



HAL
open science

Comparison of the validity, perceived usefulness, and usability of I-MeDeSA and TEMAS, two tools to evaluate alert system usability

Romaric Marcilly, Wu-Yi Zheng, Paul Quindroit, Sylvia Pelayo, Sarah Berdot, Bruno Charpiat, Jennifer Corny, Sylvain Drouot, Pauline Frery, Géraldine Leguelinel-Blache, et al.

► To cite this version:

Romaric Marcilly, Wu-Yi Zheng, Paul Quindroit, Sylvia Pelayo, Sarah Berdot, et al.. Comparison of the validity, perceived usefulness, and usability of I-MeDeSA and TEMAS, two tools to evaluate alert system usability. *International Journal of Medical Informatics*, In press, 175, pp.105091. 10.1016/j.ijmedinf.2023.105091 . hal-04097607

HAL Id: hal-04097607

<https://hal.science/hal-04097607>

Submitted on 26 May 2023

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

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

Comparison of the validity, perceived usefulness, and usability of I-MeDeSA and TEMAS, two tools to evaluate alert system usability

Romaric Marcilly^{a,b}, Wu-Yi Zheng^c, Paul Quindroit^a Sylvia Pelayo^{a,b}, Sarah Berdot^{d,e,f}, Bruno Charpiat^g, Jennifer Corny^h, Sylvain Drouotⁱ, Pauline Frery^j, Géraldine Leguelinel-Blache^{k,l}, Lisa Mondet^m, Arnaud Potier^{n,o}, Laurine Robert^{a,p}, Laurie Ferret^q, Melissa Baysari^r

^a Univ. Lille, CHU Lille, ULR 2694 - METRICS: Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

^b Inserm, CIC-IT 1403, F-59000 Lille, France

^c Black Dog Institute, Randwick, NSW, Australia

^d Assistance Publique - Hôpitaux de Paris, Département de Pharmacie, Hôpital européen Georges-Pompidou, Paris, France.

^e Inserm, Cordeliers Research Centre, Université de Paris, Sorbonne Université, Paris, France.

^f HeKA, Inria, Paris, France

^g Pharmacie, Hôpital de la Croix-Rousse, Hospices Civils de Lyon, 69004 Lyon, France

^h Service de Pharmacie, Groupe Hospitalier Paris Saint-Joseph, Paris, France

ⁱ Clinical Pharmacy Department, Hôpital Bicêtre, APHP, Paris, France

^j Hôpital Bel Air CHR Metz – Thionville, France

^k Desbrest Institute of Epidemiology and Public Health, Univ Montpellier, Inserm, Montpellier, France

^l Department of Pharmacy, CHU Nîmes, Univ Montpellier, Nîmes, France

^m Department of Pharmacy, CHU Amiens-Picardie, Amiens, France

ⁿ Service de pharmacie, CH de Luneville, 54300 Luneville, France

^o Service de pharmacie, CHRU de Nancy, 54000 Nancy, France

^p CHU Lille, Institut de Pharmacie, Lille, France

^q Department of Pharmacy, General hospital of Valenciennes, 59300, France

^r The University of Sydney, Faculty of Medicine and Health, School of Medical Sciences, Biomedical Informatics and Digital Health, Sydney, Australia

Corresponding author: Romaric Marcilly, romaric.marcilly@univ-lille.fr

Email addresses: Wu-Yi Zheng, wuyi.zheng@blackdog.org.au ; Paul Quindroit, paul.quindroit@univ-lille.fr ; Sylvia Pelayo, sylvia.pelayo@univ-lille.fr ; Sarah Berdot, sarah.berdot@aphp.fr ; Bruno Charpiat, bruno.charpiat@chu-lyon.fr ; Jennifer Corny, jennifercorny@yahoo.fr ; Sylvain Drouot, sylvain.drouot@aphp.fr ; Pauline Frery, p.frery@chr-metz-thionville.fr ; Géraldine Leguelinel-Blache geraldine.leguelinel@chu-nimes.fr ; Lisa Mondet, mondet.Lisa@chu-amiens.fr ; Arnaud Potier, apotier@ch-luneville.fr ; Laurine Robert, laurine.robert@chu-lille.fr ; Laurie Ferret ferret-l@ch-valenciennes.fr ; Melissa Baysari, melissa.baysari@sydney.edu.au

Abstract

Objective. Two tools are currently available in the literature to evaluate the usability of medication alert systems, the instrument for evaluating human factors principles in medication-related decision support alerts (I-MeDeSA) and the tool for evaluating medication alerting systems (TEMAS). This study aimed to compare their convergent validity, perceived usability, usefulness, strengths, and weaknesses, as well as users' preferences. **Method.** To evaluate convergent validity, two experts mapped TEMAS' items against I-MeDeSA's items with respect to the usability dimensions they target. To assess perceived usability, usefulness, strengths, and weaknesses of both tools, staff with expertise in their medication alerting system were asked to use French versions of the TEMAS and I-MeDeSA. After the use of each tool, participants were asked to complete the System Usability Scale (SUS) and answer questions about the understandability and usefulness of each tool. Finally, participants were asked to name their preferred tool. Numeric scores were statistically compared. Free-text responses were analyzed using an inductive approach. **Results.** Forty-five participants from 10 hospitals took part in the study. In terms of convergent validity, I-MeDeSA focuses more on the usability of the graphical user interface while TEMAS considers a wider range of usability principles. Both tools have a fair level of perceived usability (I-MeDeSA' SUS score=61.85 and TEMAS' SUS score=62.87), but results highlight that revisions are necessary to both tools to improve their usability. Participants found TEMAS more useful than I-MeDeSA ($t=-3.63$, $p=.005$) and had a clear preference for TEMAS to identify problems in formative evaluation (39 of 45; .867, $p < .001$) and to compare the usability of alert systems during the procurement process (36 of 45; .8, $p < .001$). **Conclusions.** The TEMAS is perceived as more useful and is preferred by participants. The I-MeDeSA seems more relevant for quick evaluations that focus on the graphical user interface. The TEMAS seems to be more suitable for in-depth usability evaluations of alert systems. Even if both tools are perceived to be equally usable, they suffer from wording, instructional, and organizational problems that hinder their use. The results of this study will be used to improve the design of I-MeDeSA and TEMAS.

Keywords

Ergonomics; Usability; Human factors; Decision support systems, clinical; Assessment tool.

1. Introduction

Alert systems are increasingly being used by healthcare organizations to warn users about potential medication errors, thus enhancing medication safety. The usability of these systems, defined as the "extent to which [they] can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use"[1], is a critical factor in their acceptance and safe use[2–4]. Research has shown that alerts with good usability are associated with higher user satisfaction[5], faster work, fewer prescription errors and less workload for clinicians compared to poorly designed alerts[6].

The application of usability design principles to an alert system’s design, evaluation, purchase, and implementation ensures good usability and improves the usefulness of systems and the likelihood of user acceptance. Various sets of usability design principles are available in the scientific and grey literature. They can be general and applicable to any type of interactive technology[7–9] or specific to a type of technology such as an alert system or medical device [10,11]. However, application of these principles is often limited to usability experts and not formulated, presented, and justified in a way that is accessible and understandable to designers and decision-makers.

With respect to medication alert systems, two paper-based evaluation tools have been developed to assist vendors and stakeholders in determining whether an alert system complies with usability design principles. These are the Instrument for evaluating human factors principles in medication-related decision support alerts (I-MeDeSA)[12] and the Tool for evaluating medication alerting systems (TEMAS)[13], as described in Table 1.

Table 1. Foundation, key elements, target users, and previous applications of two tools for evaluating alert usability.

