for the user to achieve their goals in typical usage situations" [8].

Usability is an intrinsic characteristic of the technology. However, due to the integration of the technology in the work system, the technology further interacts with the work system's components and therefore the usability also impacts those interactions. A work system is usually defined through five components: person, task, tools or technology, physical environment and organizational conditions [9]. The interactions between those components lead to outcomes in terms of performance, safety and health, and quality of working life [9].

The existence of usability flaws in the technology implemented will firstly impact, through its interaction with the user, the user and the tasks to perform. Those conscious or unconscious issues are referred to "usage problems". Then through the user, the other components of the work system will be impacted: those issues are referred to "negative outcomes". Figure 4 illustrates how the origin of usability flaws in the violation of usability design principles and the propagation of their consequences through the user up to the work system can be described.

The ongoing influence of the consequences of usability flaws is not causal and depends on several factors. Some factors, independent from the technology characteristics (*e.g.,* users' training experience and expertise, their workload, the fitting of their level of resilience according to the needs of the situation, their clinical skills) may either favor or mitigate the usability flaws ongoing influence. These factors impact the usability flaws ongoing influences at both levels of usage problems and outcomes. For instance, if clinicians are not trained with a technology and are overworked, it is unlikely that they can stop the impact

of the usability flaw: there is a risk of usage problems and of negative outcomes. On the contrary, if they are well-trained and have a regular workload, they may be able to stop this ongoing influence; the usability flaw may therefore have no negative consequences.

Scientific literature reports numerous case studies where the introduction of HIT in a work system leads to negative consequences for the clinicians in other components of the work system, including endangering patients [10-15]. However, to our knowledge, there is no attempt at exploring systematically the consequences of usability flaws in a HIT and at providing a comprehensive synthesis of those consequences.

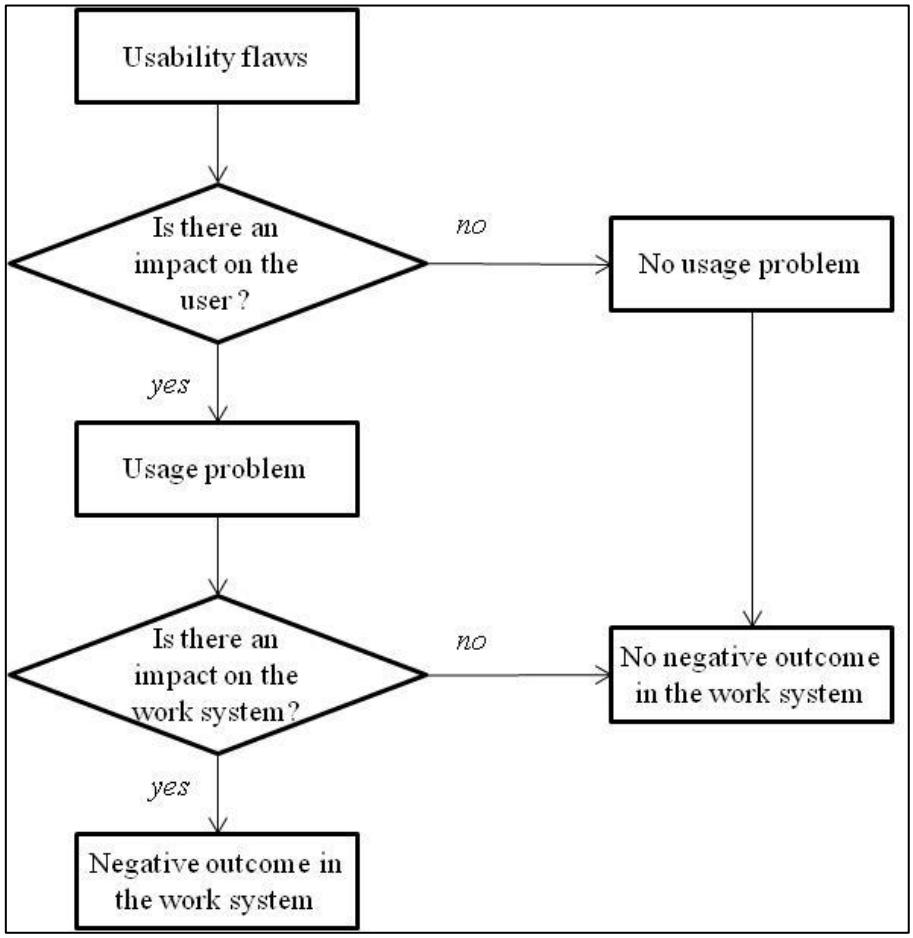


Figure 4. Schematic representation of the propagation of the usability flaws through the user up to the work system.

3.2. Study context

The present study focuses on medication alerting systems. There are many types of HIT and the promising impact such as increasing the safety of hospital drug management is not always observed [16; 17]. Their usability is often highlighted as a factor impeding their acceptance and implementation [18].

Table 7 - Summary of the general and specific types of usability flaws reported in the literature on medication alerting functions and their description. Names of the categories were adapted from [20] for general flaws and from [19] for specific flaws.

Type of usability flaws	Category	Description
General	Guidance issues	Prompting issues due for instance to unclear text, deficiency in information highlight. No visual distinction of different types and severity of alerts. Legibility issues. No feedback to inform the user (s)he just missed an alert. No grouping of same severity alerts.
	Workload issues	Too many actions for entering and obtaining information. Too much information in an alert and several alerts in the same window. Lack of concision.
	Explicit control issues	System's actions do not correspond to the action requested by the user. There is no way to undo an action.
	Adaptability issues	The alerting system does not support all user types.
	Error management issues	Problem messages are unclear.
	Consistency issues	Inconsistency of the behavior of the system for similar tasks and according to the type of data analyzed.
	Significance of codes issues	Non-intuitive wording and icons.
Specific	Low alerts' signal-to-noise ratio	Alerts may be irrelevant (regarding expertise/ward habits, existing good practices, pharmaceutical knowledge, data considered, patient case, actions engaged, clinician's interest for at risk situations, care logic) or redundant (appear very frequently/ several times during the decision making, clinically relevant solutions from the clinicians are not accepted, there is no feature to turn-off a specific alert).
	Alerts' content issues	Missing information (about actions that could be taken, patient's data, problem detected its evidence and its severity and information to interpret data within the alert) or erroneous proposed action according to the clinical context.
	System not transparent enough for the user	About the way it works: (no information on alert severity scale, on the up-to-dateness of alerts' rules) about the data it uses (every available data is not used to trigger the alert or incomplete mapping).
	Alert's appearance issues (timing and mode)	Alert appears before the decision process start, at the wrong moment or after the decision is made; alerts are not sufficiently intrusive or too intrusive; data processing is too slow.
	Tasks and control distribution issues	Alerts are not displayed to the right clinician or only to one clinician; users can enters comments on alerts that are displayed to no one; alert are not transferable from a clinician to another.
	Alerts' features issues	Missing features (to reconsider later an alert, from the alert there is no access to additional information nor action tool to solve the problem); existing features do suit users' needs.

In a previous study [19], we performed a systematic review that aimed at identifying and categorizing usability flaws reported in published studies on usability of medication alerting systems. This systematic review included 26 papers that reported:

- Original evaluation studies of medication alerting functions supporting e-prescribing; and
- Objective descriptions of usability flaws. Papers reporting evaluation of the perceived usability were excluded.

Based on an inductive content analysis, usability flaws were identified and classified. Details are described in [19]. Table 7 describes those categories.

In the present study we go one step further and perform a secondary analysis of the 26 papers included. This analysis aims at identifying and categorizing usage problems and negative outcomes reported by authors as being caused by those usability flaws. Two questions guide this analysis:

- What types of usage problems and negative outcomes originating in identified usability flaws are reported in usability studies of medication alerting functions?
- What are the cause-consequence links reported between usability flaws, usage problems and negative outcomes in medication alerting functions?

3.3. Method

The overall methodology is synthesized in Figure 5.

3.3.1. Method for the extraction of usage problems, outcomes and links

Usage problems and outcomes were collected in the 26 papers included in the aforementioned systematic review through independent readings by two experts (MCBZ and RM). All usage problems and outcomes that the authors of the 26 papers reported as being caused by a given usability flaw were searched. They were identified through the extraction of meaningful semantic units, *i.e.,* sets of words representing a single idea that is sufficiently self-explanatory to be analyzed. In order to get reliable data, only objective descriptions of usage problems and outcomes were extracted; hypotheses drawn by the authors of the studies were not analyzed.

Precisely:

- For usage problems, we targeted descriptions of how usability flaws impact the experience of the user interacting with the alerting function, including user's cognitive processes, behaviors and feelings.
- For negative outcomes, we targeted descriptions of the negative consequences of the usability flaws mediated through usage problems.

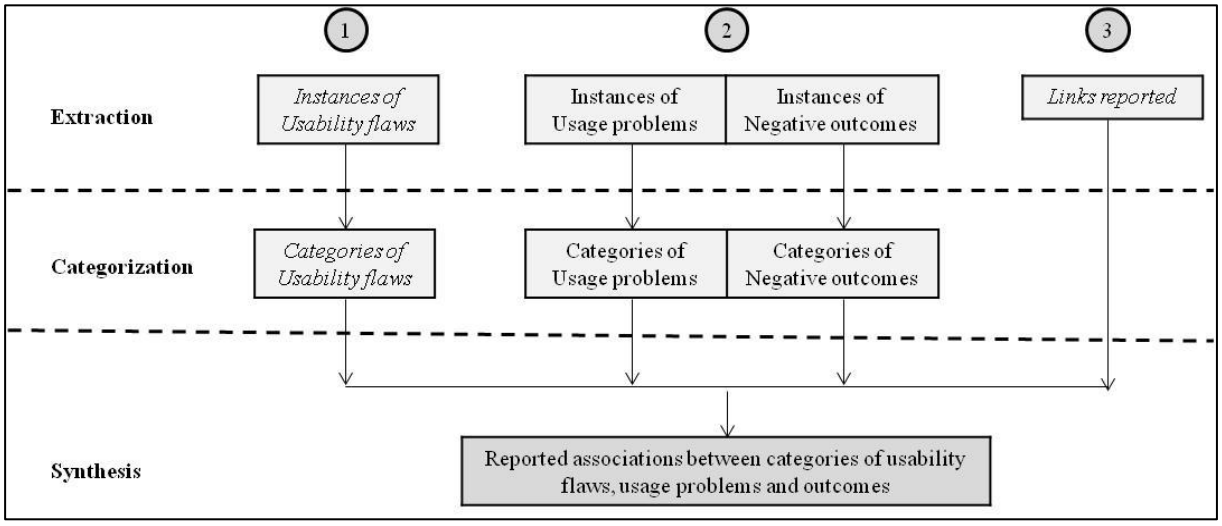


Figure 5. Graphical representation of the study plan. First, usability flaws were identified and categorized in the aforementioned systematic review [19]. Second, usage problems and negative outcomes were identified and categorized. Third, the links reported between these elements were extracted and used to synthesize the associations between categories of flaws, of usage problems and of negative outcomes.

Moreover, during the extraction process, three types of links reported by the authors were extracted: links between (a) usability flaws and usage problems, (b) links between usability flaws and negative outcomes and (c) links between usability flaws, usage problems and negative outcomes.

3.3.2. Method for the analysis of usage problems, outcomes and links

3.3.2.1. Question 1: types of usage problems and negative outcomes

Usage problems were categorized inductively through an open card sorting [21], a Human Factor method in which each item (here usage problems) is represented on a card that participants have to sort into logical categories and find a name for each category. It was performed independently by both experts (MCBZ and RM) on the whole set of meaningful semantic units representing usage problems. A conciliation meeting was organized to find an agreement on the number and titles of the categories of usage problems. During this meeting, internal consistency of categories and sub-categories was improved by developing new categories and sub-categories.

Outcomes were categorized through a closed card sorting [21]. In contrast to open card sorting, closed card sorting defines the number and the names of categories beforehand: participants have to sort cards into those categories. It was performed by both experts together (MCBZ and RM) on the whole set of meaningful semantic units representing negative outcomes. During the sorting process, experts discussed together to explain their choices. Four categories of negative outcomes were defined by both experts after a first reading of all negative outcomes extracted.

3.3.2.2. Question 2: links between usability flaws, usage problems and negative outcomes

All identified links between categories of usability flaws, usage problems and negative outcomes in the analyzed papers were then summarized. All complete links between usability flaws, usage problems and negative outcomes were analyzed along with links between usability flaws and usage problems only. When the link between usability flaw and outcome did not explicitly mention the mediating usage problem both experts, based on their experience in usability, drew independent inferences on the usage problem before discussing it together. Then, both experts assessed the plausibility of the links analyzed together. When information was missing to fully understand the cause-consequence link, this link was excluded.

3.4. Result

Out of the 26 analyzed papers, 21 report at least one instance of usage problems [10; 22-41] and 15 report at least one instance of an outcome [10; 22-33; 37; 42].

The 168 usability flaws extracted during the previous systematic review [19] were associated by the study authors with 111 usage problems and only 20 negative outcomes.

Table 8 - Categories of usage problems identified, illustrative instances and reference to the papers they are retrieved from. Users' comments are in italic font.

Usage problem	Illustrative instance from the 106 items	Reference
Behavioral issues		
Increased workload due to the alerting function	"If I have to consider every DDI, than I am busy with it, all day, and that is not my job." [32]	[22; 24; 27; 29; 31; 32; 37; 38; 40; 41]
Users do not use the system at all	"Two subjects did not use the decision support feature" [25]	[25; 28; 29]
Users voluntarily ignores the alerts	""Five nurses and two providers were observed to skip all or some of the reminders" [37]	[10; 22; 26; 28; 29; 32; 34-37; 40; 41]
Users ineffectively use the system	"The physician reported that specific features of the system (…) were hindering the use" [41]	[24; 29; 37; 41]
Users use workarounds	"Provider arbitrarily selected a date to satisfy the reminder" [37]	[25; 37; 38]
Users follow blindly the advice even if they do not understand it	MD clicks through [the alert] [accepts the advice without understanding the alert] [26]	[26]
Users are lost/stuck: they do not know how to go on	"Physicians were lost" [27]	[27; 37]
Cognitive issues		
Information involuntarily missed: they cannot access or find it	"Not having noticed the DDI alert that appeared as a second DDI alert" [32]	[10; 24; 29; 30; 32; 35]
Increased memory load while using the alerting system: users must rest on their memory	"Some prescribers relied solely on their memory of the patient profile" [29]	[29]
Users experience difficulties to understand the alert	"Had difficulty identifying the patient's risk factors for the interaction" [30]	[26; 27; 29; 30; 33; 41]

Usage problem	Illustrative instance from the 106 items	Reference
Users experience difficulties to identify alert's components (including icons, features or specific data)	"They misidentified the alert as a general guideline reminder and did not notice the dose calculations embedded in text." [25]	[25; 36; 37]
Users misinterpret of alerts' components (including icons, features or specific data)	"A user thought that the appearance of the "stamp" window implied that the patient had a chronic pain problem or diagnosis" [36]	[36;39]
Users misinterpret of alerts' content	"Misinterpretation was rife, as shown by the high numbers of wrong or inapplicable rules and reasoning." [32]	[23;32]
Users are interrupted by the alerts while making their decision or interviewing the patient	"There were several cases where inadequate alert design (…) disrupted their workflow."[22]	[22;24;29;34]
Emotional issues		
Annoyance/irritation	"Repetitive alerts are both annoying and unnecessary."[29]	[24; 26; 28; 29; 32; 41]
Frustration	"Physicians became frustrated" [27]	[22; 24; 27; 28]
Ugly experience	"Reading them is ugly" [22]	[22]
Stress, pressure	"Place prescribers under pressure"[29]	[29]
Cynicism	"This lack of information led to prescriber cynicism." [22]	[22]
Attitudinal issues		
Users question the behavior of the system: how the system is working, how it responds to users' actions	"Did it accept my changes?" [26]	[24-26; 29]
Users question the triggering and sorting model of the alerts	"I am not confident it's checking all the interactions that I want it to check." [29]	[22; 26; 29; 38]
Users question the usefulness of the alerting system	"The alerts were most likely to be helpful if they [were] presented when the users were entering orders or were otherwise at the point of making a decision about the issue in question or closely related issues." [24]	[24; 33; 36]
Users question the validity of the alert	"That's not true to my knowledge. The patient doesn't like to take it; I doubt he's taking it [from a non-VA source]. I will talk to the patient about it." [31]	[22; 29; 31; 33; 34]
Users experience alert fatigue/desensitization	"Some doctors recognized that they had become desensitized to the alerts."[40]	[22; 29; 31-34; 37; 40]
Users have negative feelings towards the system	"Justification requirement often viewed as time burden" [29]	[24; 29; 35; 36]

3.4.1. Question 1: types of usage problems and negative outcomes

3.4.1.1. Usage problems

The complete list of usage problems is provided in appendix 3. The open card sorting resulted in identifying 15 or 23 categories depending on the reviewer; however, the same themes were highlighted in

both categorizations (cf. appendix 3). After a conciliation meeting, the classification scheme was finalized to four main categories; behavioral, cognitive, emotional and attitudinal issues. These four main categories can be divided into 25 subcategories. It is fully described with instances in Table 8.

3.4.1.2. Negative outcomes

The complete list of negative outcomes is provided in appendix 3. Four categories were defined:

- "Workflow issues" include instances related to the increase of the communications between clinicians and with the patient, along with one instance of shift in responsibility of the alerts from the physicians to the pharmacists.
- "Technology effectiveness issues" include instances such as the non achievement of the expected gain in speed.
- "Medication management process issues" include instances of slowing down this process.
- "Patient safety issues" include instances of errors in ordering. However, no lethal consequence is reported even if once the dose of aspirin ordered is doubled involuntarily by the physician [26].

No disagreement was observed between both experts during the closed card sorting. A synthesis of data categorization is provided in Table 9.

Table 9 - Categories of negative outcomes identified, their description, illustrative instances and reference of the papers they are retrieved from.

Issue	Description	Illustrative instance from the 19 items	Study
Workflow issues	The communication between the clinicians and the patient are increased. Moreover, the responsibility of the alert may also be shifted.	"House staff claimed post hoc alerts unintentionally encourage house staff to rely on pharmacists for drug allergy checks, implicitly shifting responsibility to pharmacists." [10] "Pharmacists call house staff to clarify questionable orders"[10]	[10; 22; 26; 29; 31]
Technology effectiveness issues	The expected usefulness of the technology to manage the care is not noticed.	"Consequently they did not derive all the speed and accuracy benefit and did not reduce their cognitive effort the feature was in part designed to."[25]	[23-25; 29; 33; 37]
Medication management process issues	The efficiency of the medication management process is bothered by the use of the alerting system	Problems experiences with the alerting system "slowed down [users'] work" [28]	[22; 27; 28]
Patient safety issues	The use of the alerting system produces conditions for decreasing the quality of care and even endanger the patient.	"MD goes back to the medication list. Aspirin is now listed both under VA list and non-VA medication list" [double order of aspirin] [26]	[26; 30; 32; 42]

3.4.2. Question 2: links between usability flaws, usage problems and negative outcomes

Forty-seven complete cause-consequence links between usability flaws, usage problems and negative outcomes are reported. In addition, 129 links between usability flaws and usage problems with no mention of related outcomes are reported. There are also 6 links between usability flaws and outcomes for which mediating usage problems had to be inferred leading to a total of 53 cause-consequence links.

3.4.2.1. Usability flaws linked to usage problem

A total of 182 links between usability flaws and usage problems were synthesized including 129 links between just flaws and problems and the 53 associations between flaws and problems described in the 53 complete chains of links.

As shown in Table 10, all categories of usability flaws cause usage problems. A total of 81 different associations between categories of usability flaws and categories of usage problems are identified. Almost all types of flaws are not specific to one type of usage problems, with up to 14 types of usage problems linked to one category of usability flaw. Only two types of flaws, namely "consistency issues" and "error management issues" lead to only one type of usage problem.

Overall, the type of usability flaws (*i.e.,* general vs. specific) does not appear to lead to exactly the same type of usage problems:

- Specific types of usability flaws are linked to more attitudinal usability problems than the general usability flaws (15 links between categories vs. 5);
- General types of flaws are linked to a larger range of cognitive issues than specific ones (13 vs. 8);
- Both types of flaws are related more or less equally to behavioral and emotional issues (respectively 13 for general problems vs. 15 for specific problems for behavioral issues and 6 vs. 6 for emotional issues).

3.4.2.1. Complete chains

The 53 complete chains between categories of usability flaws, usage problems and negative outcomes represent 42 different types of associations. They are summarized in Table 11. Few clear schemes of the usability flaws ongoing influence appear.

All categories of usability flaws apart from "consistency" issues cause negative outcomes through usage problems. Emotional and attitudinal usage problems are never related directly to negative outcomes.

Six types of usability flaws are associated to only one type of negative outcome: "significance of codes issues" and "alert content issues" are associated only to "medication management process issues". "Error management issues" "adaptability issues" and "tasks and control distribution issues" are associated only to "workflow issues". "Explicit control issues" is only associated to "technology effectiveness issues". On

the contrary, "workload issues" are linked to all four types of outcomes. "Guidance issues" and "low signal-to-noise ratio" are associated to three types of negative outcomes.

Table 10 - Associations between usability flaws and usage problems. Black boxes represent that the link between related flaw and problem is reported in the literature review.

			Categories of usability flaws													
			Specific						General							
			Low alerts signal-to-noise ratio	Alert content issues	Alert's presentation issues	Transparency issues	Alerts' features issues	Tasks and control distribution issues	Workload issues	Guidance issues	Significance of codes issues	Adaptability issues	Explicit control issues	Consistency issues	Error management issues	Number of links with different categories
Categories of usage problems	Behavioral	Users ineffectively use the system	■	■		■		■		■		■				6
		Increased workload	■	■					■	■	■					5
		Users voluntarily ignore the alerts	■	■	■				■	■						5
		Users use workarounds		■	■		■		■				■			5
		Users are lost / stuck	■	■						■			■			4
		System not used at all	■						■							2
		Users blindly follow the alert								■						1
	Cognitive	Difficulties to understand the alert		■					■	■	■	■			■	6
		Information missed involuntarily			■		■	■	■	■						5
		Misinterpretation of alerts' content				■			■	■						3
		Users are interrupted by the alerts	■		■				■							3
		Difficulties to identify alerts' components							■		■					2
		Increased memory load		■												1
		Misinterpretation of alerts' components									■					1
	Emotional	Annoyance / irritation	■		■				■	■						4
		Frustration	■		■					■	■					4
		Ugly experience							■	■						2
		Cynicism		■												1
		Stress			■											1
	Attitudinal	Users question the validity	■	■		■			■							4
		Users question the triggering/sorting model	■	■		■								■		4
		Users question the usefulness	■	■	■											3
		Users question the behavior	■			■	■									3
		Alert fatigue / desensitization	■						■	■						3
		Users have negative feelings	■		■				■							3
Number of links with different categories			14	11	9	5	3	2	14	12	5	2	2	1	1	81

Patient safety issues mainly reside in workload and guidance issues but can also appear in transparency issues. The main usage problems associated to patient safety issues are those dealing with the understanding of the alert (misinterpretation, information missed, alert not understood but blindly followed). Medication management process issues are also caused mainly by guidance and alert content issues.

Table 11 - Synthesis of the 53 complete links between usability flaws (column), usage problems (row) and negative outcomes (cell) categories. Emotional and attitudinal usage problems are not represented because they are never related directly to negative outcomes. Negative outcomes' acronyms: W, workflow issues; T, technology effectiveness issues; M, medication management process issues; and P, patient safety issues.

			Categories of usability flaws													
			Specific						General							
			Low alerts signal-to-noise ratio	Alert content issues	Alert's presentation issues	Transparency issues	Tasks and control distribution issues	Alerts' features issues	Workload issues	Guidance issues	Significance of codes issues	Adaptability issues	Explicit control issues	Consistency issues	Error management issues
Categories of usage problems	Behavioral	Users ineffectively use the system				W				T		W			
		Increased workload	W	M					T	M					
		Users use workarounds			T										
		Users voluntarily ignore the alerts	M T	M	W				T						
		Users are lost/stuck	T	M						M			T		
		Users do not use the system at all							T						
		Users blindly follow the alert								P					
	Cognitive	Difficult to understand the alert		M W					T W M P	T M P	M	W			W
		Information missed involuntarily					W	T	P	T P					
		Users are interrupted by the alerts			T				T						
		Misinterpreation of alerts content				P			P	T P					
		Difficult to identify alerts' components													
		Increased memory load													
		Misinterpretation of alert components													

3.4.2.2. A journey through using a badly designed medication alerting system

Categorizing usage problems and negative outcomes enables a synthetic representation of the diversity of usability flaws' negative consequences on the users and the work system. However, categorizing mutes the actual experience of real clinicians: it camouflages the actual impact of the flaws on the user and the work system. Therefore the main results were gathered together in order to help understand what users are actually experiencing while using badly designed alerting system and what their consequences in the work system are. During the actual interaction of the user with an alerting system, all types of usability flaws, usage problems and outcomes are tightly intertwined. In order to present usability flaws, usage problems and negative outcomes in an integrative manner, a presentation according to clinicians' possible interactions with the alerting system has been preferred over following the aforementioned categorizations.

3.4.2.2.1. Trying to interact with the alerting system

Alert appearance and information gathering

Various problems in the presentation of the alert such as its timing, its mode (*i.e.,* intrusive or non intrusive alert) and the requirement to use the scrolling bar to see the information, prevent the user from getting the information when they need it. Additionally, clinicians may miss the whole alert. For instance, the physician does not see the alerts because they are not sufficiently noticeable [24]. Another reason for missing the alert is the moment of its appearance is too late according to clinicians' needs. For instance, the physician receives the alert once the patient is out the examination room [35]. When the alert appears too late, according to the moment the clinicians need it, they have to perform by themselves the operation the alerting system is supposed to do. For instance, "six subjects computed, estimated or used heuristic to get the dose amount at some point before the system-calculated dose presentation" [25].

Clinicians miss also the alert by accident. Indeed, it has been observed that they also "unintentionally override[s]" an alert because it appears in place of another alert that has just been dismissed [32]. Clinicians do not notice this change and dismiss the second alert too, thinking the system has not taken into account their former action. In addition, some alerting systems do not provide users with the opportunity to display the alert a second time therefore, they do not have a second chance to read it [32] and they may ultimately "forgot what alert(s) appeared" [29].

Even when clinicians see the alert, usability issues prevent them from finding the information they need regarding the alert. For instance, users "misidentified the alert as a general guideline reminder and did not notice the dose calculations embedded in text" [25]. In a simulation study, half of the participants missed the information about the duration of the patient's therapy in the alert even when this information is important for the clinical decision [30]. In the same study, users even made wrong clinical decisions because they miss patient's risks factors that were hidden in a tab.

Understanding alert's information

Even when the alert is accessible at the right time for the user and the information is seen, other problems arise. The alerts' language "which did not adequately support all prescriber types (…) is difficult for prescriber to interpret" [26; 29]. As one user states "it's hard to see what [the alert] is trying to tell you" [26]. Alerts are "not understandable by physicians" [27], "'difficult to interpret in content and purpose" [28], "precluding [the users] to understand the problem that generated the alert or how to solve [it]" [27].

Those understanding difficulties prevent from optimally using the alerting functions and prompt clinicians to ask for help: ""physicians often come and ask about an alert triggered by the combination or amiodarone and simvastatin" says a pharmacist, "the doctors don't know what the order check really means"" [22]. It is sometimes necessary for nurses and physicians to have "real time, face-to-face communication with clinical pharmacists" [29]. Conversely, sometimes "pharmacists call house staff to clarify questionable orders" because they have been alerted by the pharmacy information system about

clinical issues that physicians' and nurses' information system do not provide [30]. Consequently, those issues "slowed down their work" [28].

These difficulties in understanding the content of the alert(s) also lead to misinterpretations of the alerts. Numerous problems have been observed "as shown by the high number [of alerts handled incorrectly]" [32]. For instance, the directive in the alert explaining that clinicians have to leave a comment (*e.g.,* reason for not adhering to the system's suggestion) is sometimes so hard to interpret that users do not understand it and do not leave any comment leading to a "relatively high proportion of content free comments" [23]. There are also more severe instances in which patient safety could have been endangered. Indeed, misinterpretation issues cause "respondents [to make] a wrong selection [of drugs], because they trusted the alerting system (and followed the incorrect dose recommendation for an unfamiliar drug)" [32]. In another instance, an alert appears that says "duplicate drug order. Non-VA ASPIRIN. [Alerts] mentions 325mg…MD is looking at it also and [appears] confused. MD to observer (Obs): "What's it going to do? Is it going to switch the patient to 325mg?"" [26]. The clinicians are not sure of the meaning of the alert: the alert is actually informing them that there is a duplicate order of aspirin but they believe that there was only an automatic change of dosage and so the clinician "clicks [it] through" [26], mistakenly validating two orders of aspirin ("aspirin is now listed both under VA list and non-VA list" [26]). Another study explains that in a Danish hospital, "there are various examples of complex registrations that lead to medication errors" that are due to the fact that clinicians do not know precisely how alerts are triggered. Therefore they have to infer, sometimes wrongly, the cause of alerts' appearance and their clinical decision may be wrongly based [42].

Understanding difficulties are not only related to the content of the alert: while they are designed to help clinicians handle the alert efficiently, icons and titles are also a cause of alert's misinterpretations: "three providers misinterpret this question mark" [37], "several users (…) did not realize [the arrows under the clinical recommendations] provided additional more detailed information about the basic recommendation when clicked on" [36]. Another instance: "the appearance of the 'stamp' window implied that the patient had a chronic pain problem or diagnosis. In actually, the 'stamp' indicated that the patient had a scheduled appointment (…) and that ATHENA-OT had recommendations available." [36]. And also, "two participants misinterpreted the meaning of "when" to represent the last time the current patient received the intervention instead of the frequency the intervention is due for all patients" [39].

Memory-load and workload are increased

The lack of information on the patient in the alerts requires clinicians to rely "solely on their memory of the patient profile" [29] or at making "assumptions about patient history" [29]. In addition, the wrong setting of alerts leads also to disruptions of clinicians' workflow [22], of their "thought process" [24] making harder the decision making process and oblige them to concentrate more not to lose the track of their thought. Those recurrent interruptions ultimately hinder alert effectiveness [29].

At a behavioral level, the overall poor quality of alerts' message, their repetition and their length compel the clinicians to waste "time searching information in the [Electronic Health Record]" [29], scrolling down [32] and to read it [40]. They also "resort to trial-and-error behavior exemplified by the extra mouse clicks and keystrokes they needed for locating and executing the right action in response to the message" [27].

This increased workload also comes from the alerts' documentation tasks (especially for clinical reminders). Indeed clinicians experience "double documentation" burdens since, besides alert's documentation, they generally keep track of this information outside the alerting system [37]. The high demand on documentation leads also clinicians to satisfy alerts/reminders once "the patient has left the room", "after the clinic close" and even to delegate this task to "case mangers" [38].

Finally, the absence of features to share the information provided by the alerting system also requires the user to utilize "paper-based workarounds" to share it with other clinicians [37], *i.e.,* they have to copy the information of interest on a paper to share it with their partners.

Lost in interaction

Overall, the usage of the alerting system is greatly hindered in terms of efficiency by usability issues due to the system features [37; 41]. The inability of the alerting system to efficiently support users' cognitive activities and their clinical tasks is apparent when users are stuck in their use because there is no appropriate option available to satisfy the alert [37]. To go on despite this dead-end, clinicians develop workarounds behaviors: one clinician has been observed to "arbitrarily select[ing] a date to satisfy the reminder" because none of the options within the dialogue box match her/his intentions; another one "had to leave the reminder unsatisfied" [37].

The interface also impedes "prescribers' ability to act on alerts" [29]. For instance, the user does "not always seem to understand how to use and manage the alerts effectively" leading to "some unnecessary repetitions of alerts" [24]. Similarly, when the user is unable to satisfy an alert or reminder because of response choices provided by the system "the [clinical reminders] (…) continues to appear" [37]. When the user wants to cancel a clinical reminder without losing the data entered previously in the reminders, they "select[s] each [reminder] individually from the list rather than using the Next button to navigate through a sequence of [reminders]" [37]. The usability problem at the root of this behavior "introduces the possibility of losing data previously inputted" [37].

In summary, the improvement of the medication management process that users have the right to expect from the alerting system is not observed because the system actually impairs the ordering efficiency by increasing their workload [22; 27]. Clinicians "did not derive all the speed and accuracy benefit and did not reduce their cognitive effort the [alerting system] was in part designed to" [25].

3.4.2.2.2. Emotional and long term consequences

Users emotionally react

Daily facing poorly designed alerting systems is not the only impact on the cognitive processes and the behavior of the clinician. Users are also emotionally affected. It is unpleasant to read alerts [22] with display issues. The delay before the alerts appear "place[s] the prescribers under pressure" [29] because they are supposed to quickly make an informed decision. Unsurprisingly "the lack of information [in the alert] led to prescriber cynicism" [22]. Users become frustrated [22], overwhelmed [29] and are irritated with the repetitive appearance of the same alert: "same alert appears a 3rd time when [nurse practitioner] goes to sign the order. [Nurse practitioner] gestures to the screen, "See – three times!"" [29]. The high number of alerts has been reported by users to "drive you mad" [32]. In addition, alerts that appear again and again in spite of actions of the user to cancel them (*e.g.,* by modifying the order) may "freak someone out" [26].

Over time, the repetitive everyday use, several times a day, of the alerting system with usability flaws impacts also the attitude of the user towards the alerting system. It is foreseeable that there are "numerous complaints about getting too many alerts or alerts at an inappropriate time" [35]. Users even "complain vociferously" [24] that "it is hard to use the tool" [36]. Those complaints are completely understandable but the impact on the user is unfortunately deeper than complaints initially indicate.

Skipping alerts

The low signal-to-noise ratio of alerts compels clinicians to repetitively dismiss numerous alerts. This creates alert fatigue and desensitizes users, *i.e.,* they lose interest in the alerts. Authors noticed numerous "remarks [of the users] suggesting alert fatigue" [32]. Clinicians themselves recognized that "they had become desensitized to the alerts" [40]. One of them explains that there are "too many things popping at [him]" [34]. The alert fatigue impacts everyone since "even prescribers with a very positive view of the alert system showed signs of desensitization" [29]. In turn, alert fatigue causes voluntarily ignorance of the alerts, resulting in the clinicians overriding the alert [29]. There are numerous descriptions of how prescribers "rapidly [override] these alert types once they recognized that they had seen the alert before" [28]. Some users note that "it's gotten to the point that [they] don't hardly look at significant (interactions) anymore" [26]. They are "often inclined to rapidly click [the alerts] away (…) [to] simply skip them" [32]. If there's more that one [alert in the popup window], [they] don't read through them all" [22]. They "click off by rote and [risk] not see something that is different" [26]; they develop "a sort of mechanism" to dismiss the alerts [40]; one clinician explains that she had "memorized the location of the override button" for these situations [28]. To summarize this point, a clinician says that "once you realize that most of the information is useless or superfluous or not relevant, you stop looking at it" [34] and they rely instead on their "own clinical judgment" [29]. This unnecessary redundancy of alerts, by developing alert fatigue and alert's ignorance, ultimately "imped[ed] the medication ordering process" [22] and "lead[ing] to low

response levels to the alerts" [33]. Moreover, alerts that are ignored because of their wrong timing "encourage house staff to rely on pharmacists for drug allergy checks, implicitly shifting responsibility to pharmacists" [10].

Loss of confidence

Desensitization process goes along with a loss of confidence in the alerting function. One clinician explains "I see it does say 'active' though. Technically, the [old] medication [order] isn't 'active' because I just changed them to discontinued" [26]. Logically, the clinician wonders: "did it accept my changes?" [26]. They have doubts about the way the alerting system is working. In other cases, the clinicians are "unsure if the pharmacists review" the override justifications they entered in the system [29]. Even for the management of alert, clinicians are "uncertain how long the reminders would be turned off" [38]. Clinicians express the same kind of doubts regarding the triggering of alerts by the system: they are not sure "that the system based its recommendation on the same assumptions that [he] would have made" [25]. They say also that they are "not sure why [the alert] didn't come up this time". They are even not sure if the "order check system automatically check when [they] order medications". Moreover, the clinical validity of the alerts is also questioned because "it was unclear if the warnings were "evidence-based"" [22] and because the system provides clinicians with information which they think is not right or not updated: clinicians express doubts "on whether the system has up-to-date information" [33] and, in another case, "upon seeing [a duplicate order alert about two orders of iron], the physician stated, "that's not true to my knowledge. The patient doesn't like to take it; I doubt he's taking it"" [31]. The physician was compelled to ask the patient whether he is obtaining iron from an outside source or not [31]: using the alerting system changed his communications.

Ultimately, clinicians question seriously the usefulness of the alerting system, considering it as unhelpful [24]. They say "it's just crying wolf" [34]. Their "perceptions of the credibility and trustworthiness of the alert system" [29] is negatively impacted by the poor usability of this system. In summary, poorly designed alerting function in terms of usability encourages clinicians not to use the alerting system [38].

The instances of usage problems and negative outcomes take on their full meaning in this journey through using a badly designed medication alerting system. Altogether, those actual consequences of usability flaws draw up a negative report about alerting systems that are deficient in terms of usability. The existence of usability flaws in an alerting system used daily actually negatively impacts the users, their cognitive activities, their behaviors and their feelings. Ultimately, those usability flaws hinder also other work system's components, potentially endangering the patient.

3.5. Discussion

The present study aimed at answering two questions: "What types of usage problems and negative outcomes originating in usability flaws are reported in medication alerting functions?" And "What are the cause-consequence links reported between usability flaws, usage problems and negative outcomes in medication alerting functions?"

Results show that the consequences of usability flaws are various and are clearly identified in the literature and that they concern both the user and the other components of the work system. Indeed, 111 usages problems along with 20 negatives outcomes are identified in the 26 papers. Usage problems can be categorized into four main categories, "cognitive", "behavioral", "emotional" and "attitudinal", which can be further subdivided in 25 subcategories. For negative outcomes, four categories were used: "workflow issues", "technology effectiveness issues", "medication management process issues" and "patient safety issues".

As for the second question, 129 links between usability flaws and usage problems are reported along with 53 links between usability flaws, usage problems and negative outcomes. Even if some trends arise that highlight the important roles of workload issues, guidance issues and low signal-to-noise ratio issues along with the role of information understanding, there are only few converging lines of ongoing influences from usability flaws resulting in negatives outcomes. The clearest line is the absence of ongoing influence from attitudinal and emotion problems to outcomes. This result can be explained by the fact that feelings cannot directly impact the work system: they need to be mediated through a decision and/or a behavior. There is no other definitive clear association that appears between categories of usability flaws, categories of usage problems and categories of negative outcomes.

The instances collected, once put in relation, restitute the difficulties and even the pain the clinicians may experience while using a badly designed alerting system. This "journey" collects all the worst of observed usability flaws and their consequences. This list of actual concrete illustrations of the consequences of the usability flaws could be used to make designers and/or project managers more aware of the importance of considering usability during the design process of HIT.

Publication and selective reporting biases in the analyzed papers may have impacted the comprehensiveness and representativeness of usage problems and negative outcomes reported. Moreover, the papers analyzed in this review were selected because they present usability flaws not because they report usage problems and negative outcomes. This inclusion criterion may explain the small number (n = 20) of negative outcomes retrieved from the analyzed papers and ultimately the few propagation lines observed. Therefore, the results presented here must be handled carefully. They do not represent exhaustively the usage problems and negative outcomes that are related to medication alerting systems. They rather present the usage problems and negative outcomes that are caused by usability flaws according to the authors of the studies analyzed.

Of course, this review should be regularly up-dated with new insights from publications. To improve the collection of usability data, researchers should follow reporting guidelines [43] and take advantages of online appendices to publish complete sets of usability results. Other sources of usability data should also be explored such as incident report systems that include appropriate descriptions of the usability flaws, usage problems and negative outcomes [6; 7; 44].

Several studies have shown that implementing alerting functions actually participate in the improvement of the medication management safety [45] by improving, for instance, medication dosing [46], antibiotic use [47], clinical practice [48; 49]. And, when associated with Computerized Physician Order Entry (CPOE), they contribute to enhance the healthcare quality and safety [50]. However, those benefits are not always observed [16; 17]. Considering all reported usability issues one may reasonably think that a poor usability of those technologies is partly responsible either of reducing their impacts or of preventing those impacts' appearance. Several studies have recently shown that improving the usability of HIT (*e.g.,* CPOE or alerting systems) by applying usability design principles or by following a user-centered design process improves the efficiency of the technology [51; 52], reduces users' workload [52], and increases users' satisfaction [52; 53]. Nonetheless, usability is not the only technology characteristic that may negatively impact users' experience and outcomes. For instance, Magrabi and colleagues noticed that technical issues such as configuration issues, access oravailability issues or data capture issues may also ultimately endanger patient safety [44]. Therefore, other characteristics of the technology should be considered besides its usability to fix problems experienced by the clinicians and to prevent negative outcomes. Even so, improving alerting systems' usability may also participate to decrease usage problems and negative outcomes in the work system. The appendices published in [19] aim precisely at that.

3.6. Conclusion

This paper aimed at exploring and synthesizing consequences of usability flaws in terms of usage problems and negative outcomes. We performed a secondary analysis of 26 papers included in a systematic review on the usability flaws reported in medication related alerting functions. Results show that bad usability actually impacts the user and the work system in various ways. Only few lines of propagation of usability flaws in the work system through the user can be drawn along with noticeable tendencies. Results highlight nonetheless the large variety of difficulties that the users may experience using an alerting system badly designed and their other consequences. Improving the usability, along with technical issues, of the alerting systems may contribute to improving the expected positive clinical impact of this promising technology.

References

[1] To err is human, building a safer health system. Koln L, Corrigan J, Donaldson M, editors. 2000. Washington, D.C., National Academic Press.

[2] Kaplan B, Harris-Salamone KD. Health IT success and failure: recommendations from literature and an AMIA workshop. J Am Med Inform Assoc 2009 May;16(3):291-9.

[3] McManus J, Wood-Harper T. Understanding the sources of information systems project failure. Manag Serv 2007;38-43.

[4] International Standardization Organization. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability (Rep N° 9241-11). Geneva: International Standardization Organization; 1998.

[5] Middleton B, Bloomrosen M, Dente MA, Hashmat B, Koppel R, Overhage JM, et al. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. J Am Med Inform Assoc 2013 Jun;20(e1):e2-e8.

[6] Magrabi F, Ong MS, Runciman W, Coiera E. An analysis of computer-related patient safety incidents to inform the development of a classification. J Am Med Inform Assoc 2010 Nov;17(6):663-70.

[7] Samaranayake NR, Cheung ST, Chui WC, Cheung BM. Technology-related medication errors in a tertiary hospital: a 5-year analysis of reported medication incidents. Int J Med Inform 2012 Dec;81(12):828-33.

[8] Lavery D, Cockton G, Atkinson M. Comparison of evaluation methods using structured usability problem reports. Behaviour and Information Technology 1997;16(4):246-66.

[9] Carayon P, Schoofs HA, Karsh BT, Gurses AP, Alvarado CJ, Smith M, et al. Work system design for patient safety: the SEIPS model. Qual Saf Health Care 2006 Dec;15 Suppl 1:i50-i58.

[10] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005 Mar 9;293(10):1197-203.

[11] Kushniruk A, Beuscart-Zephir MC, Grzes A, Borycki E, Watbled L, Kannry J. Increasing the safety of healthcare information systems through improved procurement: toward a framework for selection of safe healthcare systems. Healthc Q 2010 Sep;13 Spec No:53-8.

[12] Weiner JP, Kfuri T, Chan K, Fowles JB. "e-Iatrogenesis": the most critical unintended consequence of CPOE and other HIT. J Am Med Inform Assoc 2007 May;14(3):387-8.

[13] Campbell EM, Sittig DF, Ash JS, Guappone KP, Dykstra RH. Types of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc 2006 Sep;13(5):547-56.

[14] Kushniruk A, Triola M, Stein B, Borycki E, Kannry J. The relationship of usability to medical error: an evaluation of errors associated with usability problems in the use of a handheld application for prescribing medications. Stud Health Technol Inform 2004;107(Pt 2):1073-6.

[15] Nanji KC, Rothschild JM, Boehne JJ, Keohane CA, Ash JS, Poon EG. Unrealized potential and residual consequences of electronic prescribing on pharmacy workflow in the outpatient pharmacy. J Am Med Inform Assoc 2014 May;21(3):481-6.

[16] Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. JAMA 1998 Oct 21;280(15):1339-46.

[17] Ranji SR, Rennke S, Wachter RM. Computerised provider order entry combined with clinical decision support systems to improve medication safety: a narrative review. BMJ Qual Saf 2014 Apr 12.

[18] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004 Mar;11(2):104-12.

[19] Marcilly R, Ammenwerth E, Vasseur F, Beuscart-Zephir MC. Usability flaws of medication-related alerting systems: a systematic review. IJMI (submitted). In press 2014.

[20] Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behaviour and Information Technology 1997;6(4-5):220-31.

[21] Sebillotte S. Action representation for home automation. North-Holland: Elsevier Science Publishers B.V.; 1990 p. 985-90.

[22] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. A human factors investigation of medication alerts: barriers to prescriber decision-making and clinical workflow. AMIA Annu Symp Proc 2009;548-52.

[23] Chused AE, Kuperman GJ, Stetson PD. Alert override reasons: a failure to communicate. AMIA Annu Symp Proc 2008;111-5.