	I-MeDeSA	TEMAS
Foundation of tool	Based on a set of human factors principles related to drug-drug interaction alerts in electronic health records (EHRs)[14].	Based on a set of evidence-based usability design principles for alert systems[10]
Key elements	26 items representing 9 usability principles <i>Normative approach:</i> Systems are scored on each item (0 or 1) with the maximum score (26) representing a high level of adherence to human factors principles	66 items organized into 6 sections <i>Formative approach:</i> Systems are not scored. TEMAS is a checklist to identify the usability-related strengths and weaknesses of an alert system
Target users	Not specified	Group of local experts with extensive knowledge of their alert system

Previous applications	Alerts embedded in EHRs in the United States[15], Korea[16], and Australia[17,18] Tool has been extended to enable evaluation of a broader spectrum of alerts[19].	Alert systems in Australia[13]
-----------------------	---	--------------------------------

1.1.Objective

The two tools differ in their development process and structure, but both are designed to evaluate an alert system’s usability. Despite this common goal, the tools have not been compared, making it challenging for end-users to make an informed choice about which is most suitable for alert evaluation. This study aimed to compare I-MeDeSA and TEMAS in their ability to evaluate the usability of an alert system. In particular, we evaluated their convergent validity, their perceived usability and usefulness, perceived strengths and weaknesses, as well as users’ preferences for the tools. Our goal was to provide an evidence-base for end-users to identify the most suitable tool for their current or prospective alert system assessments.

2. Materials and methods

The method in our planned study protocol [20] was modified slightly due to variations in the context of implementation. The study included two main components: (1) assessment of the convergent validity of both tools, and (2) assessment of the perceived understandability, usability and usefulness of each tool and of users’ preferences.

2.1.Convergent validity: Do TEMAS and I-MeDeSA measure the same usability dimensions?

Application of I-MeDeSA produces a score, but not TEMAS, preventing a coefficient correlation from being calculated, as is usually done to estimate convergent validity[21]. We set out to determine which usability concepts embodied in I-MeDeSA and TEMAS were common to both tools and which were unique to each. To do this, two experts with experience in usability of alerting systems (14 years of experience each) mapped TEMAS’ items against I-MeDeSA’s items with respect to the usability dimensions they target. This matching process was completed independently. An agreement score was calculated (Gwet’s AC1[22]), and disagreements were discussed until a consensus was reached. Any remaining disagreements were resolved through discussion with a third expert (8 years of experience).

2.2. Perceived usability, understandability, usefulness, and users' preferences

This part of the study was conducted in France which required the translation of I-MeDeSA and TEMAS.

2.2.1. Material: French I-MeDeSA and TEMAS

The original versions of I-MeDeSA [12] and TEMAS [13] were translated into French by a human factors researcher with experience in medication alert systems. Two human factors specialists and one clinician also with expertise in health technology, all fluent in French and in English, and with expertise in decision support, rated their level of agreement with the translation of each item. For I-MeDeSA, the experts had a good level of agreement between them (Gwet's AC2 = 0,70 [0,08; 1]). For 9 items of 26 (34,61%), at least 1 expert expressed a slight disagreement with the translation. For TEMAS, the three experts achieved a very high level of agreement between them (Gwet's AC2 = 0,94 [0,92;1]). There were 6 items out of 66 (10%) for which 1 of the 3 experts expressed a slight disagreement with the translation. Disagreements were discussed with the translator until a consensus was reached. A professional translator validated the final French translation of both tools.

2.2.2. Recruitment of study sites

Staff from French hospitals equipped with intimate knowledge of the medication alerting system implemented in their facility were identified through professional networks and at relevant conferences. They were contacted through email or by phone call and briefed on the research protocol. At each site, a group of experts in the local alerting system (including both end-users and experts) was recruited. Group sizes ranged from 3 to 6 people. There was no financial reimbursement for participation.

2.2.3. Study design and data collection

At each study site, the local experts met face-to-face, online, or in a hybrid way. During this meeting, both I-MeDeSA and TEMAS were completed and assessed. At no time before or during the meetings were they informed of the origin of the tools. After informed consent was obtained, the name and version of the alert system in use were collected along with information about the participants' background. Then, participants completed as a group the tools, I-MeDeSA and TEMAS. The running order of the tools was counterbalanced between sites to prevent an order

effect. Completion time was recorded for each tool. An observer, with expertise in human factors, took note of participants' comments and behaviors while completing the tools (e.g., quality and problems with the items/tools). After each tool's completion, participants individually completed the French version of the System Usability Scale (SUS)[23]. They also answered 2 questions about the understandability of the items and the usefulness of the tool on 5-point Likert scales. Participants were asked to explain what they found useful in the tools and in which contexts they would like to use them. After completion of both tools and related questions, participants were asked which tool they would prefer to use in two use contexts (1. comparison of alert systems; 2. identification of an alert system's weaknesses for improvement) and the perceived added value of each tool (open-ended question). All questions were answered individually.

2.2.4. Data analysis

For each tool, the average SUS score was calculated and compared to Bangor et al.'s scale to determine the level of perceived usability[24]. Inferential statistics (paired *t*-test) were performed to compare I-MeDeSA and TEMAS on each SUS item, on the full SUS score, on the understandability and usefulness ratings, and on the completion time. Answers to the preference questions were compared by one-sample binomial tests to 0.5 (representing equal preference). Significance threshold was set at 0.05. Statistics were performed with Jamovi 2.2.5. software.

Meaningful semantic units (i.e., sets of words representing a single idea that is sufficiently self-explanatory for analyses) were extracted from comments and responses to open-ended questions. Units were analyzed inductively. Two researchers independently assigned each unit a code they named and defined. The resulting codes were discussed by the two researchers until agreement was reached on a clear, unambiguous, and exhaustive classification scheme. The analyses were initially performed without the analysts knowing which tool they were classifying the data from. The names of the tools were given after the analyses were completed so that they could be compared.

In the tables of results, the number of participants whose verbatim comments contributed to a theme is given for information only; the results were not quantified.

3. Results

3.1. Convergent validity: Do TEMAS and I-MeDeSA measure the same usability dimensions?

Experts had a very high level of agreement ($AC1 = 0.99 [0.986;0.995]$): 99.06% of the pairs were the same. After consensus, 40 TEMAS items (60.61%) were not associated with any I-MeDeSA items and 9 I-MeDeSA items (34.62%) were not matched to any TEMAS items (Appendix 1 and Appendix 2). At the level of the tools' dimensions, for two TEMAS dimensions, no items were matched by I-MeDeSA items; for one I-MeDeSA dimension, no items were matched by TEMAS items (Appendix 3).

Usability concepts that are exclusively represented by I-MeDeSA deal with usability principles not specific to alerting systems (e.g., font size, background/text contrast, accessibility, number of colors, guidance of sequences of "tasks"), the presentation of multiple alerts, or visual discriminability of alerts. Concepts that are exclusive to TEMAS concern the data used and process to trigger alerts (e.g., contextualization, knowledge base and data sources, rules' engine), the alert recipient, alert's characteristics (look-and-feel, mode of appearance, content, and customization), the fit with workflow, and alerts' monitoring and management by the healthcare facility.

3.2. Perceived usability, understandability, usefulness, and preferences

3.2.1. Study sites description

Twenty-four French hospitals were contacted. Thirteen agreed to take part. In one hospital, staff needed for the evaluation session were unavailable due to workload and understaffing issues. In another, commitment to confidentiality with the alerting system's vendor prevented participation in the study. A total of 11 sites took part in the evaluation. At one site, data were lost during transfer resulting in 10 sites and 45 participants being included (Figure 1 and Table 2).