[24] Krall MA, Sittig DF. Clinician's assessments of outpatient electronic medical record alert and reminder usability and usefulness requirements. Proc AMIA Symp 2002;400-4.

[25] Horsky J, Kaufman DR, Patel VL. Computer-based drug ordering: evaluation of interaction with a decision-support system. Stud Health Technol Inform 2004;107(Pt 2):1063-7.

[26] Russ AL, Saleem JJ, McManus MS, Zillich AJ, Doebbling BN. Computerized medication alerts and prescriber mental models: observing routine patient care. 2009 p. 655-9.

[27] Khajouei R, Peek N, Wierenga PC, Kersten MJ, Jaspers MW. Effect of predefined order sets and usability problems on efficiency of computerized medication ordering. Int J Med Inform 2010 Oct;79(10):690-8.

[28] Feldstein A, Simon SR, Schneider J, Krall M, Laferriere D, Smith DH, et al. How to design computerized alerts to safe prescribing practices. Jt Comm J Qual Saf 2004 Nov;30(11):602-13.

[29] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. 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 2012 Apr;81(4):232-43.

[30] Duke JD, Bolchini D. A successful model and visual design for creating context-aware drug-drug interaction alerts. AMIA Annu Symp Proc 2011;339-48.

[31] Russ AL, Saleem JJ, McManus MS, Frankel RM, Zillich AJ. The Workflow of Computerized Medication Ordering in Primary Care is Not Prescriptive. 2010 p. 840-4.

[32] van der Sijs H, van GT, Vulto A, Berg M, Aarts J. Understanding handling of drug safety alerts: a simulation study. Int J Med Inform 2010 May;79(5):361-9.

[33] Wipfli R, Betrancourt M, Guardia A, Lovis C. A qualitative analysis of prescription activity and alert usage in a computerized physician order entry system. Stud Health Technol Inform 2011;169:940-4.

[34] Weingart SN, Massagli M, Cyrulik A, Isaac T, Morway L, Sands DZ, et al. Assessing the value of electronic prescribing in ambulatory care: a focus group study. Int J Med Inform 2009 Sep;78(9):571-8.

[35] Ash JS, Sittig DF, Dykstra RH, Guappone K, Carpenter JD, Seshadri V. Categorizing the unintended sociotechnical consequences of computerized provider order entry. Int J Med Inform 2007 Jun;76 Suppl 1:S21-S27.

[36] Trafton J, Martins S, Michel M, Lewis E, Wang D, Combs A, et al. Evaluation of the acceptability and usability of a decision support system to encourage safe and effective use of opioid therapy for chronic, noncancer pain by primary care providers. Pain Med 2010 Apr;11(4):575-85.

[37] Saleem JJ, Patterson ES, Militello L, Render ML, Orshansky G, Asch SM. Exploring barriers and facilitators to the use of computerized clinical reminders. J Am Med Inform Assoc 2005 Jul;12(4):438-47.

[38] Patterson ES, Nguyen AD, Halloran JP, Asch SM. Human factors barriers to the effective use of ten HIV clinical reminders. J Am Med Inform Assoc 2004 Jan;11(1):50-9.

[39] Saleem JJ, Patterson ES, Militello L, Anders S, Falciglia M, Wissman JA, et al. Impact of clinical reminder redesign on learnability, efficiency, usability, and workload for ambulatory clinic nurses. J Am Med Inform Assoc 2007 Sep;14(5):632-40.

[40] Baysari MT, Westbrook JI, Richardson KL, Day RO. The influence of computerized decision support on prescribing during ward-rounds: are the decision-makers targeted? J Am Med Inform Assoc 2011 Nov;18(6):754-9.

[41] Kortteisto T, Komulainen J, Makela M, Kunnamo I, Kaila M. Clinical decision support must be useful, functional is not enough: a qualitative study of computer-based clinical decision support in primary care. BMC Health Serv Res 2012;12:349-57.

[42] Hartmann Hamilton AR, Anhoj J, Hellebek A, Egebart J, Bjorn B, Lilja B. Computerised Physician Order Entry (CPOE). Stud Health Technol Inform 2009;148:159-62.

[43] Peute LW, Driest KF, Marcilly R, Bras Da Costa S, Beuscart-Zephir MC, Jaspers MW. A Framework for reporting on Human Factor/Usability studies of Health Information Technologies. Stud Health Technol Inform 2013;194:54-60.

[44] Magrabi F, Ong MS, Runciman W, Coiera E. Using FDA reports to inform a classification for health information technology safety problems. J Am Med Inform Assoc 2012 Jan;19(1):45-53.

[45] Jaspers MW, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. J Am Med Inform Assoc 2011 May 1;18(3):327-34.

[46] Oppenheim MI, Vidal C, Velasco FT, Boyer AG, Cooper MR, Hayes JG, et al. Impact of a computerized alert during physician order entry on medication dosing in patients with renal impairment. Proc AMIA Symp 2002;577-81.

[47] Thursky K. Use of computerized decision support systems to improve antibiotic prescribing. Expert Rev Anti Infect Ther 2006 Jun;4(3):491-507.

[48] Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005 Apr 2;330:765-73.

[49] Schedlbauer A, Prasad V, Mulvaney C, Phansalkar S, Stanton W, Bates DW, et al. What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior? J Am Med Inform Assoc 2009 Jul;16(4):531-8.

[50] Ammenwerth E, Schnell-Inderst P, Machan C, Siebert U. The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. J Am Med Inform Assoc 2008 Sep;15(5):585-600.

[51] Chan J, Shojania KG, Easty AC, Etchells EE. Does user-centred design affect the efficiency, usability and safety of CPOE order sets? J Am Med Inform Assoc 2011 May 1;18(3):276-81.

[52] Russ AL, Zillich AJ, Melton BL, Russell SA, Chen S, Spina JR, et al. Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation. J Am Med Inform Assoc 2014 Mar 25.

[53] Tsopra R, Jais JP, Venot A, Duclos C. Comparison of two kinds of interface, based on guided navigation or usability principles, for improving the adoption of computerized decision support systems: application to the prescription of antibiotics. J Am Med Inform Assoc 2014 Feb;21(e1):e107-e116.

Chapter 4. Usability design principles for medication alerting systems: literature synthesis and study of their coverage regarding related usability flaws

Romaric Marcilly, Elske Ammenwerth, Julie Nies, Erin Roeher, Francis Vasseur and Marie-Catherine Beuscart-Zephir

Abstract

Introduction. Medication alerting systems are a promising technology to change prescribers' behavior. However, usability flaws prevent those systems from being fully usable. Several lists of usability design principles dedicated to medication alerting systems have been published but are scattered across several supports. Moreover, their coverage regarding related usability flaws is unknown. This paper aims (i) to develop a structured and comprehensive list of usability design principles dedicated to medication alerting systems and (ii) to match this list with an exhaustive list of actual usability flaws to assess the coverage of those principles and to illustrate them with actual instances of their violation.

Method. Papers reporting sets of usability design principles were searched in the literature on design, evaluation and implementation of medication alerting systems. Principles extracted in those papers were synthesized by two usability experts. Then, a set of 107 actual usability flaws specific to medication alerting systems was matched with those synthesized principles.

Results. Eight papers report sets of usability design principles aiming to improve the usability of medication alerting systems. A total of 125 unique principles are sorted into 54 synthesized principles. Amongst those principles, 22 are matched directly by usability flaws, 9 principles need to be slightly extended to be matched and 23 are not matched at all. Besides, 9 new principles have been defined for all usability flaws to be matched. The final synthesized list of principles includes 63 usability design principles categorized into 6 themes: improve the signal-to-noise ratio, fit the clinicians' workflow, support collaborative work, display relevant information to the clinicians, make the system transparent for the user and provide useful tools to manage the alert and instantiate the decision made.

Discussion. Amongst the principles synthesized, numerous are closely interconnected. Almost all usability flaws specific to medication alerting systems are covered by the set of usability design principles synthesized from the literature (only 9 principles had to be created). Conversely, a large number of usability design principles from the literature are matched with actual usability flaws. Moreover, bringing several sets of usability design principles together increased the coverage power of those lists regarding usability flaws. The final list of usability design principles illustrated by actual related usability flaws could be used by designers and Human Factors experts as a check-list of usability points to consider when designing and/or evaluating a medication alerting system.

Keywords.

Usability; Decision support, clinical; Design principles; Human engineering;

4.1. Introduction

Decision Support functions such as medication alerting systems are promising technologies to change prescribers' behavior, help them avoid errors [1] and ultimately improve the quality of the medication management process [2]. Nonetheless, their poor usability is regularly pointed out to explain their difficulties of integration into the clinical workflow [3] hampering their acceptance and even endangering the patient [4].

Usability is "the extent to which a product can be used by specified users to achieve specified goals effectively, efficiently and satisfactorily within a specific context of use" [5]: it includes the features of the Graphical User Interface (*e.g.,* legibility of texts), the functionalities and the knowledge implemented in the system and the fitting between the system's behavior and the needs of the users [6]. An optimal usability is supposed to "increase individual effectiveness and efficiencies" along with "organizational efficiencies" to ultimately improve "patient, provider and organizational outcomes" [7].

The presence of usability flaws in the system prevents this system from being fully usable. Usability flaws are also known as usability problems [8]: both denominations are synonyms. A usability flaw "is an aspect of the system and/or a demand on the user which makes it unpleasant, inefficient, onerous or impossible for the user to achieve their goals in typical usage situations" [9]. When the system is put into use, usability flaws may bother the users through usage problems and they may ultimately lead to negative outcomes on the performance and the work system. Usability flaws are the consequences of the violation of usability design principles, *i.e.,* recommendations of good practice in terms of usability. Those principles are also known as usability heuristics or ergonomics criteria [8; 10]. The application of usability design principles during the design process is a requisite to prevent from usability flaws appearance.

Two levels of usability design principles may be distinguished according to their level of specificity:

- General design principles developed for interacting systems. Those principles are pretty much independent from the task to perform with the system. Therefore they apply to design any kind of interactive systems [8; 10].
- Specific design principles rest on the knowledge of the specificities of the interaction between a user and a system when performing a task. They are directly related to the task to perform with the system. Therefore they apply only to a specific type of system (*e.g.,* [11; 12] for medication alerting systems).

In this paper we focus on the category of specific design principles dedicated to medication alerting systems, *i.e.,* design principles that are directly related to the specific tasks to perform with a medication alerting system.

Medication alerting systems may face a very large variety of usability flaws: in a previous study we performed a systematic review [13] that identified 168 usability flaws in medication alerting systems that

may be sorted into 13 types of flaws. Precisely, amongst them, 107 are specific to medication alerting systems and therefore are violations of usability design principles dedicated to medication alerting systems.

Conversely on the other hand, literature proposes several synthetic lists of usability design principles [1] aiming to improve the usability of medication alerting systems and based on expert consensus or on targeted review of published studies [11]. However, those lists are scattered across several publications and there is no overview paper providing an exhaustive structured list of those principles. Moreover, there is no assessment of the coverage of those principles regarding related usability flaws identified in the literature. It is still unknown whether all types of usability flaws known in medication alerting systems are covered by those lists. Therefore, the capacity of those lists to fix all types of flaws, if they were applied, remains uncertain.

Finally, amongst the available lists of principles, only few provide actual instances of the usability flaws related to those principles [11; 14]. Yet, principles are likely not to be correctly interpreted when they do not contain clear information on their rationale, their benefits, their nature and their conditions and procedure of application [15]. Illustrating usability design principles with actual instances of "what should not be done" may participate in improving designers', developers' and even Human Factors (HF) experts' understanding of the related usability design principles; this may help them apply the principles purposefully with no mistake.

The aim of this paper is twofold:

- First, based on existing published papers, to develop a structured and comprehensive list of usability design principles dedicated to medication alerting systems; and,
- Second, to match this list with the exhaustive list of usability flaws identified in the aforementioned systematic review to assess the coverage of those principles and to illustrate them with actual instances of their violation.

With this aim in mind, two questions are asked:

- What are the usability design principles specific to medication alerting systems reported in literature?
- How well usability flaws reported in the literature are matched with the usability design principles?

4.2. Methods

4.2.1. Synthesizing usability design principles

Two usability experts (RM and MCBZ), with experience in the design and evaluation of decision support functions, identified papers reporting sets of usability design principles based on the references often cited in the topic "design, evaluation and implementation of medication alerting systems". This

search was completed by papers of interest identified during the process of the aforementioned systematic review [13] and by other relevant papers known by the authors. Finally references of papers included were searched.

To be included, papers must report a synthesis of design principles that included at least one usability design principle dedicated to medication alerting systems. Papers that reported opportunistically isolated principles were excluded from the search. Papers that presented a synthesis of usability design principles but that were included in the systematic review because they presented also usability flaws were excluded from the review to avoid auto-matching bias. Both experts agreed on the final set of papers to analyze.

Then, one expert (RM) looked for the usability design principles in each of those papers. Once all usability design principles extracted, duplicate principles presented by the same authors were excluded. Then, principles published by different authors but that were similar in purpose were gathered together. The final list of usability design principles was organized into categories and sub-categories by RM. The results of the categorization process were cross-checked by MCBZ. Any disagreements were solved by discussion and consensus. Finally, both experts together synthesized the sub-categories of flaws by rephrasing those principles.

4.2.2. Matching usability design principles with known usability flaws

The 107 usability flaws specific to medication alerting systems identified in the aforementioned systematic review [13] were checked against the synthesized usability design principles. A usability flaw was considered as matching a given usability design principle if it is a violation of this latter. Reciprocally, a usability design principles was matched by a usability flaw if it could have prevented from the appearance of the related flaw if applied. If a flaw could not match directly any usability design principle, then the possibility to extend an existing usability design principle was considered (*i.e.,* including other contexts) to cover a larger range of flaws. If there was no usability design principle that could be extended to cover the flaw, a new principle was defined.

The matching process was performed by both experts (MCBZ and RM) together. All matching decisions were made by consensus after discussion. The matching process intended to be as univocal as possible, *i.e.,* one flaw matching one principle.

4.3. Results

4.3.1. Characteristics of the papers analyzed

Eight papers [1;3;11;12;14;16-18] are identified as containing sets of principles related to the improvement of the usability of medication-related alerting systems. A paper of interest [18] is excluded because it is also part of the papers analyzed in the systematic review; after an in-depth reading to check

whether this paper does present principles not found elsewhere, it appears that it does not provide new insights. Main characteristics of the 7 analyzed references are presented in Table 12.

Table 12 - Main characteristics of the references reporting sets of usability design principles for medication alerting systems.

	First author	Year	Focus	Method used to provide the principles
[3]	Sittig DF	2008	Design, implementation, research	Authors experience / experts consensus
[16]	Kuperman GJ	2007	Design, implementation	Experts consensus
[14]	Horsky J	2012	Usability	Targeted review
[11]	Horsky J	2013	Usability	Targeted review
[1]	Bates DW	2003	Design, implementation, monitoring	Authors experience
[17]	Pelayo S	2011	Usability	Targeted review & analysis of cognitive and collaborative tasks
[12]	Phansalkar S	2010	Usability	Targeted review

4.3.2. List of usability design principles and matching with usability flaws

For clarity sake, the list of usability design principles is presented along with the results of the matching with usability flaws. Some titles of categories of principles have been created by the authors of this paper in order to optimize the presentation of the principles included in those categories: those titles are per se not matched by usability flaws.

A total of 132 usability design principles specific to medication alerting systems are identified. The complete list of principles extracted is presented in appendix 3. Amongst them 7 duplicates of other principles are identified, all in both Horsky's papers [11; 14]; duplicates are brought together with the original principles. The 125 remaining principles are sorted into 54 synthesized principles, themselves categorized into 6 themes. The 54 principles are organized hierarchically: some large principles are specified more precisely by several sub-principles. The 6 themes are:

- Improve the signal-to-noise ratio of the system;
- Fit the clinicians' workflow;
- Support the collaborative work;
- Display relevant information to the clinicians;
- Make the system transparent for the user to understand precisely what it does and does not;
- Provide useful tools to manage the alerts and to instantiate the decision made.

Only very few adjustments into the categories have been made after the cross-checking process: they concern mainly moving some items from the theme "provide useful tools" towards the theme "display relevant information".

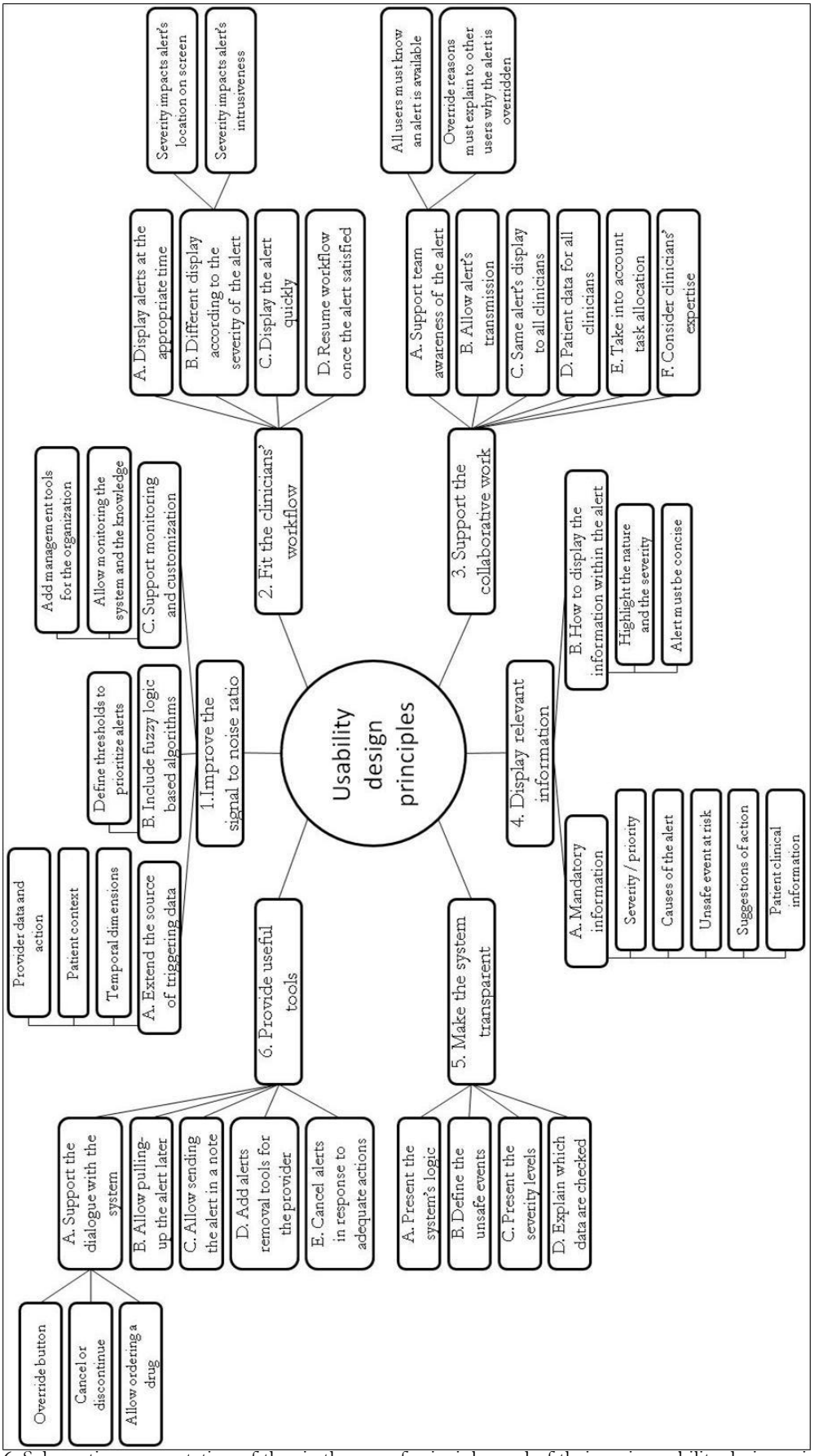


Figure 6. Schematic representation of the six themes of principles and of their main usability design principles.

The matching process highlights that 23 out of the 54 principles are not matched at all. However, amongst those principles, some are indirectly matched because one of their sub-principles is matched by a flaw or because, if it is a sub-principle, its "meta" principle is already matched. Amongst the remaining 31 principles, 22 are matched directly and 9 are matched after a slight extension. Moreover, in order to match all usability flaws with a principle, 9 new principles have been defined leading to a total number of 63 usability design principles dedicated to medication alerting systems.

An overview of the principles that are included in the 6 themes is presented in Figure 6. The 63 principles are then presented in details in Table 13 to Table 18 along with a synthesis of the related usability flaws. Those tables present in the left columns the synthesized usability design principles and a synthesis of their corresponding actual violations (usability flaws) in the right columns. Figure 7 proposes an illustration of the violation of a principle from the theme "improve the signal to noise ratio".

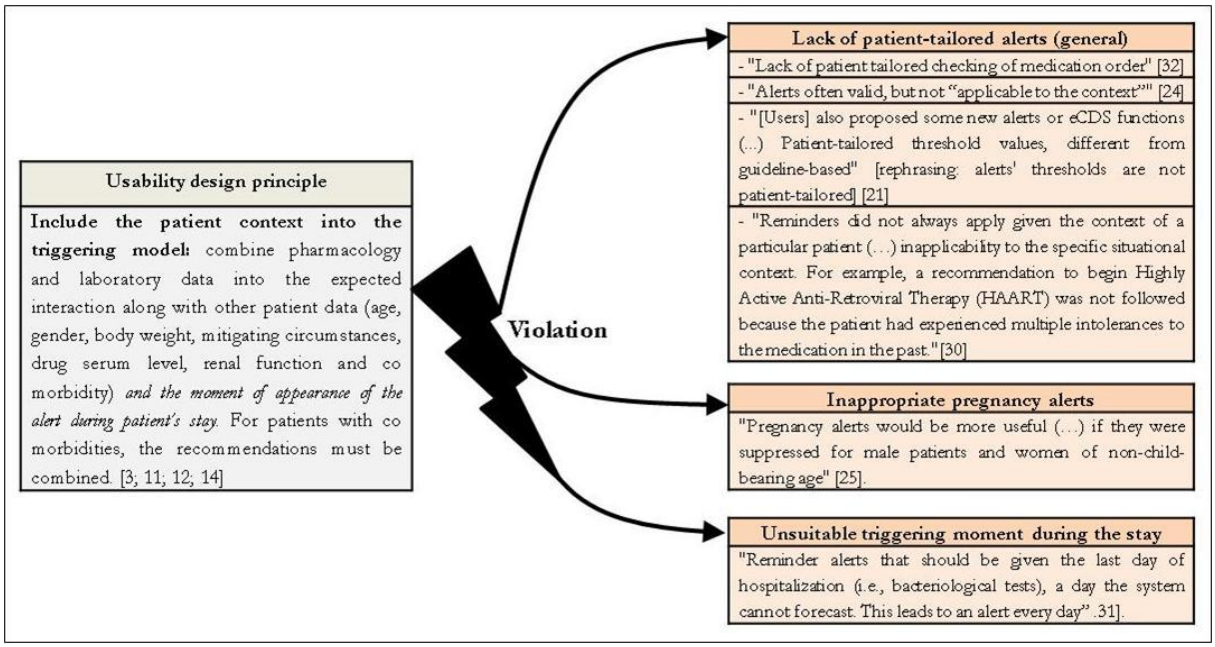


Figure 7. Schematic representation of the violation of a usability design principle from the theme "improve the signal to noise ratio" and various resulting actual instances of usability flaws.

4.3.2.1. Improve the signal to noise ratio

All in all, the principles presented in Table 13 aim to decrease the over alerting. With this aim in mind, the signal-to-noise ratio must be improved. Other main principles that emerge in this theme deal with increasing the variety of data included in the alerting system logic: move from a "boundary base" alarm strategy towards a strategy that monitor several parameters simultaneously. This implies that the system must consider various types of data such as the temporal dimensions of the unsafe events and its causes, the patient clinical context along with provider's actions. Moreover, the system must provide the provider organization with tools to customize the alerting system and the knowledge implemented in it. Finally, the usage of the alerting system must be monitored to remove irrelevant or useless alerts.

Table 13 - Usability design principles included in the theme "improving the signal-to-noise ratio" and synthesis of the related flaws. Italic fonts represent the principles (or the extension) that have been added to complete the matching process. Titles of categories of principles marked with a * are created by the authors of this paper to organize optimally the other principles.

Usability design principles	Synthesis of related flaws
1. Improve the signal-to-noise ratio by improving the sensitivity and the specificity of the alerting system in order to decrease the number of irrelevant alerts (*e.g.,* alerts with little evidentiary basis, low clinical relevance or redundant alerts [3; 12; 14; 16], *along with alerts that are only informational, i.e., that require no action) and the number of alerts per patient. This includes defining precisely the level of the triggering thresholds of each alert.*	Too many alerts [19-23], potential problems are over or under detected [24; 25], alert requiring no action (thus useless) [23], too low thresholds [21; 26], too many alerts per patient [22; 24; 27], redundancy of alerts [20; 28; 29]
A. Increase the variety of information sources that are available to the decision rules engine (*) *and check the accuracy of the information retrieved from the CPOE/EHR.*	Alert inconsistent with EHR data [24], erroneous data [30], non up-to-date patient data [31]
Include temporal dimensions into the triggering model (*):	
Consider the interval between drugs' administrations: an earlier prescribed drug may have been completely metabolized by the time a contraindicated drug is entered [11]. *More generally, the system must clearly distinguish orders specified as "now" and those specified as "future" and "standing".*	Drugs not taken at the same time (past, current, future drugs) are incriminated [25; 29]
Consider the evolution of the unsafe event at risk: if it gets worse, the level of severity of the alert should increase. [17]	/
Consider the delay of appearance of the unsafe event at risk after the drugs has been administered.	Alert irrelevant due to the rapidity of the appearance of the unsafe event [23]
Include the patient context into the triggering model: combine pharmacology and laboratory data into the expected interaction along with other patient data (age, gender, body weight, mitigating circumstances, drug serum level, renal function and co morbidity) *and the moment of appearance of the alert during patient's stay*. For patients with co morbidities, the recommendations must be combined. [3; 11; 12; 14]	Lack of patient-tailored alert [21; 24; 30; 32], pregnancy alerts for males or women of non-child-bearing age [25], an alert is supposed to appear the last day of the stay which is often unforeseeable, thus the alert appears every day [31]
Take into account provider-specific data along with the actions already taken by the provider (incl. the dose adjustment) [11; 17] [12]	The alert requires an action already taken (serum level measured, effect monitored) [23], drugs discontinued by the clinicians are still considered as active [24; 33], clinicians satisfy the intend of the reminder but that did not resolve it [30]
Take into account the clinical specialty (e.g., psychiatric and pediatrics): according to the specialty, some medicaments may be used off-label.	Drugs are intentionally combined because of the desired effect of the interaction or are used off-label [23;25]
B. Include fuzzy logic-based algorithms, multi-attribute utility model and filters into the alerting system's reasoning model to modify the activation of alerts when certain conditions apply. [11; 12]	/
Define thresholds to prioritize the alerts according to the clinical context of the patient: they may modify the severity of the expected interaction and the system should then select appropriate warning level. [11; 16]	/
Define thresholds to prioritize the alerts according to the severity of the unsafe event [11; 12]	The unsafe event's incidence is "not serious" [23; 26], no distinction between true allergy and bothersome but not serious side effect [22; 24]

Usability design principles	Synthesis of related flaws
C. Support monitoring and customization (*)	
Add management tools so that provider organizations can customize purchased drug information and so that customizations persist across version upgrades. Those local customizations aim to adapt the knowledge base to the local practice and to remove potential errors in the knowledge base. In each organization, a committee of expert physicians and pharmacists should be in charge of this customization. [16]	Alerts in conflict with local/standard practices or specialists recommendations [22; 24; 25]
Allow monitoring the alerting system and the knowledge base over time (*):	
Update regularly the drug database along with the list of accepted actions (e.g., monitoring, ordering drugs) that satisfy an alert	Actions satisfying the intend of the alert do not resolve it because they are not yet included in the logic of the system [30], not up-to-date data about drug interactions or new drugs [31]
Monitor alerts' override reasons: alerts frequently overridden and of little value should be considered for removal from the alert base or for a change of their presentation format. [14; 16]	Alerts recognized as not useful are still presented [23]

The matching process of the usability flaws with the principles included in this theme requires 4 principles synthesized from the literature be extended to suit related usability flaws. Moreover, 3 principles are created to represent other usability flaws. Those extensions and creations consist mainly in considering larger sources and types of data (*e.g.,* temporal characteristics of the alert's triggering) along with actions to perform on the data retrieved and knowledge used (*e.g.,* up-dating, checking of accuracy etc.). In the final set of principles, only 3 are not matched by any actual usability flaw. Two of them are related to the reasoning model of the alerting system; the third one deals with considering the evolution over time of the unsafe event to define the severity of the alert.

4.3.2.2. Fit clinicians' workflow

Those principles (Table 14) aim to make the alerting system comply with the tasks performed by the clinicians. The system must not disturb constantly and futilely the clinician. For this purpose the alert must be presented at the right moment during the decision making process; it implies that the decision making process is well-known and that the alert appears quickly. Once the alert is satisfied, clinicians must be able to resume their tasks. In order less severe alerts not to perturb the clinicians, the main principles require that only the most severe alerts interrupt the users; the other must be displayed in the system more discreetly.

In this theme, no usability design principle has been extended or created to suit related usability flaws. Two principles are not matched by usability flaws. The first one deals with the position of the alert on the screen that must vary according to the severity of the alert. The second one explains that users must be

able to resume their tasks as soon as the alert is satisfied. All other principles are matched by several usability flaws.

Table 14 - Usability design principles included in the theme "fit the clinicians' workflow" and synthesis of the related flaws. Titles of categories of principles marked with a * are created by the authors of this paper to organize optimally the other principles.

Usability design principles	Synthesis of related flaws
2. Fit the clinicians' workflow	
A. Alert must be displayed at the appropriate time during the decision making. [1; 14; 17]	Institutional guideline for heparin administration and drug allergy alerts show up too late in the ordering process [20; 34-36]; some alerts appear too early [29]; the reminder about the corollary PTT is presented out of the logical workflow [34]; alert appear after the patient has left the consultation room [19;29]
B. The severity of the unsafe event must be taken into account to display the alert within the technology (*e.g.,* CPOE) (*)	
Alert's location and presence on the screen must be adapted according to the severity of the unsafe event, alerts must be placed within the operator's visual field in order of importance (focal region for the more severe alerts vs. side regions for not severe alerts) and color-coding could be used to differentiate them. These messages can also be aggregated and shown together in a single display to be reviewed all at once at a convenient point in the workflow such as at the end, during order signing. [11; 12; 14; 16]	/
Alert's mode of appearance (intrusive vs. non intrusive) must be proportional to the severity of the unsafe event: the most serious warnings must be intrusive and still need an explicit response by a clinician but less important alerts are displayed less intrusively on the screen as messages not requiring any actions. Interruptive alerts should be reserved only for high severity warning and used judiciously. [3; 11; 12; 14; 16]	Pop-up alerts can be very annoying [29], on the contrary some alerts are not sufficiently visible [29]
C. Alert must be displayed quickly: screen transition time must be well under a second [1; 14]	Too long delay of appearance: up to 10-15 second computer delays between the order and the subsequent alert [24], 3 sites reported an average 8 seconds delay for the alert to load [27]
D. After the alert is satisfied, resume the workflow. [11; 16]	/

4.3.2.3. Support collaborative work

The goal of the principles in the theme "support the collaborative work" (Table 15) is to make the alerting system support the collaborative aspects of the medication management process by physicians, pharmacists and nurses. For this purpose, the system must become an actual "team player". Overall, the alerting system must suits the tasks of each clinician, provide them with the same information even if some professionals may have access to supplementary information according to their specialty. The

alerting system must also help them know how other clinicians have managed the alert for instance, in order to understand why an alert has been overridden.

In this set of usability design principles, one principle has been created and another one has been completed to match related usability flaws. The first one deals with a missing function allowing clinicians to send the alert to others clinicians. The second one is related to the possibility to target more precisely the addressee of the alert. There are also three principles not matched at all that deal with providing indication of the availability of alert's information to all clinicians and with sharing patient's information among all clinicians.

Table 15 - Usability design principles included in the theme "support the collaborative work" and synthesis of related flaws. Italic fonts represent the principles (or the extension) that have been added to complete the matching process. Titles of categories marked with a * are created by the authors to organize optimally the other principles.

Usability design principles	Synthesis of related flaws
3. Support the collaborative work. Make it a team player! [17]	/
A. Support the team awareness of the alert management (*):	
Provide all indications of the availability of information to all users [17]	/
The override reason should allow nurse and pharmacists to understand the rationale for the override [16]	Prescribers unsure whether pharmacists review the override reasons [24], override reasons are stored in databases and displayed to no one [37]
B. Allow clinicians relay the alert message to other clinicians by supporting their transmission	The system does not support the transmission of reminders [27]
C. Provide to all clinicians the same display of alert information[17]	Only the pharmacy's computer provides drug interaction and lifetime limits warning [36]
D. Share patient clinical information with all clinicians (especially with pharmacists who have generally no other source of information about the patient) [14]	/
E. Take into account clinicians tasks allocations: some alerts that do not concern physicians can be offloaded from physician workflow and redirected to support staff. [14]. *On the contrary, alerts that are dedicated to the physicians should not be displayed to non medical professionals (e.g., psychologists, physiotherapists).*	Drug administration alerts are displayed to physicians not to nurses [26], drug alerts displayed to physiotherapists or psychologists and most of the time they are not relevant [21]
F. Each clinician's expertise specificities must be considered (pharmacists may need a supplementary more pharmacology-oriented alerting system) [14]	Mismatch between the alerting system model and the pharmacist mental model about alerts' categorization according to their severity [33]

4.3.2.4. Display relevant information

Within this theme (Table 16), the principles list the information that must be mandatorily presented in the alert and how to display them within the alert. The alert must be concise but must present in its window several types of information. The most important ones are the cause of the unsafe event and its severity; the others are the description of the unsafe event, patient clinical information and suggestion of corrective actions. Links to get supplementary information on the alert must also be available. Those

principles aim to provide the clinicians with the necessary information to make informed decision.

Table 16 - Usability design principles included in the theme "display relevant information" and synthesis of related flaws. Italic fonts represent the principles (or the extension) that have been added to complete the matching process. Titles of categories marked with a * are created by the authors to organize optimally the other principles.

Usability design principles	Synthesis of related flaws
4. Display relevant information	
A. Some information are mandatory in the alert, *i.e.,*: (*)	
Display the severity / priority of the unsafe event [11; 12; 16]	No clear information on relative risk of harm [22; 24; 38], alerts not stratified by severity levels [18]
Display the cause of the unsafe event *and its characteristics: e.g.,* name of the incriminated drugs [11; 12; 16]	Missing information about the dose-dependence of the interaction [38]
Display the unsafe event (potential or currently happening) [12; 16]	No information on the potential problem [22; 24; 33; 38]
Display suggestions of action: suggest, do not impose [11; 12; 14; 16]	No guidance on actions to take [18; 24; 39; 40]
Options to proceed must suit the patient and/or the clinical situation	There is no available options to satisfy the alert that apply to the patient or the situation [27]
Include possible ancillary orders [11]	/
Include action links to make it easier for clinicians to take action [3; 11; 12; 16]	/
Include drug alternative (incl. dose and frequency) [11; 14]	No suggestion of an alternative treatment [38]
Prioritize suggestions [3]	/
Justify suggestions [17]	Erroneous message on how to solve the problem [32]
Suggestions must be locally checked [16] and must include a link to institution-specific guidelines [16]	Alerts do not always match the local practice [30]
Suggestions must be consensual [14; 16]	/
Suggestions must be evidence-based with documentations of benefits [16]	/
Display patient clinical information supporting the relevance of the advice [14]	The alert does not provide the essential patient information [24]
Summarize patient information to support optimal decision-making process [3; 11; 14]: *the summary must include all referential data to interpret the patient's case (e.g., lab thresholds to interpret the normality of the results*) and provide a link to a summary of relevant patient clinical data [11].	Missing normal range of labs to interpret some lab results in the patient data summary [38]
Include a link to a more complete documentation (monograph, evidence, extended information) [1; 11; 14; 16]	No sufficient information [18; 22; 24], no more detailed sections [38] such as no evidence to support the alert [22-24]
B. How to display the information within the alert: (*)	
Highlight the cause of the unsafe event and its severity information [11]	/
Make the message of the alert concise / succinct and actionable, no more than 10 words [1; 11; 16]	/

One new design principle about the need of taking into account the patient context to suggest actions has been defined. Two principles have also been slightly extended to be more precise about the information required. In this theme, seven principles are not matched by usability flaws. Five are sub-principles; the two other are the two unique principles of the category "how to display the information within the alert".

4.3.2.5. Make the system transparent

This set of principles (Table 17) aims to help clinicians know how the alerting system is working in order to prevent erroneous interpretation of its behavior. Concretely, the user must have an access to a precise description of the logic applied by the system, a definition and a list of the unsafe events targeted by the system along with the description of the various levels of severity an alert may take. Finally, the types of data that are checked by the alerting system must also be presented. All in all, those principles should help the user have a precise representation of what the alerting system can do and what it cannot and how it does it.

Table 17 - Usability design principles included in the theme "make the system transparent" and synthesis of the related flaws. Italic fonts represent the principles (or the extension) that have been added to complete the matching process.

Usability design principles	Synthesis of related flaws
5. Make the system transparent for the user. The system must not be a black box and its coverage must be accessible to its user [14; 16]	Systems capabilities and limitations are ambiguous, invisible [24; 41]
A. Make accessible the alerting system's algorithm / logic / formulas [14]	User uncertain how long a reminder would be turn off [30], alerts do not appear while they are expected [33], dose calculation not understood [34]
B. Make accessible a definition of what is an unsafe event, the list of the events that are checked by the alerting system and explanations of their classification as unsafe events. [12]	Physicians not confident the system is checking all the interactions that they want to check [24]
C. Make accessible the various levels of severity used by the alerting system (and their number by unsafe event) [12]	No catalog about the level of severity [18]
D. Make accessible the format of data that are checked: e.g., free-text, origin of the data (other hospital data), name of drugs vs. ATC codes.	Some patient data are not used to trigger the alert [29], prescribers confused whether the system could check non-formulary, non-VA medications [24], whether free-text entries are checked [24], no information whether the system checks ATC codes or drugs names [42]

In this theme, one principle has to be created to represent the related flaws about the transparency of the capacity of the alerting system in terms of types of data that are checked. All principles are matched by usability flaws.

4.3.2.6. Provide useful tools

This set of principles (Table 18) present the type of tools that must be available within the alert. First, there are tools that must be within the alert's dialogue box: there must be buttons to cancel/discontinue an order or add a new one, a button to override the alert; if other corrective actions are proposed, they must be easy to carry out from the alert. Those principles aim to help the user instantiate their clinical decision in reaction to the alert into easy to perform actions. The second kind of tools represents those to manage the alert: pull up the alert later, send it into a clinical note, and remove the alert for a patient.

In this set of principles, 3 principles have been created and 2 others have been extended for the related usability flaws to match them. Those principles deal:

- With the types of actions that clinicians should be able to perform when facing an alert to satisfy it (*e.g.,* several actionable buttons), override it or manage it (*e.g.,* pull up the alert later or transfer it into a note); and
- With details on how to interact with the alert and on when to use the alert's removal tool.

In this theme, 8 principles are not matched by usability flaws including sub-principles (*e.g.,* information about how to override an alert and how to enable the user to order a drug directly from the alert window) and also a high level principle (*i.e.,* how the system must behave in response to corrective action.

4.4. Discussion

The aim of this paper was twofold: to develop a structured and comprehensive list of usability design principles dedicated to medication alerting systems; and to match this list against a list of usability flaws in order to evaluate the coverage of the principles and beyond, to illustrate those principles with actual instances of their violations.

Eventually, results identified 6 main themes of principles including 63 usability design principles. The 7 papers analyzed participate differently in the 6 themes. Only one [16] takes part in the six themes of synthesized principles. All the others miss out at least one theme [12; 14; 17] or more. On the other hand, one theme out of 6 (*i.e.,* "display relevant information in the alert") is supported by the 7 papers analyzed while 2 themes are fed by three papers analyzed (*.i.e.,* "make the system transparent", "support the collaborative work"). Therefore, synthesizing together several "syntheses" of usability design principles found in the literature allowed improving the variety of the topics represented in each individual list.

Table 18 - Usability design principles included in the theme "provide useful tools" and synthesis of the related flaws. Italic fonts represent the principles (or the extension) that have been added to complete the matching process.

Usability design principles	Instance of flaw
6. Provide useful tools	
A. The dialogue box must include several actionable buttons, not only "cancel" and "order", but also appropriate suggestions for action.	No available option in the dialogue box match patient's response [27], only "continue with the drug" or "cancel the order" are available [18]
Include a button to override the alert (meaning continue ordering, ignoring the caveat) [11; 17]	/
Make vary the difficulty to override an alert according to the severity of the unsafe event: *e.g.,* require a second confirmation for the more severe alerts; Even, overriding may not be possible for critical alerts. [11]	/
Require the reason for override [11; 17]	/
In "one click": propose a list of 3-4 coded (max 5 items) override reasons that are selectable; reasons must be 1-2 word long. [11; 16]	/
Make mandatory override reason selections for the most critical alerts but otherwise make them optional. [11]	/
Distinguish clearly "cancel" and "discontinue" an order according to the status of the order [11; 17] *Provide the clinician with information on how to switch or discontinue a drug.*	No sufficient information on how to switch or discontinue an alert [39]
Include a link to order a drug suggested [11] or new drug: the one previously ordered that trigger the alert must be automatically discontinued and a pre-populated ordering screen must open. [11]	/
Clearly state that the existing order will be discontinued if the new one is finalized. [11]	/
B. Provide a way to pull up the alert later	The alert cannot be pulled up later [24; 30], clinicians do not get the warning again [26]
C. Allow sending the alert into the template of the clinical note	The recommendations do not go into a note (or its template) to document what is done with the patient [27; 39]
D. Add alerts removal tools for the provider to suppress redundant alerts when a previously tolerated medication combination for the same patient is renewed or when providers feel they have sufficient practice and knowledge about this alert or when the alert is out-dated for a specific patient. [11]	The alert is already known [23], the alert about an adverse reaction of patient to a drug is not true since the patient has been taking this drug for a long time [24], some alerts cannot be removed [27], alerts outdated for a patient (*e.g.,* past infection) cannot be removed [31]
E. Alerts must cancel or reset in response to the appropriate corrective action rather than requiring an acknowledgment from the operator followed by the corrective actions. [12]	/

The 63 principles synthesized target all components of the alerting system: the knowledge base implemented, the setting of the triggering engine, the meta-rules managing the behavior of the system and the system's graphical user interface. Numerous principles are closely interconnected. For instance, to display specific information on user's interface (*e.g.,* severity, patient data, type of unsafe event etc.), it is

necessary that the data model of the alerting system be designed to consider those data. Even if both aforementioned principles are not presented in the same themes (respectively "display relevant information" and "improve the signal-to-noise ratio"), they must be considered together during the design process. Similarly, presenting differently the alerts according to their severity (theme "fit the clinicians' workflow") requires first to define the severity of those alerts (theme "improve the signal-to-noise ratio"). The themes of the principles identified are therefore also closely intertwined. The present organization represents the perspective of the usability experts who categorized the principles; other categorizations could have been proposed if they had been carried out by designers or developers (*e.g.,* according to the technical components of the system, engine, database etc.).

As for the matching of usability flaws against those usability design principles, the process has been performed quite easily with no uncertainty; the hierarchical structure of the synthesized usability design principles allows matching usability flaws whatever their level of description. Twenty-two usability design principles were directly matched by usability flaws. Moreover, all usability flaws matched the themes of principles identified; no new theme had to be created especially for matching some usability flaws. Only 9 principles had to be fully defined and 9 others had to be slightly adapted for all the usability flaws to be matched by a principle. This means that the lists of usability design principles identified in the literature cover most of the usability flaws known in medication alerting systems; however, they were not sufficiently extended and/or precise to cover all kinds of usability flaws.

Due to the non systematic collection of usability design principles one cannot ignore that a few syntheses of usability design principles might have been missed. Consequently, the principles that had to be defined from the descriptions of the related flaws could have been already reported in other papers.

Moreover, 23 principles were not matched at all by any usability flaw. This could be explained by several reasons. Publication and reporting biases may have led to an under-report of the usability flaws identified in [13]: it is possible that some types of flaws are never reported that would have matched those principles. Moreover, we organized hierarchically the principles to enable considering usability flaws of various levels of precision. Therefore, some principles are not matched because either one of their sub-principles or the "meta" principle they belong to are themselves matched. Finally, since several principles are interdependent (*e.g.,* "use fuzzy logic-base algorithms" and "increase the variety of information sources" both in theme "improve the signal-to-noise ratio"), a flaw that is documented for one of those principles matches indirectly also the other one (*i.e.,* "use fuzzy logic-base algorithms" is indirectly matched by the flaws noticed for "increase the variety of information sources"). Therefore, most of the time, usability design principles can be considered indirectly matched by a usability flaw.

In summary, even if it is not complete, the coverage of the usability flaws by the list of design principles is quite good considering the biases that may have impacted the retrieval of the principles and of flaws. Beyond the question of the coverage, the results of the matching process provide also illustrations of the violations of the usability design principles. This type of study has already been performed in general usability [43]. However, as far as we know, it is the first time that it is carried out for

usability design principles dedicated to a type of HIT. The list of synthesized principles illustrated by actual usability flaws we finally developed (Table 13 to Table 18) may participate in helping designers, developers and even HF experts have a clearer understanding of those design principles. It may also help them identify easily and concretely what are the usability mistakes not to make for each principle or to catch them during evaluation phases.

This list could be used by designers, developers and also HF experts as a check-list of the usability points to consider when designing and/or evaluating a medication alerting system. Nonetheless, this list must be used in combination with ergonomics criteria for interactive systems [8; 10] to optimize the usability of the system. In the current state of this list, the wording used to present the design principles and the organization of those principles is usability-oriented due to the topics of the literature that has been analyzed and to the background of both experts who performed the analysis. Therefore slight wording and organization changes must be performed before the list be used by non HF experts; this requires involving experts in the design of medication alerting systems with a technical background.

The validity of the illustrated lists of usability design principle may be time-dependant. Indeed, the list of usability flaws used in the matching process may change over time depending on the publication in the literature of new studies reporting usability flaws. Moreover, the technology used in medical informatics evolves rapidly and the associated principles may change accordingly. For instance, the principles presented here are formulated for medication alerting system implemented in laptop and/or desktop; but as mobile health technologies are improved and developed, medication alerting systems are now progressively implemented on smart mobile tools. This could impact the purpose of the usability design principles dedicated to those systems. It is therefore mandatory to update regularly both usability design principles and usability flaws bases in order to provide an up-to-date knowledge fitting the characteristics of the technology designed or evaluated. Yet, the maintenance of that knowledge base may be very time-consuming and represent a challenge for the HF community in medical informatics. Increasing the usage of incident reporting systems that include usability and encourage the sharing of their data may help lighten this maintenance process by providing numerous structured usability data.

4.5. Conclusion.

This study developed a structured and comprehensive list of usability design principles specific to medication alerting systems and illustrated by actual related usability flaws. Such a list is intended to help designers and HF experts understand clearly the usability design principles to apply them purposefully with no mistake. Usability design principles reported in the literature were synthesized and matched with actual instances of usability flaws. Results show that 6 themes of usability design principles are emerging. The matching between principles and actual flaws is quite good even if some principles had to be created to match all usability flaws. The final results of this work, the illustrated list, is usable to support the design and evaluation process of medication alerting systems even if wording and structure adaptations must be

carried out to suit designers' needs. Ultimately, using this list during the design and evaluation process may help improve the usability of medication alerting systems.

Reference

[1] Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, et al. Ten commendments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003 Nov;10(6):523-30.

[2] Jaspers MW, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. J Am Med Inform Assoc 2011 May 1;18(3):327-34.

[3] Sittig DF, Wright A, Osheroff JA, Middleton B, Teich JM, Ash JS, et al. Grand challenges in clinical decision support. J Biomed Inform 2008 Apr;41(2):387-92.

[4] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004 Mar;11(2):104-12.

[5] International Standardization Organization. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability (Rep N° 9241-11). Geneva: International Standardization Organization; 1998.

[6] Beuscart-Zephir MC, Borycki E, Carayon P, Jaspers MW, Pelayo S. Evolution of human factors research and studies of health information technologies: the role of patient safety. Yearb Med Inform 2013;8(1):67-77.

[7] HIMSS UsabilityTask Force. Promoting Usability in Health Organizations: Initial Steps and Progress Toward a Healthcare Usability Maturity Model. Healthcare Information and Management Systems Society; 2011 Feb.

[8] Nielsen J. Usability Engineering. Boston: Academic Press; 1993.

[9] Lavery D, Cockton G, Atkinson M. Comparison of evaluation methods using structured usability problem reports. Behaviour and Information Technology 1997;16(4):246-66.

[10] Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behaviour and Information Technology 1997;6(4-5):220-31.

[11] Horsky J, Phansalkar S, Desai A, Bell D, Middleton B. Design of decision support interventions for medication prescribing. Int J Med Inform 2013 Jun;82(6):492-503.

[12] Phansalkar S, Edworthy J, Hellier E, Seger DL, Schedlbauer A, Avery AJ, et al. A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. J Am Med Inform Assoc 2010 Sep;17(5):493-501.

[13] Marcilly R, Ammenwerth E, Vasseur F, Beuscart-Zephir MC. Usability flaws of medication-related alerting systems: a systematic review. IJMI (submitted). In press 2014.

[14] Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. J Biomed Inform 2012 Dec;45(6):1202-16.

[15] de Souza F, Bevan N. The use of guidelines in menu interface design: evaluation of a draft standard.: North Holland; 1990 p. 435-40.

[16] Kuperman GJ, Bobb A, Payne TH, Avery AJ, Gandhi TK, Burns G, et al. Medication-related clinical decision support in computerized provider order entry systems: a review. J Am Med Inform Assoc 2007 Jan;14(1):29-40.

[17] Pelayo S, Marcilly R, Bernonville S, Leroy N, Beuscart-Zephir MC. Human factors based recommendations for the design of medication related clinical decision support systems (CDSS). Stud Health Technol Inform 2011;169:412-6.

[18] Zachariah M, Phansalkar S, Seidling HM, Neri PM, Cresswell KM, Duke J, et al. Development and preliminary evidence for the validity of an instrument assessing implementation of human-factors principles in medication-related decision-support systems--I-MeDeSA. J Am Med Inform Assoc 2011 Dec;18 Suppl 1:i62-i72.

[19] Ash JS, Sittig DF, Dykstra RH, Guappone K, Carpenter JD, Seshadri V. Categorizing the unintended sociotechnical consequences of computerized provider order entry. Int J Med Inform 2007 Jun;76 Suppl 1:S21-S27.

[20] Baysari MT, Westbrook JI, Richardson KL, Day RO. The influence of computerized decision support on prescribing during ward-rounds: are the decision-makers targeted? J Am Med Inform Assoc 2011 Nov;18(6):754-9.

[21] Kortteisto T, Komulainen J, Makela M, Kunnamo I, Kaila M. Clinical decision support must be useful, functional is not enough: a qualitative study of computer-based clinical decision support in primary care. BMC Health Serv Res 2012;12:349-57.

[22] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. A human factors investigation of medication alerts: barriers to prescriber decision-making and clinical workflow. AMIA Annu Symp Proc 2009;548-52.

[23] van der Sijs H, Aarts J, van GT, Berg M, Vulto A. Turning off frequently overridden drug alerts: limited opportunities for doing it safely. J Am Med Inform Assoc 2008 Jul;15(4):439-48.

[24] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. 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 2012 Apr;81(4):232-43.

[25] Weingart SN, Massagli M, Cyrulik A, Isaac T, Morway L, Sands DZ, et al. Assessing the value of electronic prescribing in ambulatory care: a focus group study. Int J Med Inform 2009 Sep;78(9):571-8.

[26] van der Sijs H, van GT, Vulto A, Berg M, Aarts J. Understanding handling of drug safety alerts: a simulation study. Int J Med Inform 2010 May;79(5):361-9.

[27] Saleem JJ, Patterson ES, Militello L, Render ML, Orshansky G, Asch SM. Exploring barriers and facilitators to the use of computerized clinical reminders. J Am Med Inform Assoc 2005 Jul;12(4):438-47.

[28] Feldstein A, Simon SR, Schneider J, Krall M, Laferriere D, Smith DH, et al. How to design computerized alerts to safe prescribing practices. Jt Comm J Qual Saf 2004 Nov;30(11):602-13.

[29] Krall MA, Sittig DF. Clinician's assessments of outpatient electronic medical record alert and reminder usability and usefulness requirements. Proc AMIA Symp 2002;400-4.

[30] Patterson ES, Nguyen AD, Halloran JP, Asch SM. Human factors barriers to the effective use of ten HIV clinical reminders. J Am Med Inform Assoc 2004 Jan;11(1):50-9.

[31] Wipfli R, Betrancourt M, Guardia A, Lovis C. A qualitative analysis of prescription activity and alert usage in a computerized physician order entry system. Stud Health Technol Inform 2011;169:940-4.

[32] Khajouei R, Peek N, Wierenga PC, Kersten MJ, Jaspers MW. Effect of predefined order sets and usability problems on efficiency of computerized medication ordering. Int J Med Inform 2010 Oct;79(10):690-8.

[33] Russ AL, Saleem JJ, McManus MS, Zillich AJ, Doebbling BN. Computerized medication alerts and prescriber mental models: observing routine patient care. 2009 p. 655-9.

[34] Horsky J, Kaufman DR, Patel VL. Computer-based drug ordering: evaluation of interaction with a decision-support system. Stud Health Technol Inform 2004;107(Pt 2):1063-7.

[35] Khajouei R, de Jongh D, Jaspers MW. Usability evaluation of a computerized physician order entry for medication ordering. Stud Health Technol Inform 2009;150:532-6.

[36] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005 Mar 9;293(10):1197-203.

[37] Chused AE, Kuperman GJ, Stetson PD. Alert override reasons: a failure to communicate. AMIA Annu Symp Proc 2008;111-5.

[38] Duke JD, Bolchini D. A successful model and visual design for creating context-aware drug-drug interaction alerts. AMIA Annu Symp Proc 2011;339-48.

[39] Trafton J, Martins S, Michel M, Lewis E, Wang D, Combs A, et al. Evaluation of the acceptability and usability of a decision support system to encourage safe and effective use of opioid therapy for chronic, noncancer pain by primary care providers. Pain Med 2010 Apr;11(4):575-85.

[40] van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006 Mar;13(2):138-47.

[41] Chan J, Shojania KG, Easty AC, Etchells EE. Usability evaluation of order sets in a computerised provider order entry system. BMJ Qual Saf 2011 Nov;20(11):932-40.

[42] Hartmann Hamilton AR, Anhoj J, Hellebek A, Egebart J, Bjorn B, Lilja B. Computerised Physician Order Entry (CPOE). Stud Health Technol Inform 2009;148:159-62.

[43] Nielsen J. Enhancing the explanatory power of usability heuristics. New-York: ACM; 1994 p. 152-8.

Chapter 5. Methods uncovering usability issues in medication-related alerting functions: results from a systematic review

Romaric Marcilly, Francis Vasseur, Elske Ammenwerth, Marie-Catherine Beuscart-Zéphir

Accepted for MIE 2014

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

5.1. 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 [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 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 [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?

5.2. 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 or 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.

5.3. 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 (Table 19). The agreement score between both experts on the categorization of the papers according to the type of evaluation is almost perfect (Cohen's $\varkappa = 0.94$). 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 19 - 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 evaluations	Instances 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 20 presents the number of papers reporting usability flaws, usage problems and negative outcomes according to the type of evaluation used.

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.

Table 20 - Number of papers reporting usability flaws, usage problems and 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
		Error management	0	0	1	0
	Specific	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
Outcomes			0	3	7	4

5.4. 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 [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 or 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 or simulation"). More specifically, just like other papers do [4] they question the power of detection of "user-testing or simulation" and their coverage in term usability flaws. Currently, the question of the power of "user-testing or simulation" is handled through calculating the optimal size of the sample of participants to observe in order to uncover all significant usability flaws [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.

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] Beuscart-Zephir MC, Elkin P, Pelayo S, Beuscart R. The human factors engineering approach to biomedical informatics projects: state of the art, results, benefits and challenges. Yearb Med Inform (2007): 109-27.

[4] Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 78 (2009): 340-53.

[5] Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. AMIA Annu Symp Proc (2007): 26-30.

[6] Marcilly R, Ammenwerth, E, Vasseur F, Roehrer E, Beuscart-Zephir MC. Usability flaws of medication-related alerting functions: a systematic review. Int J Med Inform. Submitted.

[7] Marcilly R, Beuscart-Zephir MC, Ammenwerth E, Pelayo S. Seeking evidence to support usability principles for medication-related clinical decision support (CDS) functions. Stud Health Technol Inform 192 (2013), 427-31.

[8] Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behaviour and Information Technology 6 (1997), 220-31.

[9] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. 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] Schmettow M, Vos W, Schraagen JM. 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] Peute LW, Driest KF, Marcilly R, Bras Da Costa S, Beuscart-Zephir MC, Jaspers MW. A framework for reporting on human factor/usability studies of health information technologies. Stud Health Technol Inform 194 (2013), 54-60.

Chapter 6. Perceived usefulness of a usability issues reporting form to help understand "usability-induced use-errors": a preliminary study

Romaric Marcilly, Cesar Boog, Nicolas Leroy, Sylvia Pelayo

Accepted for MIE 2014

This work has received research funding from the French National Research Agency under contract ANR-11-TecSan-011-01.

Abstract

The Medical Device regulation requires manufacturers to anticipate and prevent risks of use errors of their medical device. However, manufacturers experience difficulties to understand the concept of "usability-induced use-errors". Based on a "usability framework" aiming at describing the relationship between usability design principles, usability flaws, usage problems, and outcomes, a usability evaluation reporting form had been designed to support understanding the use-error concept. This paper reports the preliminary evaluation of the perceived usefulness of this form. Results show that manufacturers found helpful the presentation of the results of a usability evaluation through this form for it supports the understanding of the usability origins and the consequences of use-errors. Even if the use of this reporting form should be made easier as usability experts experience difficulties to fill it, it seems a promising way to clearly present "usability-induced use-errors" to manufacturers.

Keywords

Human engineering; Heuristic inspection; Usability; Pain monitor; Use error;

6.1. Introduction

Manufacturers are asked by medical devices' (MD) regulatory authorities to anticipate and manage risks of use-errors related to poor design decision made during the design process [1]. Yet, manufacturers experience difficulties to understand what use-errors concretely refer to and how to prevent them. One reason for those difficulties may be found in a misunderstanding of the impact of their design choices on potential use-errors [2].

A use-error is an "act or omission of an act that results in a different medical device response than intended by the manufacturer or expected by the user" [3]. They are part of the "technology-induced errors", *i.e.,* errors induced by the introduction of health information technology [4]. Amongst use-errors, a subset origins in the poor design of the device in terms of usability, namely "usability-induced use-errors".

Providing manufacturers with usability evaluation reporting forms that structure the description of potential use-errors may help them understand use-errors and their origin in usability. This paper proposes a preliminary evaluation of the perceived usefulness of such a usability study reporting form. We took the opportunity of a formative heuristic inspection of a medical device to gather perceptions of manufacturers and usability experts' on this form.

6.2. Background

The reporting form has been developed from a "usability framework" [5] aiming at describing the relationship between usability design principles, usability flaws, usage problems, and outcomes (Figure 8).

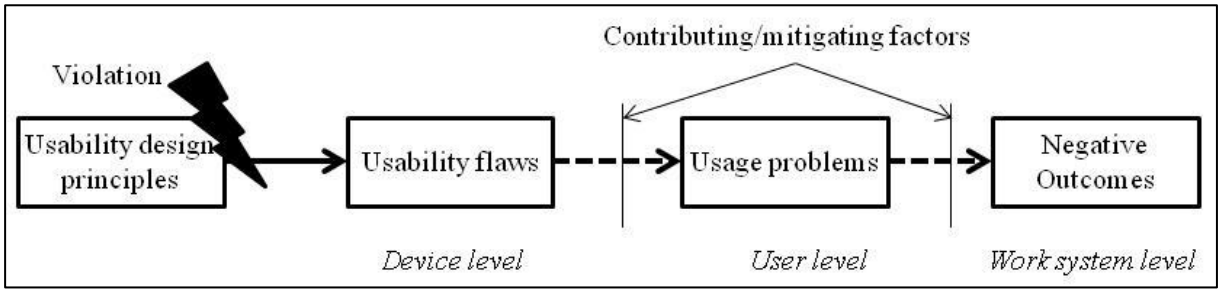


Figure 8. Schematic representation of the origin of usability flaws and of the propagation of their consequences towards the user and the work system in which it is implemented (adapted from [5])

Usability design principles (also usability heuristics or criteria) are recommendations for good practice in usability [6]. When those principles are not applied during the design of a product, usability flaws appear. Those flaws are noticed from the device perspective. When the device is put in use, flaws may negatively impact users' experience. Those negative experiences are usage problems. Finally, usability flaws may also have consequences on the work system in terms of performance, workarounds and patient-safety (*i.e.,* negative outcomes). The propagation of the infraction of the usability principles up to the work system is rather probabilistic than causal: mitigating factors independent from the device characteristics

(*e.g.,* high training, regular workload or adapted resilience) may prevent this propagation; au contraire, contributing factors (*e.g.,* lack of training or high workload) may facilitate this propagation. In the framework, "usability-induced use-errors" are a specific type of usage problems which consequences are noticeable at the outcomes level.

6.3. Study context

NIPE® is a comfort monitor displaying graphically the result of an algorithmic transformation of the cardiac frequency retrieved from patient's monitoring [7]. It supports monitoring the well-being of newborns especially for babies born with delivery complications and preterm or ill babies. Usability experts have been solicited to ensure the device be compliant with usability requirements, *i.e.,* safety use, to go to CE marking. Their intervention included applying the usability engineering process. The results of the present study were recorded during the first formative evaluation of this prototype through a heuristic evaluation.

6.4. Methods

6.4.1.1. Heuristic evaluation

Three usability experts independently performed a heuristic inspection of the NIPE® using Scapin and Bastien's list of usability principles [6]. They followed frequent and reasonably foreseeable worst use case scenarios based on the analysis of the intended context of use analysis (details in [8]). To stick to risk management requirements, for each usability flaw, the risk associated to their consequences was estimated (low, medium or high) as a function of their severity and frequency. Two involved experts (RM and CB) already took part in the intended context of use analysis; the third one (NL) read the detailed report of the analysis.

In contrast with usual heuristic inspections, this one was supported by a structured reporting form adapted by RM from the "usability framework" (Table 21).

6.4.2. Perceived usefulness evaluation

Data on perceived usefulness of the reporting form has been recorded opportunistically via handwritten notes by RM. They were collected from four representatives of the manufacturers (chief executive, designer, engineer and quality control officer). To see the benefice-risk ratio of the structured reporting form, it was necessary to evaluate its perceived usefulness but also if difficulties were encountered filling it in. Therefore, comments from two usability experts (CB & NL) were also collected to get their perception on filling in the form.

Table 21 - Instances of reports of usability flaws through the proposed reporting form.

Usability principle	Violation (usability flaw)	Potential usage problem	Potential outcome	Probability x Severity = Risk
Significance of codes	There is no information about the meaning of the curve and the smiley (Figure 9, right)	Difficult to interpret, wasting time and inability to understand	Decision based on a wrongly understood information	Moderate x Moderate = Medium
Consistency	"Instantaneous" NIPE® index corresponds to the "averaged index" curve not to the "instantaneous index" one (Figure 9, left)	Inability to understand, wrong representation of the patient status, doubt about the reliability of the device decreasing confidence	Decision based on a wrongly understood information	Moderate x Minimal = Low

Recorded comments were classified by RM as advantages and drawbacks either for filling in the form or for interpreting its content.

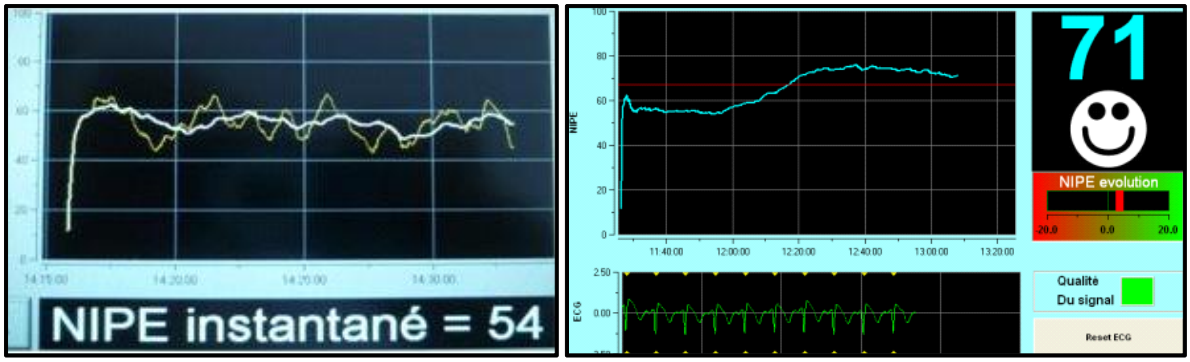


Figure 9. Screenshots of the "instantaneous" NIPE® screen (left) and of the main NIPE® screen (right).

6.5. Results

A total of 46 usability flaws are uncovered. A huge majority (40) are assigned with a low risk: they mostly deal with non prompting icons and graphical elements along with the handling of the setting features. Six usability flaws are considered with a medium risk: they are related to the difficulty to interpret the NIPE index. For instance, there is a potential risk of use-error (usage problem) related to the misinterpretation of the valence of the NIPE index which could lead to a wrongly based therapeutic decision.

As for the perception of the usefulness of the reporting form, users' main comments are synthesized in Table 22. Manufacturers' representatives consider this way to present the results is useful to understand the origins and consequences of use-errors. As for usability experts, they think it can be useful to help readers to understand the origin of the use-errors because it presents more clearly and in a more structured way the results of the inspection. However, they reported also difficulties to fill in the form, *e.g.,* it is sometimes difficult to anticipate the usage problems and outcomes. Nonetheless, they think this form

could improve the consistency of their reports. In sum, whatever the profile of the user, no perceived drawback was reported about interpreting usability results through the form.

Concerning the filling in of the report, an important insight is that the usability expert who did not analyze the context of use (NL) was unable to anticipate potential usage problems and outcomes along with their related risks.

Table 22 - Synthesis of the main comments from manufacturers' representatives and usability experts on the usefulness of the reporting form. Comments are sorted as perceived advantages or drawbacks and according to the potential function of the reporting form (support to report vs. support to interpretation).

		Usability experts	**Manufacturers' representatives**
Data reporting	Advantages	Compels to have the mind clear about the description of the usability flaw, the potential usage problems and outcomes.	Non applicable
	Drawbacks	Difficult to make easy-to-distinguish descriptions of usage problems and outcomes (hard "exercise in style"). Difficult to anticipate some usage problems/outcomes along with the risk level even more when the expert did not perform the context of use analysis. Not obvious whether the estimated risk is related to usage problem, outcome or both. It makes harder its estimation.	Non applicable
Data interpretation	Advantages	More precise description of the usability flaws: straight to the point. Clarify the consequences of the usability problems. Outcomes related to patient safety risks are clearly identifiable.	It is important distinguishing how the usability flaws can affect the users from the use-errors they can induced. Faster identification of the origins of potential use-errors.
	Drawbacks	/	/

6.6. Discussion

This paper presents a first attempt to clarify the concept of "usability-induced use-errors" during usability evaluation by turning a "usability framework" into a reporting form. This preliminary study focuses on the perceived usefulness of this form. Overall, manufacturers and usability experts consider this form can be useful to understand "usability-induced use-errors"; however, usability experts reported difficulties to fill in it. Those difficulties could be fixed by explaining more clearly how to evaluate the severity of the consequences of a usability flaw. In sum, basing the usability evaluation reporting form on a framework linking usability principles, usability flaws, usage problems and negatives outcomes seems a promising way to help manufacturers understand "usability-induced use-errors". Another important result that could have been expected is the impossibility for a usability expert to anticipate usability flaws' consequences without having performed first an analysis of the intended context of use.

There is an increasing need for standardized descriptions of usability flaws related to use-errors [9] (a) to capitalize on them and build databases allowing comparisons across systems and (b) to use those databases to design in an informed way similar medical devices. Yet, databases of structured descriptions of "usability-induced use-errors" are currently often confidential or hardly accessible. Reporting forms can structure them: *e.g.,* Peute et al. [10] developed an efficient scheme to capitalize unambiguously usability facts; however, this scheme is research-oriented and not adapted to manufacturers' needs. An important added-value of the present reporting form is that, besides providing a detailed report of potential use-errors, their underlying usability causes and their potential impact, manufacturers understand it.

References

[1] European Parliament and Council. Council Directive 2007/47/EC (2007), Official Journal L247/21.

[2] Pelayo S, Bras Da Costa S, Leroy N, Loiseau S, Beuscart-Zephir MC. Software as a medical device: regulatory critical issues. Stud Health Technol Inform 183 (2013), 337-42.

[3] International Electrotechnical Commission. Medical Devices - Application of usability engineering to medical devices (Rep. No. ISO 62366). International Electrotechnical Commission, Geneva, 2008.

[4] Kushniruk AW, Triola MM, Borycki EM, Stein B, Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. Int J Med Inform 74 (2005), 519-26.

[5] Marcilly R, Beuscart-Zephir MC, Ammenwerth E, Pelayo S. Seeking evidence to support usability principles for medication-related clinical decision support (CDS) functions. Stud Health Technol Inform 192 (2013), 427-31.

[6] Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behaviour and Information Technology 6 (1997), 220-31.

[7] de Jonckheere J, Rommel D, Nandrino J, Jeanne M, Logier R. Heart rate variability analysis as an index of emotion regulation processes: Interest of the Analgesia Nociception Index (ANI). Conf Proc IEEE Eng Med Biol Soc (2012), 3432-5.

[8] Marcilly R, Bras Da Costa S, Boog C, Beuscart-Zephir MC, de Jonckheere J, Pelayo S. Impact of the context of use analysis for the extension of an existing medical device: an analgesia monitor case study. Stud Health Technol Inform 194 (2013), 139-44.

[9] Peute LW, Driest KF, Marcilly R, Bras Da Costa S, Beuscart-Zephir MC, Jaspers MW. A framework for reporting on human factor/usability studies of health information technologies. Stud Health Technol Inform 194 (2013), 54-60.

[10] Peute MW, Khajouei R, Hassman A, Jaspers MW. Classification and priorization of usability problems using an augmented classification scheme. J Biomed Inform 44 (2011), 948-57.

Discussion and Conclusion

1. Summary of aims and results

This work focuses on usability issues of Health Information Technology (HIT). Its aim is two-fold:

- To participate in improving the accumulation of usability knowledge for HIT; and,
- To provide synthetic structured easy-to-use HIT usability knowledge with a clear coverage.

Both topics were addressed together along with questioning methods used to cumulate usability data and methods used to report them. Those aims are applied to medication alerting systems. Table 23 presents a summary of the results for the questions addressed in introduction.

The first result of this work is the "usability framework". This framework allows structuring the relations between usability design principles, usability flaws, usage problems and negative outcomes in the work system.

The reviews reported in chapters 2 to 4 identified 168 usability flaws. Amongst them, 107 were specific to medication alerting systems; they describe violations of 63 usability design principles dedicated to those systems (chapter 1). Those principles can be structured into 6 themes (improve the signal-to-noise ratio, fit clinician's workflow, support collaborative work, display relevant information, make the system transparent and provide useful tools). Violating those principles causes a large variety of difficulties. It impacts the user from a cognitive, a behavioral, an emotional and an attitudinal perspective; ultimately those violations have negative consequences on the workflow, on the effectiveness of technology, on the medication use process and, more dangerous, in terms of patient safety.

Amongst all methods applied to identify usability issues, observing in-situ the actual use of the technology is the method that is associated with the largest report of usability flaws along with usage problems and negative outcomes in the work system (chapter 5). As for reporting usability flaws and their consequences, a preliminary study has shown that the reporting form adapted from the "usability framework" is perceived as useful by both designers and HF experts; it presents clearly the origin of the usability issues and their consequences (chapter 6).

2. Main limits and strengths of the approach

The review studies carried out may have been impacted by various limits and biases: *e.g.*, publication and selective reporting biases, incompleteness or absence of significant data, non systematic searching of lists of usability design principles. Those biases and limits have already been discussed in each related chapter. It is important to be aware of them to appreciate precisely the reliability and the extent of the results presented.

Table 23 - Summary of the results for each addressed question.

Question 1. What are the usability flaws of medication alerting functions identified in published studies? *(chapter 2)*

In the 26 papers included in the systematic review, 168 usability flaws were identified (including 107 flaws specific to medication alerting systems). Those flaws represent a great diversity of the manifestations of the violations of usability design principles. They were sorted in 13 categories of general usability and specific usability flaws. Specific types of flaws refer to: issues of low signal-to-noise ratio of the alerts; incomplete content of the alerts; transparency issues; issues of alert's presentation mode and timing; missing alert features; task and control distribution issues.

Question 2. What types of usage problems and negative outcomes originating in identified usability flaws are reported in usability studies of medication alerting functions? *(chapter 3)*

A secondary analysis of the 26 papers uncovered 111 usage problems and 20 negative outcomes in the work system. The consequences of a poor usability of medication alerting systems are very varied. First, users are impacted in their cognitive, behavioral, emotional and attitudinal components. Then, the performance and the work system are impaired: the workflow is negatively impacted; the effectiveness of the technology is decreased; the medication management process is disturbed and patient safety issues arise.

Question 3. What are the cause-consequence links between usability flaws, usage problems and negative outcomes reported in medication alerting functions? *(chapter 3)*

One hundred and twenty-nine associations between usability flaws and usage problems were extracted along with 53 complete associations between usability flaws, usage problems and negative outcomes. Some trends highlight the important roles of workload issues, guidance issues and low signal-to-noise ratio issues along with the role of information understanding. However, there are only few converging lines of ongoing influence from usability flaws up to negatives outcomes: the clearest line is the absence of direct consequences of attitudinal and emotional issues on the work system. Nonetheless, a narrative analysis of the associations gathered renders the daily difficulties faced by the clinicians while using the alerting systems.

Question 4. What are the usability design principles specific to medication alerting systems that are reported in literature? *(chapter 4)*

Eight papers reporting lists of usability design principles were analyzed. A total of 125 unique design principles were extracted from those papers and synthesized into 54 principles. Those principles were structured into 6 themes: improve the signal-to-noise ratio; fit the clinician's workflow; support collaborative work; display relevant information; make the system transparent and provide useful tools.

Question 5. How well usability flaws reported in the literature are matched with the usability design principles? *(chapter 4)*

The synthesized usability design principles have been matched with the 107 usability flaws specific to those systems identified in chapter 2. The results highlight that 22 principles are matched directly by actual usability flaws, 9 principles need to be slightly extended to be matched and 23 are not matched at all. Besides, 9 new principles have been defined for all usability flaws to be matched. This leads to a final list including 63 usability design principles All usability flaws specific to medication alerting systems are covered by the final set of usability design principles. Conversely, a large number of usability design principles are matched with actual usability flaws.

Question 6. What are the methods used that detect facts on usability in medication alerting systems? What type of usability issues those methods allow to report? *(chapter 5)*

The types of methods applied to uncover usability issues in the 26 papers included in the review were collected and the studies were sorted in four kinds of evaluation ("expert evaluation", "user-testing/simulation", "on site observation" and "impact studies"). Results show that a very large set of methods are used (from heuristics inspection or focus group to shadowing or retrospective analysis). Most papers report "on site observation" evaluations. This is the only kind of evaluation for which every kind of usability flaws, usage problems and negative outcomes are reported: this type of evaluation provides a valuable insight on the question of the usability of medication alerting systems.

Question 7. How useful is a usability reporting form that distinguishes clearly usability design principles, usability flaws, usage problems and negative outcomes? *(chapter 6)*

The perceived usefulness of this reporting form was evaluated with manufacturer's representatives and Human Factors experts. Results reveal that this form is a valuable insight for the manufacturer's representatives to understand the different usability concepts. As for the Human Factors experts, filling in this form is an interesting exercise in style.

Several methodological strengths have also been highlighted all along the manuscript: *e.g.*, identification of papers and extraction/categorization of data by at least by two Human Factors (HF) experts with experience in usability of medication alerting systems. Another instance of the strength of that work deals with the robustness of the results. Indeed, the categories of flaws created are supported by instances of usability flaws from various papers: this overlapping of data ensures the reliability of this result. This good empirical support is also observed for usage problems and, at a lesser extent, for negative outcomes. Moreover, while they come from different sets of papers, usability flaws are quite easily matched with synthesized usability design principles, needing only few adjustments of existing principles. This good fitting between the two separate sets of data guarantees also a good validity of the results. Finally, at the level of the whole approach engaged, the tasks carried out do not settle for accumulating existing data to create knowledge (chapters 2 to 4). They also considered how to improve identifying (chapter 5) and reporting (chapter 6) usability issues in order to enhance the completeness and the structure of future results of usability studies.

3. Contributions of this work

Structuring usability knowledge: collection and report

The "usability framework" has been developed to structure several concepts about usability (chapter 1). This framework supported the systematic review process (i) methodologically, by guiding the selection of the papers, and (ii) by providing an interpretative frame for the results. Using this framework proved to be a significant help in order to clearly distinguish usability flaws from usage problems and negative outcomes in the paper analyzed. Indeed, in the current state of the literature, descriptions of usability issues are very confused.

In order to enhance the quality of the report of usability issues, the "usability framework" has been turned into a usability issues reporting form. Applying this form to report usability issues would enable to structure in a consistent way usability data. Using this form would ultimately support accumulating, in a structured way, usability data to turn them into knowledge.

Towards an evidence-based usability knowledge for medication alerting systems

The results of this work, in terms of usability knowledge, provide a first step to develop evidence-based usability design principles for medication alerting systems. This work makes available a list of synthesized usability design principles illustrated by their actual violations and their negative consequences. This knowledge provides evidence that violating usability design principles when designing medication alerting systems may cause detriments for the clinicians and the work system. This knowledge must be interpreted carefully. Indeed, even if results revealed that violating usability design principles gives rise to detriments for the clinicians and their work system (including the patient), this does not mean that applying those principles is a guarantee of success in the design, implementation and usage of the

technology: several technological and implementation factors may ruin a good usability (*e.g.*, wrong interconnectivity with other hospital systems).

The violations of those principles and their consequences may be underestimated due to publication and selective reporting biases. Therefore, this knowledge is incomplete. It must be strengthened (i) by examining data from other sources (*e.g.*, data from usability-oriented incident report systems) and (ii) by updating regularly the lists of reported flaws and their consequences.

According to the original concept of "evidence-based medicine" [1], evidence coming from reviews of literature is not a goal in itself: it must be delivered to practitioners (here designers and HF experts) who have to integrate it with their individual expertise to make informed design decisions. Therefore, this "evidence-based usability" knowledge must now be adapted to be given in a suitable format to designers and HF experts.

4. Transferability of this work

Transferability of the usability results

This work has been performed on medication alerting systems but its overall aim deals with the usability knowledge on all types of HIT. In chapters 2, 3 and 4, results that deal with general usability design principles and related flaws are per se extendable to any kind of computerized interactive systems.

As for usability design principles dedicated to medication alerting systems and related usability flaws, they are, per se, specific to medication alerting systems. However, part of the accumulated knowledge could potentially be extended to other types of clinical alerting systems (*e.g.*, abnormal lab results alerting systems) depending on how specific to medications is this knowledge. For instance, principles and flaws dealing with the medication part of the alert (*e.g.*, list of the mandatory information to display within the alert) cannot be transferred to other types of clinical alerting systems. On the contrary, principles and flaws related to the interruptive characteristics of the alerts could be applied to any kind of alerting system.

As for usage problems and negative outcomes related to those general and specific flaws, they depend closely upon the characteristics of the end-users and of their work system. Therefore, the relevance of extending results on usage problems and negative outcomes to other types of HIT must be evaluated on a case to case basis in regards to the specificities of the end-users and of the work system.

Transferability of the approach

This work does not aim only at providing useful data; it also aims to develop an approach to improve the accumulation of the usability knowledge. The proposed approach, that involves reviewing the literature in search of usability flaws and their consequences and then matching them with related usability design principles, may be applied to any kind of HIT as soon as usability data on this technology are available. However, this approach is very time-consuming: besides gathering usability data to develop new knowledge, the already developed usability knowledge must be regularly updated. Applying it to all kinds

of HIT seems hardly fundable. This work cannot be performed by only one HF team: it must be a collaborative work involving the whole HF community in the medical informatics field. HIT vendors should also be involved to make accessible the results of the evaluation of their HIT: this will not be an easy thing due to competitive interests.

One way to improve the efficiency of this approach would require being more pro-active and feeding gradually a usability knowledge base, instead of resting only on retrospective analyses of the literature. Such a knowledge base remains to be developed.

5. Future researches

Validating the reporting form

The usability reporting form adapted from the "usability framework" has been the object of a preliminary evaluation. The light issues identified during this evaluation (*e.g.*, misunderstanding of how to quote the risk associated to a usability issue) must be fixed. Then, a complete evaluation of its usefulness, of its impact on the usability evaluation process and on the quality of usability reports must be performed with a larger sample of participants.

Developing the usability knowledge base

This work constitutes a first contribution to developing a usability knowledge base. It has to be developed in a joint effort of the community of HF experts in medical informatics. Other databases already exist in the field of medical informatics (*e.g.*, IT Evaluation Database[4] [2]) but they provide mainly the references and summaries of studies and they are not specially designed for usability topics. The usability knowledge base must go a step further: it must be designed as a repository, as exhaustive as possible, of the usability flaws and their consequences known for HIT. Sorting keys (*e.g.*, "type of HIT", "function" or "screen") must be defined along with the feeding and updating processes must. But, first of all, it is necessary to define precisely the intended users of this knowledge base. This will impact several characteristics of the knowledge base including the classification and the wording of the knowledge within it.

HF experts and designers of HIT are the professionals targeted by the present work. For this reason, usability flaws were ordered according to usability heuristics: this kind of classification is usable by HF experts and its purpose and organization can also be explained to trained designers. However, the knowledge base may also be used by HF researchers aiming to improve designing practices: yet, heuristics do not represent the cognitive processes that are related to the usability flaws. In this case, other classification schemes should be favored (*e.g.*, extended Usability Action Framework [3]). Therefore, the final usability knowledge base should allow a dual classification of usability data: one dedicated to

[4] http://evaldb.umit.at/index.htm

designers and one for HF experts and researchers. The connections between both types of classifications must be studied.

Turning the list of illustrated usability design principles into a useful tool

Finally, the results of this work, especially those presented in chapter 4, should be used to develop a tool supporting the design and the evaluation of medication alerting systems from a usability perspective. This tool is not to stand for an actual HF expertise during the project: it is a supplementary source of knowledge to improve the effectiveness of the design and the evaluation process.

For now, a structured list of usability design principles illustrated by actual instances of their violations for which the negative consequences are known is available. This list could be turned into evidence-based heuristics. Similar heuristics have already been proposed in the field of HIT safety [4]. Designing this kind of tool has the same requisites as designing the usability knowledge base: adapting the tool to its intended users. It requires involving a panel of experts to organize and to word the components of the heuristics. Then, the usability of the tool must be evaluated and improved iteratively with actual designers and HF experts. Once the heuristics are sufficiently usable, their effectiveness will be tested during actual design/evaluation projects. Finally, a continuous improvement process will have to be applied once the final tool is used.

This kind of tools would have been very useful in our lab if it had been available during projects in which we had to support the design and to evaluate various medication alerting systems (*e.g.*, such as during the "Patient Safety through Intelligent Procedures in medication" (PSIP) project[5] [5-8]): it would have help us not reinvent the wheel at each project and save time. At a larger extent, evidence-based heuristics dedicated to the usability of HIT could provide a relevant support to designers, HF experts and researchers working in this field. There should be incentives to develop and share this kind of tools.

6. Conclusion

Usability applied to HIT is a quite recent research field that faces difficulties to become a mature science. One necessary step on this way is to be able to accumulate and reuse data from previous studies to create new knowledge. This work participates in improving the usability knowledge by reusing data from former studies. However, this task can be fully achieved for each HIT only if the whole HF community in the medical informatics field and HIT vendors for which this community works, agree to share usability data and to manage them together.

[5] http://psip.univ-lille2.fr/prototypes/public/

References

[1] Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996 Jan 13;312(7023):71-2.

[2] Ammenwerth E, de KN. An inventory of evaluation studies of information technology in health care trends in evaluation research 1982-2002. Methods Inf Med 2005;44(1):44-56.

[3] Khajouei R, Peute LW, Hasman A, Jaspers MW. Classification and prioritization of usability problems using an augmented classification scheme. J Biomed Inform 2011 Dec;44(6):948-57.

[4] Borycki E, Kushniruk A, Carvalho C. A methodology for validating safety heuristics using clinical simulations: identifying and preventing possible technology-induced errors related to using health information systems. Comput Math Methods Med 2013;2013:526419.

[5] Hackl WO, Ammenwerth E, Marcilly R, Chazard E, Luyckx M, Leurs P, et al. Clinical evaluation of the ADE scorecards as a decision support tool for adverse drug event analysis and medication safety management. Br J Clin Pharmacol 2013 Sep;76 Suppl 1:78-90.

[6] Marcilly R, Leroy N, Luyckx M, Pelayo S, Riccioli C, Beuscart-Zephir MC. Medication related computerized decision support system (CDSS): make it a clinicians' partner! Stud Health Technol Inform 2011;166:84-94.

[7] Marcilly R, Bernonville S, Riccioli C, Beuscart-Zephir MC. Patient safety-oriented usability testing: a pilot study. Stud Health Technol Inform 2012;180:368-72.

[8] Marcilly R, Beuscart-Zephir MC, Beuscart R. Integrating Human Factors in an international research project: lessons learned from the PSIP project. Stud Health Technol Inform 2013;183:162-7.

Appendices

French summary

1. Introduction

Les Technologies de l'Information en Santé (TIS) sont de plus en plus installées (*e.g.* en établissement de soin) dans le but d'améliorer la qualité des soins et la sécurité du patient. Malheureusement, ces technologies font face à des problèmes d'acceptabilité [1], peuvent être rejetées [2] et/ou causer des erreurs médicales [3] qui peuvent porter préjudice au patient voire même le tuer [4-7]. Une des origines de ces problèmes est une problématique de Facteurs Humains. Au-delà de problèmes organisationnels dus à des choix erronés d'installation, la mauvaise utilisabilité de ces technologies est souvent incriminée [5-7]. L'utilisabilité, ou encore "aptitude à l'utilisation", est "le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié" [8]. Appliquée à un outil, l'utilisabilité renvoie aux caractéristiques de cet outil qui le rendent son utilisation et son usage plus aisés pour les utilisateurs. Ces caractéristiques prennent leurs origines dans l'application de principes d'utilisabilité durant la conception et le développement d'un outil.

Pour empêcher les difficultés causées par des problèmes d'utilisabilité d'une technologie (*e.g.* erreurs d'utilisation), son utilisabilité doit être considérée tout au long de la conception, du développement et de l'évaluation de cette technologie. Cette nécessité est de plus en plus reconnue de manière légale : ce besoin est maintenant inclus dans les textes régissant le processus de marquage « Conformité Européenne » (CE) des dispositifs médicaux, ce qui comprend aussi certaines catégories de TIS comme les Systèmes Interactifs d'Aide à la Décision (SIAD) [9].

L'amélioration de l'utilisabilité des TIS est au cœur de la pratique ainsi que des recherches de l'équipe d'experts en Facteurs Humains du Centre d'Investigation Clinique pour les Innovations Technologiques (CIC-IT) de Lille. Dans cette optique, le processus de Conception Centrée sur l'Utilisateur (CCU) [10], allié aux méthodes Facteurs Humains [11;12], est régulièrement appliqué pour concevoir et évaluer différentes TIS [13-19]. Néanmoins, au sein du CIC-IT mais aussi dans d'autres structures de recherche et de développement, appliquer ces processus et méthodes est reconnu comme indispensable mais non suffisant pour concevoir des technologies « utilisables », *i.e.* sans problèmes d'utilisabilité. Il est en effet aussi nécessaire d'appliquer des connaissances en Facteurs Humains pour concevoir efficacement les technologies. Dans tous les projets de (re-)conception ou d'évaluation, deux types de connaissance sont nécessaires :

- La connaissance de la CCU et des méthodes associées ;
- La connaissance des principes d'utilisabilité à appliquer pour la technologie considérée.

La connaissance de la CCU et des méthodes associées est bien décrite durant les formations en Facteurs Humains mais aussi dans différents manuels et documents de référence, dont la norme harmonisée IEC 62366 qui adapte le processus d'ingénierie de l'utilisabilité à la prévention des risques d'erreurs d'utilisation des dispositifs médicaux [20].

La connaissance des principes d'utilisabilité est moins stable, moins structurée et aussi moins accessible pour plusieurs raisons :

- Les technologies évoluent rapidement ; il est difficile de trouver des principes d'utilisabilité confirmés pour les technologies complètement innovantes.
- Les principes d'utilisabilité ont besoin d'un certain niveau de généralité pour englober les catégories les plus importantes d'outils. Les systèmes informatiques interactifs sont une des catégories d'outils pour lesquels des principes d'utilisabilité ont été capitalisés depuis une quarantaine d'années.
- Les principes d'utilisabilité sont essentiellement basés sur l'application de connaissances théoriques provenant de la psychologie cognitive et de la physiologie, aux spécificités d'un type d'outil. Ces principes identifient les capacités et les limites de l'Humain qui doivent être prises en compte lors de la conception de l'outil. Cependant, pour un type d'outil donné, l'ensemble des principes d'utilisabilité est régulièrement modifié en réponse aux reports de problèmes d'utilisabilité identifiés pour des outils similaires soumis à une évaluation d'utilisabilité ou déjà installés.

Il est donc crucial de capitaliser des données empiriques sur les problèmes d'utilisabilité d'un type de technologie pour améliorer la précision et l'efficacité des principes d'utilisabilité.

En ce qui concerne les TIS, l'accumulation de ces données d'utilisabilité et la définition de principes d'utilisabilité ne fait que débuter. Il existe quelques publications fournissant des principes d'utilisabilité pour quelques TIS. Cependant ces principes sont éparpillés sur plusieurs types de documents pas toujours accessibles ; ils ne sont pas structurés et formulés de manière à être facilement compréhensibles par les concepteurs et parfois même par des experts en Facteurs Humains. Par ailleurs, la couverture de ces principes en termes de problèmes d'utilisabilité qu'ils peuvent traiter est très rarement décrite.