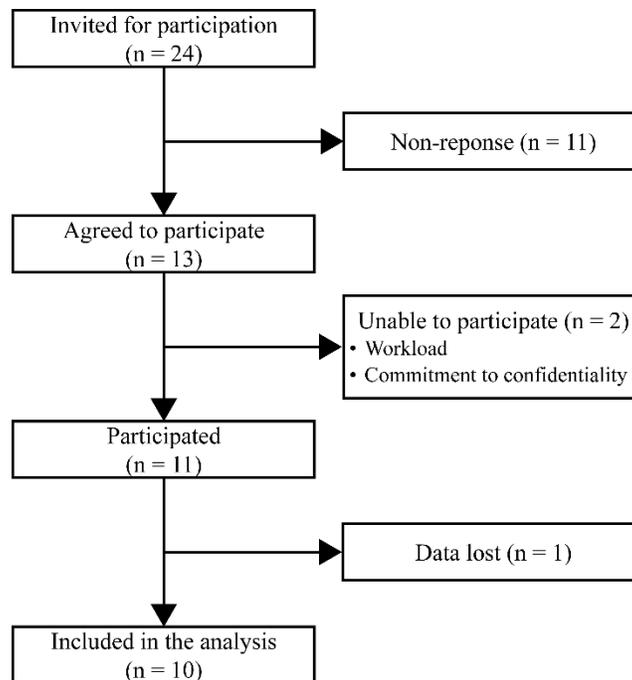


Figure 1. Study sites inclusion process.

Table 2. Characteristics of the study sites.

Hospital	Number of beds	Alert system	Number of participants and profile
Centre Hospitalier Régional de Metz – Thionville	3000	VIDAL sentinel	4 pharmacists including the one in charge of the alert system parameterization
Centre Hospitalier Universitaire de Lille	2741	PharmaClass v1.8	3 pharmacists, experts in the system and involved in the design of alerts
Hospices Civils de Lyon	5362	EASILY 6.1	5 pharmacists including the expert in parameterization of the protocols
Hôpital Bicêtre Assistance Publique – Hôpitaux de Paris	987	PharmaClass v1.9	6 pharmacists with expertise in the system
Groupement Hospitalier de Territoire Lorraine Sud	3083	PharmaClass v1.9.0	5 pharmacists, expert in the system and involved in the design of alerts
Centre Hospitalier de Valenciennes	1824	Millenium 2015	3 pharmacists, users of the system with expertise in health information system
Centre Hospitalier Universitaire de Nîmes	2000	Pharma Version 5.9.10827.1034	4 pharmacists, users of the system
Centre Hospitalier Universitaire d'Amiens	1705	PharmaClass v1.8	5 pharmacists, expert users of the system and involved in the design of the alerts

Groupe Hospitalier Privé Saint Joseph	587	Lumio/GHPSJ (Homegrown)	4 pharmacists with various expertise in the system
Hôpital Européen Georges Pompidou Assistance Publique – Hôpitaux de Paris	750	DXCARE 7.7.9.5.2. including VIDAL 2021.4.0.	6 – 3 pharmacists and 3 physicians expert users of the system

3.2.2. Usability and understandability

3.2.2.1. Overall

The average SUS scores for I-MeDeSA and TEMAS were 61.85 and 62.87 respectively (Student's $t = -.12$, $p = .909$) which represent an "OK" level of perceived usability[24]. The completion time for I-MeDeSA was significantly shorter than for TEMAS (22min48s vs. 46min, Student's $t = -6.93$, $p < .001$). At the items' level, completion time was longer for I-MeDeSA than for TEMAS (respectively 53s and 38s per item; Student's $t = 2.28$, $p = .048$).

Table 3 synthesizes the main perceived usability characteristics of I-MeDeSA and TEMAS. Participants found I-MeDeSA easy and fast to complete, because of the smaller number of items. TEMAS was perceived as time-consuming and complex to complete because of the large number of items. For both tools, participants found the response mode easy, although the I-MeDeSA response mode was reported as unintuitive by some participants. Participants found the TEMAS "partial" response option helpful. In terms of comprehension, some participants found the instructions for using I-MeDeSA unclear; participants also expressed difficulty in understanding how to complete the TEMAS summary table.

Table 3. Main usability features of I-MeDeSA and TEMAS. The + and – signs express the valence of the usability feature (Complete table in Appendix 4).

Usability Features			I-MeDeSA	TEMAS
Ease of use	Overall use	Speed of completion	+ "Quick to complete", "[it is] useful because it is succinct and direct" (n = 5)	- "This tool has too many items", "very exhaustive (time consuming) questionnaires", "the questionnaire is too long for my taste" (n = 11)
		Ease of use	+ "Simple, clear", "Useful: its simplicity", "Useful: its ease of use" (n = 6)	- "Its complexity makes the tool not useful", "it's quite heavy to handle" (n = 4)
	Responding	Simplicity of answer	+ "Simplicity of the scoring: present/absent" (n = 1)	
		Adaptation of the response	- "For some items we had trouble giving a binary answer", "a binary	+ "The 'Partial' checkbox is quite useful here", "'partial' was quite

		mode to the context		answer is not appropriate in my opinion" (n = 8)		handy", "and the 'partial' option is handy" (n = 5)
		Intuitiveness of the response mode	-	"If we had a "yes/no" checkbox it would be easier, wouldn't it? It would avoid having to remember that 0 is no and 1 is yes" (n = 1)		
		Connection between the synthesis table and the items		NA	-	"It would have been good to put these boxes at the end of each section, we should have answered section by section because at the end we lost everything we had said before" (n = 7)
Understandability	Instructions for use		-	"The instructions are not clear" (n = 1)	-	
	Summary table			NA	-	"The summary table of points to be clarified would require more explanation for its completion", "what should we do? should we fill it out?", "The expectations of free part on the points requiring particular attention should be better clarified" (n = 3)
	Items	Clarity	-	"Some ambiguous questions," "Questions not clear enough for me to answer", "Not all items in the tool are intelligible to me" (n = 7)	-	"Difficulty in interpreting items and examples", "Some terms could be defined before doing the questionnaire to avoid ambiguities in the understanding of the items", "Questions sometimes ambiguous" (n = 7)
		Simplicity	-	"Complexity of multi-component questions", "Question wording sometimes long and complex" (n = 7)	-	"Sometimes the questions don't get to the point and make the use and understanding more complex than necessary" (n = 1)
					+	"Simple items", "questions are clear, simple to understand, relevant to practice" (n = 4)
	Examples	Help	+	"The idea of using examples is very good" (n = 1)	+	"The fact that there is an example is an important help", "there are twisted these questions, you have to read the example to understand" (n = 5)
					-	"Sometimes we understand the item and the example differently. So, some examples need to be revised in wording", "Revise examples to fit well with the item (and vice versa)" (n = 7)

		Influence	-	"The examples presented with the questions sometimes focus the attention of the participants, who then think only about the application of the example to their alerting system and no longer about the application of the item" (n = 1)	-	"The examples direct and reduce the field of answer [...] it helps but it directs", "The example influences the comprehension of the item. Without reading it, I understand the item a little differently" (n = 11)
--	--	-----------	---	---	---	---

3.2.2.2. Items

Participants raised concerns about the usability of 12 I-MeDeSA items (46.15%) and 35 TEMAS items (53.03%), some items suffering from several problems (Table 4, Appendix 5 and Appendix 6). Understandability issues were raised for 12 I-MeDeSA items (46.15%) and 26 TEMAS items (39.39%). Overall, participants did not rate the understandability of I-MeDeSA and TEMAS items differently (3.44 vs. 3.30, Student's $t = 35.12$, $p = .734$, respectively). Some participants found TEMAS items simple to understand. However, participants also pointed out that some items in both tools were too long, too technical, and complex in phrasing or wording. This made the items confusing or difficult to understand, requiring the items to be read several times in order to understand them. Some participants did not understand the difference between some items: they found items 4i and 6i of I-MeDeSA redundant, as well as items A11 and A12 of TEMAS. Participants generally found the examples in TEMAS and I-MeDeSA relevant and helpful for understanding the items. However, they mentioned that some examples illustrating TEMAS items did not appear to fit the items. For both tools, participants found that some examples focused their attention on particular areas and led them to respond more to the examples than to the items more broadly.

Participants questioned the relevance of some items for their alerting systems (3 items for I-MeDeSA, 11.54%; 10 for TEMAS, 15.15%). In particular, some items assume that the alerting system is integrated into another software (e.g., EHR) and is to be used by the prescriber, while some systems under evaluation were stand-alone and used by pharmacists.