Ces diverses barrières empêchent les experts en Facteurs Humains d'utiliser de manière optimale la connaissance disponible sur les principes d'utilisabilité des différentes TIS : à chaque nouveau projet, ces experts perdent du temps à chercher, à comprendre et à s'approprier cette connaissance. A plus large échelle, ces barrières empêchent l'accumulation d'une connaissance sur les principes d'utilisabilité pour les TIS ce qui ralentit grandement les progrès des recherches en utilisabilité appliquée aux TIS.

Ce travail de thèse a deux objectifs :

- Participer à l'amélioration de la capitalisation de la connaissance sur l'utilisabilité des TIS ;
- Fournir une liste synthétique et structurée de principes d'utilisabilité pour les TIS tout en explicitant leur couverture en termes de problèmes d'utilisabilité.

Ces deux thèmes sont extrêmement intriqués et seront traités ensemble. Par ailleurs, traiter ces thèmes demande aussi de questionner les méthodes actuellement utilisées pour accumuler et reporter la connaissance en utilisabilité.

Accumuler des données d'utilisabilité demande d'abord de distinguer clairement les données qui sont du ressort de l'utilisabilité de l'outil de celles qui concernent l'usage ou l'impact de celui-ci. Pour définir précisément ce que sont ces données d'utilisabilité et les concepts attenants, un « cadre de travail en utilisabilité » a été développé. Dans ce cadre, 4 concepts sont mis en relation (Figure 10):

- Les principes d'utilisabilité. Ces principes sont des recommandations en termes d'utilisabilité pour la conception et l'évaluation de technologies.
- Les problèmes d'utilisabilité. Ce sont des violations des principes d'utilisabilité. Ces éléments sont décrits objectivement au niveau de la technologie.
- Les problèmes d'usage. Ces éléments représentent la manière dont l'utilisateur de la technologie vit les problèmes d'utilisabilité lors de son utilisation de la technologie. Ils concernent les processus cognitifs, les comportements et les sentiments des utilisateurs.
- Les conséquences négatives sur le système de travail. Ces éléments sont la résultante des problèmes d'utilisabilité au niveau du système de travail et de la performance en termes de sécurité, de communication, d'utilité de la technologie, etc.

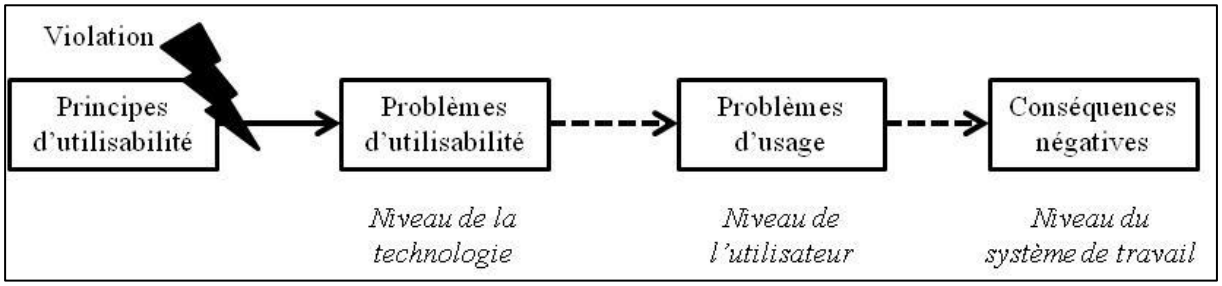


Figure 10. Le « cadre de travail en utilisabilité », représentation schématique de l'origine des problèmes d'utilisabilité dans la violation des principes d'utilisabilité et de la propagation de leurs conséquences à l'utilisateur (« problèmes d'usage ») et au système de travail (« conséquences négatives »), adaptée de [21]. Les flèches pointillées représentent la possibilité de propagation alors que la flèche pleine représente la conséquence automatique.

La propagation des conséquences des problèmes d'utilisabilité n'est pas causale (Figure 11) : elle dépend de plusieurs facteurs. Certains facteurs, indépendants des caractéristiques de la technologie (*e.g.* formation des utilisateurs, expérience et expertise, adéquation entre le niveau de résilience des utilisateurs et les demandes de la situation, habiletés cliniques) peuvent soit favoriser soit empêcher la propagation des problèmes d'utilisabilité. Ces facteurs influencent la propagation à la fois au niveau des problèmes d'usage et aussi au niveau des conséquences négatives sur le système de travail. Par exemple, si les cliniciens ne sont pas formés à une technologie et sont aussi débordés de travail, il est probable que les problèmes d'utilisabilité les affectent plus et que des problèmes d'usage et des conséquences négatives pour le système de travail soient générés. Au contraire, s'ils sont bien formés et avec une charge de travail gérable, les utilisateurs risquent moins d'être affectés par les problèmes d'utilisabilité et ainsi les problèmes d'usage et les conséquences négatives ont une moindre probabilité d'apparition.

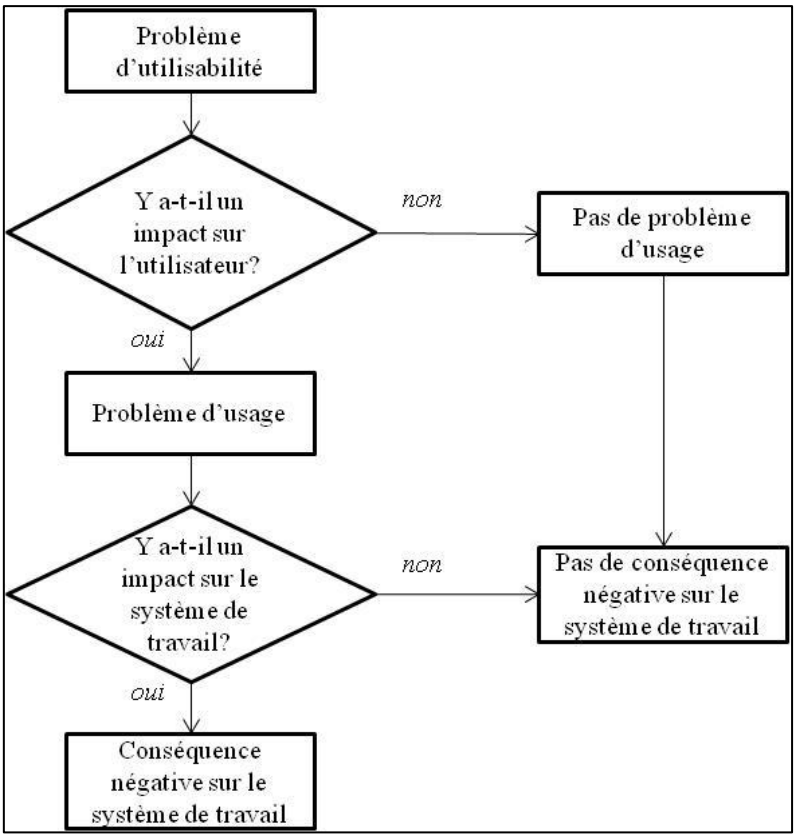


Figure 11. Représentation schématique de la propagation des problèmes d'utilisabilité à travers l'utilisateur et jusqu'au système de travail.

Rechercher les conséquences des violations des principes d'utilisabilité pour développer une connaissance en utilisabilité fait écho au concept de « médecine fondée sur les preuves » adapté à la l'utilisabilité. La « médecine fondée sur les preuves » a été mise en lumière par les travaux de Sackett [22] : cet auteur définit la « médecine fondée sur les preuves » comme « l'utilisation consciencieuse, explicite et judicieuse des meilleures données disponibles pour la prise de décisions concernant les soins à prodiguer à chaque patient, [...] une pratique d'intégration de chaque expertise clinique aux meilleures données cliniques externes issues de recherches systématiques ». Par analogie, l'« utilisabilité fondée sur les preuves » pourrait être définie comme « l'utilisation consciencieuse, explicite et judicieuse des meilleures données disponibles pour la prise de décisions concernant la conception et l'évaluation des technologies en termes d'utilisabilité » (adaptation de l'auteur). Les données d'utilisabilité proviennent surtout d'études observationnelles : les résultats issus de ces méthodes sont considérés comme étant des preuves de moindres qualité que les résultats issus de revues systématiques d'essais cliniques randomisés [23]. Même si les méthodes d'utilisabilité ne permettent pas de fournir la meilleure preuve possible, cela reste néanmoins la meilleure preuve disponible puisqu'il n'est pas possible de faire des essais cliniques randomisés dont l'objet est l'utilisabilité d'une TIS.

Ce travail de thèse propose une approche afin de rechercher la connaissance en utilisabilité :

- Rassembler les données d'utilisabilité rapportées dans la littérature en réalisant une revue systématique des études d'utilisabilité. Cette recherche vise les éléments du « cadre de travail en utilisabilité » qui sont directement liés à l'utilisabilité et qui sont peu influencés par des considérations théoriques : les « problèmes d'utilisabilité ». Les principes d'utilisabilité sont eux-aussi directement liés à l'utilisabilité mais leur formulation peut être facilement influencée par les conceptions théoriques de leur auteur.
- Rechercher les conséquences négatives des problèmes d'utilisabilité pour l'utilisateur et pour le système de travail. Pour ce faire, une analyse secondaire des publications incluses dans la revue systématique est à réaliser. Les liens entre les problèmes d'utilisabilité et leurs conséquences pour l'utilisateur et le système de travail doivent être identifiés.
- Recenser les principes d'utilisabilité puis les associer aux problèmes d'utilisabilité pertinents.

Le résultat de ce processus est une liste des erreurs d'utilisabilité à ne pas faire, complétées par leurs conséquences et les principes de conception qui pourraient les corriger.

L'approche décrite ci-dessus est appliquée à un type particulier de TIS, les SIAD et plus précisément les systèmes d'alerte médicamenteux. Ces systèmes présentent en temps réel au clinicien la connaissance clinique et pharmacologique au moment de sa prise de décision. Ces systèmes aident à améliorer la prescription de médicaments [24]. En général, ils améliorent la qualité et la sécurité des soins [25]. Cependant certaines études ne montrent pas cet impact positif [26;27] ; d'autres études montrent aussi que ces systèmes rencontrent des problèmes d'acceptation et d'utilisation [28-30]. Des problèmes d'utilisabilité sont souvent incriminés [31;32]. Les débats sur l'utilité et l'acceptation de ces systèmes ont donné lieu à une grande quantité de publications fournissant suffisamment de données pour réaliser une revue systématique.

Pour s'assurer que les données collectées dans la revue systématique sont valides, les méthodes qui ont permis de relever ces données (*e.g.* tests utilisateurs, observation) sont étudiées. Une étude est destinée à savoir si ces méthodes ont un impact sur les problèmes d'utilisabilité reportés. Par ailleurs, un problème relevé de manière récurrente dans les revues systématiques est le manque de complétude et la désorganisation des reports d'évaluation. Pour améliorer l'accumulation des connaissances sur l'utilisabilité des TIS, un formulaire de report des problèmes d'utilisabilité a été développé à partir du « cadre de travail en utilisabilité ». Une étude préliminaire de l'utilité perçue de ce formulaire a été réalisée.

Cette thèse répond à 7 questions :

- Quels sont les problèmes d'utilisabilité des systèmes d'alerte médicamenteux reportés dans les études publiées ? *(Question traitée dans le chapitre 2)*

- Quels sont les types de problèmes d'usage et de conséquences négatives pour le système de travail dont l'origine se trouve dans les problèmes d'utilisabilité des systèmes d'alerte médicamenteux ? *(Question traitée dans le chapitre 3)*
- Quelles sont les liens de causes à effets reportés dans la littérature entre les problèmes d'utilisabilité, les problèmes d'usage et les conséquences négatives pour le système de travail ? *(Question traitée dans le chapitre 3)*
- Quels sont les principes d'utilisabilité spécifiques aux systèmes d'alerte médicamenteux identifiés dans la littérature ? *(Question traitée dans le chapitre 4)*
- A quel point les problèmes d'utilisabilité reportés dans la littérature correspondent à ces principes d'utilisabilité ? *(Question traitée dans le chapitre 4)*
- Quelles sont les méthodes appliquées pour détecter les problèmes d'utilisabilité des systèmes d'alerte médicamenteux ? Quels types de faits d'utilisabilité (problèmes d'utilisabilité, problème d'usage, conséquences négatives) ces méthodes permettent-elles de reporter ? *(Question traitée dans le chapitre 5)*
- Le formulaire de report des problèmes d'utilisabilité est-il utile et en quoi? *(Question traitée dans le chapitre 6)*

2. Quels sont les problèmes d'utilisabilité des systèmes d'alerte médicamenteux reportés dans les études publiées ?

Dans le but de répondre à cette question, une revue systématique de la littérature a été réalisée. Pour être inclus dans la revue, les articles doivent être anglophones ou francophones et publiés après 1980. Trois critères d'éligibilité ont été définis.

- Les études publiées doivent concerner les systèmes d'alerte médicamenteux qui supportent la prescription de médicaments. Ces systèmes doivent être utilisés dans les hôpitaux ou en médecine de ville. Les systèmes d'alerte utilisés par les patients ont été exclus de la revue.
- Les études publiées doivent être des études d'utilisabilité, sociotechniques ou d'impact visant l'utilisabilité des systèmes. Seules les publications avec des méthodes et des résultats suffisamment décrits sont incluses.
- Seules les publications reportant des problèmes d'utilisabilité décrits de manière objective sont incluses. Ainsi, les articles reportant l'utilisabilité perçue de ces systèmes ont été exclus.

Des publications dans des journaux scientifiques ainsi que des actes de conférences ont été recherchés dans 3 bases de données, PubMed, Scopus et Ergonomics Abstracts, ainsi que dans les listes de références des papiers inclus. Les termes clefs utilisés sont présentés dans le Tableau 1.

La sélection des publications, l'extraction des données et leur analyse ont été réalisées par deux à trois experts en Facteurs Humains. Les principales données extraites étaient les problèmes d'utilisabilité. Ces

problèmes furent analysés par catégorisation suivant une liste de principes d'utilisabilité généraux [33]. Les problèmes spécifiques aux systèmes d'alerte médicamenteux furent ensuite classés de manière inductive.

Tableau 1 - Termes utilisés pour la recherche des publications.

	PubMed	**Scopus**	**Ergonomics Abstracts**
Termes liés aux systèmes d'alerte	Medical order entry systems, Medication alert system, Computerized physician order entry system, CPOE, Decision Support Systems, Clinical, Clinical decision support systems, CDSS	Medical order entry, Medication alert, Computerized physician order entry, CPOE, Clinical decision support, CDSS	
Termes liés aux facteurs humains	User-computer interface, Human engineering, Risk factors, Humans, Usability	User-computer interface, Human engineering, Risk factor, Human factor, Usability, Human-computer interaction	Non applicable

Le processus de sélection des articles est décrit dans la Figure 12. Parmi les 6380 publications identifiées, 26 remplissent tous les critères d'éligibilité.

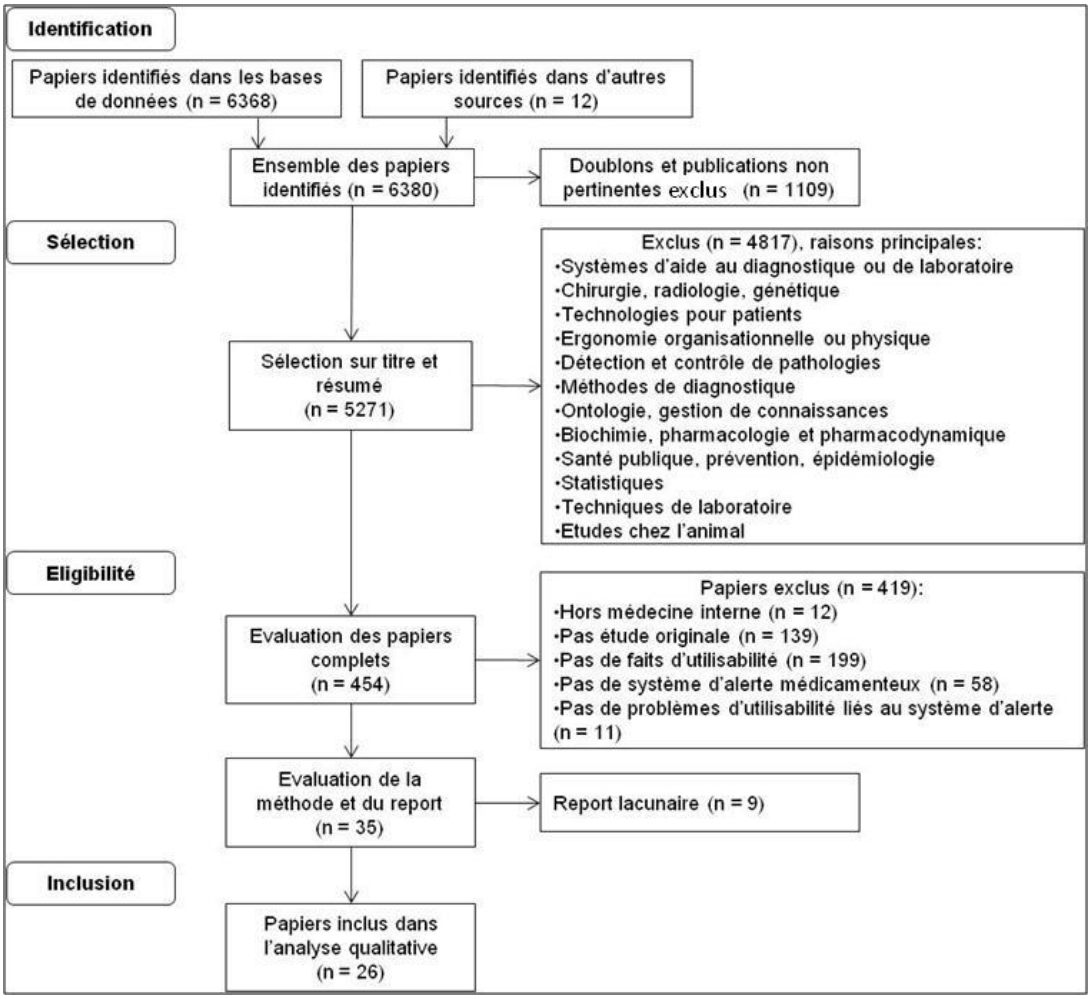


Figure 12. Processus de sélection des publications à inclure dans la revue systématique.

L'analyse des papiers inclus dans la revue a permis d'identifier 168 problèmes d'utilisabilité. Ces problèmes ont été classés en 13 catégories. Ces catégories sont présentées dans le Tableau 2.

Tableau 2 - Catégories de problèmes d'utilisabilité généraux et spécifiques aux systèmes d'alerte médicamenteux.

Problèmes d'utilisabilité généraux	
Problèmes de guidage	
Problèmes d'incitation: texte non clair, problèmes de mise en saillance des informations, information excentrée.	[34-41]
Pas de distinction par le format (couleur, forme) des niveaux de sévérité différents, des types d'alerte différents ou des types de message (erreur système vs. information médicale).	[42-45]
Problèmes de lisibilité: interligne insuffisant, utilisation de majuscules, éléments trop petits.	[38;42;43]
Utilisateur non informé qu'il a manqué une alerte.	[35]
Trop de distinction des informations par leur localisation: pas de regroupement d'alertes de même sévérité.	[45]
Problèmes de charge de travail	
Action minimale : trop d'actions à effectuer pour entrer ou obtenir une information (*e.g.* onglets).	[35;38;40;43;45-47]
Densité informationnelle: trop d'informations de différentes sortes dans la fenêtre, plusieurs alertes dans la même fenêtre, contenu de l'alerte présenté en un seul paragraphe.	[34;36;37;42;43;48]
Manque de concision	[39;44;49]
Problèmes de signifiance des codes	
Termes non intuitifs	[37;47;50-52]
Icônes non intuitifs	[37;53]
Problèmes d'homogénéité	
Comportement du système inconsistant d'une utilisation à une autre ou en fonction des données utilisées.	[36;54]
Problèmes de contrôle explicite	
Actions explicites des utilisateurs: la réponse du système ne correspond pas à l'action émise par l'utilisateur.	[53]
Contrôle utilisateur: aucun moyen d'annuler une action.	[46]
Problèmes d'adaptabilité	
Manque de flexibilité: le système ne prend pas en compte tous les types d'utilisateurs.	[43]
Problèmes de management de l'erreur	
Qualité des messages d'erreur: messages non clairs.	[43]
Problèmes d'utilisabilité spécifiques aux systèmes d'alerte médicamenteux	
Faible ratio signal / bruit	
Alertes non pertinentes au regard de l'expertise des utilisateurs, des habitudes du service, des bonnes pratiques, des connaissances pharmacologiques, des données utilisées, du cas patient, des actions des utilisateurs, de ce que le clinicien considère comme étant risqué, de la logique de soin.	[34-36;40;42-44;46;50-52;55]
Pas de description précise de la cause.	[35;36;42-44;49;51-53;56]

Les alertes sont redondantes: l'alerte apparaît très fréquemment / plusieurs fois durant la prise de décision, les actions correctives effectuées par le clinicien ne sont pas acceptées par le système, il n'y a pas de fonction permettant d'éteindre une alerte.	[39;42;43;46;49;52;53]
Problèmes dans le contenu de l'alerte	
L'information nécessaire pour prendre une décision est absente de l'alerte: les suggestions d'action, les données cliniques pertinentes du patient, le problème détecté par le système d'alerte, les preuves appuyant ce problème, la sévérité du problème, les informations nécessaires pour interpréter l'alerte.	[36-38;42;43;45;51]
Les suggestions de l'alerte sont erronées: les actions proposées ne prennent pas en compte le contexte clinique.	[50;53]
Le système n'est pas suffisamment transparent pour l'utilisateur	
Manque de transparence à propos de la manière dont le système fonctionne : il n'y a pas d'information à propos de l'échelle de sévérité des alertes, à propos des mises à jour des connaissances.	[34;36;43;45;46;48]
Manque de transparence à propos des données que le système utilise: toutes les données disponibles ne sont pas utilisées pour déclencher les alertes ou le système n'est pas conçu pour prendre en compte ces données.	[42;43;52;54;57]
Problèmes d'apparition des alertes: moment et mode de déclenchement	
L'alerte n'apparaît pas au bon moment pour aider le processus de prise de décision: avant que ce processus ne commence, au mauvais moment, ou après que la décision est prise.	[34;39;48;49;52;58;59]
Le traitement des données est lent.	[43;53]
Le mode de présentation des alertes ne correspond pas au processus de prise de décision: alerte pas suffisamment visible ou trop intrusive.	[52]
Problèmes de distribution des tâches	
L'alerte n'est pas présentée au "bon" clinicien ou uniquement au pharmacien.	[35;40;59]
Le système d'alerte permet aux utilisateurs d'entrer des commentaires qui ne sont présentés à personne.	[41;43]
Les alertes ne sont pas transférables d'un clinicien à un autre.	[53]
Problèmes des outils dans les alertes	
Outils manquants: il n'est pas possible de revoir une alerte ultérieurement, pas d'accès à des informations complémentaires à partir de l'alerte, pas d'outil dans l'alerte pour la résoudre rapidement.	[35;37;43;45;46;53]
Outils non adaptés aux besoins des utilisateurs.	[53]

Les biais de publication et de report ont peut-être impacté la représentativité des résultats : certains types de problèmes, faciles à décrire en quelques mots ou avec une impression d'écran (*e.g.* guidage), sont peut-être reportés plus facilement et donc plus fréquemment que des problèmes liés au comportement du système qui requièrent une longue description pour être compris (*e.g.* gestion des erreurs). Cependant toutes les catégories de problèmes identifiés sont alimentées par des exemples provenant de différents articles, ce qui assure une bonne fiabilité des résultats.

Malgré l'inclusion uniquement de publications de bonne qualité en méthode et en report, cette revue systématique a mis en exergue des lacunes dans le report de certaines données non essentielles pour cette étude (*e.g.* les caractéristiques du système d'alerte, contexte d'utilisation). Ce problème de report a déjà été

identifié dans d'autres études [60]. Il est absolument nécessaire d'améliorer le report des études d'utilisabilité.

Les problèmes d'utilisabilité généraux identifiés sont connus pour rendre l'utilisation du système difficile [33]. En ce qui concerne les problèmes d'utilisabilité spécifiques, certains d'entre eux sont discutés ponctuellement dans la littérature en raison de leur impact négatif sur l'utilité du système. Néanmoins, il n'existe aucune étude qui recense précisément les conséquences de ces problèmes.

3. Quels sont les types de problèmes d'usage et de conséquences négatives pour le système de travail dont l'origine se trouve dans les problèmes d'utilisabilité des systèmes d'alerte médicamenteux ?

Cette partie vise à explorer et à synthétiser les conséquences recensées des problèmes d'utilisabilité en termes de problèmes d'usage et de conséquences négatives pour le système de travail. Dans ce but, une analyse secondaire des 26 articles inclus dans la revue systématique présentée ci-dessus a été réalisée.

Les problèmes d'usage et les conséquences négatives sur le système de travail ont été extraits par deux experts en Facteurs Humains. Seuls les problèmes d'usage et les conséquences négatives reportés comme étant causés par des problèmes d'utilisabilité ont été considérés. De manière à n'avoir que des données fiables, seules les descriptions objectives et suffisamment évidentes à interpréter ont été extraites : les hypothèses esquissées par les auteurs des articles n'ont pas été analysées. Les problèmes d'usage et les conséquences négatives ont été identifiés et catégorisées. Enfin, les liens entre ces éléments ont été extraits et utilisés pour synthétiser les associations entre les catégories de problèmes d'utilisabilité, de problèmes d'usage et de conséquences négatives pour le système de travail.

Les problèmes d'usage ont été catégorisés de manière inductive par un tri de carte ouvert [61] ; les conséquences négatives ont été catégorisées par un tri de carte fermé [61].

Cent-onze problèmes d'usage et 20 conséquences négatives sur le système de travail ont été extraits. Les résultats montrent qu'une mauvaise utilisabilité cause véritablement des difficultés de nature variée. Cela impacte l'utilisateur dans ses processus cognitifs, son comportement, ses émotions et ses attitudes (Tableau 3). Au final, les problèmes d'utilisabilité impactent aussi négativement les processus de travail, l'efficacité de la technologie installée, le circuit du médicament et, plus important, la sécurité du patient (Tableau 4).

Tableau 3 - Catégories de problèmes d'usage liés à des problèmes d'utilisabilité et reportés dans la littérature.

Type de problème d'usage	Référence
Problèmes comportementaux	
Augmentation de la charge de travail due au système d'alerte.	[35;40;42;43;46;49;50;52;53;62]
Les utilisateurs n'utilisent pas du tout le système.	[39;43;48]
Les utilisateurs ignorent volontairement les alertes.	[35;37;39;40;42-44;49;53;56;59;63]
Les utilisateurs utilisent de manière inefficace le système d'alerte.	[40;43;52;53]
Les utilisateurs ont des comportements de détour.	[46;48;53]
Les utilisateurs suivent aveuglément des alertes sans les comprendre.	[63]
Les utilisateurs sont coincés ou perdus dans leur utilisation.	[50;53]
Problèmes cognitifs	
L'information est ratée involontairement.	[35;38;43;52;56;59]
Augmentation de la charge en mémoire due à l'utilisation du système : les utilisateurs se reposent sur leur mémoire.	[43]
Les utilisateurs ont des difficultés pour comprendre l'alerte.	[34;38;40;43;50;63]
Les utilisateurs ont des difficultés à identifier les éléments de l'alerte (*e.g.* icones)	[37;48;53]
Les utilisateurs mésinterprètent les éléments de l'alerte.	[37;47]
Les utilisateurs mésinterprètent l'alerte.	[35;41]
Les utilisateurs sont interrompus par les alertes pendant qu'ils prennent leur décision ou s'entretiennent avec le patient.	[42-44;52]
Problèmes émotionnels	
Énervement	[35;39;40;43;52;63]
Frustration	[39;42;50;52]
Expérience désagréable	[42]
Stress, pression	[43]
Cynisme	[42]
Problèmes d'attitude	
Les utilisateurs remettent en cause le comportement du système d'alerte.	[43;48;52;63]
Les utilisateurs remettent en cause le modèle de déclenchement et classement des alertes.	[42;43;46;63]
Les utilisateurs remettent en cause l'utilité du système.	[34;37;52]
Les utilisateurs remettent en cause la validité des alertes.	[34;42-44;62]
Les utilisateurs ressentent de la fatigue due à l'alerte et sont désensibilisés.	[34;35;42-44;49;53;62]
Les utilisateurs ont une mauvaise opinion du système d'alerte.	[37;43;52;56]

Tableau 4 - Catégories des conséquences négatives liées à des problèmes d'utilisabilité reportées dans la littérature.

Type de problème	Description	Référence
Problèmes liés au processus de travail	Les communications entre les cliniciens et entre les cliniciens et les patients sont augmentées et la responsabilité de l'alerte est mue d'un clinicien vers un autre.	[42;43;59;62;63]
Problèmes d'efficacité de la technologie	L'utilité attendue du système d'alerte n'est pas observée.	[34;41;43;48;52;53]
Problèmes du circuit du médicament	L'efficacité du circuit du médicament est altérée.	[39;42;50]
Problèmes de sécurité du patient	L'utilisation du système d'alerte produit les conditions de la diminution de la qualité des soins et peut mettre le patient en danger.	[35;38;57;63]

4. Quelles sont les liens de causes à effets reportés dans la littérature entre les problèmes d'utilisabilité, les problèmes d'usage et les conséquences négatives pour le système de travail ?

Durant le processus d'extraction des problèmes d'utilisabilité, de problèmes d'usage et de conséquences négatives pour le système de travail, trois types de liens rapportés par les auteurs des études ont été extraits : liens entre les problèmes d'utilisabilité et les problèmes d'usage, liens entre les problèmes d'utilisabilité et les conséquences négatives sur le système de travail, et liens complets entre les problèmes d'utilisabilité, les problèmes d'usage et les conséquences négatives sur le système de travail. Toutes les associations identifiées ont été analysées et résumées.

Quarante-sept associations entre problèmes d'utilisabilité, problèmes d'usage et conséquences négatives sont reportées dans la littérature. Pour 6 liens entre problèmes d'utilisabilité et conséquences négatives, il a fallu inférer le problème d'usage. Ainsi le nombre total de liens entre les trois types d'éléments est porté à 53. En plus, 129 liens entre problèmes d'utilisabilité et problèmes d'usage sont reportés.

En tout, 182 liens (129 + 53) entre problèmes d'utilisabilité et problèmes d'usage ont été analysés. Le Tableau 5 montre que tous les types de problèmes d'utilisabilité causent des problèmes d'usage. Presque tous les problèmes d'usage ne sont pas spécifiques à un type de problème d'utilisabilité. Seuls les problèmes d'homogénéité et les problèmes de gestion des erreurs ne donnent lieu qu'à un seul type de problèmes d'usage.

Les 53 liens complets entre les catégories de problèmes d'utilisabilité, de problèmes d'usage et de conséquences négatives pour le système de travail sont résumés dans le Les catégorisations effectuées permettent d'avoir une représentation synthétique de la diversité des conséquences des problèmes d'utilisabilité. Seulement, cette présentation ne permet pas de rendre compte des difficultés effectivement vécues par l'utilisateur et le système de travail. Pour mettre en exergue ces difficultés, une revue narrative des problèmes a été réalisée.

Tableau 6. Quelques rares lignes de propagation se dégagent des résultats, la principale étant l'absence de lien direct entre les problèmes émotionnels et attitudinaux, et les conséquences négatives sur le système de travail.

Tableau 5 - Associations entre les problèmes d'utilisabilité et les problèmes d'usage. Les cases noires représentent la présence d'un lien entre les problèmes d'utilisabilité et les problèmes d'usage reportés dans la littérature.

			Catégories de problèmes d'utilisabilité													
			Spécifiques						Généraux							
			Faible ratio signal / bruit	Problème de contenu de l'alerte	Problème d'apparition des alertes	Système non transparent	Problème des outils dans l'alerte	Problème de distribution des tâches	Problème de charge de travail	Problème de guidage	Problème de signifiance des codes	Problème d'adaptabilité	Problème de contrôle explicite	Problème d'homogénéité	Problème de management de l'erreur	Nombre de liens avec les catégories
Catégories de problèmes d'usage	Comportementaux	Utilisation inefficace	■	■		■		■		■		■				6
		Charge de travail augmentée	■	■					■	■	■					5
		Ignorance volontaire des alertes	■	■	■				■	■						5
		Comportements de détours		■	■		■		■				■			5
		Utilisateurs coincés / perdus	■	■						■			■			4
		Système inutilisé	■						■							2
		Utilisateurs suivant le système sans comprendre								■						1
	Cognitifs	Difficultés de compréhension de l'alerte		■					■	■	■	■			■	6
		Information ratée involontairement			■		■	■	■	■						5
		Mésinterprétation de l'alerte				■			■	■						3
		Utilisateurs interrompus	■		■				■							3
		Difficultés d'identification des éléments dans l'alerte							■		■					2
		Augmentation de la charge en mémoire		■												1
		Mésinterpration des éléments de l'alerte									■					1
	Emotionels	Enervement	■		■				■	■						4
		Frustration	■		■					■	■					4
		Expérience désagréable							■	■						2
		Cynisme		■												1
		Stress, pression			■											1
	Attitudinaux	Remise en cause de la validité	■	■		■			■							4
		Remise en cause du modèle de déclenchement/tri	■	■		■								■		4
		Remise en cause de l'utilité	■	■	■											3
		Remise en cause du comportement	■			■	■									3
		Alerte fatigue / désensibilisation	■						■	■						3
		Mauvaise opinion sur le système	■		■											3
Nombre de liens avec les catégories			14	11	9	5	3	2	14	12	5	2	2	1	1	81

Les catégorisations effectuées permettent d'avoir une représentation synthétique de la diversité des conséquences des problèmes d'utilisabilité. Seulement, cette présentation ne permet pas de rendre compte des difficultés effectivement vécues par l'utilisateur et le système de travail. Pour mettre en exergue ces difficultés, une revue narrative des problèmes a été réalisée.

Tableau 6 - Synthèse des 53 liens complets entre les catégories de problèmes d'utilisabilité (en colonne), de problèmes d'usage (en ligne) et de conséquences négatives pour le système de travail (dans les cellules). Les problèmes émotionnels et attitudinaux ne sont pas représentés car ils ne sont jamais associés directement à des conséquences négatives pour le système de travail. Acronymes : W, problèmes liés au processus de travail ; T, problèmes d'efficacité de la technologie ; M, problèmes liés au circuit du médicament ; et P, problèmes de sécurité du patient.

			Catégories de problèmes d'utilisabilité												
			Spécifiques						Généraux						
			Faible ratio signal / bruit	Problème de contenu de l'alerte	Problème d'apparition des alertes	Système non transparent	Problème de distribution des tâches	Problème des outils dans l'alerte	Problème de charge de travail	Problème de guidage	Problème de signifiance des codes	Problème d'adaptabilité	Problème de contrôle explicite	Problème d'homogénéité	Problème de management de l'erreur
Catégories de problèmes d'usage	Comportementaux	Utilisation inefficace				W				T		W			
		Charge de travail augmentée	W	M					T	M					
		Comportements de détours			T										
		Ignorance volontaire des alertes	M T	M	W				T						
		Utilisateurs coincés / perdus	T	M						M			T		
		Système inutilisé							T						
		Utilisateurs suivant le système sans comprendre								P					
	Cognitifs	Difficultés de compréhension de l'alerte		M W					T W M P	T M P	M	W			W
		Information ratée involontairement					W	T	P	T P					
		Utilisateurs interrompus			T				T						
		Mésinterprétation de l'alerte				P			P	T P					
		Difficultés d'identification des éléments dans l'alerte													
		Augmentation de la charge en mémoire													
		Mésinterpration des éléments de l'alerte													

Revue narrative

Problèmes d'interaction avec le système d'alerte

Les problèmes de présentation de l'alerte comme le moment et le mode d'apparition (*i.e.* intrusif vs. non intrusif) ainsi que la nécessité d'utiliser les barres de défilement vertical pour prendre des informations empêchent l'utilisateur d'obtenir l'information dont il a besoin. Les cliniciens peuvent même manquer toute l'alerte parce qu'elle n'est pas suffisamment visible [52] ou parce qu'elle arrive trop tard [56]. Une alerte qui apparait trop tard oblige les utilisateurs à faire eux-mêmes les opérations que le système était censé faire. Par exemple, « 6 sujets calculèrent, estimèrent ou utilisèrent une heuristique pour connaitre la dose avant que le système ne présente cette information » [46].

Les cliniciens peuvent manquer l'alerte involontairement parce qu'elle est apparue à la place d'une autre sans qu'ils se soient aperçus du changement [35]. Par ailleurs, plusieurs systèmes ne permettent pas de revoir ultérieurement l'alerte : ainsi, les cliniciens ne peuvent pas la relire [35] et peuvent, au final, oublier son contenu [29].

Problèmes de compréhension de l'alerte

Même lorsque les cliniciens voient l'alerte, l'information n'est pas toujours aisée à trouver. Certains utilisateurs n'ont pas remarqué le calcul de dose incorporé dans le texte de l'alerte [46]. Dans une autre étude, la moitié des participants ont raté l'information sur la durée de la thérapie du patient [15]. Dans cette même étude, des utilisateurs ont pris une décision clinique erronée car ils ont manqué l'information sur les facteurs de risque du patient, cachée dans un onglet.

Même lorsque l'alerte est accessible au bon moment pour l'utilisateur et que l'information est vue, les utilisateurs rencontrent d'autres problèmes. Un utilisateur rapporte, par exemple, qu'il « est difficile de voir ce que [l'alerte] essaie de dire » [63]. L'alerte n'est pas compréhensible par les médecins [50], difficile à interpréter [39], ce qui empêche les utilisateurs de comprendre les problèmes qui ont déclenché l'alerte et comment les résoudre [50].

Ces problèmes de compréhension empêchent d'utiliser de manière optimale le système d'alerte et obligent le médecin à demander de l'aide au pharmacien [42]. Il est parfois nécessaire pour les infirmières et les médecins d'avoir des discussions en face-à-face avec les pharmaciens [29]. A l'inverse, les pharmaciens sont parfois obligés de téléphoner aux cliniciens car ils ont été alertés par leur système d'alerte alors que celui des médecins et infirmières n'avait rien détecté [15]. Le vocabulaire utilisé dans les alertes ne convient pas aux différents profils d'utilisateurs et ces derniers ont de grandes difficultés à le comprendre [43;63]. En somme, le travail des cliniciens est ralenti par les conséquences de ces défauts des systèmes d'alerte [39].

Ces difficultés de compréhension conduisent à des mésinterprétations, et donc à des actions erronées [35]. Une alerte apparait disant qu'il y a une double prescription d'aspirine ; le texte de l'alerte n'est pas clair et le médecin ne la comprend pas mais clique néanmoins sur « ok », validant la double prescription d'aspirine [63]. Dans une autre étude, il est rapporté que dans un hôpital Danois, plusieurs erreurs de prescription de médicaments étaient dues au fait que les prescripteurs n'avaient pas une idée exacte de ce sur quoi le déclenchement des alertes était basé : ils faisaient donc des hypothèses, parfois fausses, sur les causes du déclenchement des alertes, conduisant ainsi à des prises de décision mal fondées [57]. Les icones et les titre d'alertes sont aussi parfois mésinterprétés [37;47;53].

Augmentation de la charge en mémoire et de la charge de travail

Le manque d'information sur le patient dans les alertes oblige les cliniciens à se reposer sur leur mémoire du cas patient ou à faire des hypothèses sur l'histoire du patient [29].

Un paramétrage inexact des alertes conduit à de nombreuses interruptions du travail des cliniciens [42], de leur processus de raisonnement [52] ce qui rend difficile leur prise de décision, les oblige à se concentrer plus et, parfois, leur fait perdre le fil de leur raisonnement. Cela va à l'encontre de l'efficacité recherchée des systèmes d'alerte [29].

D'un point de vue comportemental, la mauvaise qualité des messages d'alertes, leur répétition et leur longueur obligent les médecins à perdre du temps pour rechercher des informations complémentaires [29],

à faire défiler le texte des l'alertes [35] ou simplement à les lire [49]. Ils procèdent aussi par essai-erreur (notable à cause d'un grand nombre de cliques) pour exécuter l'action suggérée par l'alerte [50].

L'augmentation de la charge de travail est aussi due aux tâches de documentation des alertes qui sont parfois redondantes par rapport à la documentation que font habituellement les cliniciens hors du système d'alerte [53]. La lourdeur de la documentation oblige parfois à accomplir cette tâche après le départ du patient, en fin de journée, ou encore à faire appel à d'autres personnels (« case managers ») pour remplir cette tâche [46]. Enfin, l'absence d'outils pour partager les alertes contraint les utilisateurs à encore utiliser le papier pour transmettre l'information [53].

Difficultés d'interactions

En général, l'efficacité des systèmes d'alerte est grandement gênée par l'utilisabilité de ces systèmes [40;53]. L'incapacité de ces systèmes à soutenir correctement les activités cognitives des utilisateurs ainsi que leurs tâches est perceptible lorsque les utilisateurs sont bloqués dans leur utilisation [53]. Pour avancer malgré un manque de suggestions adaptées, les cliniciens ont dû développer des comportements de détour : par exemple, certains sélectionnent de manière arbitraire une option juste pour pouvoir continuer le processus, d'autres ne sélectionnent rien [53].

L'interface gêne aussi la capacité des prescripteurs à agir sur l'alerte [29]: certains utilisateurs ne semblent pas comprendre comment gérer les alertes de manière efficace ce qui fait se répéter les alertes [52]. De même, quand les utilisateurs ne peuvent pas satisfaire une alerte à cause des suggestions proposées, les alertes continuent forcément d'apparaitre [53]. Dans un cas, si l'utilisateur veut annuler un mémo clinique sans perdre les données rentrées précédemment dans d'autres mémos, il doit alors sélectionner les mémos les uns après les autres au lieu d'utiliser le bouton « suivant » pour naviguer entre les mémos. Ce problème peut causer la perte de données rentrées [53].

En somme, l'amélioration du circuit du médicament que les utilisateurs attendent n'est pas observée car les systèmes d'alerte diminuent l'efficacité de la prescription en augmentant la charge de travail des utilisateurs [42;50] : les cliniciens ne retirent pas toute la vitesse et la précision que le système d'alerte est supposé leur apporter et leur effort cognitif n'est pas diminué, au contraire [46].

Problèmes émotionnels

Faire face, tous les jours, à un système d'alerte peu utilisable n'impacte pas uniquement les processus cognitifs et le comportement des utilisateurs. Ces derniers sont aussi affectés émotionnellement. Il ne leur est pas plaisant des lire les alertes [42] avec des problèmes d'affichage. Le fait que les alertes n'apparaissent pas immédiatement met sous pression les prescripteurs [29] car ils sont supposés prendre rapidement leur décision. Sans surprise, le manque d'information dans les alertes ainsi que leur répétition rendent les utilisateurs cyniques [42], frustrés [42], accablés [29] et énervés : « la même alerte apparait pour la troisième fois au moment où une infirmière prescriptrice allait signer la prescription. Celle-ci montre l'écran et s'écrit « regardez, trois fois ! » » [29]. Le trop grand nombre d'alertes « rend fou » d'après un utilisateur [35]. En plus, il y a des alertes qui apparaissent encore et encore alors même que les utilisateurs essaient de les

annuler (en changeant la prescription par exemple) : cela aussi peut véritablement « rendre dingue » d'après un utilisateur [63].

Avec le temps, la répétition de ces alertes tous les jours, plusieurs fois par jour, impacte aussi l'attitude des utilisateurs vis-à-vis du système d'alerte. Il y a de nombreuses plaintes, évidemment [56], même de fortes plaintes [52] à propos des difficultés à utiliser le système [37]. Ces plaintes sont compréhensibles mais l'impact va bien plus loin.

Désensibilisation et outrepassement des alertes

Le faible ratio signal-bruit des alertes génère une grande quantité d'alertes et crée de la fatigue due à l'alerte et aussi de la désensibilisation. Cet état de fatigue et la désensibilisation sont largement observés [35;44] et même reconnus par les utilisateurs [49]. Cet état contraint les utilisateurs à volontairement ignorer les alertes : ils les outrepassent [29]. Dès qu'ils ont reconnu une alerte déjà vue, ils l'outrepasse [39]. Certains en sont arrivés au point qu'ils ne regardent même plus véritablement les alertes [63], ils sont tentés de les passer rapidement [35], et s'il y a plus d'une alerte dans la fenêtre, ils ne les lisent pas [42]: ils cliquent rapidement, quitte à ne pas voir une autre alerte [63]. Certains utilisateurs ont développé un mécanisme pour passer les alertes [49] ; une clinicienne explique qu'elle a mémorisé l'emplacement du bouton sur l'écran pour passer l'alerte [39]. En résumé, un clinicien explique qu' « une fois que vous avez compris que la plupart des informations est inutile ou superflue ou non pertinente, vous arrêtez de les regarder » [44] : il préfère se reposer sur son propre jugement clinique [29]. La redondance non nécessaire des alertes, en créant un état de fatigue et d'ignorance volontaire des alertes, nuit au final au processus de prescription [42] et amène à un bas taux de réponse aux alertes [34]. Une autre conséquence est que les cliniciens font confiance aux pharmaciens pour gérer les alertes, plaçant implicitement la responsabilité des alertes sur ces derniers [64].

Perte de confiance

Il existe aussi une perte de confiance dans le système d'alerte. Un clinicien explique qu'il vient de passer un médicament en inactif dans le système d'information mais que l'alerte lui dit néanmoins le contraire. En toute logique, le clinicien se demande si le système prend véritablement en compte ses actions [63]. Dans un autre cas, les cliniciens doutent que les pharmaciens à qui ils écrivent des commentaires reçoivent ces commentaires [29]. Les cliniciens ne savent pas non plus pour combien de temps les alertes sont éteintes lorsqu'ils parviennent à les éteindre [46]. Ils ont aussi quelques doutes sur les données sur lesquelles le système se base ainsi que sur les algorithmes et formules utilisées [46]. La validité de l'alerte est aussi remise en cause : les recommandations sont-elles basées sur des preuves ? [42] Les informations utilisées sont-elles récentes? [34]. Parfois même, l'alerte fournit des informations que le clinicien sait fausses sur le patient [62], ce qui oblige ce dernier à réinterroger le patient pour être certain du traitement qu'il reçoit. Enfin, les cliniciens remettent sérieusement en cause l'utilité du système d'alerte, le trouvant parfois inutile [52] car il ne fait que « crier au loup » [44]. Ils perçoivent négativement la

crédibilité et le sérieux de ces systèmes [29] à cause de l'ensemble de ces problèmes d'utilisabilité. En résumé, les problèmes d'utilisabilité dissuadent les cliniciens d'utiliser les systèmes d'alertes [46].

Afin de prévenir ces problèmes d'usage et conséquences négatives sur le système de travail, il est nécessaire d'empêcher l'apparition de problèmes d'utilisabilité. Pour ce faire, il faut appliquer durant la conception et l'évaluation des TIS des principes d'utilisabilité.

5. Quels sont les principes d'utilisabilité spécifiques aux systèmes d'alerte médicamenteux qui correspondent aux problèmes d'utilisabilité identifiés ?

Pour répondre à cette question, deux experts en Facteurs Humains ont recherché les articles reportant des ensembles de principes d'utilisabilité. Les publications qui contenaient aussi des problèmes d'utilisabilité et qui avaient été, à ce titre, incluses dans la revue systématique, ont été exclues. Les principes d'utilisabilité ont été extraits de ces articles par un expert. Une fois les principes en doublons enlevés, la liste finale des principes a été organisée en catégories et sous-catégories. A l'intérieur de ces (sous-catégories, les principes similaires ont été synthétisés ensemble.

Huit articles ont été identifiés [29;31;45;65-69]. Un article a dû être exclu [45] car il faisait partie des articles de la revue systématique. En tout, 132 principes ont été identifiés dont 7 doublons. Les 125 principes restant ont été synthétisés en 54 principes eux-mêmes organisés hiérarchiquement en 6 thèmes : améliorer le ratio signal-bruit, être en adéquation avec l'activité des cliniciens, supporter le travail collaboratif, afficher des informations pertinentes, rendre le système transparent et fournir des outils utiles pour interagir avec l'alerte et appliquer la décision prise. Une vue d'ensemble des principes est présentée Figure 13.

6. A quel point les problèmes d'utilisabilité reportés dans la littérature correspondent à ces principes d'utilisabilité ?

Les 107 problèmes d'utilisabilité spécifiques aux systèmes d'alerte médicamenteux identifiés durant la revue systématique [70] ont été associés aux principes d'utilisabilité synthétisés. Un problème était considéré comme associé à un principe lorsqu'il en était la violation. Inversement, un principe était considéré comme associé au problème si son application avait pu empêcher l'apparition de ce problème. Si un problème ne trouvait pas correspondance directe avec un principe, alors la possibilité d'étendre un principe existant (*e.g.* inclure des contextes d'application supplémentaires) était considérée. S'il n'y avait aucun principe portant sur le thème du problème alors un nouveau principe était défini. La mise en correspondance a été réalisée par les deux experts en Facteurs Humains.

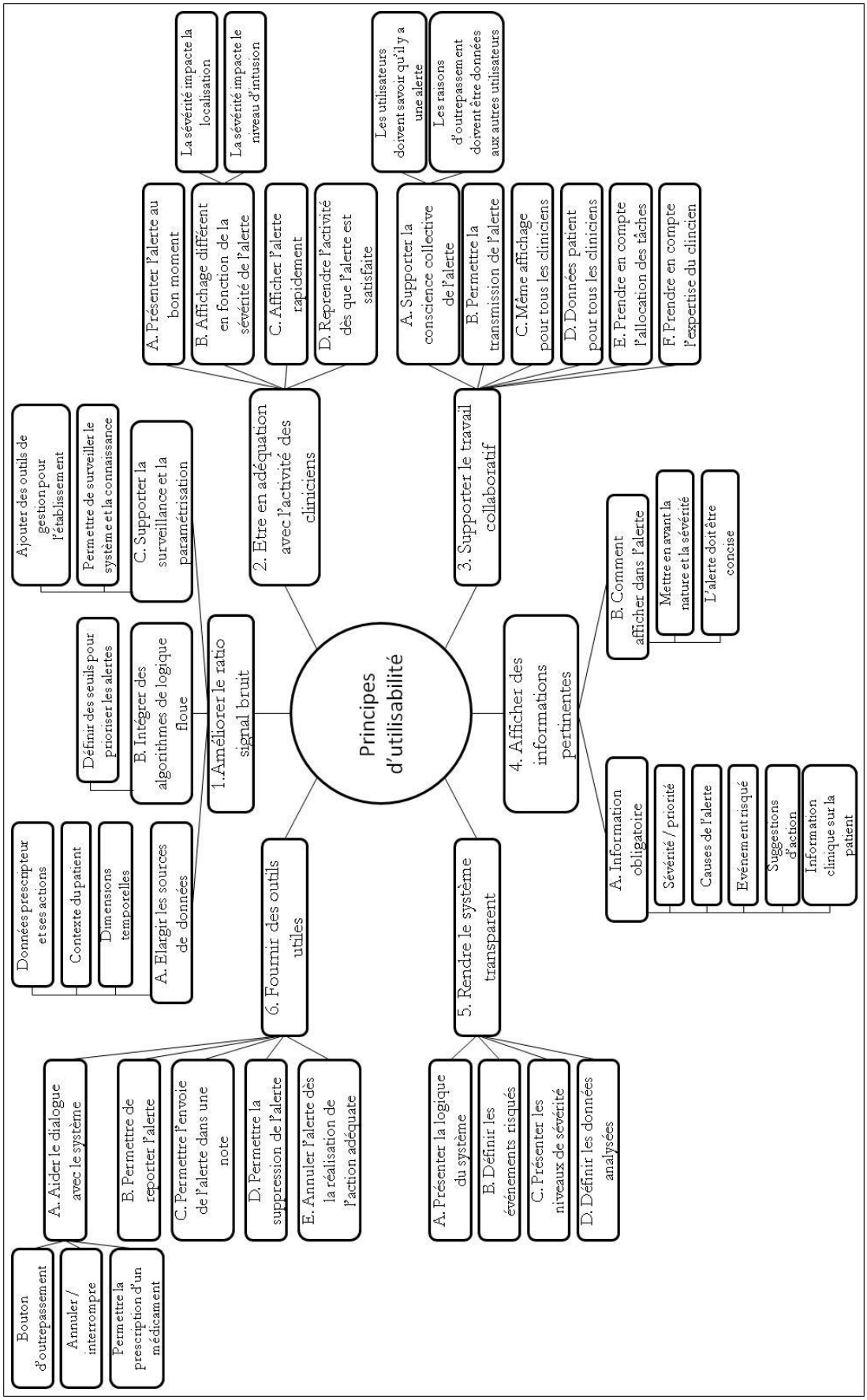


Figure 13. Représentation schématique des 6 thèmes et des principes d'utilisabilité les composant.

Les résultats montrent que 23 des 54 principes ne correspondent à aucun problème d'utilisabilité reporté dans la littérature. Cependant, parmi ces principes, certains sont indirectement associés parce qu'un de leurs sous-principes, ou son principe supérieur hiérarchiquement, est lui-même associé à un problème d'utilisabilité. Parmi les 31 principes associés, 22 le sont directement et 9 demandent une légère extension de contexte. Par ailleurs, pour associer tous les principes d'utilisabilité avec un principe, 9 principes non trouvés dans la littérature ont dus être définis portant à 63 le nombre final de principes d'utilisabilité spécifiques aux systèmes d'alerte médicamenteux.

Améliorer le ratio signal-bruit

Dans ce thème, 4 principes ont été étendus pour correspondre à des problèmes d'utilisabilité. Par ailleurs, 3 principes ont été créés. Ces changements concernent surtout l'élargissement des sources de données à prendre en considération ainsi que les actions à mettre en œuvre sur les données récupérées (*e.g.* vérification, mise à jour). Au final, seulement 3 principes ne sont associés à aucun problème d'utilisabilité : parmi ceux-ci, deux sont liés au modèle de raisonnement du système et le dernier concerne la prise en compte de l'évolution de l'événement négatif pour définir la sévérité de l'alerte.

Etre en adéquation avec l'activité des cliniciens

Aucun principe appartenant à ce thème n'a été étendu ou créé pour correspondre à un problème d'utilisabilité. Deux principes ne sont associés à aucun problème d'utilisabilité : le premier renvoie à la position de l'alerte sur l'écran qui doit varier avec la sévérité de l'alerte ; le second concerne le fait que l'on doit donner à l'utilisateur la possibilité de reprendre son activité après qu'il a satisfait l'alerte. Tous les autres principes sont associés à plusieurs problèmes d'utilisabilité.

Supporter le travail collaboratif

Un principe a été créé et un autre a été étendu pour correspondre aux problèmes d'utilisabilité. Le premier renvoie à la mise à disposition d'une fonction de transfert de l'alerte entre cliniciens. Le second est lié à la nécessité de viser précisément le destinataire de l'alerte. Dans ce thème, 3 principes ne rencontrent aucun problème d'utilisabilité : faire du système d'alerte un soutien de l'équipe clinique, fournir une indication de la disponibilité d'une alerte à tous les cliniciens et partager les informations cliniques du patient entre tous les cliniciens.

Afficher les informations pertinentes

Un principe de ce thème a été créé : prendre en compte le contexte du patient pour suggérer des actions à l'utilisateur. Deux principes ont été légèrement adaptés afin d'être plus précis sur les informations requises dans le corps de l'alerte. Dans ce thème, 7 principes ne sont pas associés à des problèmes d'utilisabilité. Cinq d'entre eux sont des sous-principes ; les deux autres sont les seuls principes de la catégorie « comment afficher dans l'alerte » l'information.

Rendre le système transparent

Dans ce thème, un principe a dû être créé pour rendre compte des problèmes d'utilisabilité liés à la transparence du système d'alerte : expliquer à l'utilisateur quelles sont les données analysées par le système. Tous les principes trouvent une correspondance dans les problèmes d'utilisabilité.

Fournir des outils utiles

Dans cet ensemble de principes, 3 principes ont été créés et 2 autres ont été étendus. Ces principes concernent :

- Les types d'actions que les cliniciens doit pouvoir effectuer à partir de l'alerte pour la satisfaire (*e.g.* disposer plusieurs boutons d'action), l'outrepasser ou la gérer (*e.g.* reporter l'alerte à plus tard, la transférer dans une note) ;
- Les détails de l'interaction de l'utilisateur avec l'alerte ainsi que ceux concernant les conditions d'utilisation d'un outil de suppression d'une alerte par l'utilisateur.

Dans ce thème, 8 principes ne correspondent à aucun problème d'utilisabilité : des sous-principes (*e.g.* information sur comment outrepasser l'alerte ou comment permettre à l'utilisateur de prescrire un médicament directement de la fenêtre de l'alerte) mais aussi un principe (*i.e.* comment le système doit se comporter en réponse aux actions correctrices de l'utilisateur).

7. Quelles sont les méthodes appliquées pour détecter les problèmes d'utilisabilité des systèmes d'alerte médicamenteux ? Quels types de faits d'utilisabilité ces méthodes permettent-elles de reporter ?

Une analyse secondaire des publications incluses dans la revue systématique a été conduite. En plus des données sur les problèmes d'utilisabilité, les problèmes d'usage et les conséquences négatives sur le système de travail, deux experts en Facteurs Humains ont extrait les caractéristiques des méthodes utilisées.

Ensuite, en se basant sur ces méthodes ainsi que sur l'objectif affiché de ces publications, ces dernières ont été triées en 4 catégories en fonction du type d'évaluation auquel elles se rattachent :

- Les évaluations expertes : audits d'utilisabilité réalisés en laboratoire par plusieurs experts sans recourir à des utilisateurs finaux.
- Les tests utilisateurs ou simulations : observations en laboratoire d'utilisateurs finaux en train d'utiliser la technologie évaluée en suivant des scénarios et en « pensant à haute voix ».
- Les observations sur site : observations de la réelle utilisation de la technologie par des utilisateurs finaux dans leur système de travail.
- Les analyses d'impact : analyses rétrospectives (i) des résultats d'activités réalisées avec la technologie ou (ii) de l'expérience des utilisateurs.

Les évaluations expertes concernent 4 publications, les tests utilisateurs / simulations 6 publications, les observations sur site 11 publications et les études d'impact 7 publications. Une publication a utilisé à la fois une évaluation experte et un test utilisateurs / simulation, et une autre a appliqué un test utilisateur / simulation ainsi que des observations sur site. Ces deux articles ne différencient pas nettement l'origine des résultats ; ils n'ont donc pas été inclus dans l'analyse suivante.

Le Tableau 7 montre que seules les observations sur site reportent tous les types de problèmes d'utilisabilité ainsi que les problèmes d'usage et les conséquences négatives. Cette couverture complète est principalement due à l'utilisation d'annexes en ligne par l'article [17] pour fournir une liste complète des problèmes observés.

Tableau 7 - Nombre de publications reportant des problèmes d'utilisabilité (rangés en sous-catégories), des problèmes d'usage et des conséquences négatives sur le système de travail en fonction du type d'évaluation.

			Evaluations expertes	**Tests utilisateurs / simulations**	**Observations sur site**	**Etudes d'impact**
Problèmes d'utilisabilité	Généraux	Guidage	1	3	1	4
		Charge de travail	1	3	6	3
		Signifiance des codes	0	2	1	2
		Homogénéité	1	0	1	0
		Contrôle explicite	0	0	2	0
		Adaptabilité	0	0	1	0
		Gestion de l'erreur	0	0	1	0
	Spécifiques	Faible ratio signal-bruit	0	2	9	5
		Problèmes de contenu de l'alerte	1	2	4	1
		Problèmes de transparence	2	0	5	2
		Problèmes d'apparition de l'alerte	1	0	5	2
		Problèmes de distribution des tâches et du contrôle	0	1	3	2
		Problèmes des outils de l'alerte	1	1	3	0
Problèmes d'usage			0	4	10	5
Conséquences négatives			0	3	7	4

Les résultats montrent une large palette de méthodes utilisées. Tous les types d'évaluation ne reportent pas tout à fait les mêmes types de problèmes d'utilisabilité et tous ne reportent pas des problèmes d'usage et des conséquences négatives sur le système de travail. Les observations sur site qui ne sont pas à proprement parlé reconnues comme des méthodes d'évaluation de l'utilisabilité représentent la plus grande partie des publications incluses et fournissent des données sur tous les éléments d'utilisabilité. Ce résultat montre que la détection de problèmes d'utilisabilité n'est pas qu'une question d'effectif de

participants lors de l'évaluation, comme cela est souvent discuté [71], mais aussi une question de connaissance sur l'usage réel.

8. Le formulaire de report des problèmes d'utilisabilité est-il utile et en quoi?

De manière à améliorer la précision et la complétude du report des problèmes d'utilisabilité, un formulaire de report de ces problèmes a été développé en le basant sur le « cadre de travail en utilisabilité ». Ce chapitre porte sur l'évaluation préliminaire de l'utilité perçue de ce formulaire par des concepteurs / développeurs ainsi que par des experts en Facteurs Humains. Cette évaluation a été réalisée durant une inspection heuristique d'un moniteur du confort de nouveau-nés [72].

Trois experts en Facteurs Humains ont réalisé l'inspection de manière indépendante en utilisant la liste des principes d'utilisabilité de Scapin et Bastien [73]. Un des évaluateurs n'avait pas effectué préalablement l'analyse du contexte d'usage prévu. Afin de suivre les recommandations en termes de gestion du risque, pour chaque problème d'utilisabilité, le risque associé à leur conséquence était estimé en fonction de leur sévérité et de leur fréquence. Le report des problèmes d'utilisabilité a été fait à travers le formulaire à évaluer (Tableau 8).

Les données à propos de l'utilité perçue ont été récoltées de manière opportuniste par RM auprès de 4 représentants de l'industriel (PDG, concepteur, ingénieur et responsable qualité). Afin de connaitre le ratio bénéfice-risque de ce formulaire, il fallait aussi connaitre les difficultés potentielles pour le remplir. Les commentaires des deux experts en Facteurs Humains non impliqués dans le développement du formulaire ont aussi été recueillis. Les commentaires ont été classés en tant qu'avantages / inconvénients pour le remplissage / l'interprétation du contenu.

Tableau 8 - Exemples de reports de problèmes d'utilisabilité à travers le formulaire développé.

Principe d'utilisabilité	**Problème d'utilisabilité**	**Potentiel problème d'usage**	**Potentielle conséquence négative**	**Probabilité x sévérité = risque**
Signifiance des codes	Il n'y a aucune information sur la signification de la courbe et du "smiley" (Figure 14, droite).	Difficile à interpréter, perte de temps et incapacité à comprendre.	Prise de décision basée sur une information mal comprise	Modéré x modéré = médium
Homogénéité	L'index NIPE « instantanée » correspond à la courbe « index moyenné » et non à la courbe « index instantané » (Figure 14, gauche)	Incapacité à comprendre, mauvaise représentation de l'état du patient, doutes sur la fiabilité du dispositif, diminution de la confiance.	Prise de décision basée sur une information mal comprise	Modéré x minimal = bas

Les principaux commentaires recueillis sont synthétisés dans le Tableau 9. Pour les représentants industriels, ce formulaire est utile pour comprendre les origines et les conséquences des erreurs d'utilisation liées à l'utilisabilité, avis partagé par les experts en utilisabilité. Cependant ces derniers ont rencontré des difficultés pour remplir ce formulaire. L'expert qui n'avait pas analysé le contexte d'usage

s'est trouvé incapable d'anticiper les problèmes d'usage et leurs conséquences négatives et d'estimer leur risque. Néanmoins, quel que soit le profil, aucun inconvénient pour l'interprétation n'a été relevé.

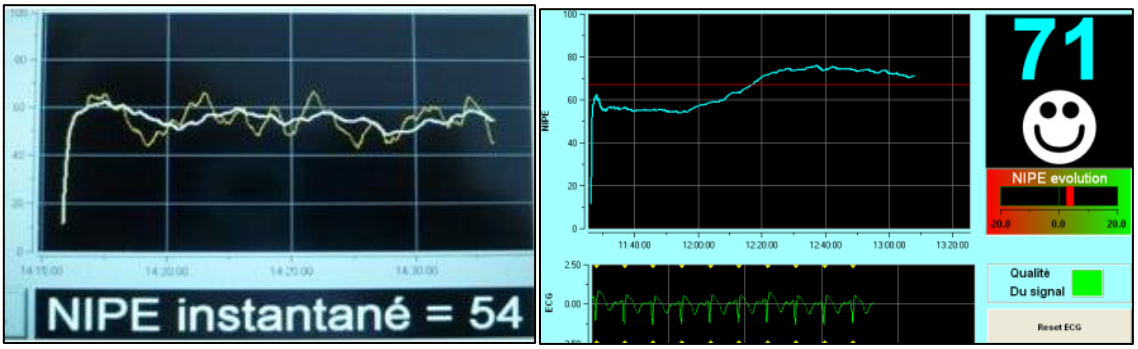


Figure 14. Capture d'écran de l'index NIPE instantané (gauche) et de l'écran principal (droite)

Les résultats de cette étude préliminaire sont encourageants : ce formulaire de report semble être intéressant pour permettre aux industriels de comprendre les erreurs d'utilisation liées aux problèmes d'utilisabilité. Un autre résultat important concerne la nécessité pour les évaluateurs de réaliser l'analyse du contexte d'usage afin de pouvoir réaliser de manière complète et précise l'évaluation ergonomique.

Tableau 9 - Synthèse des principaux commentaires recueillis.

		Experts en utilisabilité	**Représentants industriels**
Remplissage	Avantages	Oblige à décrire clairement les différents types de problèmes.	Non applicable
	Inconvénients	Difficile de faire des descriptions séparées des problèmes d'usage et des conséquences négatives sur le système de travail (« difficile exercice de style »). Difficile d'anticiper certains problèmes/conséquences ainsi que le niveau de risque d'autant plus quand l'expert n'a pas analysé le contexte d'usage. Ce à quoi renvoie le risque n'est pas évident ce qui rend son estimation plus difficile.	Non applicable
Interprétation	Avantages	Cela permet une description plus précise des problèmes d'utilisabilité, plus directe. Cela clarifie les conséquences des problèmes d'utilisabilité. Les risques pour le patient sont clairement identifiables.	Il est important de distinguer (i) comment les problèmes d'utilisabilité peuvent impacter les utilisateurs (ii) des erreurs d'utilisations que ces problèmes peuvent induire. Identification plus rapide de l'origine des potentielles erreurs d'utilisation.
	Inconvénients	/	/

9. Discussion

Synthèse des objectifs et des résultats

Ce travail de thèse s'intéresse à la problématique de l'utilisabilité des Technologies de l'Information en Santé (TIS). Son objectif est double:

- Participer à l'amélioration de la capitalisation de la connaissance sur l'utilisabilité des TIS ;
- Fournir une liste synthétique et structurée de principes d'utilisabilité pour les TIS tout en explicitant leur couverture en termes de problèmes d'utilisabilité.

Ces deux thèmes ont été traités conjointement avec l'étude des méthodes utilisées pour accumuler et reporter les données d'utilisabilité. Les systèmes d'alerte médicamenteux ont servi de domaine d'étude.

Le premier résultat de ce travail est le « cadre de travail en utilisabilité ». Ce cadre permet d'expliciter les liens entre les principes d'utilisabilité, les problèmes d'utilisabilité, les problèmes d'usage et les conséquences négatives sur le système de travail (chapitre 1). Les études reportées dans les chapitres 2 à 4 ont permis d'identifier 168 problèmes d'utilisabilité. Parmi eux, 107 sont spécifiques aux systèmes d'alerte médicamenteux ; ils sont des violations de 63 principes d'utilisabilité dédiés à ces systèmes. Ces principes d'utilisabilité sont organisés en 6 thèmes (améliorer le ratio signal-bruit, être en adéquation avec l'activité des cliniciens, supporter le travail collaboratif, afficher les informations pertinentes, rendre le système transparent, fournir des outils utiles). Les violations des ces principes ont des conséquences variées. En effet, elles impactent les cliniciens d'un point de vue cognitif, comportemental, émotionnel et attitudinal. Au final, ces violations altèrent le flux de travail, l'efficacité de la technologie, le circuit du médicament, et peuvent même affecter la sécurité du patient.

Par ailleurs, parmi toutes les méthodes employées pour identifier des problèmes d'utilisabilité, l'observation in-situ est la méthode qui est associée avec la plus grande variété reportée de problèmes d'utilisabilité, de problèmes d'usage et de conséquences négatives pour le système de travail (chapitre 5). En ce qui concerne le report des problèmes d'utilisabilité et de leurs conséquences, une étude préliminaire a montré que la grille de report adapté du « cadre de travail en utilisabilité » est perçue comme utile par les concepteurs et les experts en Facteurs Humains pour présenter clairement l'origine des problèmes d'utilisabilité et leurs conséquences (chapitre 6).

Principales limites et forces de l'approche

Les revues de la littérature réalisées ont pu être impactées par des biais et des limites qui ont été discutées dans les chapitres concernés: *e.g.*, biais de publication et de report sélectif, incomplétude ou absence de données, recherche non systématique des principes d'utilisabilité. Il est important d'avoir conscience de ces limites pour comprendre précisément la fiabilité et la portée des résultats présentés.

Ces études représentent aussi plusieurs forces méthodologiques discutées tout au long du manuscrit : *e.g.*, l'identification des publications, l'extraction et la catégorisation des données ont été réalisées par au

moins deux experts en Facteurs Humains avec de l'expérience en conception et évaluation de systèmes d'alertes médicamenteux. Un autre point fort de ce travail est en lien avec la robustesse des résultats. En effet, les catégories de problèmes d'utilisabilité créées sont assises sur des données issues de plusieurs publications : ce recouvrement des données assure une bonne fiabilité des résultats. Ce bon support empirique est aussi observé pour les problèmes d'usage et, de manière moins importante, pour les conséquences négatives sur le système de travail. Par ailleurs, les listes de problèmes d'utilisabilité et de principes d'utilisabilité s'emboitent relativement bien alors même qu'elles sont issues de deux ensembles différents de publications. Cette bonne correspondance garantit aussi une bonne validité des résultats Finalement, au niveau de l'approche globale adoptée, l'objectif d'améliorer l'accumulation des connaissances en utilisabilité des TIS n'a pas été traité que dans une seule direction. Evidemment, les études réalisées ont visé à accumuler les données déjà existantes pour créer de la connaissance (chapitres 2 à 4). Cependant, elles ont aussi cherché à améliorer l'identification (chapitre 5) et le report (chapitre 6) de ces données de manière à pouvoir accroitre la quantité et la qualité des données d'utilisabilité obtenues dans de futures études.

Contributions issues de ce travail

Structurer la connaissance en utilisabilité: collecter et reporter les données

Le « cadre de travail en utilisabilité » a été développé pour structurer différents concepts tournant autour de l'utilisabilité (chapitre 1). Ce cadre a aidé le processus de revue systématique tout d'abord méthodologiquement en guidant la sélection des publications et aussi en aidant l'interprétation des données récoltées. Utiliser ce cadre de travail s'est révélé être une aide importante pour distinguer clairement les problèmes d'utilisabilité de leurs conséquences dans les publications analysées. En effet, dans l'état actuel de la littérature, la description des problèmes d'utilisabilité est mélangée avec la description de leurs conséquences.

Afin d'améliorer la qualité des reports des problèmes d'utilisabilité, ce cadre a été transformé en grille de report. Utiliser cette grille pour reporter des problèmes d'utilisabilité devrait permettre de structurer de manière homogène les données d'utilisabilité. Au final, cela devrait contribuer à accumuler des données d'utilisabilité utilisables pour en faire de la connaissance.

Vers une "utilisabilité basée sur les preuves" pour les systèmes d'alerte médicamenteux

Les résultats de ce travail, en termes de connaissance en utilisabilité, fournissent une première étape pour développer des principes d'utilisabilité basés sur des preuves pour les systèmes d'alertes médicamenteux. Ce travail rend disponible une liste de principes d'utilisabilité illustrés par des cas réels de leur violation et les conséquences négatives de ces dernières. Cette connaissance fournit une preuve que violer les principes d'utilisabilité lors de la conception de systèmes d'alerte peut être préjudiciable pour les cliniciens et leur système de travail (dont la qualité des soins). Cette connaissance doit être interprétée avec précautions. En effet, même si les résultats révèlent que violer les principes d'utilisabilité peut nuire à l'usage et au système de travail, cela ne signifie pas pour autant qu'appliquer ces principes est une garantie

de résolution de succès de la conception : plusieurs facteurs autres que l'utilisabilité peuvent intervenir durant la conception et l'implémentation de la technologie et venir gâcher une bonne utilisabilité (*e.g.*, mauvaise inter-connectivité avec d'autres technologies).

Les violations des principes d'utilisabilité et leurs conséquences ont pu être sous-estimées à cause de biais de publication et de report sélectif. Ainsi, la connaissance constituée est certainement incomplète. Elle doit être renforcée en intégrant d'autres sources de données que celles issues de la littérature (*e.g.*, données de systèmes de report d'incidents intégrant une composante utilisabilité) et elle doit être actualisée très régulièrement.

D'après la définition originale de la « médecine basée sur les preuves » [22], rassembler des preuves de la littérature n'est pas un but en soi : il faut délivrer cette connaissance aux praticiens (ici, concepteurs et experts en Facteurs Humains) pour qu'ils l'intègrent dans leur pratique personnelle afin de prendre des décisions éclairées. Ainsi, la connaissance rassemblée dans ces études doit maintenant être adaptée de manière à être présentée correctement aux concepteurs et experts Facteurs Humains.

Généralisation de ce travail

Généralisation des résultats d'utilisabilité

Ce travail a été réalisé pour les systèmes d'alerte médicamenteux mais son but concerne plus largement la connaissance en utilisabilité appliquée à tous les types de TIS. Dans les chapitres 2, 3 et 4, les résultats en lien avec les principes et les problèmes d'utilisabilité dits « généraux » sont par définition applicables à n'importe quels types de systèmes informatiques interactifs.

En ce qui concerne les principes et problèmes dédiés aux systèmes d'alerte médicamenteux, ils sont, par définition, propres à ces systèmes. Néanmoins, une partie de la connaissance accumulée pourrait être possiblement étendue à d'autres types de systèmes d'alerte cliniques (*e.g.*, systèmes d'alerte de résultats anormaux de biologie) en fonction de son niveau de spécificité au médicament. Par exemple, les principes et problèmes liés directement à la dimension « médicament » de l'alerte (*e.g.*, listes des informations à afficher dans l'alerte) ne peuvent pas être transférées à d'autres types de systèmes d'alerte. A l'opposé, les principes et les problèmes liés à l'interruption par les alertes pourraient tout à fait être considérés pour d'autres types de systèmes d'alerte.

En ce qui concerne les problèmes d'usage et les conséquences négatives sur le système de travail qui sont liés aux principes et problèmes généraux et spécifiques, ils sont fortement dépendants des caractéristiques des utilisateurs et du système de travail. Ainsi, la pertinence de les généraliser à d'autres types de technologies doit être évaluée au cas par cas en fonction des spécificités des utilisateurs et des systèmes de travail.

Généralisation de l'approche

Ce travail ne visait pas uniquement à fournir de nouvelles connaissances en utilisabilité. Il se voulait aussi proposer une approche pour améliorer l'accumulation des données en utilisabilité des TIS. L'approche proposée (rechercher les problèmes d'utilisabilité et leurs conséquences puis les croiser avec

les principes d'utilisabilité correspondants existants) peut s'appliquer à n'importe quel type de TIS du moment que des données d'utilisabilité sont accessibles.

Cependant, cette approche est très demandeuse en ressources humaines : en plus de rassembler des données d'utilisabilité pour développer de la connaissance, il est nécessaire d'actualiser régulièrement la connaissance déjà développée. Il semble peu raisonnable d'imaginer que cela peut être réalisé pour l'ensemble des TIS par une seule équipe de chercheurs en Facteurs Humains. Ce travail doit être collaboratif et impliquer l'ensemble de la communauté Facteurs Humains en informatique médicale. Les industriels du domaine devraient être eux aussi impliqués de manière à pouvoir accéder et utiliser aux données issues des évaluations de leurs technologies. Néanmoins, impliquer ces derniers risque d'être compliqué pour des raisons de forte concurrence industrielle.

Il serait possible d'améliorer l'efficacité de cette démarche si, au lieu de se contenter de réaliser des analyses rétrospectives de données, on tendait vers une démarche proactive en alimentant en continu une base de connaissance en utilisabilité. Une telle base reste à développer.

Futures recherches

Valider la grille de report

La grille de report adapté du « cadre de travail en utilisabilité » a fait l'objet d'une évaluation préliminaire. Les quelques problèmes d'utilisation de la grille identifiés durant cette évaluation doivent être résolus. Ensuite, une évaluation plus complète de l'impact de cette grille sur le processus d'évaluation sur la qualité des reports devra être réalisée avec un plus large échantillon de participants.

Développer une base de connaissance en utilisabilité

Cette base de connaissance doit être développée dans un effort conjoint de la communauté Facteurs Humains en informatique médicale. D'autres bases de données existent déjà dans le domaine de l'évaluation des technologies en santé (*e.g.*, « IT Evaluation Database »[6] [74]) mais ces bases fournissent essentiellement les références de publications intéressantes ainsi que leurs résumés. Par ailleurs, ces bases n'ont pas été développées en tenant compte des spécificités de l'utilisabilité. La base de connaissances en utilisabilité doit aller plus loin : elle doit être conçue comme un entrepôt, aussi complet que possible, contenant les problèmes d'utilisabilité des TIS et leurs conséquences. Des clefs de tri de ces connaissances doivent être développées pour permettre un tri par type de TIS ou par fonction par exemple. Par ailleurs, il est nécessaire de définir et planifier le mode d'alimentation et d'actualisation de cette base. Mais tout d'abord, il est indispensable de connaitre les futurs utilisateurs de cette base ; cela va influencer le type de classification ainsi que la formulation de cette connaissance.

Ce travail visait à fournir une connaissance structurée aux experts en Facteurs Humains ainsi qu'aux concepteurs. Pour cette raison, les problèmes d'utilisabilité à travers une heuristique ergonomique car ce type de classification est à la fois facilement utilisable par des experts en Facteurs Humains et relativement

[6] http://evaldb.umit.at/index.htm

accessible pour des concepteurs initiés. Cependant, la base de connaissance pourrait aussi s'adresser à des chercheurs en Facteurs Humains. Dans ce cas, la classification par heuristique est insuffisante car elle ne permet pas de représenter les processus cognitifs liés aux problèmes d'utilisabilité : d'autres types de classifications doivent alors être préférés (*e.g.*, "extended Usability Action Framework" [75]).

Ainsi dans sa forme finale, la base de connaissance en utilisabilité appliqué aux TIS devrait permettre une double classification de la connaissance: pour concepteurs et pour experts ou chercheurs en Facteurs Humains.

Transformer la liste illustrée des principes d'utilisabilité en un outil utile

Enfin, les résultats présentés dans le chapitre 4 devraient être utilisés pour développer un outil aidant la conception et l'évaluation de systèmes d'alerte médicamenteux du point de vue de l'utilisabilité. Cet outil n'a pas pour but de remplacer l'implication de compétences en Facteurs Humains tout au long du processus de conception : il serait plutôt une source complémentaire de connaissance pour améliorer l'efficacité de la conception et de la l'évaluation.

Une manière intéressante d'utiliser cette liste serait de la transformer en une liste d'« heuristiques d'utilisabilité basées sur des preuves » comme il en existe déjà dans le domaine des erreurs d'utilisation des TIS [76]. Pour réaliser cet outil, il est nécessaire de regrouper un panel d'experts pour organiser et formuler les heuristiques (différemment en fonction des utilisateurs). Cet outil devra ensuite faire l'objet d'une série d'évaluations pour améliorer son utilisabilité. Enfin, son efficacité devra être évaluée lors de projets réels de conception et/ou d'évaluation. Finalement, une approche d'amélioration continue devra être mise en place.

Ce type d'outil aurait été très utile dans notre laboratoire s'il avait été disponible lors de nos précédents projets de conception et évaluation de divers systèmes d'alertes médicamenteux. (*e.g.*, le projet Européen "Patient Safety through Intelligent Procedures in medication (PSIP)"[7] [16-18;77]): il nous aurait permis d'éviter de « réinventer la roue » à chaque projet. Plus largement, des « heuristiques d'utilisabilité basées sur des preuves » dédiés aux TIS pourraient fournir une aide efficace aux concepteurs et aux experts en Facteurs Humains du domaine. Le développement et le partage de tels outils devraient être encouragés.

10. Conclusion

L'utilisabilité appliquée aux TIS est un thème de recherche relativement récent qui fait face a des difficultés de maturation. Une étape nécessaire est de capitaliser et de réutiliser les données produites dans des études antérieures de manière à créer de nouvelles connaissances. Ce travail participe à sa mesure à l'amélioration des connaissances en utilisabilité en réutilisant les données d'études précédentes. Si ce travail devait être étendu à l'ensemble des TIS, cette tâche ne pourrait être accomplie que si l'ensemble de la communauté des Facteurs Humains en informatique médicale ainsi que les industriels du domaine s'accordaient à partager et à travailler ensemble sur les données d'utilisabilité.

[7] http://psip.univ-lille2.fr/prototypes/public/

Références

[1] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004 Mar;11(2):104-12.

[2] Kaplan B, Harris-Salamone KD. Health IT success and failure: recommendations from literature and an AMIA workshop. J Am Med Inform Assoc 2009 May;16(3):291-9.

[3] Kushniruk A, Triola M, Stein B, Borycki E, Kannry J. The relationship of usability to medical error: an evaluation of errors associated with usability problems in the use of a handheld application for prescribing medications. Stud Health Technol Inform 2004;107(Pt 2):1073-6.

[4] Han YY, Carcillo JA, Venkataraman ST, Clark RS, Watson RS, Nguyen TC, et al. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. Pediatrics 2005 Dec;116(6):1506-12.

[5] Magrabi F, Ong MS, Runciman W, Coiera E. An analysis of computer-related patient safety incidents to inform the development of a classification. J Am Med Inform Assoc 2010 Nov;17(6):663-70.

[6] Magrabi F, Ong MS, Runciman W, Coiera E. Using FDA reports to inform a classification for health information technology safety problems. J Am Med Inform Assoc 2012 Jan;19(1):45-53.

[7] Samaranayake NR, Cheung ST, Chui WC, Cheung BM. Technology-related medication errors in a tertiary hospital: a 5-year analysis of reported medication incidents. Int J Med Inform 2012 Dec;81(12):828-33.

[8] International Standardization Organization. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability (Rep N° 9241-11). Geneva: International Standardization Organization; 1998.

[9] Council Directive 2007/47/EC, 2007:247:0021:0055, European Parliament Council, (2007).

[10] International Standardization Organization. Ergonomics of human system interaction - Part 210: Human centered design for interactive systems (Rep N°9241-210). Geneva: International Standardization Organization; 2010.

[11] Beuscart-Zephir MC, Elkin P, Pelayo S, Beuscart R. The human factors engineering approach to biomedical informatics projects: state of the art, results, benefits and challenges. Yearb Med Inform 2007;109-27.

[12] International Standardization Organization. Ergonomics of human-system interaction -- Usability methods supporting human-centred design (Rep N° 16982). Geneva: International Standardization Organization; 2002.

[13] Beuscart-Zephir MC, Pelayo S, Degoulet P, Anceaux F, Guerlinger S, Meaux JJ. A usability study of CPOE's medication administration functions: impact on physician-nurse cooperation. Stud Health Technol Inform 2004;107(Pt 2):1018-22.

[14] Beuscart-Zephir MC, Anceaux F, Menu H, Guerlinger S, Watbled L, Evrard F. User-centred, multidimensional assessment method of Clinical Information Systems: a case-study in anaesthesiology. Int J Med Inform 2005 Mar;74(2-4):179-89.

[15] Beuscart-Zephir MC, Pelayo S, Bernonville S. Example of a Human Factors Engineering approach to a medication administration work system: potential impact on patient safety. Int J Med Inform 2010 Apr;79(4):e43-e57.

[16] Hackl WO, Ammenwerth E, Marcilly R, Chazard E, Luyckx M, Leurs P, et al. Clinical evaluation of the ADE scorecards as a decision support tool for adverse drug event analysis and medication safety management. Br J Clin Pharmacol 2013 Sep;76 Suppl 1:78-90.

[17] Marcilly R, Leroy N, Luyckx M, Pelayo S, Riccioli C, Beuscart-Zephir MC. Medication related computerized decision support system (CDSS): make it a clinicians' partner! Stud Health Technol Inform 2011;166:84-94.

[18] Marcilly R, Bernonville S, Riccioli C, Beuscart-Zephir MC. Patient safety-oriented usability testing: a pilot study. Stud Health Technol Inform 2012;180:368-72.

[19] Pelayo S, Bras Da CS, Leroy N, Loiseau S, MacKeon D, Trancard D, et al. Application of the medical device directive to software: methodological challenges. Stud Health Technol Inform 2013;192:437-41.

[20] International Electrotechnical Commision. Medical devices - Application of usability engineering to medical devices (Rep N° 62366). Geneva: International Standardization Organization; 2007.

[21] Marcilly R, Beuscart-Zephir MC, Ammenwerth E, Pelayo S. Seeking Evidence to Support Usability Principles for Medication-Related Clinical Decision Support (CDS) Functions. Stud Health Technol Inform 2013;192:427-31.

[22] Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996 Jan 13;312(7023):71-2.

[23] Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. JAMA 1995 Dec 13;274(22):1800-4.

[24] Jaspers MW, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. J Am Med Inform Assoc 2011 May 1;18(3):327-34.

[25] Gandhi TK, Weingart SN, Seger AC, Borus J, Burdick E, Poon EG, et al. Outpatient prescribing errors and the impact of computerized prescribing. J Gen Intern Med 2005 Sep;20(9):837-41.

[26] Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. JAMA 1998 Oct 21;280(15):1339-46.

[27] Ranji SR, Rennke S, Wachter RM. Computerised provider order entry combined with clinical decision support systems to improve medication safety: a narrative review. BMJ Qual Saf 2014 Apr 12.

[28] Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. AMIA Annu Symp Proc 2007;26-30.

[29] Kuperman GJ, Bobb A, Payne TH, Avery AJ, Gandhi TK, Burns G, et al. Medication-related clinical decision support in computerized provider order entry systems: a review. J Am Med Inform Assoc 2007 Jan;14(1):29-40.

[30] van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006 Mar;13(2):138-47.

[31] Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, et al. Ten commendments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003 Nov;10(6):523-30.

[32] Seidling HM, Phansalkar S, Seger DL, Paterno MD, Shaykevich S, Haefeli WE, et al. Factors influencing alert acceptance: a novel approach for predicting the success of clinical decision support. J Am Med Inform Assoc 2011 Jul;18(4):479-84.

[33] Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behaviour and Information Technology 1997;6(4-5):220-31.

[34] Wipfli R, Betrancourt M, Guardia A, Lovis C. A qualitative analysis of prescription activity and alert usage in a computerized physician order entry system. Stud Health Technol Inform 2011;169:940-4.

[35] van der Sijs H, van GT, Vulto A, Berg M, Aarts J. Understanding handling of drug safety alerts: a simulation study. Int J Med Inform 2010 May;79(5):361-9.

[36] Russ A.L., Saleem J.J., McManus M.S., Zillich A.J., Doebbling B.N. Computerized medication alerts and prescriber mental models: observing routine patient care. 2009 p. 655-9.

[37] Trafton J, Martins S, Michel M, Lewis E, Wang D, Combs A, et al. Evaluation of the acceptability and usability of a decision support system to encourage safe and effective use of opioid therapy for chronic, noncancer pain by primary care providers. Pain Med 2010 Apr;11(4):575-85.

[38] Duke JD, Bolchini D. A successful model and visual design for creating context-aware drug-drug interaction alerts. AMIA Annu Symp Proc 2011;339-48.

[39] Feldstein A, Simon SR, Schneider J, Krall M, Laferriere D, Smith DH, et al. How to design computerized alerts to safe prescribing practices. Jt Comm J Qual Saf 2004 Nov;30(11):602-13.

[40] Kortteisto T, Komulainen J, Makela M, Kunnamo I, Kaila M. Clinical decision support must be useful, functional is not enough: a qualitative study of computer-based clinical decision support in primary care. BMC Health Serv Res 2012;12:349-57.

[41] Chused AE, Kuperman GJ, Stetson PD. Alert override reasons: a failure to communicate. AMIA Annu Symp Proc 2008;111-5.

[42] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. A human factors investigation of medication alerts: barriers to prescriber decision-making and clinical workflow. AMIA Annu Symp Proc 2009;548-52.

[43] Russ AL, Zillich AJ, McManus MS, Doebbeling BN, Saleem JJ. 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 2012 Apr;81(4):232-43.

[44] Weingart SN, Massagli M, Cyrulik A, Isaac T, Morway L, Sands DZ, et al. Assessing the value of electronic prescribing in ambulatory care: a focus group study. Int J Med Inform 2009 Sep;78(9):571-8.

[45] Zachariah M, Phansalkar S, Seidling HM, Neri PM, Cresswell KM, Duke J, et al. Development and preliminary evidence for the validity of an instrument assessing implementation of human-factors principles in medication-related decision-support systems--I-MeDeSA. J Am Med Inform Assoc 2011 Dec;18 Suppl 1:i62-i72.

[46] Patterson ES, Nguyen AD, Halloran JP, Asch SM. Human factors barriers to the effective use of ten HIV clinical reminders. J Am Med Inform Assoc 2004 Jan;11(1):50-9.

[47] Saleem JJ, Patterson ES, Militello L, Anders S, Falciglia M, Wissman JA, et al. Impact of clinical reminder redesign on learnability, efficiency, usability, and workload for ambulatory clinic nurses. J Am Med Inform Assoc 2007 Sep;14(5):632-40.

[48] Horsky J, Kaufman DR, Patel VL. Computer-based drug ordering: evaluation of interaction with a decision-support system. Stud Health Technol Inform 2004;107(Pt 2):1063-7.

[49] Baysari MT, Westbrook JI, Richardson KL, Day RO. The influence of computerized decision support on prescribing during ward-rounds: are the decision-makers targeted? J Am Med Inform Assoc 2011 Nov;18(6):754-9.

[50] Khajouei R, Peek N, Wierenga PC, Kersten MJ, Jaspers MW. Effect of predefined order sets and usability problems on efficiency of computerized medication ordering. Int J Med Inform 2010 Oct;79(10):690-8.

[51] van der Sijs H, Aarts J, van GT, Berg M, Vulto A. Turning off frequently overridden drug alerts: limited opportunities for doing it safely. J Am Med Inform Assoc 2008 Jul;15(4):439-48.

[52] Krall MA, Sittig DF. Clinician's assessments of outpatient electronic medical record alert and reminder usability and usefulness requirements. Proc AMIA Symp 2002;400-4.