Participants also criticized the validity of some items. They were not convinced that the recommendations underlying the items would improve an alert system's usability (2 items for I-MeDeSA, 7.69%; 2 for TEMAS, 3.03%). Finally, participants explained that the organization of 3 TEMAS items (4.54%) was nonsensical because items that are connected to each other are not

always presented close to each other, and items that depend on the answer to another item are not presented conditionally.

Table 4. Number (and percentage) of I-MeDeSA's and TEMAS' items with a problem highlighted by the participants according to the type of problem. Some items suffered from multiple types of issues.

	I-MeDeSA	TEMAS
Number of items where a concern was raised	12 (46.15%)	35 (53.03%)
Type of issue		
Understandability	12 (46.15%)	26 (39.39%)
Relevance	3 (11.54%)	10 (15.15%)
Validity	2 (7.69%)	2 (3.03%)
Organization	0	3 (4.54%)

3.2.3. Perceived usefulness

Participants rated the usefulness of I-MeDeSA significantly lower than TEMAS (3.34 vs. 4.24 respectively, Student's $t = -3.63$, $p = .005$). Participants found both I-MeDeSA and TEMAS useful in exploring and getting to know their alert system better, identifying flaws and ways to fix the system (Appendix 7). They also indicated that TEMAS is an educational tool that increased their awareness of best practices in alert system usability. Both tools' perceived usefulness was negatively impacted by the irrelevance of some items for the alert systems under evaluation.

3.2.4. Perceived added value and preference

Table 5 depicts participants' perception of the added value of I-MeDeSA and TEMAS. Some participants indicated that I-MeDeSA is simpler and more understandable than TEMAS; others reported the opposite. The added value of I-MeDeSA over TEMAS, as perceived by participants, was that I-MeDeSA is shorter and easier to implement, provides a score, and addresses more focused topics ("appearance" of the alerts). The latter point was also seen as a limitation. The perceived added value of TEMAS compared to I-MeDeSA was that it allows for a broader assessment of the alert system by addressing more comprehensive and relevant themes while having more specific and detailed items. Participants also found TEMAS to have fewer inconsistencies than I-MeDeSA, to allow more precise responses (because of the "partial" response option), and to provide a better synthesis of results (because of the final table).

Table 5. Perceived added value for I-MeDeSA and TEMAS (excerpts; complete version in Appendix 8).

I-MeDeSA's added value

Shorter and simpler (n = 21)	"Even though, I-MeDeSA questions are clearer", "[I-MeDeSA is] quicker to fill out", "[I-MeDeSA is] shorter; allows for more recurrent use", "At first glance [TEMAS] is longer so it can be discouraging", "[TEMAS is] heavier and longer"
More understandable (n = 8)	"[I-MeDeSA is] easier to understand questions.", "[for I-MeDeSA] Simplicity and lack of ambiguity in wording", "TEMAS has items that are sometimes not very explicit despite the examples"
Theme is more precise (n = 8)	"Presence of items on the font or the rules of legibility (mix of lower / upper case) that we do not find on TEMAS", "[I-MeDeSA] goes into much more detail regarding the display of alerts (shapes, colors)", "[I-MeDeSA] allows a better approach to the user experience and the ergonomics of a system"
Score (n = 1)	"The score"
Conditional items (n = 1)	"You see, there are questions [in I-MeDeSA] that should not be answered: it depends on the previous question"
TEMAS' added value	
More complete and relevant themes (n = 44) vs too precise I-MeDeSA themes (n = 12)	"[TEMAS] addresses more items and therefore better source for thinking about a rigorous procurement process", "I feel like [TEMAS] covers a wider range of items to incorporate into a CDSS (clinical decision support system). I would say it allows for more relevant developments to be suggested that would not have been spontaneously thought of", "[TEMAS allows] projecting into the use in routine care in collaborative work mode and workflow", "[I-MeDeSA proposes] less exhaustive questions, too focused on the same parameters of the tool: many items on the visual aspect"
More precise, detailed (n = 28)	"[TEMAS is a] more detailed tool to see the improvements and expectations of a software", "in the end [TEMAS'] items are more detailed and that allows in my opinion a better evaluation of the software", "[TEMAS] goes into more detail, especially on the content of the alerts"
More understandable (n = 19)	"[In TEMAS] questions are better turned", "Items are more synthetic [in TEMAS], I find them more understandable especially since they are accompanied by examples", "Items are easier to understand [in TEMAS]", "[I-MeDeSA] is more complicated to understand. The sentences used are complex/more difficult to understand than those in TEMAS. The examples given by I-MeDeSA are confusing: they stop the thought on a particular character"
Better synthesis (n = 15)	"[TEMAS'] final synthesis allowing to prioritize the improvements to be made", "the final synthesis table [of TEMAS] can be used to identify the areas of improvement of a CDSS" (clinical decision support system)
More precise response mode (n = 9)	"TEMAS allows a less clear-cut answer to the items", "The possibility of answering "partial" and commenting is great [in TEMAS]", "Possibility of nuance with partial answer [in TEMAS]", "the big advantage over I-MeDeSA is when you have several types of alerts, you can nuance by partial"
Simpler and more concrete (n = 7)	"TEMAS (...) less complex", "[TEMAS is] easier to use because more concrete", "Simplicity of use [of TEMAS]"
TEMAS more consistent (n = 1)	"Less inconsistencies [in TEMAS]"

Participants reported that both tools could be used to develop or improve their alert system, to evaluate it, especially the changes made during re-engineering, or to help them choose a new alert

system by allowing them to determine the specifications of an "ideal" system to choose (e.g., tenders) or to compare systems on the criteria of both tools (Appendix 9). When asked which tool they would like to use in the future, most participants answered they would prefer TEMAS over I-MeDeSA: 36 of 45 (.8, $p < .001$) choose TEMAS to compare alert systems and 39 of 45 (.867, $p < .001$) said they would use TEMAS to identify an alert system's weaknesses and areas for improvement.

4. Discussion

This study aimed to compare I-MeDeSA and TEMAS from various angles: their convergent validity by estimating their overlap in terms of the usability concepts covered by each tool, their perceived usability and usefulness, perceived strengths and weaknesses, as well as users' explicit preferences for a tool. The intended aim of this study was not to provide a conclusion on which is the better tool. Instead, data collected could be used to assist users to determine which tool to use depending on the purpose of the assessment, and to guide improvements to the TEMAS and I-MeDeSA.

Unsurprisingly, results from the current study suggest that I-MeDeSA fundamentally differs from TEMAS, focusing on usability of the graphical user interface, while the TEMAS takes into account a wider range of relevant usability principles derived from a literature review [10]. Consequently, participants perceived the TEMAS as more comprehensive, allowing a broader assessment than the I-MeDeSA. Even if both tools are perceived as accurate, understandable, and as having a fair level of usability (characterized as "ok" according to Bangor *et al.*[24]), it is apparent from the results that revisions are necessary to both tools. Clarity of items and examples are crucial as misunderstanding could lead to erroneous assessments which will negatively impact decisions made by organizations on procurement/reconfiguration of alerting systems. Both TEMAS and I-MeDeSA included items where wording and phrasing misled users (Appendices 5 and 6). Most of the problematic items would be clearer and more understandable if they were shorter and used less technical language. Relevance of items is also paramount. Users should only be presented with items that are relevant to their system (e.g., standalone) or the way they use it (e.g., by pharmacist only) especially in the case of TEMAS whose length has been the subject of criticism. On the contrary, the brevity of I-MeDeSA was seen as an advantage, perhaps owing to the fact that hospital staff are often time-poor. Participants identified irrelevant items mainly in the TEMAS but also, to

a lesser extent in the I-MeDeSA. Item presentation should be based on a branching system: depending on the characteristics of the alert system or the use context previously entered, or on a response to a previous item, only relevant items should be presented. A concise and well-structured tool with appropriate branching to skip irrelevant items would be best suited to hospital staff with little time to perform the evaluation. In moving forward, we suggest each problematic item be reviewed jointly with user representatives to ensure that it is clear, unambiguous, and relevant. Addressing these points will contribute to improving the usability of both tools.