[53] Saleem JJ, Patterson ES, Militello L, Render ML, Orshansky G, Asch SM. Exploring barriers and facilitators to the use of computerized clinical reminders. J Am Med Inform Assoc 2005 Jul;12(4):438-47.

[54] Chan J, Shojania KG, Easty AC, Etchells EE. Usability evaluation of order sets in a computerised provider order entry system. BMJ Qual Saf 2011 Nov;20(11):932-40.

[55] Russ A.L., Saleem J.J, McManus M.S., Frankel R.M., Zillich A.J. The Workflow of Computerized Medication Ordering in Primary Care is Not Prescriptive. 2010 p. 840-4.

[56] Ash JS, Sittig DF, Dykstra RH, Guappone K, Carpenter JD, Seshadri V. Categorizing the unintended sociotechnical consequences of computerized provider order entry. Int J Med Inform 2007 Jun;76 Suppl 1:S21-S27.

[57] Hartmann Hamilton AR, Anhoj J, Hellebek A, Egebart J, Bjorn B, Lilja B. Computerised Physician Order Entry (CPOE). Stud Health Technol Inform 2009;148:159-62.

[58] Khajouei R, de Jongh D, Jaspers MW. Usability evaluation of a computerized physician order entry for medication ordering. Stud Health Technol Inform 2009;150:532-6.

[59] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005 Mar 9;293(10):1197-203.

[60] Peute LW, Spithoven R, Bakker PJ, Jaspers MW. Usability studies on interactive health information systems; where do we stand? Stud Health Technol Inform 2008;136:327-32.

[61] Sebillotte S. Action representation for home automation. Amsterdam: Elsevier Press; 1990 p. 985-90.

[62] Russ AL, Saleem JJ, McManus MS, Frankel RM, Zillich AJ. The Workflow of Computerized Medication Ordering in Primary Care is Not Prescriptive. 2010 p. 840-4.

[63] Russ AL, Saleem JJ, McManus MS, Zillich AJ, Doebbling BN. Computerized medication alerts and prescriber mental models: observing routine patient care. 2009 p. 655-9.

[64] Chan J, Shojania KG, Easty AC, Etchells EE. Does user-centred design affect the efficiency, usability and safety of CPOE order sets? J Am Med Inform Assoc 2011 May 1;18(3):276-81.

[65] Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. J Biomed Inform 2012 Dec;45(6):1202-16.

[66] Horsky J, Phansalkar S, Desai A, Bell D, Middleton B. Design of decision support interventions for medication prescribing. Int J Med Inform 2013 Jun;82(6):492-503.

[67] Pelayo S, Marcilly R, Bernonville S, Leroy N, Beuscart-Zephir MC. Human factors based recommendations for the design of medication related clinical decision support systems (CDSS). Stud Health Technol Inform 2011;169:412-6.

[68] Phansalkar S, Edworthy J, Hellier E, Seger DL, Schedlbauer A, Avery AJ, et al. A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. J Am Med Inform Assoc 2010 Sep;17(5):493-501.

[69] Sittig DF, Wright A, Osheroff JA, Middleton B, Teich JM, Ash JS, et al. Grand challenges in clinical decision support. J Biomed Inform 2008 Apr;41(2):387-92.

[70] Marcilly R, Ammenwerth E, Vasseur F, Beuscart-Zephir MC. Usability flaws of medication-related alerting systems: a systematic review. IJMI (submitted). In press 2014.

[71] Schmettow M, Vos W, Schraagen JM. With how many users should you test a medical infusion pump? Sampling strategies for usability tests on high-risk systems. J Biomed Inform 2013 Aug;46(4):626-41.

[72] de Jonckheere, Rommel D, Nandrino J, Jeanne M, Logier R. Heart rate variability analysis as an index of emotion regulation processes: Interest of the Analgesia Nociception Index (ANI). Conf Proc IEEE Eng Med Biol Soc 2012 Aug;2012:3432-5.

[73] Nielsen J. Usability engineering. 1993. London, Academic Press.

[74] Ammenwerth E, de KN. An inventory of evaluation studies of information technology in health care trends in evaluation research 1982-2002. Methods Inf Med 2005;44(1):44-56.

[75] Khajouei R, Peute LW, Hasman A, Jaspers MW. Classification and prioritization of usability problems using an augmented classification scheme. J Biomed Inform 2011 Dec;44(6):948-57.

[76] Borycki E, Kushniruk A, Carvalho C. A methodology for validating safety heuristics using clinical simulations: identifying and preventing possible technology-induced errors related to using health information systems. Comput Math Methods Med 2013;2013:526419.

[77] Marcilly R, Beuscart-Zephir MC, Beuscart R. Integrating Human Factors in an international research project: lessons learned from the PSIP project. Stud Health Technol Inform 2013;183:162-7.

Curriculum Vitae and Portfolio

Curriculum Vitae

Identity	Romaric Marcilly
Birth date and place	Born on September 4th 1980, in Cormeilles-en-Parisis, France, Val d'Oise.
Since June 2008	Research engineer – Human Factors (Lille' CIC IT)
Since December 2011	PhD student in Medical Informatics (Lille 2 University)
2008	Master in Occupational Psychology and Human Factors (Lille 3 University, France)
2003	Degree (D.E.A.) in Psychology, specialty visual perception and ecological psychology (Lille 3 University, France)
2001	Graduate in Psychology (Lille 3 and Reims University, France)
1998	Baccalauréat, literature stream, specialty mathematics and latin

Research topics

- Prevention of Adverse Drug Events (ADE) in hospital settings. How to improve patient safety through innovative technologies? Usual ADE prevention technologies, such as alerting systems, face some limits due to usability and knowledge issues. Using scorecards that display monthly statistics about the occurrence of the ADE seems to be a promising technology to improve clinicians ADE awareness, ultimately leading them to change their prescribing practices. The appropriation by clinicians of this tool and the evaluation of its impact must be studied in-depth. I work on this topic with the Lille former PSIP[8] team (CERIM, Lille 2 university)
- Improving the quality and the completeness of the usability reports in order to get sufficient data to be able to look for evidence supporting usability design principles. I work on this topic with Dr. L. Peute and Pr. M. Jaspers (AMC, Amsterdam). For usability to become a mature science it is necessary to be able to reuse data from former studies to develop usability knowledge. Yet, usability reports are often weak in terms of completeness and quality of data reported. Improving usability reports quality is the first mandatory step to be able to look for usability evidence.
- "Usability-induced use errors", how to prevent them? Usability issues may cause use errors. The concept of "usability induced use errors" proposed in this PhD must be clarified in regards to other concepts such as "technology induced errors" and "use errors". To prevent their appearance it is necessary to convince designers to consider usability issues during the design process: however, designers do not often understand the usefulness of considering usability. In this topic I explore with the HF experts of Lille' CIT-IT different ways to convince designers: specific usability reports for instance.

[8] Patient Safety through Innovative Procedures in medication: http://psip.univ-lille2.fr/prototypes/public/

Portfolio

Courses /seminars	Year(s)
Seminar "good collaborative practice": how private and public partners may work together in a research project, what is the legal frame for such collaborations, what are the requisites and the good practices in terms of data sharing and innovation property.	2013
Training course « describe and analyze multi-factorial data » with R software environment: principal component analysis, canonical factor analysis, factor regression analysis, common factor analysis.	2013
Training course « small samples statistics » with R software environment: correlations, descriptive parameters, Khi square, chart design etc.	2013
Training course « Wikipedia »: how to assess the validity of the information in Wikipedia, how to create topics and/or complete existing topics. Creation of the topic "usability goals".	2012
Seminar of the PhD students in Human Factors from Northern France (Lille, Valenciennes): current research topics in Human Factors.	2012-2014

International conferences in Medical Informatics	Year(s)
Congress of the European Federation for Medical Informatics (EFMI): MIE	2014, 2012, 2011
World congress of Medical Informatics (MedInfo)	2013
Congress Context Sensitive in Health Informatics	2013
Congress ITCH (health informatics)	2013, 2011
International Symposium on Human Factors in Health Informatics	2012
Congress on Healthcare Systems Ergonomics and Patient Safety (HEPS)	2011
Congress on the Detection and Prevention of the Adverse Drug Events: information technologies and Human Factors	2011

National conferences in Human Factors	Years
Congress of the French Speaking Ergonomics Society (SELF)	2011, 2010

Organization of workshops	**Year**
MIE: "Enhancing the Reporting on Human Factor/Usability studies of Health Information Technologies". LW Peute, R **Marcilly**, KF Driest, S Pelayo, MC Beuscart-Zephir and MWM. Jaspers	2014
ITCH: "Medication-related CDS functions usability features: from literature to the design and evaluation". R **Marcilly**, S Pelayo, R Beuscart & MC Beuscart-Zephir	2012

Parameters of esteem	**Year**
Nominated for distinguished paper at the MedInfo 2013	2013
Young Researcher Award of the 2d International Workshop on Patient Safety though Intelligent Procedures in medication	2011
Best Poster Award of the XXIII International Conference of the European Federation for Medical Informatics (MIE 2011)	2011

Working groups and membership
Human Factors Engineering in Health Informatics Working Group (HFEHI)
Evaluation of Health Information Systems Working Group (HISEval)
French Speaking Society of Ergonomics

Expert activities: reviewing	**Year**
Context Sensitive in Health Informatics (conference)	2013
International Journal for Quality in Healthcare	2013

Teaching experience	**Year(s)**
Lille 3 University	
Master degree in Occupational Psychology and Evaluation, psychology students (initial training) • Evaluations methods: heuristics inspection, user testing, critical incidents technique	2012-2014
Bachelor degree in Psychology, psychology students (initial training) • First year: scientific methodology (questionnaire, testing, standardization, experimental plan etc.) • Second year: psychophysiology: brains differences across species, reaction time • Third year: Psychometrics (standardization, reliability, reliability, sampling, sensitivity, specificity etc.)	2004-2007
University Institutes of Technology, caseworker students (initial training) • Developmental psychology especially adolescent psychology (physiological, cognitive, affective, social and moral development)	2006
Lille 2 University	
Master degree in Human Factors, continuing education for occupational physicians and nurses, human resource managers etc. • Cognitive activity and workflow analysis • Human error and systems' reliability	2011-2014
Master degree in Public Health, interns in medicine (initial training) • Human factors issues of Decision Support Systems	2014
Lille 1 University	
Master degree in Engineering in Multimedia Technology, computer engineers, pedagogy scientists (initial training and continuing education) • User-centered design process • Usability design principles for software and internet interfaces • Usability evaluation methods: heuristics inspection, user testing	2011-2014

List of publications and conferences

To be submitted as soon as the paper of chapter 2 is accepted

1. **Marcilly** R., Ammenwerth E., Roehrer E., Pelayo S., Vasseur F., Beuscart-Zéphir M.-C. Impact of usability flaws in medication-related alerting systems on usage and work system.
2. **Marcilly** R., Ammenwerth E., Nies J., Roehrer E., Vasseur F., Beuscart-Zéphir M.-C. Usability design principles for medication alerting systems: literature synthesis and study of their coverage regarding related usability flaws.

Submitted

1. **Marcilly** R., Ammenwerth E., Vasseur F., Roehrer E., Beuscart-Zéphir M.-C. Usability flaws of medication-related alerting functions: a systematic review. Submitted to the Journal of Biomedical Informatics (September 2014)

Accepted

1. **Marcilly** R., Vasseur F., Ammenwerth E., Beuscart-Zéphir M.-C. Methods uncovering usability issues in medication-related alerting functions: results from a systematic review. *MIE 2014* (accepted)
2. **Marcilly** R., Boog C., Leroy N., Pelayo S. Perceived usefulness of a usability issues reporting form to help understand "usability-induced use-errors": a preliminary study. *MIE 2014* (accepted)
3. Lamer A., **Marcilly** R., Jeanne M., Logier R. Automatic Scanning of Free-Text Entries. *MIE 2014* (poster accepted)
4. Peute L., Driest K., **Marcilly** R., Bras da Costa S., Beuscart-Zephir M. C., Jaspers M.. A framework for reporting on Human Factor/Usability studies of Health Information Technologies. Stud Health Technol Inform. 2013;194:54-60.
5. Hackl W.O., Ammenwerth E., **Marcilly** R., Chazard E., Luyckx M., Leurs P., Beuscart R. Clinical evaluation of the ADE scorecards as a decision support tool for adverse drug event analysis and medication safety management. British Journal of Clinical Pharmacology. 2013, 76 (Suppl 5): 78-90.
6. Hackl W.O, Ammenwerth E, **Marcilly** R, Chazard E, Luyckx M, Leurs P, Beuscart R. Time-series study of the impact of ADE Scorecards on Adverse Drug Events: Preliminary results. In: Ammenwerth E, Hörbst A, Hayn D, Schreier G (eds.). eHealth2013 - Health Informatics meets eHealth. Big Data: eHealth von der Datenanalyse bis zum Wissensmanagement. Proceedings of eHealth2013 in Wien, 23.-24.5.2013. OCG. 2013. pp. 25-30.
7. **Marcilly** R., Beuscart-Zephir M-C., Ammenswerth E. & Pelayo S. Seeking Evidence to Support Usability Principles for Medication-Related Clinical Decision Support (CDS) Functions. Stud Health Technol Inform. 2013;192:427-31.
8. **Marcilly** R, Beuscart-Zephir MC, Beuscart R. Integrating Human Factors in an international research project: lessons learned from the PSIP project. Stud Health Technol Inform. 2013;183:162-7.
9. **Marcilly** R, Bras Da Costa S, Boog C, Beuscart-Zephir MC, De Jonckheere J, Pelayo S. Impact of the context of use analysis for the extension of an existing medical device: an analgesia monitor case study. Stud Health Technol Inform 2013;194:139-44..
10. **Marcilly** R, Bernonville S, Riccioli C, Beuscart-Zephir M-C. (2012). Patient safety-oriented usability testing: a pilot study. Stud Health Technol Inform, 180:368-72.
11. **Marcilly** R., Pelayo S., Beuscart-Zéphir M.-C., Hackl W.O. (2011). Utilisation de tableaux de bord pour la prévention des Evénements Indésirables Médicamenteux: premiers résultats. In F. Jeffroy

& A. Garrigou (Ed). L'ergonomie à la croisée des risques. *SELF'2011, Congrès International d'Ergonomie*. Paris, France.

12. Pelayo S., **Marcilly** M., Bernonville S., Leroy N. & Beuscart-Zephir M.-C. (2011). Human Factors based recommendations for the design of medication related clinical decision support systems (CDSS) Communication proposed to the XXIII International Conference of the European Federation for Medical Informatics, Oslo, Norway, august 28-31
13. **Marcilly** R., Hackl W. O, Pelayo S., Hassler S., Riccioli C., Beuscart-Zéphir M.-C. (2011). Human-Centered Design of a Scorecard tool for Adverse Drug Events, *MIE 2011*, Olso.
14. Chazard E., Baceanu A., **Marcilly** R., Bernonville S., Ficheur G., Beuscart R. (2011). A Web tool for automated Adverse Drug Events detection: the ADE Scorecards, *23rd International conference of the European Federation for Medical Informatics (MIE),* Oslo.
15. **Marcilly** R., Pelayo S., Hackl O. W., Beuscart-Zéphir M.-C., Ammenwerth E. (2011). Scorecards as "team ADE awareness" support, *5th International Symposium on Human factors Engineering in Health Informatics*.
16. Chazard, E., Baceanu A., Ficheur G., **Marcilly** R., Beuscart R. (2011). Les « ADE Scorecards », outil de détection et visualisation des effets indésirables médicamenteux. *24ème congrès nationale EMOIS (Evaluation Management Organisation Information Santé)*, Nancy.
17. **Marcilly** R., Hackl W., Beuscart-Zéphir M.-C. & Pelayo S. (2011). Scorecards: a method to prevent Adverse Drug Events? *International Conference HEPS (Healthcare Systems Ergonomics and Patient Safety congress)*.
18. Bernonville S., **Marcilly** R., Messai R., Leroy N., Przewozny E., Souf N., Beuscart-Zéphir M.-C. (2011) Implementation of a taxonomy aiming to support the design of a contextualised Clinical Decision Support System. *Studies in Health Technology and Informatics*. 166:74-83.
19. **Marcilly** R., Leroy N., Luyckx M., Pelayo S., Riccioli, C. & Beuscart-Zéphir M.-C. (2011). Medication related Computer Decision Support System (CDSS): make it a clinicians' partner! *Studies in Health Technology and Informatics*. 166:84-94.
20. **Marcilly** R., Hackl W. O., Luyckx M., Ammenwerth E. (2011). Scorecards: a new method to prevent Adverse Drug Events? Preliminary results from a clinical field study. *Studies in Health Technology and Informatics*. 166:234-245.
21. Chazard E., Baceanu A., Ficheur G., **Marcilly** R., Beuscart R. (2011). Les « ADE Scorecards », un outil de détection et de visualisation des effets indésirables liés aux médicaments. *Journées Francophones d'Informatique Médicale 2011*.
22. Bernonville S., Messai R., **Marcilly** R., Leroy N., Przewozny E., Souf N. & Beuscart-Zéphir M.-C. (2001). Développement et exploitation d'une taxonomie visant l'aide à la conception d'un système d'aide à la décision médicamenteuse « contextualisant ». *Journées Francophones d'Informatique Médicale 2011*.
23. **Marcilly**, R., Charzard, E., Beuscart-Zéphir, M.-C., Hackl W., Baceanu, A., Kushniruk, A., Borycki, E. (2011). Design of Adverse Drug Events-Scorecards. *Studies in Health Technology and Informatics*. 164:377-381.
24. Hackl W., **Marcilly** R., Baceanu A., Chazard E. & Beuscart-Zephir MC. (2010). Prevention of adverse drug events (ADE) by ADE-Scorecards: an interrupted time series study. *Congress of the German Society for Medical Informatics and Epidemiology*. Mannheim, Allemagne.
25. **Marcilly** R., Pelayo S., Bernonville S., Leroy N. & Forrierre J. (2010). Prise en Compte du Contexte pour la Conception d'un Système d'Aide à la Décision Médicamenteuse. *45ème Congrès de la Société d'Ergonomie de Langue Française,* Liège, Belgique.
26. Marciniak B., **Marcilly** R., Aldegheri J. & Anceaux F. (2009). Impact of the Expertise on the Gathering of Information Contained in the Anesthetic File. *Proceedings of the 2009 Annual Meeting of the American Society of Anesthesiologists*, New Orleans, LA.
27. **Marcilly** R., Aldegheri J. & Marciniak B. (2009). L'influence de l'expertise sur l'analyse de

l'information contenue dans la feuille d'anesthésie. *Congrès Annuel de la Société Française d'Anesthésie et de Réanimation*, Paris

28. Watbled L., **Marcilly** R. & Guerlinger S. (2009). Intérêt de la démarche ergonomique dans la conception et l'implémentation d'un système informatique en santé : exemple du dossier d'anesthésie électronique, *Premier forum international Ergonomie-Design*, Lyon, France.
29. Leroy N., Luyckx M., Lecocq P., **Marcilly** R. & Beuscart-Zéphir MC. (2009). Contribution of human factors for the review of automatically detected ADE. *Studies in Health Technology and Informatics*. 148:170-80.

Supplementary material

Complete queries

PubMed

("medical order entry systems"[MH] OR "medication alert system"[TIAB] OR "computerized physician order entry system"[TIAB] OR "CPOE"[TIAB] OR "Decision Support Systems, Clinical"[MH] OR "clinical decision support systems"[TIAB] OR "CDSS"[TIAB]) AND (English[lang] OR French[lang]) AND ("1980"[PDAT] : "2012"[PDAT]) AND ("User-computer interface"[MeSH Terms] OR "Human engineering"[MeSH Terms] OR "Risk factors"[MeSH Terms] OR "Humans"[MeSH Terms] OR "usability"[TIAB])"

Scopus

pub-date > 1979 and ("medical order entry" OR "medication alert" OR "computerized physician order entry" OR "CPOE" OR "clinical decision support" OR "CDSS") AND ("User-computer interface" OR "Human engineering" OR "Risk factor" OR "Human factor" OR "usability" OR "Human-computer interaction")) [Journals(Medicine and Dentistry)]

Ergonomics Abstracts

"medical order entry" OR "medication alert" OR "computerized physician order entry" OR "computerized provider order entry" OR "CPOE" OR "clinical decision support" OR "CDSS"

Classification of usage problems

Names (and numbers) of the categories and subcategories of usage problems used to classify the meaningful semantic units (left) and corresponding previous categorizations defined through the open card sorting (right). Some "open card sorting" categories are presented several times because they correspond to several final categories.

Final categorization			Open card sorting resulting categories	
Categories and subcategories (25)			Reviewer 1 (15)	Reviewer 2 (23)
Attitudinal issues	Questioning	The behavior of the system: how the function is working, how it responds to users' actions	Confusion on system behavior	Questioning function's behavior
		The triggering and sorting model of the alerts	Confusion on coverage of the function.	Questioning the coverage of the function
		The usefulness of the alerting system	Negative attitude toward usefulness of the function	Disappointed by the function
		The validity of the alert	No confidence in the validity, skepticism & cynicism	- Questioning the validity - Confidence issue
	Alert fatigue / desensitization		Alert fatigue & override	Alert fatigue
	Negative feelings towards the function		Negative attitude	
Behavioral issues	Increased workload		Increased workload	Increased workload
	Users' actions	Users do not use the function at all	/	User does not use the system
		Users voluntarily ignore the alerts	Alert fatigue & override	/
		Users ineffectively use the system	/	Difficult to : - To apply the action - To use the function
		Users use workarounds	/	Workarounds
		Users follow blindly the advice	/	/
	User lost		User lost	Lost in use
Cognitive issues	Information involuntarily missed: they cannot access or find it		Missed information	- User knows he missed an information - Difficult to found the information
	Increased memory load while using the alerting system: users must rest on their memory		- Increased workload - Missing information - Impede cognitive process, rely on his own information	- Missing information - Increased workload
	Difficulties	Difficulties to understand the alert	Difficult to understand the alert	Information hard to interpret
		Difficulties to identify alert's components	/	/
	Misinterpretation	Misinterpretation of alerts' components	Misidentification	Misinterpretation
		Misinterpretation of alerts' content	Misinterpretation	Misinterpretation
	User interrupted by the alerts		/	- Attentionnal distraction - Decision making bothered
Emotional issues	Annoyance/irritation		- Frustration - Irritation	- Irritation - Complain
	Frustration		- Frustration - Irritation	Frustration
	Ugly experience		/	/
	Stress, pressure		/	Stress
	Cynicism		No confidence in the validity, skepticism & cynicism	/

List of usability flaws reported in the literature classified

The related reference list is presented in chapter 2.

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[42]	"Interface, which did not adequately support all prescriber types."	Adaptability	Flexibility	/	/
[26]	"erroneous system messaging [about how to solve the problem]"	Compatibility	Alert content issues	Alert proposes erroneous actions to take	No detail
[42]	"guidance on actions to take" "Prescribers wanted alert to provide (…) advice on how to respond to the alert"	Compatibility	Alert content issues	Missing information	About the actions that could be taken
[42]	"Handbooks, Micromedex, and literature are consulted"; "Pharmacists are consulted real-time via phone or face-to-face; *e.g.*,, [Observer to NP: 'What do you do if you are unfamiliar with an order check or have a question about it?'] NP: "Sometimes, I order the medication, because I know I can back out [cancel medication] after deeply looking into it. Sometimes, I do not order the medication. I get a book and look it up. If I don't find it there, then I call the clinical pharmacist." " [rephrasing: the alert content lacks of information]	Compatibility	Alert content issues	Missing information	About the evidence of the alert
[42]	"The alert did not provide essential patient information for the prescriber, even though it existed elsewhere in the EHR. For example, decision-making for some drug interaction alerts (*e.g.*,, amiloride and lisinopril, which can cause hyperkalemia) depend on patient labs (*e.g.*,, potassium). This missing, or more accurately, 'hidden' patient data triggered varying responses."	Compatibility	Alert content issues	Missing information	About the patient
[43]	"[missing] information in the alert text (…) (recommendations to adjust doses, to measure serum levels, monitor patient parameters, or prescribe alternative drugs)	Compatibility	Alert content issues	Missing information	About the actions that could be taken
[43]	"[missing] evidence of the DDI (…) [as a reason to turn off an alert]" [Inferred: there is no evidence mentioned in the alert, or not sufficiently serious]	Compatibility	Alert content issues	Missing information	About the evidence of the alert
[45]	"There were cases in which clinicians reached a dead end within the CR system, with no reasonable option to proceed." "We observed instances in which none of the available options to satisfy the CR applied to the patient or situation."	Compatibility	Alert content issues	Alert proposes erroneous actions to take	Actions proposed do not suit the clinical context
[45]	"Options within the dialogue box of the CRs do not always match the patient's response or there is not an appropriate option for indicating why the provider has decided not to order a test, for example."	Compatibility	Alert content issues	Alert proposes erroneous actions to take	Actions proposed do not suit the clinical context
[42; 49]	"Alert is not evidence-based, does not provide a reference to evidence that does exist, and/or the actual or perceived level of evidence is low." [49]"The alerts themselves do not present the evidence nor do they provide links to any supporting documentation" [49]"Quality/strength of alert evidence; guidance on actions to take: Prescribers wanted alert to provide: references or links to evidence (...) some alerts were not evidence-based" [42]. "The actual or perceived level of evidence is low"; "unclear if the warnings were "evidence-based""; "extend of evidence is fair, stemming mainly from a few case reports"	Compatibility	Alert content issues	Missing information	About the evidence of the alert
[42; 49]	"Alert does not provide clear information on relative risk of harm for a given patient"[49]"Although some alerts are categorized by risk, (*e.g.*,, some are marked as "significant" and others are "critical"), this notation was not always sufficient for prescribers"[49]; "Risk rating/degree of risk: Prescribers wanted a more clear indication of risk and suggested that alerts display a quantitative rating" [Inferred: there is no clear risk information in the alert] [42]	Compatibility	Alert content issues	Missing information	About the severity of the problem
[42; 47; 49]	"Alert does not provide information on why it was triggered and/or the potential problem"[49] "One physician said, "I wish it stated what the problem truly is, and simply. For example, simvastatin and diltiazem. [The alert] just says the drug names, not the problem.""[49][47] "Alerts did not adequately indicate alert triggers, explain drug interactions, or describe the problem (*e.g.*,, "potentially causes hypotension")"[42]	Compatibility	Alert content issues	Missing information	About the problem detected
[51]	"Another in-clinic user stated that the system would be more helpful if it gave more detailed information on "how to switch or discontinue drugs.”"	Compatibility	Alert content issues	Missing information	About the actions that could be taken

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[51]	"They wanted the CDSS to tell them whether to initiate or discontinue therapy or increase or decrease dosing rather than provide a detailed explanation of how to evaluate these potential actions."	Compatibility	Alert content issues	Missing information	About the actions that could be taken
[53]	"The most common feature suggestions were to provide the normal ranges for labs (6 subjects)" [rephrasing: it lacks of the normal ranges for labs data in the medication alerts]	Compatibility	Alert content issues	Missing information	To interpret data within the alert
[53]	"The most common feature suggestions were (...) to provide more details about severity and clinical effect (3 subjects)."	Compatibility	Alert content issues	Missing information	About the severity of the problem
[53]	"The most common feature suggestions were (…) to offer alternative treatment options (4 subjects)…"	Compatibility	Alert content issues	Missing information	About the actions that could be taken
[53]	"The most common feature suggestions were (…) to display whether a DDI is dose-dependent or idiosyncratic (4 subjects)…"	Compatibility	Alert content issues	Missing information	About the problem detected
[53]	"Two areas of minor criticism were (…) the lack of a more detailed reference section."	Compatibility	Alert content issues	Missing information	About the evidence of the alert
[61]	"Alerts are not stratified by levels of severity"	Compatibility	Alert content issues	Missing information	About the severity of the problem
[61]	"All alerts only present information, and no alerts require action"	Compatibility	Alert content issues	Missing information	About the actions that could be taken
[42]	"Alerts cannot be pulled up later, as needed, hindering alert resolution" ; "Sometimes, prescribers wanted a way to retrieve an alert that had been displayed, but the alert system did not support this function."	Compatibility	Alert's features issues	Some alert-related features are missing	Alert displayed only once - no way to reconsider it later
[44]	""You do not get the warning again, and there is no button to get it [the second alert] back""	Compatibility	Alert's features issues	Some alert-related features are missing	Alert displayed only once - no way to reconsider it later
[46]	"In borderline cases, providers did not want to order medications yet but wanted to reconsider the action at the next visit."	Compatibility	Alert's features issues	Some alert-related features are missing	Alert displayed only once - no way to reconsider it later
[51]	"I would like more of the recommendations to go into a note to document what I have done with the patient"	Compatibility	Alert's features issues	Some alert-related features are missing	Alert displayed only once - no way to reconsider it later
[45]	"The use of CRs automatically generates text that is added to the progress note, but that text is not integrated with the template information and is generally added to the bottom of the note."	Compatibility	Alerts' features issues	Inappropriate feature	Reminder not integrated in the progress note template
[61]	"Directly from the alert, the user can choose to continue with or cancel the order of the offending drug; no other actionable options are available."	Compatibility	Alerts' features issues	Some alert-related features are missing	No feature to instantiate the decision within the alert
[61]	"The user can link to outside sources of information from elsewhere in the system when connected to a workstation, but there is no link within the alert"	Compatibility	Alerts' features issues	Some alert-related features are missing	No access to additional information
[27]	"A major problem in Medicator is that the alert screen "medication dose units control" shows up too late in the ordering process"	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears after the decision is made
[42]	"Technology lags/down- times" "Even 10-15 second computer delays between the order and subsequent alert"	Compatibility	Alert's presentation issues	Data processing is too slow	/
[45]	"The CRs appear on this cover sheet but are delayed in loading and displaying (Site 3 reported an average delay of 8 seconds for the CRs to load)."	Compatibility	Alert's presentation issues	Data processing is too slow	/
[52]	"There is a reminder to add the corollary PTT check order. However, it is presented out of the logical workflow as an addition to the calculated dose alert and not at the end (Task 10), when the user reviews order completeness."	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears before the decision making process start

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[52]	"The institutional guideline for heparin administration (…) triggered later in the process when the planning stage of the order is generally completed"	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears after the decision is made
[54]	"Too many alerts or alerts at an inappropriate time: "Now we get alerts when we go to charting, which in my workflow is the last step. It's after the patient's gone. Now I get warned they've got some drug interaction. Great!"	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears after the decision is made
[55]	"CPOE provides feedback on drug allergies, but only after medications are ordered."	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears after the decision is made
[56]	"Some doctors explained that the alerts appeared after they'd already made their prescribing decision and often provided them with information that they already knew."	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears after the decision is made
[57]	"Intrusive alerts presented at the wrong time in the workflow"	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears at the wrong time
[60]	"Alerts that are triggered by charting tasks rather than by ordering tasks may not be seen in the exam room workflow, as many clinicians complete their charting outside of the exam room, often after the patient has left."	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears after the decision is made
[60]	"However, even some of these users acknowledged that "pop-up" alerts can be very annoying, ..."	Compatibility	Alert's presentation issues	Display mode does not suit the decision making process	Alert too intrusive
[60]	"They said that pop-up alerts particularly were annoying or unhelpful if they popped up "too early" in the encounter, or on the wrong screen."	Compatibility	Alert's presentation issues	Alert does not appear at the right moment	Alert appears before the decision making process start
[60]	"Some acknowledged they were unlikely to respond, or perhaps even be aware of alerts, unless they were intrusive." [rephrasing: alerts not sufficiently intrusive in this case]	Compatibility	Alert's presentation issues	Display mode does not suit the decision making process	Alert not sufficiently intrusive
[42]	"Prescribers unsure if pharmacists review these (override justifications) or find them useful"	Compatibility	Tasks and control distribution issues	Entered comments are displayed to no-one	/
[44]	"Alerts regarding drug administration times should be handled by nurses [while they are displayed to physicians]"	Compatibility	Tasks and control distribution issues	Alert not displayed to the right clinician	/
[45]	"CR system is currently insufficient for supporting transmission of reminder results from nursing intake to provider examination"	Compatibility	Tasks and control distribution issues	Alert not transferable from a clinician to another	/
[55]	"The CPOE system does not display information available on other hospital systems. For example, only the pharmacy's computer provides drug interaction and lifetime limit warnings."	Compatibility	Tasks and control distribution issues	Alert displayed only to the pharmacist	/
[59]	""Mainly drug interaction alerts, and I [nurse] do not prescribe.""	Compatibility	Tasks and control distribution issues	Alert not displayed to the right clinician	/
[59]	"Most of the drug alerts are not relevant for physiotherapists' works"	Compatibility	Tasks and control distribution issues	Alert not displayed to the right clinician	/
[59]	"Reminders do not support psychologists' work."	Compatibility	Tasks and control distribution issues	Alert not displayed to the right clinician	/
[62]	"The comments are merely stored in the database and displayed to no one, vital patient care instructions may be overlooked."	Compatibility	Tasks and control distribution issues	Entered comments are displayed to no-one	/
[24]	"Alerts once entered in the system can be outdated. The processes how to keep them up-to-date is not yet implemented (*i.e.,* for patients who were carrier of methicillin-resistant staphylococcus aureus (MRSA) and who are now readmitted to the hospital)."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With data considered
[24]	"Reminder alerts that should be given the last day of hospitalization (*i.e.,* bacteriological tests), a day the system cannot forecast. This leads to an alert every day."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With the care logic

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[26]	"lack of patient tailored checking of medication order"	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With patient clinical case
[42]	"Alerts sometimes inconsistent with EHR data: *E.g.*,, Alert: 'metformin – no serum creatinine within 60 days'. NP: "This is often inaccurate."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With data considered
[42]	"Alert's applicability to a certain patient/situation" "Alerts often valid, but not "applicable to the context""	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With patient clinical case
[42]	"Same alerts appear for a patient across one or more med renewals; *E.g.*,, Alert says 'Previous adverse reaction to antidepressants.'…Phys types in override reason, 'Has been on venlafaxine for 5 yrs now.'"	Compatibility	The alerts signal-to-noise ratio is low	Alerts are redundant	Many alerts appear very frequently
[42]	"Number of alerts is problematic"; "Potential problems are overdetected or underdetected"	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[43]	"The fact that only specialists are prescribing a specific drug or the combination of drugs is mentioned [as a reason to turn off an alert]" [Inferred: irrelevance of the alert] "The fact that the drugs are intentionally combined because of a desired effect of the DDI or the fact that they are generally combined for other reasons is mentioned [as a reason to turn off an alert]" [Inferred: alert is irrelevant with clinical practice]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With expertise/ward's habits
[43]	"The incidence of adverse events due to the DDI is mentioned [as a reason to turn off an alert]" [Inferred: Irrelevance of the alerts due to their incidence]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With clinicians' priority for very at risk situations
[43]	"The rapidity of the adverse effect is mentioned [as a reason to turn off an alert]" [Inferred: problem of relevance of the alert]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With clinicians' priority for very at risk situations
[43]	"The fact that the alert is know is mentioned [as a reason to turn off an alert]" [alert is irrelevant]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With expertise/ward's habits
[43]	"The fact that effects are monitored or serum level measured is mentioned [as a reason to turn off an alert]" [Inferred: monitoring: the alert is triggered while the actions to take are already taken]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	Because the monitoring is already engaged
[43]	"The quantity or number of alerts (generated or overriden) is mentioned [as a reason to turn off an alert]" [Inferred: there are too many alerts]	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[43]	"The respondents mention that the alert does not need any action or that they never perform any action [as a reason to turn off an alert]" [Inferred: the alert is useless since no action has to be taken]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With patient clinical case
[43]	"DDI is mentioned to be (…) not serious [as a reason to turn off an alert]" [Inferred: the alert is not sufficiently relevant]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With clinicians' priority for very at risk situations
[44]	"DDIs that should be suppressed because of low incidence of adverse events"	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With clinicians' priority for very at risk situations
[44]	"Too low dose limits"	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[45]	"Clinicians (…) faced with a long list of them for each patient."	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[45]	"Clinicians reported that they faced situations in which CRs could not be removed and therefore continued to appear."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are redundant	No feature to turn off a specific alert for a specific patient, the alert still re-appears
[46]	"Ordering new medications satisfied the intent of the reminders but did not resolve them because they were not yet included in the logic."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are redundant	Clinically relevant solutions not accepted

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[46]	"Reminders did not always apply given the context of a particular patient."; "The third barrier to following the advice of a clinical reminder was inapplicability to the specific situational context. For example, a recommendation to begin Highly Active Anti-Retroviral Therapy (HAART) was not followed because the patient had experienced multiple intolerances to the medication in the past."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With patient clinical case
[46]	"Reminders did not always match local practice"	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With expertise/ward's habits
[47]	"VA's alert system did not match his/her (pharmacist) mental model of how an alert system should be designed (...) sometimes significant order checks really aren't significant (…) Now critical interactions (…) they are so many that are significant."	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[42; 47]	"I see it does say "active" though. Technically, the [old] medication [order] isn't "active" because I just changed them to "discontinued" [47] "Some meds should have been discontinued but were 'active', leading to extra alerts" [42]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With data considered
[48]	"When the physician used CPOE in the office area, a medication alert appeared saying 'duplicate order', which indicated that the patient was getting iron from both VA and non-VA sources (*i.e.,* over-the-counter or non-VA pharmacy). Upon seeing this alert, the physician stated, "That's not true to my knowledge. The patient doesn't like to take it; I doubt he's taking it [from a non-VA source]."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With data considered
[49]	"Some medication alerts may not be supported by pharmacy data. One pharmacist stated,"Sometimes, a doctor will call me about an interaction. I check in Micromedex®. [Micromedex® is a well-respected medication interaction database.10] Sometimes, there is no interaction shown in Micromedex® [for that alert]. I tell the doctor, 'I don't know. There is no information on it in the [Micromedex®] database.'"	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With pharmaceutical knowledge
[49]	"Low alert signal to noise ratio: numerousness of alerts"	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[42; 49]	"Alert system does not distinguish between true allergies and bothersome, but non serious, side effects"[49]"System does not specifically target true allergies: Prescribers wanted alerts to distinguish between: allergy to exact medication versus drug class, serious reactions versus bothersome side effects"[42]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With data considered
[42; 49]	"Alert conflicts with VAMC practices that are in place and/or standard medication practices." [49] "Alert conflicts with common medication practices" [42]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With expertise/ward's habits
[42; 49]	"Repeated alerts within the same encounter or over multiple encounters for a given patient"[49]; "alerts were sometimes excessively redundant. (...) The observer noted: we have now seen the same alert 4 times in the last 10 min or less." [49] "Redundant alerts within a given patient encounter. Often triggered by renewing interacting medication pairs. *E.g.,* Same alert appears a 3rd time"[42]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are redundant	Many alerts appear several times during the same decision making process
[54]	"There are too many [alerts]."	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[56]	"Most prescribers believed that most (alerts) were redundant."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are redundant	Many alerts appear very frequently
[56]	"Most prescribers believed (…) that they received too many alerts …"	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[57]	"Alerts (…) repeatedly encountered"	Compatibility	The alerts signal-to-noise ratio is low	Alerts are redundant	Many alerts appear very frequently
[58]	"Within-class interactions typically reflect an out-of-date medication list – such as an antibiotic interacting with another antibiotic – rather than a true interaction."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With data considered
[58]	"Alerts are often fired for drug combinations that conform to clinical guidelines and are recommended by specialist colleagues. Examples include aspirin with angiotensin converting enzyme inhibitors in patients with heart disease or diabetes."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With good practices
[58]	"Appropriate polypharmacy is not acknowledged" "In psychiatric care, mood stabilizers are often used intentionally in combination or "augmentation" therapy with antidepressants."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With expertise/ward's habits
[58]	"Pregnancy alerts would be more useful (…) if they were suppressed for male patients and women of non-child-bearing age."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With patient clinical case

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[58]	"Since most pediatric drugs are used off-label, pediatricians find it difficult to interpret the validity of many alerts."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With patient clinical case
[58]	"According to clinicians, the sensitivity of alerts was often set too high while the specificity was too low."	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[59]	"excess alerts - *e.g.*,, asthma and opiate, warfarin and paracetamol"	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[59]	""Too low triggering threshold with drug interaction alerts""	Compatibility	The alerts signal-to-noise ratio is low	No detail	/
[59]	"They also proposed some new alerts or eCDS functions (...) Patient-tailored threshold values, different from guideline-based" [rephrasing: alerts' thresholds are not patient-tailored]	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With patient clinical case
[60]	"The system should distinguish between orders specified as "now" and those specified as "future" or "standing" and not consider them to be duplicates "It prompts you either way. I mean, you specifically made it standing or future.""	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	With data considered
[60]	"However, even some of these users acknowledged that "pop-up" alerts can be very annoying, especially when they are "not right" for some reason."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are irrelevant	Other (no detail)
[60]	"Repetitive alerts are both annoying and unnecessary."	Compatibility	The alerts signal-to-noise ratio is low	Alerts are redundant	Many alerts appear very frequently
[24]	"Some express doubts on whether the system has up-to-date information (for instance for weight-based drug dosage alerts in pediatrics or drug interactions in cardiology where they often introduce new drugs).	Compatibility	Transparency	No transparent about the way it works	No information about up-to-dateness of alert's rules
[25]	"The reason for this is that the registration of an allergy is based on the ATC code, and the same drug can be registered under several codes if the drug has various indications." "The older version of the CPOE system only warns the user if the same drug code appears twice and not if the same drug (registered under different ATC codes) appears twice."	Compatibility	Transparency	No transparency on used data	Incomplete mapping
[28]	"Decision support was mostly invisible"	Compatibility	Transparency	No transparency on the way it works	Other
[42]	"Ambiguity about alert management; need for closed-loop feedback; *E.g.*,, Phys: "I want [the alert system] to have user control. I am not confident it's checking all the interactions that I want it to check.""	Compatibility	Transparency	No transparency on the way it works	Other
[42]	"System capabilities and limitations are ambiguous"	Compatibility	Transparency	No transparency on the way it works	Other
[42]	"No alerts for free text medications" "Free-text entry (*e.g.*,, 'OK') may not be effective (in the overide justification logic)"	Compatibility	Transparency	No transparency on used data	Every available data is not used to trigger the alert
[42]	Prescribers unaware that they could turn off some alerts"	Guidance	Prompting	/	/
[46]	"Users were uncertain how long the reminders would be turned off for each dialog option."	Compatibility	Transparency	No transparency on the way it works	Other
[47]	"A specific medication alert did not appear, even though the pharmacist was expecting it to come up for a patient. (...) Pharmacist: "The patient is 85yrs old. It's not good if he is on a full dose of aspirin and Plavix. These are both anti-platelets [medications]..." [later, while ordering medications for this patient:] An order check appears [after the] pharmacist change[s] [the] aspirin dose. [Alert] says duplicate drug class - hydrocodone. This was the only order check alert listed in the window. Pharmacist: "This is because the patient is on hydrocodone. It pulled-up that one, but not the Plavix [alert] which is interesting."" "Programmer and prescriber mental model mismatches were evident when a prescriber expected an alert to appear, but the system did not display that alert: (…) physician [is] ordering [renewing] naproxen… No order check [alerts] appear."	Compatibility	Transparency	No transparency on the way it works	Other

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[42; 49]	"Alert system does not adequately reveal its capabilities/limitations to the prescriber; full functionality of the alert system is ambiguous." [49] "System capabilities and limitations are ambiguous: Prescribers confused about whether the system could evaluate non-formulary, non-VA medications" [42]	Compatibility	Transparency	No transparency on used data	Every available data is not used to trigger the alert
[52]	"Three subjects expressed their need for better understanding of the dose calculation by the system. (...) users would not be able to "validate" the system's reasoning without resorting to calculation that is more complicated" [rephrasing: users are not sure that the system based its recommendation on the same assumptions they would have made]	Compatibility	Transparency	No transparency on the way it works	Other
[60]	"There are instances where users expect the system to "be aware" of and utilize patient information that exists in the database."	Compatibility	Transparency	No transparency on used data	Every available data is not used to trigger the alert
[61]	"No catalog exists explaining the levels of alert severity" (system 3)	Compatibility	Transparency	No transparency on the way it works	No information on the severity scale
[61]	"No catalog is visible for the user that explains in detail the levels of severity" (System 1)	Compatibility	Transparency	No transparency on the way it works	No information on the severity scale
[28]	"Inconsistent error prevention - able to prevent errors due to medication dose and frequency range, but not due to drug formulation (which is visible and changeable) or selection of incorrect patient or medication."	Consistency	/	/	/
[47]	"[Non VA system X] took up to 2-3 minutes to do the [medication] order check; (…) I haven't noticed much delay [with the VA system]"	Consistency	/	/	/
[42]	"Language of alerts is difficult for prescriber to interpret: *E.g.,* Alert says remote order checking unavailable. (See Figure 1.) Phys: "I do not know what that means.""	Error management	Quality of error messages	Weak content of the information within the alert	/
[45]	"However, there was no means to cancel the CR without losing the data already inputted for the previous CRs."	Explicit control	Explicit user actions	"Cancel" delete also data previously entered in CR	/
[46]	"There is no way to "undo" an action."	Explicit control	User control	No way to undo an action	/
[42]	"Difficult to distinguish different alert types (*e.g.,* duplicate drug versus duplicate drug class alerts)"	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the type of the alert
[42; 49]	"Overuse of pop-ups, other, non-medication related pop-ups contribute to prescriber alert desensitization" [49]"Overuse of other, non-alert pop-up windows may contribute to desensitization" [42]	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the type of message
[57]	"Cost and health maintenance alerts that were intrusive to the workflow"	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the type of message
[58]	"There was insufficient discrimination between alerts of varying severity."	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the severity of the alert
[61]	"No shapes or icons are used to convey alert priority." (System 3)	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the severity of the alert
[61]	"Alerts are color-coded by severity, not by the type of alert" (System 1)	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the type of the alert
[61]	"Alerts are color-coded by severity, not by type" (System 3)	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the type of the alert
[61]	"alerts are not grouped according to severity"	Guidance	Grouping/distinction	No distinction by the location	Too much distinction by the location of same severity alerts