During the study, participants appreciated the opportunity to use both tools to better understand their alert systems and discover potential problems one may encounter while using these systems. Thus, tools such as the I-MeDeSA or TEMAS can be helpful in the selection, evaluation or re-engineering of medication alert systems. Participants cited relatively identical contexts in which the I-MeDeSA and TEMAS could be useful (exploring and getting to know their alert system better, identifying flaws and ways to fix them). However, survey results suggest the TEMAS is the preferred tool to identify problems in formative evaluation or to compare the usability of alerting systems in the procurement process.

This study is the first to directly compare two existing tools to assess the usability of medication alert systems. To date, the literature has identified problems with the I-MeDeSA [16–18] and TEMAS [13], but not compared them. The comparison was performed using data from a larger sample size than other studies involving TEMAS and I-MeDeSA (45 participants from 10 French hospitals vs 18 participants for TEMAS [13] and at most 4 participants for I-MeDeSA [19]). This large sample size and the variety of alert systems evaluated (7 different systems) ensure a good level of reliability of the results.

Our findings suggest that the I-MeDeSA is suitable and preferred for quick evaluations that focus on the graphical user interface of alerting systems. Due to its short length, the I-MeDeSA can be used frequently during iterations of interface development for example. On the contrary, TEMAS appears to be more suitable for in-depth and multidimensional usability evaluations of alert systems. The length of the tool may prevent it from being used in its full version very frequently or repeatedly: however, depending on the purpose of the evaluation, some sections of TEMAS can be used iteratively.

In their current form, neither I-MeDeSA nor TEMAS are without flaws. The results identified areas for improvement. For the I-MeDeSA, the main opportunities for improvement relate to rewording

and shortening some items, and to clarify instructions for use and scoring. For the TEMAS, rewording and shortening of some items should be a focus, along with ensuring consistency between item's definition and example, clarifying the instructions for completing the summary table, and organizing the items so that related items are presented together and only items relevant to the alert system being evaluated and its context of use are shown.

Even though the TEMAS was preferred by a majority of participants, the results of this study highlight that revisions are needed to improve its usability and usefulness for end-users. For example, in order to present only items relevant to alert systems and their contexts of use, TEMAS could be developed online in such a way that branching logic could be applied: only relevant items will be displayed, which could in some cases reduce the length of this tool.

This study has potential biases and limitations that need to be discussed. Three authors were involved in the design of TEMAS. Methodological precautions were taken to limit the risk of bias: participants were never informed of the origin of the tools before and during the evaluation process, and blinded data were analyzed. The comparison was conducted on the French versions of the tools. Although comparing the two tools in the French context may limit the generalizability of the results, it also contributes to the cross-cultural validation of the tools. Finally, differences in the fundamental elements of the tools (e.g., scoring only in I-MeDeSA) prevented the calculation of correlations, thus limiting direct statistical comparisons. However, the multidimensionality of the comparison (including convergent validity estimation, perceived usefulness and ease of use comparison, and user preferences) and the analysis of qualitative (e.g., expression of views) and quantitative (e.g., perceived usability score) data provide rich data on the strengths and weaknesses of the tools.

5. Conclusions

This study compared the two existing tools for assessing the usability of medication alert systems, I-MeDeSA and TEMAS. Our results suggest that the two tools may be useful in different contexts, depending on the goals, time constraints and focus of the alert system evaluation. Both tools were perceived to be equally usable but also suffer from some wording, instructional, and organizational problems that impede ease of use. Even though TEMAS was perceived as being more useful and was preferred by participants, improvements are required to maximize usability of this tool. After

revisions are made based on participant feedback, further evaluations will investigate the impact of the use of these tools on the improvement of medication alert systems.

6. Ethics approval and permissions

According to French law, studies in human sciences in the field of health do not require approval by an ethics committee (decree n°2017–884, article 2). Nevertheless, this study was performed in accordance with the 1964 Helsinki declaration and later amendments. Written informed consent was obtained from all study participants.

7. Funding

There was no funding for this study.

8. Author statement

Romarc Marcilly: conceptualization of the study and of the methodology; management of the conduct of the study; data collection; data analysis; writing original draft. Wu-Yi Zheng: conceptualization of the study and of the methodology; data analysis; writing original draft. Melissa Baysari conceptualization of the study and of the methodology; data analysis; writing original draft. Paul Quindroit and Sylvia Pelayo: material preparation; review and editing. Sarah Berdot, Bruno Charpiat, Jennifer Corny, Sylvain Drouot, Pauline Frery, Géraldine Leguelinel-Blache, Lisa Mondet, Arnaud Potier, Laurine Robert, and Laurie Ferret: local organization of data collection; review and editing.

9. Declaration of interest

Three authors were involved in designing the tool TEMAS.

10. Acknowledgments

The authors thank the participants : Dr Jonathan Abisoror, Dr Anais Barbier, Dr Aurélie Barrail-Tran, Dr Rémy Basso Boccabella, Dr Stéphanie Bauler, Dr Charlotte Besson, Dr Sandrine Bulcourt, Dr Marie Camille Chaumais, Dr Alexane Crou, Dr Emmanuelle Divoux, Dr Florent Dubois, Dr Edith Dufay, Dr Marion Hennion-Coussemaq, Dr Anais Huguet, Pr Alexandre Karras, Dr Raphael Kavafian, Dr Michele Knight, Dr Benjamin Lagraulet, Dr Laure Lalande, Dr Camille Mabile, Dr Aurélien Mary, Dr Sophie Membré, Pr Tristan Mirault, Dr Anaëlle Olivo, Dr Hugo Palisson, Dr Hélène Perraud, Dr Thuy Tan Phan Thi, Pr Jacques Pouchot, Dr Raphaël Rainone, Dr Hélène Richard, Dr Chloé Rousseliere, Dr Brigitte Sabatier, Dr Pierre Sabatier, Dr Aurélie Terrier-Lenglet, Dr Julien Voyat, Dr Ameline Weber.

11. Summary table

Problem. For medication alerts to be effective, they must be usable, but there are no standardized ways to evaluate alert usability.

What is Already Known. I-MeDeSA and TEMAS are available tools for evaluation of medication alert systems. These tools have not been compared, making it challenging to make an informed choice about which tool is most suitable for usability evaluation.

What This Paper Adds. This study compared I-MeDeSA and TEMAS in terms of convergent validity, perceived usefulness and usability, and strengths and weaknesses. Results provide evaluators of alerts with evidence to guide selection of a tool to suit their needs.

12. References

1. International Organization for Standardization. ISO 9241-210:2010. Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems. International Organization for Standardization; 2010.
2. Westerbeek L, Ploegmakers KJ, de Bruijn G-J, Linn AJ, van Weert JCM, Daams JG, et al. Barriers and facilitators influencing medication-related CDSS acceptance according to clinicians: A systematic review. *International Journal of Medical Informatics*. 2021;152:104506.
3. Marcilly R, Ammenwerth E, Roehrer E, Pelayo S, Vasseur F, Beuscart-Zéphir M-C. Usability Flaws in Medication Alerting Systems: Impact on Usage and Work System. *Yearbook of Medical Informatics*. 2015;24:55–67.
4. 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;18:479–84.
5. Tsopra R, Jais J-P, 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;21:e107–16.
6. 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;21:e287-296.
7. Scapin DL, Bastien JMC. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behaviour & Information Technology*. 1997;16:220–31.
8. Nielsen J. 10 Usability Heuristics for User Interface Design [Internet]. Nielsen Norman Group. [cited 2023 Jan 26]. Available from: <https://www.nngroup.com/articles/ten-usability-heuristics/>
9. Shneiderman B, Plaisant C, Cohen M, Jacobs SM, Elmqvist N. Designing the user interface: strategies for effective human-computer interaction. Sixth Edition. Boston: Pearson; 2017.
10. Marcilly R, Ammenwerth E, Roehrer E, Niès J, Beuscart-Zéphir M-C. Evidence-based usability design principles for medication alerting systems. *BMC Med Inform Decis Mak*. 2018;18:69.
11. Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. *J Biomed Inform*. 2003;36:23–30.
12. 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. *Journal of the American Medical Informatics Association*. 2011;18:i62–72.