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[61]	"All alerts are labeled with 'Warning'"	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the severity of the alert
[61]	"Color coding within the alert is not utilized in any way"	Guidance	Grouping/distinction	No distinction by the format	Other (no information)
[61]	"Neither color, signal words, nor shapes or icons are used to signify the alert's priority; all alerts possess the same visual characteristics and appear as drugs orders are placed" (System 2)	Guidance	Grouping/distinction	No distinction by the format	No distinction based on the severity of the alert
[44]	"I did not look through the new screen, and then I hit the button and suddenly it was gone"	Guidance	Immediate feedback	No feedback to inform the user that (s)he has skip an alert	/
[42]	"Lack of spacing between alert text"	Guidance	Legibility	Not sufficient inter-line space	/
[42; 49]	"According to participants, alert displays were problematic since much of the alert text was in all capital letters"[49] "Stylistic interface features support or hinder alert effectiveness: all capital letters" [42]	Guidance	Legibility	Font in capital letters	/
[53]	"The most commonly cited design concerns were (…) small size of the patient data box (3 subjects)"	Guidance	Legibility	Small size of elements	/
[24]	"Some alerts are out of the visual focus region when using the system"	Guidance	Prompting	Alert far from the center of the screen	/
[26]	"Unclear information or guidance" in the messages	Guidance	Prompting	/	/
[42]	"Salience: Alert visibility and distinction: Prescriber wanted more visual emphasis on high risk alerts" [rephrasing: no sufficient visual emphasis on high risk alerts]	Guidance	Prompting	Highlight deficiency	About serious alerts
[44]	"The interviews suggest that both unclear alert texts and texts read incompletely play a role."	Guidance	Prompting	/	/
[47]	"The programmers' mental model, as reflected in the system image, did not adequately match prescribers' mental models (…): Physician (MD) orders [VA] aspirin - 162 mg. An order check [alert] appears. Says duplicate drug order. Non-VA ASPIRIN. [Alert] mentions 325mg....MD is looking at it also and [appears] confused"	Guidance	Prompting	/	/
[51]	"participants commonly made specific clinical suggestions, such as clarifying the wording of recommendations, cautions, and data table notation" [rephrasing: it is not clear enough]	Guidance	Prompting	/	/
[53]	"The most commonly cited design concerns were (…) lack of salience of the clinical effect (3 subjects)."	Guidance	Prompting	Highlight deficiency	About clinical effect
[53]	"Two areas of minor criticism were the layout of buttons and checkboxes" [rephrasing: they are not salient enough to be seen]	Guidance	Prompting	Highlight deficiency	About buttons and checkboxes
[53]	"Many [users] missed a question regarding data in the alert that should reduce the level of clinical concern ('attenuating information')" [rephrasing: this information is missed because of the organization of the information in the alert: it is outside the center of the alert]	Guidance	Prompting	Significant information far from the center of the alert	/
[57]	"Prescribers reported that alerts presenting during medication order entry were often (…) difficult to interpret in content and purpose"	Guidance	Prompting	/	/
[59]	"Reminders' texts are sometimes too strict in the short version. If you don't move the cursor over the text and see the whole reminder, the wording doesn't work.'"	Guidance	Prompting	/	/
[59]	"Reminders' position on the left side of the screen"	Guidance	Prompting	Alert far from the center of the screen	/
[60]	"It was somewhat surprising that users did not always seem to understand how to use and manage the alerts effectively."[rephrasing: alert's management is not intuitive]	Guidance	Prompting	Other	/
[62]	"All alerts include an acknowledgement comment field, only some alerts are marked as requiring acknowledgement. Even those alerts which require acknowledgement only require that the acknowledgement button be pressed, not that a comment be left, although clinicians may have misinterpreted the directive."	Guidance	Prompting	Unclear instructions for alert justifications	/
[45]	"When defaulting past the cover sheet, feedback for the presence of due CRs is signified by a question mark icon in the upper right corner of the display. We observed three providers misinterpret this question mark to indicate that the patient had no CRs due, when in actuality it meant the system was still evaluating data to determine which CRs were due."	Significance of codes	Non intuitive icons	/	/

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[51]	"Several users in the simulation-based testing did not notice the arrows under the clinical recommendations or did not realize they provided additional, more detailed information about the basic recommendation when clicked on."	Significance of codes	Non intuitive icons	/	/
[26]	"Use of abbreviations and expressions that were not understandable by physicians, confusing terminology in labeling of buttons"	Significance of codes	Non intuitive wording	/	/
[43]	"One surgical resident did not understand the text of the sequence-dependent alerts well, considered the administration of drugs was out of the control of physicians, and thought these alerts therefore irrelevant"	Significance of codes	Non intuitive wording	/	/
[50]	"Two participants misinterpreted the meaning of "When" to represent the last time the current patient received the intervention instead of the frequency the intervention is due for all patients."	Significance of codes	Non intuitive wording	/	/
[51]	"A user thought that the appearance of the "stamp" window implied that the patient had a chronic pain problem or diagnosis. In actuality, the "stamp" indicated that the patient had a scheduled appointment within a 5-day window and that ATHENA-OT had recommendations available should the provider consider OT for that patient."	Significance of codes	Non intuitive wording	/	/
[60]	"Yeah, you see it, but you have to do something to even find out what it means. "" [Inference: passive alerts are not sufficiently informational to support a quick triage]	Significance of codes	Non intuitive wording	/	/
[56]	"Every doctor reported that the alerts contained too much text and should be shortened."	Workload	Concision	/	/
[57]	"Prescribers reported that alerts presenting during medication order entry were often long (…) to interpret in content and purpose" " difficulty with (…) the length of the text."	Workload	Concision	/	/
[58]	"Pregnancy alerts would be more useful if they showed the category of alert (A, B, C, etc.) rather than narrative information"	Workload	Concision	/	/
[24]	"Visualization of drug-drug interaction alerts where one drug has interactions with several others" [Inference: too much information in the alert]	Workload	Information density	Several alerts in a single window	/
[42]	"Extraneous information decreases alert value: Prescribers wanted a brief description of the problem: *e.g.,* "it needs to be 10 words or less""	Workload	Information density	Too many information of different kinds in the window	/
[42; 47; 49]	"Multiple alerts were grouped together in one pop-up window." [49] "Prescriber confusion when multiple alerts were presented [in the same screen]" [42] "Multiple alerts presented in one pop-up" [42] ""There are times when there are multiple flags [order checks] in the same box [pop-up alerts]""[47]	Workload	Information density	Several alerts in a single window	/
[51]	""It is hard to use the tool when sitting with a patient because it is in paragraph form. It would be better if factoids or outlines and standardized approaches are numbered or outlined and in lists.""	Workload	Information density	Alert content is displayed in a one-paragraph format	/
[51]	"The page is too convoluted. When there are 10 different things on the screen, providers aren't going to read any of it."	Workload	Information density	Too many information of different kinds in the window	/
[52]	"The institutional guideline for heparin administration, however was embedded in the same alert containing the calculated dose, triggered later in the process when the planning stage of the order is generally completed"	Workload	Information density	Too many information of different kinds in the window	/
[42]	"Time needed to resolve alerts: Justification requirement often viewed as time burden" [rephrasing: justify the irrelevance of an alert requires several actions that take time]	Workload	Minimal action	Data entry	Text entry is required to justified the irrelevance of an alert
[42]	"(…) scrolling were problematic" [rephrasing: user needs to scroll to see the whole information]	Workload	Minimal action	Display	Seeing all information requires vertical scrolling
[44]	"Necessity for scrolling down the whole alert text to find the conclusion"	Workload	Minimal action	Display	Seeing all information requires vertical scrolling

	Usability flaw as described in the papers	Category	Sub-category 1	Sub-category 2	Sub-category 3
[45]	"Five nurses and two providers were observed to skip all or some of the reminders and explained that this was because they perceived that they did not have enough time to "satisfy" the reminders by entering data."; "Completing the CRs creates "double documentation" burdens for some providers (…) as they generally keep track of this information without the CRs (*e.g.*,, in a health maintenance list within the progress note)."[Inference: satisfy the reminders requires time]	Workload	Minimal action	Data entry	Text entry is required to justified the irrelevance of an alert
[46]	"At all sites, at least one provider never satisfied reminders that were not clinically relevant, which required data entry such as when a patient received a vaccine at another hospital."	Workload	Minimal action	Data entry	Text entry is required to justified the irrelevance of an alert
[50]	"In the current design, this data [summary of patient information] is only visible by switching tabs when completing CRs."	Workload	Minimal action	Display	Information across several tabs
[53]	"Some had difficulty identifying the patient's risk factors for the interaction. Successfully answering this question required clicking on the Risk Factors tab to reveal the data (see Figure 3)." "The most commonly cited design concerns were the unnecessary use of tabs (7 subjects)"	Workload	Minimal action	Display	Information across several tabs
[59]	""Reminders' texts are sometimes too strict in the short version. If you don't move the cursor over the text and see the whole reminder, the wording doesn't work.""	Workload	Minimal action	Display	The long version alert must be opened to get sufficient information
[61]	"The user must scroll through the screen to see all alerts"	Workload	Minimal action	Display	Seeing all information requires vertical scrolling

Links between usability flaws, usage problems and negative outcomes in the work system.

The related reference list is presented in chapter 3.

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[22] [29]	"Alert does not provide clear information on relative risk of harm for a given patient" "Although some alerts are categorized by risk, (e.g., some are marked as "significant" and others are "critical"), this notation was not always sufficient for prescribers" ; "Risk rating/degree of risk: Prescribers wanted a more clear indication of risk and suggested that alerts display a quantitative rating" [Inferred: there is no clear risk information in the alert]	Compatibility	Alert content issues	"If there's more than one [alert in a popup window], I don't read through them all, honestly"	Behavioral	Voluntarily ignoring alerts	"Ordering decisions were hampered"; "alerts (…) impeded medication ordering processes.";	Medication management process issues
[22] [29]	"Alert does not provide clear information on relative risk of harm for a given patient" "Although some alerts are categorized by risk, (e.g., some are marked as "significant" and others are "critical"), this notation was not always sufficient for prescribers" ; "Risk rating/degree of risk: Prescribers wanted a more clear indication of risk and suggested that alerts display a quantitative rating" [Inferred: there is no clear risk information in the alert]	Compatibility	Alert content issues	"It's gotten to the point that I don't hardly look at significant (interactions) anymore"	Behavioral	Voluntarily ignoring alerts	"Ordering decisions were hampered"; "alerts (…) impeded medication ordering processes.";	Medication management process issues
[22] [29]	"Alert is not evidence-based, does not provide a reference to evidence that does exist, and/or the actual or perceived level of evidence is low." "The alerts themselves do not present the evidence nor do they provide links to any supporting documentation" "Quality/strength of alert evidence; guidance on actions to take: Prescribers wanted alert to provide: references or links to evidence (...) some alerts were not evidence-based" . "The actual or perceived level of evidence is low"; "unclear if the warnings were "evidence-based""; "extend of evidence is fair, stemming mainly from a few case reports"	Compatibility	Alert content issues	"This lack of information led to prescriber cynicism."	Emotional	Cynicism		
[27]	"erroneous system messaging [about how to solve the problem]"	Compatibility	Alert content issues	"Resorted to trial-and-error behavior exemplified by the extra mouse clicks and keystrokes they needed for locating and executing the right action in response to the message"	Behavioral	Increased workload	"Impaired the ordering efficiency"	Medication management process issues
[27]	"erroneous system messaging [about how to solve the problem]"	Compatibility	Alert content issues	"Physicians were lost »	Behavioral	Users lost	"Impaired the ordering efficiency"	Medication management process issues
[27]	"erroneous system messaging [about how to solve the problem]"	Compatibility	Alert content issues	"Precluding the physicians to understand the problem that generated the alert or how to solve the problem."	Cognitive	Difficulties to understand the alert	"Impaired the ordering efficiency"	Medication management process issues
[29]	"The alert did not provide essential patient information for the prescriber, even though it existed elsewhere in the EHR. For example, decision-making for some drug interaction alerts (e.g., amiloride and lisinopril, which can cause hyperkalemia) depend on patient labs (e.g., potassium). This missing, or more accurately, 'hidden' patient data triggered varying responses."	Compatibility	Alert content issues	"Spent time searching for information in the EHR afterwards to validate their decision."	Behavioral	Increased workload		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[29]	"The alert did not provide essential patient information for the prescriber, even though it existed elsewhere in the EHR. For example, decision-making for some drug interaction alerts (e.g., amiloride and lisinopril, which can cause hyperkalemia) depend on patient labs (e.g., potassium). This missing, or more accurately, 'hidden' patient data triggered varying responses."	Compatibility	Alert content issues	"Others overrode the alert"	Behavioral	Voluntarily ignoring alerts		
[29]	"The alert did not provide essential patient information for the prescriber, even though it existed elsewhere in the EHR. For example, decision-making for some drug interaction alerts (e.g., amiloride and lisinopril, which can cause hyperkalemia) depend on patient labs (e.g., potassium). This missing, or more accurately, 'hidden' patient data triggered varying responses."	Compatibility	Alert content issues	"Made assumptions about the patient history"; "Some prescribers relied solely on their memory of the patient profile"	Cognitive	Increased memory load		
[36]	"Another in-clinic user stated that the system would be more helpful if it gave more detailed information on "how to switch or discontinue drugs.""	Compatibility	Alert content issues	"The system would be more helpful" [currently not sufficiently helpful]	Attitudinal	Questioning usefulness		
[37]	"There were cases in which clinicians reached a dead end within the CR system, with no reasonable option to proceed." "We observed instances in which none of the available options to satisfy the CR applied to the patient or situation."	Compatibility	Alert content issues	"The existence of dead-end scenarios is a barrier to the effective use of the reminders since it impedes the ability to satisfy them"	Behavioral	Ineffective use		
[37]	"There were cases in which clinicians reached a dead end within the CR system, with no reasonable option to proceed." "We observed instances in which none of the available options to satisfy the CR applied to the patient or situation."	Compatibility	Alert content issues	"Additionally, we observed one nurse and one provider become "stuck" trying to satisfy a CR because the appropriate option was not available to them."	Behavioral	Users lost		
[37]	"Options within the dialogue box of the CRs do not always match the patient's response or there is not an appropriate option for indicating why the provider has decided not to order a test, for example."	Compatibility	Alert content issues	"Finally, two providers reported using workarounds to satisfy reminders because of the inflexibility of the dialogue options". "Provider arbitrarily selected a date to satisfy the reminder"; "The provider had to leave the reminder unsatisfied."	Behavioral	Workarounds		
[29]	"Alerts cannot be pulled up later, as needed, hindering alert resolution" ; "Sometimes, prescribers wanted a way to retrieve an alert that had been displayed, but the alert system did not support this function."	Compatibility	Alert features issues	"Prescribers sometimes forgot what alert(s) appeared "	Cognitive	Information involuntarily missed	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[37]	"The use of CRs automatically generates text that is added to the progress note, but that text is not integrated with the template information and is generally added to the bottom of the note."	Compatibility	Alert features issues	"This integration of the clinical topic addressed by the reminder into the template aligns it more directly into the nurse's workflow, thus facilitating its effective use." [currently not effective use]	Behavioral	Ineffective use		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[10]	"CPOE provides feedback on drug allergies, but only after medications are ordered."	Compatibility	Alert presentation issues	"Some house staff ignored allergy notices (…) and, most important, post hoc timing of allergy information."	Behavioral	Voluntarily ignoring alerts	"House staff claimed post hoc alerts unintentionally encourage house staff to rely on pharmacists for drug allergy checks, implicitly shifting responsibility to pharmacists."	Workflow issues
[24]	"They said that pop-up alerts particularly were annoying or unhelpful if they popped up "too early" in the encounter, or on the wrong screen."	Compatibility	Alert presentation issues	"They said that pop-up alerts particularly were (…) unhelpful" ; ""The alerts were most likely to be helpful if they presented when the users were entering orders or were otherwise at the point of making a decision about the issue in question or closely related issues."	Attitudinal	Questioning usefulness		
[24]	"Some acknowledged they were unlikely to respond, or perhaps even be aware of alerts, unless they were intrusive." [Inference: alerts not sufficiently intrusive in this case]	Compatibility	Alert presentation issues	"Some acknowledged they were unlikely to respond, or perhaps even be aware of alerts, unless they were intrusive." [missing an alert]	Cognitive	Information involuntarily missed		
[24]	"They said that pop-up alerts particularly were annoying or unhelpful if they popped up "too early" in the encounter, or on the wrong screen."	Compatibility	Alert presentation issues	"Partly because of the disruption of the thought process"	Cognitive	Users interrupted		
[24]	"However, even some of these users acknowledged that "pop-up" alerts can be very annoying, ..."	Compatibility	Alert presentation issues	""Pop-up" alerts can be very annoying"; ""Well, I think anything that keeps recurring to me is [annoying].”"	Emotional	Annoyance		
[24]	"They said that pop-up alerts particularly were annoying or unhelpful if they popped up "too early" in the encounter, or on the wrong screen."	Compatibility	Alert presentation issues	""Pop-up" alerts can be very annoying"; ""Well, I think anything that keeps recurring to me is [annoying].”"	Emotional	Annoyance		
[25]	"The institutional guideline for heparin administration (…) triggered later in the process when the planing stage of the order is generally completed"	Compatibility	Alert presentation issues	"Six subjects computed, estimated or used a heuristic to get the dose amount at some point before the system-calculated dose presentation"	Behavioral	Workarounds	"Consequently they did not derive all the speed and accuracy benefit and did not reduce their cognitive effort the feature was in part designed to."	Technology effectiveness issues
[28]	"Intrusive alerts presented at the wrong time in the workflow"	Compatibility	Alert presentation issues	"Several prescribers described rapidly overriding these alert types once they recognized that they had seen the alert before. One prescriber noted that she had "memorized the location of the override button” for these situations."	Behavioral	Voluntarily ignoring alerts		
[28]	"Intrusive alerts presented at the wrong time in the workflow"	Compatibility	Alert presentation issues	"These alerts that caused frustrations"	Emotional	Frustration		
[29]	"Technology lags/down- times" "Even 10-15 second computer delays between the order and subsequent alert"	Compatibility	Alert presentation issues	"There were several cases where inadequate alert design (…) disrupted their workflow."	Cognitive	Users interrupted	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[29]	"Technology lags/down- times" "Even 10-15 second computer delays between the order and subsequent alert"	Compatibility	Alert presentation issues	"View as “very troublesome”"	Emotional	Annoyance		
[29]	"Technology lags/down- times" "Even 10-15 second computer delays between the order and subsequent alert"	Compatibility	Alert presentation issues	"Place prescribers under pressure"	Emotional	Stress		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	Category	Sub-cat.	Excerpt	Category	Sub-cat.	Excerpt	Category
[35]	"Too many alerts or alerts at an inappropriate time: "Now we get alerts when we go to charting, which in my workflow is the last step. It's after the patient's gone. Now I get warned they've got some drug interaction. Great!"	Compatibility	Alert presentation issues	"There were numerous complaints about getting too many alerts or alerts at an inappropriate time"	Attitudinal	Negative feelings		
[35]	"Too many alerts or alerts at an inappropriate time: "Now we get alerts when we go to charting, which in my workflow is the last step. It's after the patient's gone. Now I get warned they've got some drug interaction. Great!"	Compatibility	Alert presentation issues	"Or alerts at an inappropriate time" [too late, once the patient is out, the information is missed]	Cognitive	Information involuntarily missed		
[37]	"The CRs appear on this cover sheet but are delayed in loading and displaying (Site 3 reported an average delay of 8 seconds for the CRs to load)."	Compatibility	Alert presentation issues	"We observed providers and nurses use strategies such as clicking on the progress notes tab before the reminders have displayed (…) or report setting a default tab to bypass the cover sheet, thereby lessening the visibility of the reminders"	Behavioral	Workarounds		
[32]	""You do not get the warning again, and there is no button to get it [the second alert] back""	Compatibility	Alerts features issues	"You do not get the warning again" [missed information]	Cognitive	Information involuntarily missed		
[28]	"Prescribers reported that alerts presenting during medication order entry were often (…) difficult to interpret in content and purpose"	Workload	Concision	"Several prescribers described rapidly overriding these alert types once they recognized that they had seen the alert before. One prescriber noted that she had "memorized the location of the override button" for these situations."	Behavioral	Voluntarily ignoring alerts		
[28]	"Prescribers reported that alerts presenting during medication order entry were often (…) difficult to interpret in content and purpose"	Workload	Concision	"Difficult to interpret in content and purpose"; "Some reported difficulty with (…) the length of the text."	Cognitive	Difficulties to understand the alert	"Slowed down their work"	Medication management process issues
[34]	"Pregnancy alerts would be more useful if they showed the category of alert (A, B, C, etc.) rather than narrative information…"	Workload	Concision	"There are too many things popping up at me"; "(alert fatigue)"	Attitudinal	Alert fatigue		
[34]	"Pregnancy alerts would be more useful if they showed the category of alert (A, B, C, etc.) rather than narrative information…"	Workload	Concision	"It's just crying wolf"; "Difficult to interpret the validity of many alerts"	Attitudinal	Questioning the validity		
[34]	"Pregnancy alerts would be more useful if they showed the category of alert (A, B, C, etc.) rather than narrative information…"	Workload	Concision	"Many clinicians ignore potentially serious alerts"; "Once you realize that most of the information is useless or superfluous or not relevant, you stop looking at it."	Behavioral	Voluntarily ignoring alerts		
[34]	"Pregnancy alerts would be more useful if they showed the category of alert (A, B, C, etc.) rather than narrative information…"	Workload	Concision	"Trivial alerts often interrupt workflow."	Cognitive	Users interrupted		
[40]	"Every doctor reported that the alerts contained too much text and should be shortened."	Workload	Concision	"If you spend time to read it"	Behavioral	Increased workload		
[40]	"Every doctor reported that the alerts contained too much text and should be shortened."	Workload	Concision	""It pops up so often which can be a very bad thing because you're dismissing it so often that you develop this sort of mechanism""; Most doctors either admitted to not reading the warnings"; "it's so much easier to click that button"	Behavioral	Voluntarily ignoring alerts		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[37]	"However, there was no means to cancel the CR without losing the data already inputted for the previous CRs." [inference: the system does not act as the user asks]	Explicit control	Explicit user actions	[inference: user cannot perform the action as he intends to]	Behavioral	Users lost	"Introduces the possibility of losing data previously inputted."	Technology effectiveness issues
[37]	"However, there was no means to cancel the CR without losing the data already inputted for the previous CRs." [inference: the system does not act as the user asks]	Explicit control	Explicit user actions	"A workaround is to select each CR individually from the list rather than using the Next button to navigate through a sequence of CRs."	Behavioral	Workarounds		
[29]	"Interface, which did not adequately support all prescriber types."	Adaptability	Flexibility	"Prescribers unaware that they could turn off some alerts"; "Prescribers' ability to act on alerts was impeded by the alert interface"	Behavioral	Ineffective use	"We call the pharmacist and spend time looking up information.";" physicians often come and ask [the pharmacists] about an alert triggered by the combination of amiodarone and simvastatin"; "physicians and nurse practitioners found it helpful, and sometimes necessary, to have real-time, face-to-face communication with clinical pharmacists to be able to interpret and resolve alerts"	Workflow issues
[29]	"Interface, which did not adequately support all prescriber types."	Adaptability	Flexibility	"Prescribers were sometimes unsure (…) what an alert was attempting to convey"; "Language of alerts is difficult for prescriber to interpret "I do not know what that means."	Cognitive	Difficulties to understand the alert	"We call the pharmacist and spend time looking up information.";" physicians often come and ask [the pharmacists] about an alert triggered by the combination of amiodarone and simvastatin"; "physicians and nurse practitioners found it helpful, and sometimes necessary, to have real-time, face-to-face communication with clinical pharmacists to be able to interpret and resolve alerts"	Workflow issues
[22] [29]	"Overuse of Pop-ups, other, non-medication related pop-ups contribute to prescriber alert desensitization" "- Overuse of other, non-alert pop-up windows may contribute to desensitization"	Guidance	Grouping/distinction	"Prescriber desensitization"; "Even prescribers with a very positive view of the alert system showed signs of desensitization"	Attitudinal	Alert fatigue		
[28]	"Cost and health maintenance alerts that were intrusive to the workflow"	Guidance	Grouping/distinction	"Viewed as (…) most annoying"	Emotional	Annoyance		
[29]	"Difficult to distinguish different alert types (e.g., duplicate drug versus duplicate drug class alerts)"	Guidance	Grouping/distinction	"Difficult to distinguish different alert types (e.g., duplicate drug versus duplicate drug class alerts)"	Cognitive	Difficulties to understand the alert		
[32]	"I did not look through the new screen, and then I hit the button and suddenly it was gone"	Guidance	Immediate feedback	« The physician may overlook the second alert, thinking that the override button for the (first) alert has not worked properly"; "thereby unintentionally override the second alert."	Cognitive	Information involuntarily missed	"wrong selection (followed incorrect dose recommendation)"	Patient safety issues

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[22] [26] [29]	"Multiple alerts were grouped together in one pop-up window." "Prescriber confusion when multiple alerts were presented [in the same screen]" "Multiple alerts presented in one pop-up" ""There are times when there are multiple flags [order checks] in the same box [pop-up alerts]""	Workload	Information density	"If there's more than one [alert in a popup window], I don't read through them all, honestly"	Behavioral	Voluntarily ignoring alerts	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[22] [26] [29]	"Multiple alerts were grouped together in one pop-up window." "Prescriber confusion when multiple alerts were presented [in the same screen]" "Multiple alerts presented in one pop-up" ""There are times when there are multiple flags [order checks] in the same box [pop-up alerts]""	Workload	Information density	"It's gotten to the point that I don't hardly look at significant (interactions) anymore"	Behavioral	Voluntarily ignoring alerts	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[22] [26] [29]	"Multiple alerts were grouped together in one pop-up window." "Prescriber confusion when multiple alerts were presented [in the same screen]" "Multiple alerts presented in one pop-up" ""There are times when there are multiple flags [order checks] in the same box [pop-up alerts]""	Workload	Information density	"Prescribers were sometimes unsure (…) what an alert was attempting to convey"; "Language of alerts is difficult for prescriber to interpret "I do not know what that means."	Cognitive	Difficulties to understand the alert	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[22] [26] [29]	"Multiple alerts were grouped together in one pop-up window." "Prescriber confusion when multiple alerts were presented [in the same screen]" "Multiple alerts presented in one pop-up" ""There are times when there are multiple flags [order checks] in the same box [pop-up alerts]""	Workload	Information density	"Prescribers were sometimes unsure (…) what an alert was attempting to convey"; "Language of alerts is difficult for prescriber to interpret "I do not know what that means."	Cognitive	Difficulties to understand the alert	"We call the pharmacist and spend time looking up information.";" physicians often come and ask [the pharmacists] about an alert triggered by the combination of amiodarone and simvastatin"; "physicians and nurse practitioners found it helpful, and sometimes necessary, to have real-time, face-to-face communication with clinical pharmacists to be able to interpret and resolve alerts"	Workflow issues
[22] [26] [29]	"Multiple alerts were grouped together in one pop-up window." "Prescriber confusion when multiple alerts were presented [in the same screen]" "Multiple alerts presented in one pop-up" ""There are times when there are multiple flags [order checks] in the same box [pop-up alerts]""	Workload	Information density	"Many prescribers overwhelmed by (…) multiple alerts presented in one pop-up"	Emotional	Annoyance		
[22] [26] [29]	"Multiple alerts were grouped together in one pop-up window." "Prescriber confusion when multiple alerts were presented [in the same screen]" "Multiple alerts presented in one pop-up" ""There are times when there are multiple flags [order checks] in the same box [pop-up alerts]""	Workload	Information density	"Reading them is ugly"	Emotional	Ugly experience		
[25]	"The institutional guideline for heparin administration, however was embedded in the same alert containing the calculated dose, triggered later in the process when the planning stage of the order is generally completed"	Workload	Information density	"Two subjects did not use the decision support feature because they misidentified the alert as a general guideline reminder and did not notice the dose calculations embedded in text."	Behavioral	System not used at all		
[25]	"The institutional guideline for heparin administration, however was embedded in the same alert containing the calculated dose, triggered later in the process when the planning stage of the order is generally completed"	Workload	Information density	"They misidentified the alert as a general guideline reminder and did not notice the dose calculations embedded in text."	Cognitive	Difficulties to identify alerts' components		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[29]	"Extraneous information decreases alert value: Prescribers wanted a brief description of the problem: e.g. "it needs to be 10 words or less"" [Inference: the problem description is too long]	Workload	Information density	"Sometimes, when alerts contain a lot of text, prescribers rely instead on their "own clinical judgment"	Behavioral	System not used at all	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[33]	"Visualization of drug-drug interaction alerts where one drug has interactions with several others" [Inference: too much information in the alert]	Workload	Information density	"They already find it difficult to understand visualization"	Cognitive	Difficulties to understand the alert		
[36]	""It is hard to use the tool when sitting with a patient because it is in paragraph form. It would be better if factoids or outlines and standardized approaches are numbered or outlined and in lists.""	Workload	Information density	"It is hard to use the tool when sitting with a patient"	Attitudinal	Negative feelings		
[36]	"The page is too convoluted. When there are 10 different things on the screen, providers aren't going to read any of it."	Workload	Information density	"Providers aren't going to read any of it."	Behavioral	Voluntarily ignoring alerts		
[22] [29]	"According to participants, alert displays were problematic since much of the alert text was in all capital letters" "Stylistic interface features support or hinder alert effectiveness: all capital letters …"	Guidance	Legibility	[inference: hard to read the text]	Cognitive	Difficulties to understand the alert	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[22] [29]	"According to participants, alert displays were problematic since much of the alert text was in all capital letters" "Stylistic interface features support or hinder alert effectiveness: all capital letters …"	Guidance	Legibility	"Reading them is ugly"	Emotional	Ugly experience		
[29]	"Lack of spacing between alert text"	Guidance	Legibility	[inference: hard to read the text]	Cognitive	Difficulties to understand the alert	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[22]	"Low alert signal to noise ratio: numerousness of alerts"	Compatibility	Low signal-to-noise ratio	"Prescriber desensitization"; "Even prescribers with a very positive view of the alert system showed signs of desensitization"	Attitudinal	Alert fatigue		
[22]	"Some medication alerts may not be supported by pharmacy data. One pharmacist stated,"Sometimes, a doctor will call me about an interaction. I check in Micromedex®. [Micromedex® is a well-respected medication interaction database] Sometimes, there is no interaction shown in Micromedex® [for that alert]. I tell the doctor, 'I don't know. There is no information on it in the [Micromedex®] database.'"	Compatibility	Low signal-to-noise ratio	"Some prescribers were skeptical of alerts because it was unclear if the warnings were "evidence-based"; also, prescribers sometimes questioned the quality and strength of evidence."	Attitudinal	Questioning the validity		
[22] [29]	"Alert conflicts with VAMC practices that are in place and/or standard medication practices." "Alert conflicts with common medication practices"	Compatibility	Low signal-to-noise ratio	"Influence prescribers' perceptions of the credibility and trustworthiness of the alert system as a whole"	Attitudinal	Questioning the validity		
[22] [29]	"Repeated alerts within the same encounter or over multiple encounters for a given patient"; "alerts were sometimes excessively redundant. (...) The observer noted: we have now seen the same alert 4 times in the last 10 min or less." "Redundant alerts within a given patient encounter. Often triggered by renewing interacting medication pairs. E.g., Same alert appears a 3rd time"	Compatibility	Low signal-to-noise ratio	"If there's more than one [alert in a popup window], I don't read through them all, honestly"	Behavioral	Voluntarily ignoring alerts	"Ordering decisions were hampered"; "alerts (…) impeded medication ordering processes.";	Medication management process issues

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[22] [29]	"Repeated alerts within the same encounter or over multiple encounters for a given patient"; "alerts were sometimes excessively redundant. (...) The observer noted: we have now seen the same alert 4 times in the last 10 min or less." "Redundant alerts within a given patient encounter. Often triggered by renewing interacting medication pairs. E.g., Same alert appears a 3rd time"	Compatibility	Low signal-to-noise ratio	"It's gotten to the point that I don't hardly look at significant (interactions) anymore"	Behavioral	Voluntarily ignoring alerts	"Ordering decisions were hampered"; "alerts (…) impeded medication ordering processes.";	Medication management process issues
[22] [29]	"Repeated alerts within the same encounter or over multiple encounters for a given patient"; "alerts were sometimes excessively redundant. (...) The observer noted: we have now seen the same alert 4 times in the last 10 min or less." "Redundant alerts within a given patient encounter. Often triggered by renewing interacting medication pairs. E.g., Same alert appears a 3rd time"	Compatibility	Low signal-to-noise ratio	"NP gestures to the screen, "See – three times!""	Emotional	Annoyance		
[22] [29]	"Repeated alerts within the same encounter or over multiple encounters for a given patient"; "alerts were sometimes excessively redundant. (...) The observer noted: we have now seen the same alert 4 times in the last 10 min or less." "Redundant alerts within a given patient encounter. Often triggered by renewing interacting medication pairs. E.g., Same alert appears a 3rd time"	Compatibility	Low signal-to-noise ratio	"A nurse practitioner voiced her frustration"	Emotional	Frustration		
[24]	"Repetitive alerts are both annoying and unnecessary."	Compatibility	Low signal-to-noise ratio	"Users complained vociferously »	Attitudinal	Negative feelings		
[24]	"Repetitive alerts are both annoying and unnecessary."	Compatibility	Low signal-to-noise ratio	""Pop-up" alerts can be very annoying"; ""Well, I think anything that keeps recurring to me is [annoying].""	Emotional	Annoyance		
[24]	"However, even some of these users acknowledged that "pop-up" alerts can be very annoying, especially when they are "not right" for some reason."	Compatibility	Low signal-to-noise ratio	""Pop-up" alerts can be very annoying"; ""Well, I think anything that keeps recurring to me is [annoying].""	Emotional	Annoyance		
[26]	"VA's alert system did not match his/her (pharmacist) mental model of how an alert system should be designed (...) sometimes significant order checks really aren't significant (…) Now critical interactions (…) they are so many that are significant."	Compatibility	Low signal-to-noise ratio	"Sometimes significant order checks really aren't significant (…) Now critical [interactions] (…) they are so many that are significant"	Attitudinal	Questioning triggering/sorting model		
[26]	"VA's alert system did not match his/her (pharmacist) mental model of how an alert system should be designed (...) sometimes significant order checks really aren't significant (…) Now critical interactions (…) they are so many that are significant."	Compatibility	Low signal-to-noise ratio	"It's gotten to the point that I don't hardly look at significant (interactions) anymore"	Behavioral	Voluntarily ignoring alerts	"Ordering decisions were hampered"; "alerts (…) impeded medication ordering processes.";	Medication management process issues
[26] [29]	"I see it does say "active" though. Technically, the [old] medication [order] isn't "active" because I just changed them to "discontinued" "Some meds should have been discontinued but were 'active', leading to extra alerts"	Compatibility	Low signal-to-noise ratio	""Did it accept my changes?"" [doubts]	Attitudinal	Questioning the behavior		
[26] [29]	"I see it does say "active" though. Technically, the [old] medication [order] isn't "active" because I just changed them to "discontinued" "Some meds should have been discontinued but were 'active', leading to extra alerts"	Compatibility	Low signal-to-noise ratio	"This order check [appearing now for a second time] might freak someone out"	Emotional	Annoyance		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[28]	"Alerts (…) repeatedly encountered"	Compatibility	Low signal-to-noise ratio	"Several prescribers described rapidly overriding these alert types once they recognized that they had seen the alert before. One prescriber noted that she had "memorized the location of the override button" for these situations."	Behavioral	Voluntarily ignoring alerts		
[29]	"Same alerts appear for a patient across one or more med renewals; E.g., Alert says 'Previous adverse reaction to antidepressants.'…Phys types in override reason, 'Has been on venlafaxine for 5 yrs now.'"	Compatibility	Low signal-to-noise ratio	"Phys types in override reason, 'Has been on venlafaxine for 5 yrs now.'"	Behavioral	Voluntarily ignoring alerts		
[29]	"Number of alerts is problematic"; "Potential problems are overdetected or underdetected"	Compatibility	Low signal-to-noise ratio	"Many prescribers overwhelmed by (…) multiple alerts presented in one pop-up"	Emotional	Annoyance		
[31]	"When the physician used CPOE in the office area, a medication alert appeared saying 'duplicate order', which indicated that the patient was getting iron from both VA and non-VA sources (*i.e.* over-the-counter or non-VA pharmacy). Upon seeing this alert, the physician stated, "That's not true to my knowledge. The patient doesn't like to take it; I doubt he's taking it [from a non-VA source]."	Compatibility	Low signal-to-noise ratio	"That's not true to my knowledge. The patient doesn't like to take it; I doubt he's taking it [from a non-VA source]."	Attitudinal	Questioning the validity		
[31]	"When the physician used CPOE in the office area, a medication alert appeared saying 'duplicate order', which indicated that the patient was getting iron from both VA and non-VA sources (*i.e.* over-the-counter or non-VA pharmacy). Upon seeing this alert, the physician stated, "That's not true to my knowledge. The patient doesn't like to take it; I doubt he's taking it [from a non-VA source]."	Compatibility	Low signal-to-noise ratio	"The physician proceeded with the order and went back to the exam room to talk with the patient, where s/he confirmed that the patient was not obtaining iron from an outside source."	Behavioral	Increased workload	"The physician (…) talk[s] with the patient, where s/he confirmed that the patient was not obtaining iron from an outside source."	Workflow issues
[32]	"too low dose limits"	Compatibility	Low signal-to-noise ratio	"Could further add to alert fatigue"; "Remarks suggesting alert fatigue"	Attitudinal	Alert fatigue		
[32]	Furthermore, residents referred to other aspects that could further add to alert fatigue: (…) DDIs that should be suppressed because of low incidence of adverse events"	Compatibility	Low signal-to-noise ratio	"Could further add to alert fatigue"; "Remarks suggesting alert fatigue"	Attitudinal	Alert fatigue		
[32]	Furthermore, residents referred to other aspects that could further add to alert fatigue: (…) DDIs that should be suppressed because of low incidence of adverse events"	Compatibility	Low signal-to-noise ratio	"If I have to consider every DDI, than I am busy with it, all day, and that is not my job."	Behavioral	Increased workload		
[32]	Furthermore, residents referred to other aspects that could further add to alert fatigue: (…) DDIs that should be suppressed because of low incidence of adverse events"	Compatibility	Low signal-to-noise ratio	"I am often inclined to rapidly click them away; you get all those DDIs reported; you simply skip them."	Behavioral	Voluntarily ignoring alerts		
[32]	There are so many drug–drug interactions that are irrelevant, that I am often inclined to rapidly click them away [resident in internal medicine].	Compatibility	Low signal-to-noise ratio	"I am often inclined to rapidly click them away; you get all those DDIs reported; you simply skip them."	Behavioral	Voluntarily ignoring alerts		
[32]	Furthermore, residents referred to other aspects that could further add to alert fatigue: (…) DDIs that should be suppressed because of low incidence of adverse events"	Compatibility	Low signal-to-noise ratio	"All those drug–drug interactions and all those things you get reported drive you mad."	Emotional	Annoyance		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[33]	"Reminder alerts that should be given the last day of hospitalization (*i.e.*, bacteriological tests), a day the system cannot forecast. This leads to an alert every day."	Compatibility	Low signal-to-noise ratio	"This leads to an alert everyday and therefore to (…) alert fatigue."	Attitudinal	Alert fatigue		
[33]	"Alerts once entered in the system can be outdated. The processes how to keep them up-to-date is not yet implemented (*i.e.* for patients who were carrier of methicillin-resistant staphylococcus aureus (MRSA) and who are now readmitted to the hospital)."	Compatibility	Low signal-to-noise ratio	"Some brought up issues make the CPOE less utile"	Attitudinal	Questioning usefulness		
[33]	"Reminder alerts that should be given the last day of hospitalization (*i.e.*, bacteriological tests), a day the system cannot forecast. This leads to an alert every day."	Compatibility	Low signal-to-noise ratio	[inference: alerts voluntarily ignored]	Behavioral	Voluntarily ignoring the alert	"low compliance rate"	Lowering alerting system expected effectiveness
[34]	"According to clinicians, the sensitivity of alerts was often set too high while the specificity was too low."	Compatibility	Low signal-to-noise ratio	"There are too many things popping up at me"; "(alert fatigue)"	Attitudinal	Alert fatigue		
[34]	"Alerts are often fired for drug combinations that conform to clinical guidelines and are recommended by specialist colleagues. Examples include aspirin with angiotensin converting enzyme inhibitors in patients with heart disease or diabetes."	Compatibility	Low signal-to-noise ratio	"There are too many things popping up at me"; "(alert fatigue)"	Attitudinal	Alert fatigue		
[34]	"Appropriate polypharmacy is not acknowledged" "In psychiatric care, mood stabilizers are often used intentionally in combination or "augmentation" therapy with antidepressants."	Compatibility	Low signal-to-noise ratio	"There are too many things popping up at me"; "(alert fatigue)"	Attitudinal	Alert fatigue		
[34]	"Pregnancy alerts would be more useful (…) if they were suppressed for male patients and women of non-child-bearing age."	Compatibility	Low signal-to-noise ratio	"There are too many things popping up at me"; "(alert fatigue)"	Attitudinal	Alert fatigue		
[34]	"Within-class interactions typically reflect an out-of-date medication list – such as an antibiotic interacting with another antibiotic – rather than a true interaction."	Compatibility	Low signal-to-noise ratio	"There are too many things popping up at me"; "(alert fatigue)"	Attitudinal	Alert fatigue		
[34]	"According to clinicians, the sensitivity of alerts was often set too high while the specificity was too low."	Compatibility	Low signal-to-noise ratio	"It's just crying wolf"; "Difficult to interpret the validity of many alerts"	Attitudinal	Questioning the validity		
[34]	"Alerts are often fired for drug combinations that conform to clinical guidelines and are recommended by specialist colleagues. Examples include aspirin with angiotensin converting enzyme inhibitors in patients with heart disease or diabetes."	Compatibility	Low signal-to-noise ratio	"It's just crying wolf"; "Difficult to interpret the validity of many alerts"	Attitudinal	Questioning the validity		
[34]	"Appropriate polypharmacy is not acknowledged" "In psychiatric care, mood stabilizers are often used intentionally in combination or "augmentation" therapy with antidepressants."	Compatibility	Low signal-to-noise ratio	"It's just crying wolf"; "Difficult to interpret the validity of many alerts"	Attitudinal	Questioning the validity		
[34]	"Since most pediatric drugs are used off-label, pediatricians find it difficult to interpret the validity of many alerts."	Compatibility	Low signal-to-noise ratio	"It's just crying wolf"; "Difficult to interpret the validity of many alerts"	Attitudinal	Questioning the validity		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[34]	"Within-class interactions typically reflect an out-of-date medication list – such as an antibiotic interacting with another antibiotic – rather than a true interaction."	Compatibility	Low signal-to-noise ratio	"It's just crying wolf"; "Difficult to interpret the validity of many alerts"	Attitudinal	Questioning the validity		
[34]	"According to clinicians, the sensitivity of alerts was often set too high while the specificity was too low."	Compatibility	Low signal-to-noise ratio	"Many clinicians ignore potentially serious alerts"; "Once you realize that most of the information is useless or superfluous or not relevant, you stop looking at it."	Behavioral	Voluntarily ignoring alerts		
[34]	"Alerts are often fired for drug combinations that conform to clinical guidelines and are recommended by specialist colleagues. Examples include aspirin with angiotensin converting enzyme inhibitors in patients with heart disease or diabetes."	Compatibility	Low signal-to-noise ratio	"Many clinicians ignore potentially serious alerts"; "Once you realize that most of the information is useless or superfluous or not relevant, you stop looking at it."	Behavioral	Voluntarily ignoring alerts		
[34]	"Appropriate polypharmacy is not acknowledged" "In psychiatric care, mood stabilizers are often used intentionally in combination or "augmentation" therapy with antidepressants."	Compatibility	Low signal-to-noise ratio	"Many clinicians ignore potentially serious alerts"; "Once you realize that most of the information is useless or superfluous or not relevant, you stop looking at it."	Behavioral	Voluntarily ignoring alerts		
[34]	"Pregnancy alerts would be more useful (…) if they were suppressed for male patients and women of non-child-bearing age."	Compatibility	Low signal-to-noise ratio	"Many clinicians ignore potentially serious alerts"; "Once you realize that most of the information is useless or superfluous or not relevant, you stop looking at it."	Behavioral	Voluntarily ignoring alerts		
[34]	"Within-class interactions typically reflect an out-of-date medication list – such as an antibiotic interacting with another antibiotic – rather than a true interaction."	Compatibility	Low signal-to-noise ratio	"Many clinicians ignore potentially serious alerts"; "Once you realize that most of the information is useless or superfluous or not relevant, you stop looking at it."	Behavioral	Voluntarily ignoring alerts		
[34]	"According to clinicians, the sensitivity of alerts was often set too high while the specificity was too low."	Compatibility	Low signal-to-noise ratio	"Trivial alerts often interrupt workflow."	Cognitive	Users interrupted		
[34]	"Alerts are often fired for drug combinations that conform to clinical guidelines and are recommended by specialist colleagues. Examples include aspirin with angiotensin converting enzyme inhibitors in patients with heart disease or diabetes."	Compatibility	Low signal-to-noise ratio	"Trivial alerts often interrupt workflow."	Cognitive	Users interrupted		
[34]	"Appropriate polypharmacy is not acknowledged" "In psychiatric care, mood stabilizers are often used intentionally in combination or "augmentation" therapy with antidepressants."	Compatibility	Low signal-to-noise ratio	"Trivial alerts often interrupt workflow."	Cognitive	Users interrupted		
[34]	"Pregnancy alerts would be more useful (…) if they were suppressed for male patients and women of non-child-bearing age."	Compatibility	Low signal-to-noise ratio	"Trivial alerts often interrupt workflow."	Cognitive	Users interrupted		
[34]	"Within-class interactions typically reflect an out-of-date medication list – such as an antibiotic interacting with another antibiotic – rather than a true interaction."	Compatibility	Low signal-to-noise ratio	"Trivial alerts often interrupt workflow."	Cognitive	Users interrupted		
[35]	"There are too many [alerts]."	Compatibility	Low signal-to-noise ratio	"Physicians ignoring alerts because there are too many"	Behavioral	Voluntarily ignoring alerts		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	Category	Sub-cat.	Excerpt	Category	Sub-cat.	Excerpt	Category
[37]	"Clinicians (…) faced with a long list of them for each patient."	Compatibility	Low signal-to-noise ratio	"Become desensitized to the CRs"	Attitudinal	Alert fatigue		
[37]	"Clinicians reported that they faced situations in which CRs could not be removed and therefore continued to appear."	Compatibility	Low signal-to-noise ratio	"CRs could not be removed"	Behavioral	Users lost	"CRs (…) continued to appear"	Technology effectiveness issues
[38]	"Reminders did not always apply given the context of a particular patient." ; "The third barrier to following the advice of a clinical reminder was inapplicability to the specific situational context. For example, a recommendation to begin Highly Active Anti-Retroviral Therapy (HAART) was not followed because the patient had experienced multiple intolerances to the medication in the past."	Compatibility	Low signal-to-noise ratio	"Clinical reminders were not [always] used"	Behavioral	System not used at all		
[40]	"Most prescribers believed that most (alerts) were redundant."	Compatibility	Low signal-to-noise ratio	"Some doctors recognized that they had become desensitized to the alerts."	Attitudinal	Alert fatigue		
[40]	"Most prescribers believed (…) that they received too many alerts …"	Compatibility	Low signal-to-noise ratio	""It pops up so often which can be a very bad thing because you're dismissing it so often that you develop this sort of mechanism""; Most doctors either admitted to not reading the warnings"; "it's so much easier to click that button"	Behavioral	Voluntarily ignoring alerts		
[41]	"excess alerts - e.g., asthma and opiate, warfarin and paracetamol"	Compatibility	Low signal-to-noise ratio	"The physician reported that specific features of the system (…) were hindering the use"	Behavioral	Ineffective use		
[41]	""Too low triggering threshold with drug interaction alerts""	Compatibility	Low signal-to-noise ratio	"I never bother to read them"	Behavioral	Voluntarily ignoring alerts		
[41]	"excess alerts - e.g., asthma and opiate, warfarin and paracetamol"	Compatibility	Low signal-to-noise ratio	"Irritating drug interaction and contraindication alerts"	Emotional	Annoyance		
[22] [29]	"(…) scrolling were problematic" [Inference: user needs to scroll to see the whole information]; "poor screen display: alert display does not support alert resolution and/or prescriber workflow" "they have to manipulate the alert to see all of the information"	Workload	Minimal action	"There were several cases where inadequate alert design prompted prescribers to take extra steps in the medication ordering process. For example, prescribers sometimes had to manipulate the alert to see all of the information"	Behavioral	Increased workload	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues
[22] [29]	"(…) scrolling were problematic" [Inference: user needs to scroll to see the whole information]; "poor screen display: alert display does not support alert resolution and/or prescriber workflow" "they have to manipulate the alert to see all of the information"	Workload	Minimal action	"There were several cases where inadequate alert design (…) disrupted their workflow."	Cognitive	Users interrupted	"Hindering alert resolution"; "Hinder alert effectiveness"	Technology effectiveness issues