13. Zheng WY, Van Dort B, Marcilly R, Day R, Burke R, Shakib S, et al. A Tool for Evaluating Medication Alerting Systems: Development and Initial Assessment. *JMIR Med Inform.* 2021;9:e24022.
14. 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;17:493–501.
15. Phansalkar S, Zachariah M, Seidling HM, Mendes C, Volk L, Bates DW. Evaluation of medication alerts in electronic health records for compliance with human factors principles. *Journal of the American Medical Informatics Association.* 2014;21:e332–40.
16. Cho I, Lee J, Han H, Phansalkar S, Bates DW. Evaluation of a Korean version of a tool for assessing the incorporation of human factors into a medication-related decision support system: the I-MeDeSA. *Appl Clin Inform.* 2014;5:571–88.
17. Baysari MT, Lowenstein D, Zheng WY, Day RO. Reliability, ease of use and usefulness of I-MeDeSA for evaluating drug-drug interaction alerts in an Australian context. *BMC Med Inform Decis Mak.* 2018;18:83.
18. Lowenstein D, Zheng WY, Burke R, Kenny E, Sandhu A, Makeham M, et al. Do user preferences align with human factors assessment scores of drug-drug interaction alerts? *Health Informatics J.* 2020;26:563–75.
19. Snyder ME, Jaynes H, Gernant SA, DiIulio J, Militello LG, Doucette WR, et al. Alerts for community pharmacist-provided medication therapy management: recommendations from a heuristic evaluation. *BMC Med Inform Decis Mak.* 2019;19:135.
20. Marcilly R, Zheng WY, Beuscart R, Baysari MT. Comparison of the validity, perceived usefulness and usability of I-MeDeSA and TEMAS, two tools to evaluate alert system usability: a study protocol. *BMJ Open.* 2021;11:e050448.
21. Furr RM. *Psychometrics - an introduction*. Third edition. Los Angeles: SAGE; 2018.
22. Gwet KL. *Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among raters*. Fourth edition. Gaithersburg, Md: Advances Analytics, LLC; 2014.
23. Gronier G, Baudet A. Psychometric Evaluation of the F-SUS: Creation and Validation of the French Version of the System Usability Scale. *International Journal of Human–Computer Interaction.* 2021;37:1571–82.
24. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction.* 2008;24:574–94.

13. Appendices

Appendix 1. Number of items (percent) associated with items from the other tool based on the number of items they are associated with.

Number of items from the other tool matched	TEMAS	I-MeDeSA
0	40 (60.61%)	9 (34.62%)
1	14 (21.21%)	8 (30.76%)
2	9 (13.63%)	4 (15.38%)
3	2 (3.03%)	1 (3.84%)
4	1 (1.51%)	1 (3.84%)
5	/	2 (7.69%)
9	/	1 (3.84%)

Appendix 2. Results of the matching between I-MeDeSA et TEMAS items

TEMAS items	I-MeDeSA items
C1. Does the alerting system trigger alerts at the appropriate stage in a clinician's workflow?	2iii) Is the alert linked with the medication order by an appropriate timing?
C4. Do alerts appear in a central location, over the CPOE/EMR screen?	3i) Is the area where the alert is located distinguishable from the rest of the screen?
C5. Does the alerting system allow quick and easy responses to alerts?	2ii) If available, is the response to the alert, indicating the user's intended action (e.g. Accept, Cancel/Override), provided along with the alert, as opposed to being located in a different window or in a different area on the screen?
C7. Are alerts of different severity distinguishable from one another?	6i) Are the different severities of alerts easily distinguishable from one another?
C13. Does the alerting system present the most critical information at the top of the alert?	2iv) Does the layout of critical information contained within the alert facilitate quick uptake by the user? Critical information should be placed on the first line of the alert or closest to the left side of the alert box. Critical information should be labelled appropriately and must consist of: (1) the interacting drugs, (2) the risk to the patient, and (3) the recommended action.
D1. Does the alert include information on why the alert was triggered?	2iv) Does the layout of critical information contained within the alert facilitate quick uptake by the user? Critical information should be placed on the first line of the alert or closest to the left side of the alert box. Critical information should be labelled appropriately and must consist of: (1) the interacting drugs, (2) the risk to the patient, and (3) the recommended action. 7ii) Does the alert possess a statement of the nature of the hazard describing why the alert is shown? 7iia) If yes, are the specific interacting drugs explicitly indicated?
D2. Does the alert include information on the nature of the unsafe event and its likelihood?	2iv) Does the layout of critical information contained within the alert facilitate quick uptake by the user? Critical information should be placed on the first line of the alert or closest to the left side of the alert box. Critical information should be labelled appropriately and must consist of: (1) the interacting drugs, (2) the risk to the patient, and (3) the recommended action. 7ii) Does the alert possess a statement of the nature of the hazard describing why the alert is shown? 7iia) If yes, are the specific interacting drugs explicitly indicated? 7iv) Does the alert possess a consequence statement telling the user what might happen if the instruction information is ignored?
D3. Does the alert use colour and a signal word to indicate severity of the unsafe event?	4i) Is the prioritisation of alerts indicated appropriately by colour? 4iii) Are signal words appropriately assigned to each existing level of alert? 7i) Does the alert possess a signal word to indicate the priority of the alert (e.g. 'note', 'warning', or 'danger')?
D5. Does the alert include relevant patient information and provide a link for users to obtain further patient information?	8i) Are the informational components needed for decision-making on the alert present either within or in close spatial and temporal proximity to the alert?
D6. Does the alert provide clinically appropriate recommendations and suggest alternatives (i.e. drug, dose and frequency)?	2iv) Does the layout of critical information contained within the alert facilitate quick uptake by the user? Critical information should be placed on the first line of the alert or closest to the left side of the alert box. Critical information should be labelled appropriately and must consist of: (1) the interacting drugs, (2) the risk to the patient, and (3) the recommended action. 7iii) Does the alert possess an instruction statement telling the user how to avoid the danger or the desired action?