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	Category	Sub-cat.	Excerpt	Category	Sub-cat.	Excerpt	Category
[22] [29]	"(…) scrolling were problematic" [Inference: user needs to scroll to see the whole information]; "poor screen display: alert display does not support alert resolution and/or prescriber workflow" "they have to manipulate the alert to see all of the information"	Workload	Minimal action	"View as "very troublesome""	Emotional	Annoyance		
[29]	"Time needed to resolve alerts: Justification requirement often viewed as time burden" [Inference: justify an the irrelevance of an alert requires several actions that take time]	Workload	Minimal action	"Justification requirement often viewed as time burden"	Attitudinal	Negative feelings		
[30]	"Some had difficulty identifying the patient's risk factors for the interaction. Successfully answering this question required clicking on the Risk Factors tab to reveal the data (see Figure 3)." "The most commonly cited design concerns were the unnecessary use of tabs (7 subjects)"	Workload	Minimal action	"Had difficulty identifying the patient's risk factors for the interaction"	Cognitive	Difficulties to understand the alert	"Successfully answering this question (…) required clicking on the Risk Factors tab" [error in answering a clinical question]	Patient safety issues
[32]	"Necessity for scrolling down the whole alert text to find the conclusion"	Workload	Minimal action	"Could further add to alert fatigue"; "Remarks suggesting alert fatigue"	Attitudinal	Alert fatigue		
[32]	"Necessity for scrolling down the whole alert text to find the conclusion"	Workload	Minimal action	"Necessity for scrolling down" [increased workload]	Behavioral	Increased workload		
[32]	"Necessity for scrolling down the whole alert text to find the conclusion"	Workload	Minimal action	« The physician may overlook the second alert, thinking that the override button for the (first) alert has not worked properly"; "thereby unintentionally override the second alert."	Cognitive	Information involuntarily missed	"wrong selection (followed incorrect dose recommendation)"	Patient safety issues
[32]	"Necessity for scrolling down the whole alert text to find the conclusion"	Workload	Minimal action	"Misinterpretation (…) high numbers of wrong or inapplicable rules and reasoning."	Cognitive	Misinterpretation of alerts' content	"wrong selection (followed incorrect dose recommendation)"	Patient safety issues
[37]	"Five nurses and two providers were observed to skip all or some of the reminders and explained that this was because they perceived that they did not have enough time to "satisfy" the reminders by entering data."; "Completing the CRs creates "double documentation" burdens for some providers (…) as they generally keep track of this information without the CRs (e.g., in a health maintenance list within the progress note)."[Inference: satisfy the reminders requires time]	Workload	Minimal action	"Completing the CRs creates "double documentation" burdens for some providers (…) as they generally keep track of this information without the CRs"	Behavioral	Increased workload		
[37]	"Five nurses and two providers were observed to skip all or some of the reminders and explained that this was because they perceived that they did not have enough time to "satisfy" the reminders by entering data."; "Completing the CRs creates "double documentation" burdens for some providers (…) as they generally keep track of this information without the CRs (e.g., in a health maintenance list within the progress note)."[Inference: satisfy the reminders requires time]	Workload	Minimal action	"Five nurses and two providers were observed to skip all or some of the reminders"	Behavioral	Voluntarily ignoring alerts		
[38]	"At all sites, at least one provider never satisfied reminders that were not clinically relevant, which required data entry such as when a patient received a vaccine at another hospital." [data entry cumbersome] "additional workload"	Workload	Minimal action	"Lack of time to follow documentation procedures within the CR noting why the reminder's advice was not followed."	Behavioral	Increased workload		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[38]	"At all sites, at least one provider never satisfied reminders that were not clinically relevant, which required data entry such as when a patient received a vaccine at another hospital." [data entry cumbersome] "additional workload"	Workload	Minimal action	"At least one provider at all eight sites conducted documentation activities, including satisfying clinical reminders (…), after the patient had left the room. (…) at least one provider conducted documentation activities after the clinic closed, normally within 36 hours. (…) case managers documented clinical reminders (…) after the physicians did not."	Behavioral	Workarounds		
[41]	"Reminders' texts are sometimes too strict in the short version. If you don't move the cursor over the text and see the whole reminder, the wording doesn't work.'"	Workload	Minimal action	"If you don't move the cursor over the text and see the whole reminder (…)"	Behavioral	Increased workload		
[41]	"Reminders' texts are sometimes too strict in the short version. If you don't move the cursor over the text and see the whole reminder, the wording doesn't work.'"	Workload	Minimal action	"(…) the wording doesn't work"	Cognitive	Difficulties to understand the alert		
[36]	"Several users in the simulation-based testing did not notice the arrows under the clinical recommendations or did not realize they provided additional, more detailed information about the basic recommendation when clicked on." [inference: the arrow is not intuitive enough]	Significance of codes	Non intuitive icons	"Several users did not notice the arrows under the clinical recommendations or did not realize they provided additional more detailed information about the basic recommendation when clicked on."	Cognitive	Difficulties to identify alerts' components		
[37]	"When defaulting past the cover sheet, feedback for the presence of due CRs is signified by a question mark icon in the upper right corner of the display. We observed three providers misinterpret this question mark to indicate that the patient had no CRs due, when in actuality it meant the system was still evaluating data to determine which CRs were due." [inference: non intruitive enough symbol]	Significance of codes	Non intuitive icons	"We observed three providers misinterpret this question mark"	Cognitive	Difficulties to identify alerts' components		
[24]	"Yeah, you see it, but you have to do something to even find out what it means. "" [Inference: passive alerts are not sufficiently informational to support a quick triage]	Significance of codes	Non intuitive wording	"Users said they would like more information available with less effort so they can evaluate and act on alerts and reminders more easily."	Behavioral	Increased workload		
[27]	"Use of abbreviations and expressions that were not understandable by physicians, confusing terminology in labelling of buttons"	Significance of codes	Non intuitive wording	"Precluding the physicians to understand the problem that generated the alert or how to solve the problem."	Cognitive	Difficulties to understand the alert	"Impaired the ordering efficiency"	Medication management process issues
[27]	"Use of abbreviations and expressions that were not understandable by physicians, confusing terminology in labelling of buttons"	Significance of codes	Non intuitive wording	"Physicians became frustrated"	Emotional	Frustration		
[36]	"A user thought that the appearance of the "stamp" window implied that the patient had a chronic pain problem or diagnosis. In actuality, the "stamp" indicated that the patient had a scheduled appointment within a 5-day window and that ATHENA-OT had recommendations available should the provider consider OT for that patient." [inference: stamp wording not intuitive enough]	Significance of codes	Non intuitive wording	"A user thought that the appearance of the "stamp" window implied that the patient had a chronic pain problem or diagnosis" [misinterpretation]	Cognitive	Misinterpretation of alerts' components		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[39]	"Two participants misinterpreted the meaning of "When" to represent the last time the current patient received the intervention instead of the frequency the intervention is due for all patients."	Significance of codes	Non intuitive wording	"Two participants misinterpreted the meaning of "When""	Cognitive	Misinterpretation of alerts' components		
[23]	All alerts include an acknowledgement comment field, only some alerts are marked as requiring acknowledgement. Even those alerts which require acknowledgement only require that the acknowledgement button be pressed, not that a comment be left, although clinicians may have misinterpreted the directive.	Guidance	Prompting	"although clinicians may have misinterpreted the directive."	Cognitive	Misinterpretation of alerts' content	"The absence of a specific order critique in this alert, combined with the required acknowledgement, may be related to the relatively high proportion of content free comments"	Technology effectiveness issues
[24]	"It was somewhat surprising that users did not always seem to understand how to use and manage the alerts effectively."[Inference: alert's management is not intuitive]	Guidance	Prompting	"Not always seem to understand how to use and manage the alerts effectively"	Behavioral	Ineffective use	"This resulted in some unnecessary repetition of alerts"	Technology effectiveness issues
[24]	"It was somewhat surprising that users did not always seem to understand how to use and manage the alerts effectively."[Inference: alert's management is not intuitive]	Guidance	Prompting	"This (…) contributed to user frustration."	Emotional	Frustration		
[26]	"The programmers' mental model, as reflected in the system image, did not adequately match prescribers' mental models (…): Physician (MD) orders [VA] aspirin - 162 mg. An order check [alert] appears. Says duplicate drug order. Non-VA ASPIRIN. [Alert] mentions 325mg....MD is looking at it also and [appears] confused"	Guidance	Prompting	"MD clicks through [the alert] [accepts order]" (accepts without understanding the alert)	Behavioral	Advice blindly followed	"MD goes back to the medication list. Aspirin is now listed both under VA list and non-VA medication list" [double order of aspirin]	Patient safety issues
[26]	"The programmers' mental model, as reflected in the system image, did not adequately match prescribers' mental models (…): Physician (MD) orders [VA] aspirin - 162 mg. An order check [alert] appears. Says duplicate drug order. Non-VA ASPIRIN. [Alert] mentions 325mg....MD is looking at it also and [appears] confused"	Guidance	Prompting	"This led to confusion during the prescribing process"; "MD to Observer: What's it going to do? Is it going to switch the patient to 325mg?"	Cognitive	Difficulties to understand the alert	"MD goes back to the medication list. Aspirin is now listed both under VA list and non-VA medication list" [double order of aspirin]	Patient safety issues
[27]	"unclear information or guidance" in the messages	Guidance	Prompting	"Resorted to trial-and-error behavior exemplified by the extra mouse clicks and keystrokes they needed for locating and executing the right action in response to the message"	Behavioral	Increased workload	"Impaired the ordering efficiency"	Medication management process issues
[27]	"unclear information or guidance" in the messages	Guidance	Prompting	"Physicians were lost »	Behavioral	Users lost	"Impaired the ordering efficiency"	Medication management process issues
[27]	"unclear information or guidance" in the messages	Guidance	Prompting	"Precluding the physicians to understand the problem that generated the alert or how to solve the problem."	Cognitive	Difficulties to understand the alert	"Impaired the ordering efficiency"	Medication management process issues
[28]	"Prescribers reported that alerts presenting during medication order entry were often (…) difficult to interpret in content and purpose"	Guidance	Prompting	"Several prescribers described rapidly overriding these alert types once they recognized that they had seen the alert before. One prescriber noted that she had "memorized the location of the override button" for these situations."	Behavioral	Voluntarily ignoring alerts		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[28]	"Prescribers reported that alerts presenting during medication order entry were often (…) difficult to interpret in content and purpose"	Guidance	Prompting	"Difficult to interpret in content and purpose"; "Some reported difficulty with (…) the length of the text."	Cognitive	Difficulties to understand the alert	"Slowed down their work"	Medication management process issues
[29]	"Prescribers unaware that they could turn off some alerts"	Guidance	Prompting	"Prescribers unaware that they could turn off some alerts"; "Prescribers' ability to act on alerts was impeded by the alert interface"	Behavioral	Ineffective use		
[29]	"Salience: Alert visibility and distinction: Prescriber wanted more visual emphasis on high risk alerts" [Inference: no sufficient visual emphasis on high risk alerts]	Guidance	Prompting	"Difficult to distinguish different alert types (e.g., duplicate drug versus duplicate drug class alerts)"	Cognitive	Difficulties to understand the alert		
[30]	"Two areas of minor criticism were the layout of buttons and checkboxes" ; "Some users did not see the tab or did not realize it was selectable."[Inference: they are not sallient enough to be seen]	Guidance	Prompting	"Had difficulty identifying the patient's risk factors for the interaction"	Cognitive	Difficulties to understand the alert	"Successfully answering this question (…) required clicking on the Risk Factors tab" [error in answering a clinical question]	Patient safety issues
[30]	"Many [users] missed a question regarding data in the alert that should reduce the level of clinical concern ('attenuating information')" [rephrasing: this information is missed because of the organization of the information in the alert: it is outside the center of the alert]	Guidance	Prompting	"Many missed a question regarding data in the alert that should reduce the level of clinical concern ('attenuating information')."	Cognitive	Information involuntarily missed	"the patient had been taking both medications for many years and most physicians would be reassured by this fact. However, only 58% of subjects made this connection." [error in understanding patient data]	Patient safety issues
[32]	"The interviews suggest that both unclear alert texts and texts read incompletely play a role."; alert presentation was unclear (two DDI alerts in one screen provoking oversight)	Guidance	Prompting	« The physician may overlook the second alert, thinking that the override button for the (first) alert has not worked properly"; "thereby unintentionally override the second alert."	Cognitive	Information involuntarily missed	"wrong selection (followed incorrect dose recommendation)"	Patient safety issues
[32]	"The interviews suggest that both unclear alert texts and texts read incompletely play a role."; alert presentation was unclear (two DDI alerts in one screen provoking oversight)	Guidance	Prompting	"Misinterpretation (…) high numbers of wrong or inapplicable rules and reasoning."	Cognitive	Misinterpretation of alerts' content	"wrong selection (followed incorrect dose recommendation)"	Patient safety issues
[33]	"Some alerts are out of the visual focus region when using the system"	Guidance	Prompting	[inference: alert unnoticed]	Cognitive	Information involuntarily missed	"Leading to low response levels to the alerts"	Technology effectiveness issues
[41]	"Reminders' texts are sometimes too strict in the short version. If you don't move the cursor over the text and see the whole reminder, the wording doesn't work.'"	Guidance	Prompting	"If you don't move the cursor over the text and see the whole reminder (…)"	Behavioral	Increased workload		
[41]	"Reminders' position on the left side of the screen" [Inference: position outside their focal visual field]	Guidance	Prompting	"The physician reported that specific features of the system (…) were hindering the use"	Behavioral	Ineffective use		
[41]	"Reminders' texts are sometimes too strict in the short version. If you don't move the cursor over the text and see the whole reminder, the wording doesn't work.'"	Guidance	Prompting	"(…) the wording doesn't work"	Cognitive	Difficulties to understand the alert		

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[22] [26] [29]	"Language of alerts is difficult for prescriber to interpret: E.g., Alert says remote order checking unavailable. (See Figure 1.) Phys: "I do not know what that means.""	Error management	Quality of error messages	"Prescribers were sometimes unsure (…) what an alert was attempting to convey"; "Language of alerts is difficult for prescriber to interpret "I do not know what that means."	Cognitive	Difficulties to understand the alert	"We call the pharmacist and spend time looking up information.";" physicians often come and ask [the pharmacists] about an alert triggered by the combination of amiodarone and simvastatin"; "physicians and nurse practitioners found it helpful, and sometimes necessary, to have real-time, face-to-face communication with clinical pharmacists to be able to interpret and resolve alerts"	Workflow issues
[10]	"The CPOE system does not display information available on other hospital systems. For example, only the pharmacy's computer provides drug interaction and lifetime limit warnings."	Compatibility	Tasks and control distribution issues	"The CPOE system does not display information available on other hospital systems. For example, only the pharmacy's computer provides drug interaction and lifetime limit warnings." [missing information]	Cognitive	Information involuntarily missed	"Pharmacists call house staff to clarify questionable orders"	Workflow issues
[29]	"Prescribers unsure if pharmacists review these (override justifications) or find them useful"	Compatibility	Tasks and control distribution issues	"Unsure if pharmacists review these [override justifications]"	Attitudinal	Questioning the behavior		
[37]	"CR system is currently insufficient for supporting transmission of reminder results from nursing intake to provider examination"	Compatibility	Tasks and control distribution issues	"Paper-based workarounds were employed to alert providers to positive findings from the screening, including printing the nursing note for the provider (Site 1), placing a post-it note or using handwritten abbreviations on the patient's routing form (Sites 2 and 4), and marking a check box next to the appropriate item on a "green card" that the patient then gave to the provider (Site 3)."	Behavioral	Workarounds		
[22] [29]	"Alert system does not adequately reveal its capabilities/limitations to the prescriber; full functionality of the alert system is ambiguous." "System capabilities and limitations are ambiguous: Prescribers confused about whether the system could evaluate non-formulary, non-VA medications"	Compatibility	Transparency issues	""I am not confident it's checking all the interactions that I want it to check.""; "No, I'm not clear. I'm not clear what the triggers are".	Attitudinal	Questioning triggering/sorting model		
[22] [29]	"Alert system does not adequately reveal its capabilities/limitations to the prescriber; full functionality of the alert system is ambiguous." "System capabilities and limitations are ambiguous: Prescribers confused about whether the system could evaluate non-formulary, non-VA medications"	Compatibility	Transparency issues	"Prescribers unaware that they could turn off some alerts"; "Prescribers' ability to act on alerts was impeded by the alert interface"	Behavioral	Ineffective use	"We call the pharmacist and spend time looking up information.";" physicians often come and ask [the pharmacists] about an alert triggered by the combination of amiodarone and simvastatin"; "physicians and nurse practitioners found it helpful, and sometimes necessary, to have real-time, face-to-face communication with clinical pharmacists to be able to interpret and resolve alerts"	Workflow issues

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[24]	"There are instances where users expect the system to "be aware" of and utilize patient information that exists in the database."	Compatibility	Transparency issues	"There are instances where users expect the system to "be aware" of and utilize patient information that exists in the database, to more accurately target eligible patients."	Attitudinal	Questioning triggering/sorting model		
[25]	"Three subjects expressed their need for better understanding of the dose calcultation by the system. (...) users would not be able to "validate" the system's reasonning without resorting to calculation that is more complicated" [Inference: users are not sure that the system based its recommendation on the same assumptions they would have made]	Compatibility	Transparency issues	"Subjects wanted to be sure that the system based its recommendation on the same assumptions they would have made"	Attitudinal	Questioning the behavior		
[26]	"A specific medication alert did not appear, even though the pharmacist was expecting it to come up for a patient. (...) Pharmacist: "The patient is 85yrs old. It's not good if he is on a full dose of aspirin and Plavix. These are both anti-platelets [medications]..." [later, while ordering medications for this patient:] An order check appears [after the] pharmacist change[s] [the] aspirin dose. [Alert] says duplicate drug class - hydrocodone. This was the only order check alert listed in the window. Pharmacist: "This is because the patient is on hydrocodone. It pulled-up that one, but not the Plavix [alert] which is interesting."" "Programmer and prescriber mental model mismatches were evident when a prescriber expected an alert to appear, but the system did not display that alert: (…) physician [is] ordering [renewing] naproxen… No order check [alerts] appear."	Compatibility	Transparency issues	« The pharmacist was expecting it to come up for a patient."	Attitudinal	Questioning triggering/sorting model		
[29]	"Ambiguity about alert management; need for closed-loop feedback; E.g., Phys: "I want [the alert system] to have user control. I am not confident it's checking all the interactions that I want it to check.""	Compatibility	Transparency issues	""I am not confident it's checking all the interactions that I want it to check.""; "No, I'm not clear. I'm not clear what the triggers are".	Attitudinal	Questioning triggering/sorting model		
[29]	"System capabilities and limitations are ambiguous"	Compatibility	Transparency issues	""I am not confident it's checking all the interactions that I want it to check.""; "No, I'm not clear. I'm not clear what the triggers are".	Attitudinal	Questioning triggering/sorting model		
[29]	"System capabilities and limitations are ambiguous"	Compatibility	Transparency issues	"Prescribers unaware that they could turn off some alerts"; "Prescribers' ability to act on alerts was impeded by the alert interface"	Behavioral	Ineffective use	"We call the pharmacist and spend time looking up information.";" physicians often come and ask [the pharmacists] about an alert triggered by the combination of amiodarone and simvastatin"; "physicians and nurse practitioners found it helpful, and sometimes necessary, to have real-time, face-to-face communication with clinical pharmacists to be able to interpret and resolve alerts"	Workflow issues

	Usability flaw			Usage problem			Negative outcome	
	Excerpt	**Category**	**Sub-cat.**	**Excerpt**	**Category**	**Sub-cat.**	**Excerpt**	**Category**
[33]	"some express doubts on whether the system has up-to-date information (for instance for weight-based drug dosage alerts in pediatrics or drug interactions in cardiology where they often introduce new drugs)."	Compatibility	Transparency issues	"Some express doubts on whether the system has up-to-date information "	Attitudinal	Questioning the validity		
[38]	"Users were uncertain how long the reminders would be turned off for each dialog option."	Compatibility	Transparency issues	"Users were uncertain how long the reminders would be turned off for each dialog option"	Attitudinal	Questioning the behavior		
[42]	"The reason for this is that the registration of an allergy is based on the ATC code, and the same drug can be registered under several codes if the drug has various indications." "The older version of the CPOE system only warns the user if the same drug code appears twice and not if the same drug (registered under different ATC codes) appears twice."[Inference: there would be no alert for a medication]	Compatibility	Transparency issues	[inference: users are unaware of gaps in the data]	Cognitive	Misinterpretation of alerts' content	"There are various examples of complex registration that lead to medication errors"	Patient safety issues

List of usability design principles extracted from the literature

The related reference list is presented in chapter 4.

	Usability design principle
[1]	Condensate and simplify the information even if there is some utility in having the complete back-up guideline as reference material.
[1]	Speed is everything: DS must not take too long to appear (if not, useless and decreasing satisfaction of use): avoid "screen flips"
[1]	Anticipate needs and deliver in real time: the application must anticipate users' (latent and more obvious) needs and bring information at the time they need it. Because of time pressure, users can ill afford to spend more time seeking bits of information.
[1]	Fit into the users' workflow: success of alerts depends on their integration with practices. Therefore understanding clinicians' workflow is critical
[14]	Managing the specificity, sensitivity and other performance characteristics of alerts is necessary to help prevent excessive alerting and is often possible by analyzing detailed logs generated by automatic system audits.
[14]	Increasing the variety of information sources that are available to the decision rules engines to access and consider is likely to greatly increase the specificity and credibility of clinical alerts in the future and increase the response of clinicians to potentially risky medication.
[14]	Alerts should be sensitive to clinical context by incorporating more patient-specific data into trigger rules,
[14]	Reasons for override may also be prompted routinely so that knowledge engineers trying to determine why some alerts are consistently ignored can review override reasons, and analyze them in conjunction with activity logs
[14]	Filtering alerts by increasing the specificity of trigger rules may help to decrease the number of interruptive messages with little evidentiary basis or clinical relevance or those that are redundant.
[14]	Drug interaction alerts should be primarily patient-specific by taking into account age, gender, body weight, allergies, mitigating circumstances, drug serum levels, renal function and comorbidity. For example, standard alerts related to abnormal renal function should be suppressed for patients on dialysis although it is a nontrivial task to ensure that relevant EHR data are complete and updated
[14]	Time intervals between interacting drugs should also be considered as earlier drugs might have been completely metabolized.
[14]	If a potential interaction did not result in problems for a specific patient in the past, physicians should be able to suppress the alert for subsequent dose adjustments to avoid redundant messages.
[14]	Systems should be able to, where appropriate, suppress alerts at the time of renewal of previously tolerated medication combinations for the same patient.
[14]	Error-producing conditions may exist in commercially available and institutionally developed databases and customization or periodic reviews are necessary. A committee of physicians that includes domain experts and pharmacists for drug-related alerts should periodically revise rules with a focus on frequently overridden alerts and suggest safe and effective ways for either suppressing alerts of low value or changing their presentation format.
[14]	Combining pharmacology and laboratory data into decision rules provides a powerful tool to guide initial drug choice (*i.e.,* drugs where there are laboratory-based indications and contraindications), drug dosing (renal or hepatic, blood level-guided adjustments), laboratory monitoring (laboratory signals of toxicity, baseline and ongoing monitoring), laboratory result interpretation (drug interfering with test) and for broader quality improvement (surveillance for unrecognized toxicity, monitoring clinician response delays).
[14]	Pharmacy systems receiving electronic prescriptions should also have their own automatic drug–drug and drug–allergy checking in addition to decision support built into the ordering system.
[14]	The networked systems, however, should share the same clinical context so that pharmacists can better evaluate the appropriateness of each prescription: a summary screen containing key patient information.
[14]	Messages that prompt for routine actions (*e.g.*,, periodic lab tests for patients with chronic conditions) should be offloaded from physician workflows entirely and redirected to support staff.
[14]	The algorithm used by the system should be made accessible on demand by a link to promote trust in the reasoning process.
[14]	Systems need to avoid the impression of a "black box" giving advice that cannot be subjectively evaluated: an explanation of medical logic, including formulas for calculating values, should be accessible on demand so that the justification for alerting is transparent and verifiable.
[14]	A link to further evidence may also be included as clinicians often contend that more information should be accessible.
[14]	In some instances, when an alternative medication or test is offered, a link (or a button) may be added that closes the dialog and populates appropriate fields in an open order form with the suggested values.
[14]	Tight integration with clinical workflows/clinical goals/clinical environments and presentation of relevant advice at the time and place of decision making: meet clinicians' expectations of flexibility, individuality of advice, and reliability.
[14]	carefully calibrate intrusiveness to be proportional to their level of importance
[14]	However, this option (interruptive alerts) should be reserved only for high severity warnings and used judiciously

	Usability design principle
[14]	The usefulness of concurrent alerts needs to be evaluated: those that do not absolutely contribute to improving the prescribing process should be suppressed or deemphasized; Those that remain need to be prioritized by severity; Advisory messages of lower importance should be displayed in a more subtle format to avoid excessive alerting: displayed in non-intrusive, asynchronous presentation formats. These can be text messages in sidebars that can be read without explicit acknowledgment or for the moment ignored
[14]	The usefulness of concurrent alerts needs to be evaluated: those that do not absolutely contribute to improving the prescribing process should be suppressed or deemphasized; Those that remain need to be prioritized by severity; Advisory messages of lower importance should be displayed in a more subtle format to avoid excessive alerting: displayed in non-intrusive, asynchronous presentation formats. These can be text messages in sidebars that can be read without explicit acknowledgment or for the moment ignored
[14]	Reminders to take certain actions at the present time or in the near future (*e.g.*,, schedule a mammogram) can take several forms. One approach is to simply add flags to patient names in lists. Short descriptive messages may also be placed in designated locations on the screen when an individual patient record is opened.
[14]	[Clinicians] need to rapidly receive advice and take an appropriate action at a convenient point in the workflow without extraneous effort or delay. the acceptable screen transition time is well under a second,
[14]	A broader strategy to avoid distrust in the relevance of decision support is to avoid recommendations that are controversial. Rather, advice should be given only for aspects of care in which there is little disagreement on appropriate management.
[14]	Systems should therefore formulate advisory messages in the manner of highlighting potential or actual problems that require attention and suggest therapeutic opportunities rather than imposing strict, inflexible and unsolicited dictates. However, merely giving an assessment without recommending an action and providing a convenient way to either carry out or disregard it is generally not an effective way to change behavior.
[14]	Physicians may strongly resist a suggestion not to carry out an action when an acceptable alternative is not offered.
[14]	Suggesting equally effective and appropriate alternative actions is complex and not always possible. As the actual indications may differ from those considered by the decision logic, contextual information from the record needs to be evident to support the relevance of the advice.
[14]	'Intermediate states" (or state variables) are inferences from primary data that can be conceptualized in clinically relevant terms and further used in decision logic. Clinicians often find these state variables to be more intuitive and convenient for reasoning than single data points. They may be monitored and automatically updated over time to reflect changes in laboratory results, medications, problems, procedures, passages of time and other data. More complex patient states (*e.g.*, "patient is on anticoagulation therapy") can be created with sophisticated data-driven derivations that trigger more extensive and more specific interventions.
[14]	Further, when clinical data are aggregated from multiple sources they may need to be "normalized" into a common representational format (common or converted units of measurement, reference range, etc.) and analogous data reconciled by identifying their "source of truth"
[14]	Contextual information should account for the relationships and correlates between clinically dependent data. For example, an intervention may suggest lowering a drug dose when kidney function worsens and prompt for corollary lab or other orders, show allergies, renal function, microbiology results, sensitivities and the unit in which the patient is located when ordering antibiotics and suggest the best and least expensive brand or a generic and its dose.
[11]	"Discontinue" or "D/C," should not be used to cancel an order that has not yet been completely entered—the one that triggered the alert
[11]	Rule that triggered the alert and medical consequence are briefly described and a link to detailed explanation (monograph) is attached.
[11]	Clinical context shows relevant values from the patient record with a link to access further details.
[11]	An alternative drug to the one being ordered may be suggested as a third option. An action link with the drug name may be placed on the dialog box separated by enough blank space from the accept-override button pair not to visually compete with the primary actions. Clicking the link should close the dialog box and open a standard ordering form with the appropriate fields prepopulated with new values.
[11]	Triggering medical logic needs to be apparent and outlined in a few words accompanied by a link to further evidence
[11]	The most common override reasons should be selectable from a list of no more than three or four with a single click as fast and convenient processing is essential for promoting use.
[11]	Many systems allow overriding in all instances but differentiate the level of required effort to override severe interactions by requiring a secondary confirmation action. The hard stop is designed as a persistent checkbox selection to discontinue the existing drug. A less restrictive option may allow deselecting the checkbox but still require a subsequent confirmation to override.
[11]	Rules that trigger alerts can also be filtered and prioritized to suppress low-severity warnings by using more sophisticated algorithms that integrate patient context and provider-specific data into the decision logic
[11]	Drug–drug interactions may be stratified
[11]	An effective filtering method is to add to the decision logic, along with general drug–drug interaction rules, additional data from the EHR and thus making the rules more patient-specific. For example, a system could automatically prioritize recommendations according to a multi-attribute utility model by combining patient and provider-specific data [4]. Age, gender, body weight, mitigating circumstances, drug serum levels, renal function and comorbidity [68] may modify the severity of expected interaction for that patient and the system then selects appropriate warning level.
[11]	Time intervals between interacting drugs should also be considered as earlier-prescribed drugs may have completely metabolized by the time a contraindicated drug is entered

	Usability design principle
[11]	Redundant alerts can be suppressed when dose adjustments are entered for a specific patient and at times when a previously tolerated medication combination for the same patient is renewed
[11]	The dialog box may be designed as a binary choice between (a) ordering the new drug while simultaneously discontinuing the existing drug of the interacting pair, and (b) canceling the new order
[11]	Physicians may be allowed to turn off individual alerts, with caveats, based on their practice, knowledge and comfort level
[11]	Clinicians may suppress alerts for medications that a patient had previously received and tolerated [74]. However, suppressing a drug–drug interaction alert after it has been overridden only once per patient, for example, was not favored by prescribers in one study and even less by pharmacists [60].
[11]	The first choice should close the dialog box, create a discontinue order for the existing drug and open a pre-populated entry from for the new one.
[11]	The second choice should close the dialog box and place the focus back on the drug ordering screen
[11]	Alerts need to clearly state that the existing order will be discontinued if the new one is finalized.
[11]	Clinicians choosing to override an alert may be asked to give a reason for not following the advice
[11]	Override reason selections can be made mandatory for the most critical alerts but otherwise optional
[11]	Override reasons in a selection lists have 1–2 words; lists should contain less than five items.
[11]	Buttons are labeled with unambiguous verbs [Order] and [Cancel].
[11]	Override is possible with one extra click (uncheck "Discontinue") for second tier alerts; not possible (if so designed) for critical alerts.
[11]	Reason for override is selectable by one click but not mandatory.
[11]	The alert box should include an immediately actionable item [4,83] as physicians may resist suggestions not to carry out an action when an alternative is not offered
[11]	Appropriate contextual information from the patient record should be made available on demand (*e.g.,*, via a link) as the actual indications for a patient may be different from those considered by the decision logic
[11]	reduce the number of disruptive alerts of low clinical value
[11]	degree of alert intrusiveness can be adjusted according to their level of importance, allowing only the most severe warnings to interrupt work
[11]	to limit unnecessary interruptions is to assign alerts to interaction severity categories, or "tiers," and to control how they are presented to clinicians. The most serious warnings still need an explicit response by a clinician but less important alerts are displayed less intrusively on the screen as messages not requiring any actions. Research evidence suggests that this approach may improve compliance rate for higher-severity alerts
[11]	Judicious use of interruptive alerts should be considered, reserving this option only for interactions of the highest severity
[11]	The relative priority of concurrent alerts needs to be evaluated and those that do not absolutely contribute to improving the prescribing process should be suppressed or shown as low-importance messages
[11]	Alerts with lower urgency should be clearly noticeable, placed near the order for which they were triggered (*i.e.,* not at the bottom of the screen) as spatial proximity of screen items visually implies their relatedness
[11]	Messages about possible interactions that are considered merely informational (*i.e.,* with the lowest severity rating) can be placed in regions on the screen that are not in the focused visual field of the clinician at the moment the order is entered. They can be in areas dedicated to warnings, in sidebars or in the main body section and expanded on demand [61].
[11]	These messages can also be aggregated and shown together in a single display to be reviewed all at once at a convenient point in the workflow such as at the end, during order signing. The messages can also be sorted and prioritized
[11]	Filtering of alerts means that rules triggering specific intervention modes (*e.g.,*, interruptive dialogs or non-intrusive messages) are modified not to activate when certain conditions apply.
[11]	clear response options with controls placed close to relevant text
[11]	Size conveys importance and hierarchy – dose and frequency are secondary to the drug name; labels are less important than content.
[11]	The two drug names are the most important information on the screen; they are therefore the most prominent screen artifacts to draw visual attention.
[11]	Brief instructions (monitor ECG daily) are included.
[11]	Ancillary order (ECG 12-lead) is included but not mandatory.
[11]	most important attribute is the severity level. It needs to be the most visually prominent item on the screen and clearly communicated by dedicated "code" words reserved for each level, such as "critical," "significant," "caution," "recommendation," or "note,"
[11]	Accurate suggestions of drug alternatives need to include dose and frequency but those may depend on clinical context.
[11]	The dialog box, at minimum, offers a way to continue ordering (*i.e.,* override the warning) or to cancel the order in progress (*i.e.,* accept the suggestion) by clicking respective buttons
[11]	The text of the supporting information should be visually distinct (*i.e.,* deemphasized) from the main message of the alert so that it can be easily ignored when not needed. It can be printed in smaller characters on the side or at the bottom of the alert box.

	Usability design principle
[16]	Knowledge developers should avoid over-alerting by improving specificity of alerts and by improving allergy data quality.
[16]	Analyses of override reasons should occur as part of system quality improvement efforts, and contribute to further reduction of non-essential alerts.
[16]	Organizations should customize duplicate checking to decrease the number of clinically insignificant alerts.
[16]	Vendors should implement knowledge management tools for their customers' use.
[16]	Healthcare organizations (of sufficient size) that purchase CPOE products should be able to review embedded drug information resources (*e.g.*,, drug- drug interactions, drug-lab interactions, etc.) to determine where errors are present (invariably some will be present), and to decide which specific alerting rules they want to implement.
[16]	Application vendors' tools should allow provider organizations to make local customizations to the knowledge base, and these customizations should survive vendor product upgrades.
[16]	Drug information knowledge base vendors should work with CPOE and pharmacy system vendors to implement knowledge management tools so that provider organizations can customize purchased drug information, and so customizations persist across version upgrades.
[16]	Once over-alerting is under control, systems should ask the clinician to provide a coded override reason whenever he or she overrides drug-allergy alert. The override reason should allow nurses and pharmacists to understand the rationale for the override.
[16]	advanced Checking of Drug–Disease Interactions and Contraindications: in situations in which the alerting system does not cover the domain completely, clinicians should be aware of what is not covered by the alerting system.
[16]	The way alerts are presented to providers should be improved in part through differential display based on the severity of the anticipated event
[16]	Different alerts should be presented depending on whether the patient has mild renal insufficiency, untreated uremia, or is actively undergoing dialysis.
[16]	Degree of severity should influence interruptive versus non-interruptive notification methods.
[16]	The alert screen should include a link to information describing institution-specific guidelines for restricted medications
[16]	The CDS UI should present information clearly and concisely, allow clinicians to act on alerts directly from the alert screen when possible, and then return clinicians to their previously intended workflows.
[16]	formulary decision support: CDS is not effective when a recommendation is controversial or not valued by clinicians. Screen recommendations with expert local clinicians first.
[16]	drug-drug interaction checking: Alerts should present the names of the interacting drugs, a brief (one-line) description of the inter-action, optional links to more detailed information, and a menu for potentially appropriate actions in response to the alert.
[16]	Links should exist to enable clinicians to review the evidence basis for automated drug information, both as bibliographic references and as text summaries of evidence.
[16]	advanced Guidance for Medication-associated Laboratory Testing: Third, ideally, the knowledge bases upon which monitoring recommendations are made should be evidence-based, with documentation of benefits.
[16]	Vendors should work to create concise and actionable alert messages (*e.g.*,, "starting amiodarone doubles previously stable serum digoxin levels" rather than "digoxin and amiodarone interact").
[17]	Adapting its behavior according to a subset of relevant actions taken by clinicians
[17]	Adapting its behavior to the evolution of the outcome at risk over time
[17]	Make the DS a team player
[17]	Provide an indication for all the professionals of the availability of information
[17]	Incorporate function to support team awareness about the alert management
[17]	Have the same display of information for all professionals
[17]	Require provider's documentation of the reason for not following suggestion
[17]	Give access upon request to extended information
[17]	Make the DS a clinicians' partner
[17]	Provide DS automatically as part of clinicians workflow
[17]	Deliver DS at the time and location of decision-making
[17]	Incorporating function supporting the display between DS and clinicians (de-activation; acknowledgment)
[17]	Provide justification for the suggestion
[12]	underlying philosophy should seek to minimize the overall number of alerts in the system and the frequency with which they activate
[12]	to curb false alarm rates, one should move from boundary based alarm strategies to intelligent alarm monitoring systems that monitor several parameters simultaneously and use fuzzy logic-based algorithms to initiate an alert
[12]	visual alerts should be prioritized that goes hand in hand with hazard matching as a warning implementation strategy
[12]	when developing an alert philosophy, one must carefully consider the threshold of the alert

	Usability design principle
[12]	prioritization of alerts should include three levels: low, medium, and high and should be coded using word, color, shape, position on screen, and other indicators known to influence urgency
[12]	A well-documented alarm philosophy is necessay to guide decision-making and ensure consistency in alerting: it should include a catalog of unsafe events, an indication of the level of priority of an alert, a description of the logic underspinning the classification of an event as unsafe and a description of a specific alert indicating each unsafe event
[12]	philosophy requires an explicit definition of what is meant by safety-critical event since this can vary based on the judgement of the user
[12]	alert philosophy should specify as a minimum which categories of problems should be included in the alerting system, and how many priorities there should be for each category of risk.
[12]	alerts should cancel and reset in response to the appropriate corrective action rather than requiring an acknowledgment from the operator followed by the corrective action
[12]	corrective actions should be easy to perform
[12]	visual alerts must be placed within an operator's visual field in order of importance, so that the highest priority alerts are located in the stationary field, with lower priority alerts in the eye field and head field
[12]	low priority alerts: these should be avoided or classified as information only indicators. Although from a safety point of view more alerts are seen as safer, in practice the reverse is true.
[12]	alerts which require acknoledgment before the user moves on should be kept to a minimum
[12]	warning label should have four information components: a signal word to indicate the priority of the alert, a statement of the nature of the hazard, an instruction statement, and a consequence statement: instruction and hazard statements are the most important to include.
[3]	Make it easier for clinicians to take action on the information provided
[3]	Increase sensitivity to the needs of the clinical scenarios
[3]	Intrusiveness must be proportional to the importance of the information
[3]	Summarize patient-level information: to take all key data needed for optimal decision-making process.
[3]	Prioritize and filter recommendations to the users
[3]	Combine recommendations for patients with co morbidities

TOWARDS A USABILITY KNOWLEDGE BASE TO SUPPORT HEALTH INFORMATION TECHNOLOGY DESIGN AND EVALUATION: APPLICATION TO MEDICATION-RELATED ALERTING SYSTEMS

ROMARIC MARCILLY

Introduction.

Health Information Technology (HIT) is increasingly implemented to improve healthcare quality and patient safety. However, some usability issues may reduce their impact and even induce new problems (including patient safety issues). To avoid those negative outcomes, amongst other actions, HIT usability must be improved. This action requires applying validated usability knowledge. However, usability knowledge applied to HIT is scattered across several sources, is not structured and is hardly usable. Moreover, its coverage regarding related usability flaws is not known. This work has two aims: (i) to participate in improving the accumulation of usability knowledge for HIT and (ii) to provide synthetic structured easy-to-use HIT usability knowledge with a clear coverage. Those aims are applied to medication alerting systems.

Method.

Two independent analyses of the literature have been performed. On the one hand, usability flaws and their consequences for the clinicians and the work system have been searched and organized; on the other hand, existing usability design principles specific to medication alerting systems have been synthesized. Results of both analyses have been matched together.

Results.

A systematic review identified 13 types of usability flaws in medication alerting systems. Consequences on the clinicians and the work system are varied: they greatly impede the clinicians and negatively impact the work system (*e.g.*, alert fatigue, alert misinterpretation). Sixty-three usability design principles dedicated to medication alerting systems are identified. They represent six themes: improve the signal-to-noise ratio, fit clinicians' workflow, support collaborative work, display relevant information, make the system transparent and provide useful tools. The matching between usability flaws and principles is quite good.

Discussion.

As a result of this work, a list of usability design principles illustrated by actual instances of their violation has been developed. It may help designers and Human Factors experts understand and apply usability design principles when designing and evaluating medication alerting systems. Usability applied to HIT is a recent research field that suffers from a deficit of structured knowledge. This work shows that it is possible to accumulate and structure usability knowledge. It could be carried on by developing a usability knowledge base dedicated to HIT in order to strive towards "evidence-based usability".