D8. Does the alert include links to references and guidelines?	8i) Are the informational components needed for decision-making on the alert present either within or in close spatial and temporal proximity to the alert?
D9. Does the alerting system monitor whether alert recommendations are followed?	9ii) Is the system able to monitor and alert the user to follow through with corrective actions?
D10. Are users notified if alert recommendations are not followed?	9ii) Is the system able to monitor and alert the user to follow through with corrective actions?
E1. Does the alerting system inform users about the alerting algorithm/logic/formulas implemented within the system? E3. Does the alerting system inform users of the severity levels in use? E4. Does the alerting system provide an explanation to users on how severity is classified?	1i) Does the system provide a general catalogue of unsafe events, correlating the priority level of the alert with the severity of the consequences?
F1. Does the alert provide a function for the user to modify an order from within the alert?	2ii) If available, is the response to the alert, indicating the user's intended action (e.g. Accept, Cancel/Override), provided along with the alert, as opposed to being located in a different window or in a different area on the screen? 9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action?
F2. Does the alert provide a function for the user to discontinue the pre-existing order from within the alert?	2ii) If available, is the response to the alert, indicating the user's intended action (e.g. Accept, Cancel/Override), provided along with the alert, as opposed to being located in a different window or in a different area on the screen? 9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action?
F3. Does the alert provide a function for the user to cancel the new order from within the alert?	2ii) If available, is the response to the alert, indicating the user's intended action (e.g. Accept, Cancel/Override), provided along with the alert, as opposed to being located in a different window or in a different area on the screen? 9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action?
F4. Does the alert provide a function for the user to order further tests or request monitoring of patients from within the alert?	9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action? 9ia) If yes, does the alert utilise intelligent corrective actions that allow the user to complete a task?
F5. Does the alerting system allow users to delay an alert so that it can be actioned at a later time?	9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action? 9ia) If yes, does the alert utilise intelligent corrective actions that allow the user to complete a task?
F6. Does the alerting system allow users to forward the alert to another clinician?	9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action? 9ia) If yes, does the alert utilise intelligent corrective actions that allow the user to complete a task?
F7. Does the alerting system allow alert content to be directly entered into patient records?	9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action? 9ia) If yes, does the alert utilise intelligent corrective actions that allow the user to complete a task?
F8. Does the alerting system allow users to override the alert?	2ii) If available, is the response to the alert, indicating the user's intended action (e.g. Accept, Cancel/Override), provided along with the alert, as opposed to being located in a different window or in a different area on the screen? 9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action?
F9. Does the alerting system provide users with a list of override reasons to select from?	9i) Does the system possess corrective actions that serve as an acknowledgment of having seen the alert while simultaneously capturing the user's intended action?
F11. Does the alerting system allow users to update patient information from within the alert?	9ia) If yes, does the alert utilise intelligent corrective actions that allow the user to complete a task's

Appendix 3. Number of matches between TEMAS and I-MeDeSA according to I-MeDeSA' and TEMAS' dimensions.

		I-MeDeSA									Sum
		Alarm Philosophy	Placement	Visibility	Prioritisation	Colour	Learnability & confusability	Text-based information	Proximity of task components	Corrective actions	
TEMAS	Optimise the signal-to-noise ratio	-	-	-	-	-	-	-	-	-	0
	Support collaborative work	-	-	-	-	-	-	-	-	-	0
	Fit the clinicians' workflow and mental model	-	3	1	-	-	1	-	-	-	5
	Display relevant data within the alert	-	3	-	2	-	-	7	2	2	16
	Ensure the system rules are transparent to the user	3	-	-	-	-	-	-	-	-	3
	Include actionable tools within the alert	-	4	-	-	-	-	-	-	14	18
	Sum	3	10	1	2	0	1	7	2	16	42

Appendix 4. Usability features of I-MeDeSA and TEMAS. The plus and minus signs represent the valence of the features.

		I-MeDeSA		TEMAS		
Ease of use	Overall use	Speed of completion	+	"Quick to complete", "quick to use", "has fewer questions than TEMAS". It is "Useful because it is succinct and direct", "Accurate/concise" (n = 5)	-	"This tool has too many items", "very exhaustive (time consuming) questionnaires", "very long", "long questionnaire", "a lot of technical questions", "the questionnaire is relatively long", "but there are too many items", "a bit long", "the questionnaire is too long for my taste", "it's a big chunk TEMAS", "this one is long" (n = 11)
		Ease of use	+	"Simple, clear", "Relevance of items/simplicity", "Useful: its simplicity", "Useful: its ease of use", "Simple tool", "Simple to use, it's easy and quite clear" (n = 6)	-	"Its complexity makes the tool not useful", "it's too complex", "it's quite heavy to handle", "I found it too complex" (n = 4)
	Responding	Simplicity of answer	+	"Simplicity of the scoring: present/absent" (n = 1)		
		Adaptation of the response mode to the context	-	"Lack of a "partial/NA" checkbox", "For some items we had trouble giving a binary answer", "No option to put other than yes/no", "Add NA in the instructions!", "a binary answer is not appropriate in my opinion", "Is it okay to put 0.5?", "Is it possible to score NA?", Rereading instructions to see if it is possible to score NA, several groups scored NA (n = 8)	+	"The 'Partial' checkbox is quite useful here", "'partial' was quite handy", "and the 'partial' option is handy", "They put 'partial' when they didn't know too much, "when you don't really know" (n = 5)
		Intuitiveness of the response mode	-	"If we had a "yes/no" checkbox it would be easier, wouldn't it? It would avoid having to remember that 0 is no and 1 is yes" (n = 1)		
		Connection between the synthesis table and the items			-	"We don't remember the questions here", "It is difficult to come back to these points afterwards", "We should answer these questions at the end of each section", "It would have been good to put these boxes at the end of each section, we should have answered section by section because at the end we lost everything we had said before", "We lost what we said at the time", "the free text fields at the end (...) seem unnecessary at this point. Or pull them up at the end of each part (A,

						B, C etc) so it's "fresh" when responding", "These points should be at the end of each section, not at the end" (n = 7)
Understandability	Instructions for use		-	"The instructions are not clear" (n = 1)	-	
	Summary table				-	"The summary table of points to be clarified would require more explanation for its completion", "what should we do? should we fill it out?", "The expectations of free part on the points requiring particular attention should be better clarified" (n = 3)
	Items	Clarity	-	"Some terms unclear," "Ambiguous questions," "Some ambiguous questions," "Questions not clear enough for me to answer," "Very or even too technical questions," "Not all items in the tool are intelligible to me," "Some items are difficult to understand" (n = 7)	-	"Unclear term "clinician"", "Difficulty in interpreting items and examples", "items and examples difficult to understand", "rather technical vocabulary", "Some terms could be defined before doing the questionnaire to avoid ambiguities in the understanding of the items", "Questions sometimes ambiguous", Sometimes it was necessary to reread the question several times (n = 7)
		Simplicity	-	"Some questions too long, too complex", "Questions too complex", "Complexity of multi-component questions", "Very complex and sometimes too specific question", "Complexity of items", "Phrases used are complex to understand, making it difficult to answer", "Question wording sometimes long and complex" (n = 7)	-	"Sometimes the questions don't get to the point and make the use and understanding more complex than necessary" (n = 1)
					+	"Simple, clear and relevant questions", "Simple items", questions are clear, simple to understand, relevant to practice", "Very easy to understand" (n = 4)
	Examples	Help	+	"The idea of using examples is very good" (n = 1)	+	"Relevant examples", "[usefulness of] having an example for each item", "The examples provided sometimes help to answer", "The fact that there is an example is an important help", "there are twisted these questions, you have to read the example to understand" (n = 5)
					-	"Sometimes unhelpful examples", "sometimes we understand the item and the example differently. So some examples need to be revised in wording", "Revise examples to fit well with the item (and vice versa)", "[Items] with an often-inappropriate example that does not improve understanding", "Bad example, it misleads", "No on the item but yes on the example", "Yes for the item, No for example" (n = 7)
		Influence	-	"The examples presented with the questions sometimes focus the attention of the participants, who then think only about the application of the example to their alerting system and no longer about the application of the item" (n = 1)	-	"For some items, the example hides the item", "the examples are sometimes misleading", "the example directs the answer", "the example reduces the field of reflection", "The examples direct and reduce the field of answer [...] it helps but it directs", "The example influences the comprehension of the item. Without reading it, I understand the item a little

						differently", Participants say that they sometimes disregard the example so as not to concentrate only on it, "you have to ignore the example", Examples are sometimes too different from the items, "Influential examples", "But, in TEMAS, the examples can lead to losing sight of the question posed in the item. Examples can lead to digression" (n = 11)
--	--	--	--	--	--	---

Appendix 5. List of problems identified for I-MeDeSA's items (excerpts)

<p>1a. Does the system provide a general catalog of unsafe events, correlating the priority level of the alert with the severity of the consequences?</p>	<p><u>Understandability:</u> Participants did not understand the idea of "general catalog" and debated its meaning among themselves: "This is the general catalog that we don't quite understand", "is it a list of alerts?" (n = 3)</p>
<p>2ii. Is the response to the alert provided along with the alert, as opposed to being located in a different window or in a different area on the screen?</p>	<p><u>Understandability:</u> Participants are unsure of what "the response to the alert" means: "Response" is unclear ""What do they mean by 'response'? Description or course of action? To me it's the course of action but I don't know if that's what's implied" "They mean description rather than course of action?" "what is the alert response?" "we don't understand the question", "We did not fully understand what "alert response" was", "Question 2ii to be clarified: what do you mean by "response" to alert", "Notion "alert response" subject to interpretation (2ii)", "The way the question is phrased does not allow them to answer yes or no", "The way the sentence is phrased does not allow for a yes or no answer. How can you answer yes or no when the sentence makes two opposing propositions"? (n = 10)</p>
<p>2iii. Is the alert linked with the medication order by appropriate timing? (ie, a DDI alert appears as soon as a drug is chosen and does not wait for the user to complete the order and then alert him/her about a possible interaction)</p>	<p><u>Relevance:</u> Some of the alert systems evaluated do not alert the prescriber: "Our system alerts the pharmacist in real time and not the prescriber and is therefore not directly linked to the time of prescribing" (n = 3)</p> <p><u>Understandability:</u> The example is problematic for participants: "Ambiguous example" "The example is poorly chosen: we don't wait for the prescription to be completed to inform the user"(n = 2)</p>
<p>2iv. Does the layout of critical information contained within the alert facilitate quick uptake by the user? Critical information should be placed on the first line of the alert or closest to the left side of the alert box. Critical information should be labeled appropriately and must consist of: (1) the interacting drugs, (2) the risk to the patient, and (3) the recommended action. (Note that information contained within resources such as an 'infobutton' or link to a drug monograph does NOT equate to information contained within the alert.)</p>	<p><u>Understandability:</u> Participants have difficulty understanding the question and must reread it because of its length and technicality: "It's hard to understand the question because it's so technical" "Oh la la, there are too many words", "a bit cumbersome this wording" "it's long said like that" "the question is too long" (n = 7)</p>
<p>3i. Is the area where the alert is located distinguishable from the rest of the screen? This might be achieved through the use of a different background color, a border color, highlighting, bold characters, occupying the majority of the screen, etc.</p>	<p><u>Relevance:</u> Some of the alert systems evaluated are not integrated into other software: "Not concerned because our software only deals with alerts" "The question is not thought for our software" "Our alert software is independent of the computerized patient record". "There is nothing else on the screen other than the alert" (n = 4)</p> <p><u>Understandability:</u> Participants wondered what "screen" the question was about: "Which screen" (n = 1)</p>
<p>4v. In the case of multiple alerts, are the alerts placed on the screen in the order of their importance? The highest priority alerts should be visible to the user without having to scroll through the window.</p>	<p><u>Understandability:</u> Participants reread the question several times to understand it (n = 1)</p>
<p>5i. Does the alert utilize color-coding to indicate the type of unsafe event? (ie, drug-drug interaction (DDI) vs allergy alert)</p>	<p><u>Understandability:</u> Participants were hesitant about the interpretation of "unsafe event": "'type of unsafe event' is not very precise" (n = 1)</p> <p><u>Relevance:</u> Participants point out that they "don't have different types of events, [they] only have one" (n = 1)</p>
<p>6i. Are the different severities of alerts easily distinguishable from one another? For example, do major alerts possess visual characteristics that are distinctly different from minor alerts? The use of a signal word to identify the severity of an alert is not considered to be a visual characteristic.</p>	<p><u>Understandability:</u> Participants feel that this question is redundant with others (e.g., 4i): "I feel like it's the same question all the time" "We already had it earlier" (n = 2)</p>
<p>7ii. A statement of the nature of the hazard describing why the alert is shown. This may be a generic statement in which the interacting classes are listed, or an explicit explanation in which the specific DDIs are clearly indicated.</p>	<p><u>Understandability:</u> For participants, the wording of the question "is not clear". Some did not understand the phrase "nature of the hazard": "Not explicitly described. The first sentence is problematic: what is the definition of "nature of the hazard"?" Some participants changed their answer after reading the note associated with this question. (n = 3)</p> <p><u>Validity:</u> Participants discuss the merits of the recommendation. They were concerned that it would result in "a big block of text" in the alert window (n = 1)</p>
<p>7iii. If yes, does the order of recommended tasks reflect the order of required actions?</p>	<p><u>Understandability:</u> Participants have trouble understanding the question: "the question is twisted" "we have trouble understanding it" (n = 1)</p>
<p>8i. Are the informational components needed for decision making on the alert present either within or in close spatial and temporal proximity to the alert? For example, is the user able to access relevant information directly from the alert, that is, a drug monograph, an 'infobutton,'</p>	<p><u>Understandability:</u> Participants have difficulty understanding the question, especially "in close spatial and temporal proximity": "the question is not clear" "no I don't find it clear" "what does it mean 'in close spatial and temporal proximity'?" for me it is completely abstract, it leaves me confused", "Some items are not clear enough (ex: 8i)" (n = 4)</p>

or a link to a medical reference website providing additional information?	
9ia. If yes, does the alert utilize intelligent corrective actions that allow the user to complete a task? For example, if warfarin and ketoconazole are co-prescribed, the alert may ask the user to 'Reduce the warfarin dose by 33–50% and follow the patient closely.' An intelligent corrective action would be 'Continue with warfarin order AND reduce dose by 33–50%.' Selecting this option would simultaneously over-ride the alert AND direct the user back to the medication order where the user can adjust the dose appropriately.	<p><u>Understandability:</u> Participants felt that the question was too long: "There is too much information in this question". They have difficulty understanding the notion of "intelligent corrective actions": "What does "intelligent" mean here?" They find the example unclear and hiding the item: "I don't understand this example"; "It is not clear" (n = 6)</p> <p><u>Validity:</u> Participants questioned the validity of this recommendation: "Even if the system had this functionality, its systematic implementation would lead to dangerous situations. The singularity of the patient's situation is not taken into account. This functionality should be avoided." "The freedom of choice of a pharmaceutical intervention is not considered, which is very damaging" (n = 1)</p>

Appendix 6. List of problems identified for TEMAS' items (excerpts)

A3	Does the alerting system draw information from multiple sources to trigger alerts? The system draws information from patient records and laboratory results to trigger alerts.	<u>Understandability:</u> Participants thought that "information from multiple sources" referred to pharmaceutical reference databases (e.g. "Vidal", "Theriac") and not patient information. Some asked what sources were being referred to: "Multiple sources of what?" (n = 5)
A4	Does the alerting system overcome missing data and reconcile multiple entries to trigger relevant alerts? The alerting system uses allergy data from different sources for the same patient to trigger alerts	<u>Understandability:</u> Participants asked "but here what is missing data?" They hesitate in their interpretation: "I don't understand it that way" "I don't either" They found the example not concrete enough about what missing data is: "i.e. a concrete example?" Some participants understand "multiple entries" as having to enter data multiple times. A participant noticed that there was two question in one: "No for 'missing data' but Yes for multiple inputs because it is possible to combine multiple sources of information" (n = 7)
A5	Does the alerting system distinguish between past, current and future orders? The system only triggers drug-drug interaction alerts for orders that are active at the same time.	<u>Understandability:</u> Some participants felt that the example did not fully represent the item: "between the item and the example it is not the same thing: there is no notion of past or future prescriptions in the example" (n = 1)
A8	Does the alerting system consider all components of the order before triggering an alert? The system considers the medication dose, route, and duration before triggering an alert.	<u>Understandability:</u> Participants wondered what "all components" meant (n = 1)
A12	Does the alerting system consider the impact of the unsafe event on the patient when determining if an alert is triggered? The system triggers an alert when the unsafe event has the potential to cause harm to the patient.	<u>Understandability:</u> Participants did not easily understand the question: "I have to reread this!", "I don't really understand the question", "oh, you have to reread it", "I don't understand it". Some participants found this question redundant with A11: "Same as before? It is useless", "According to us it is redundant with A11" (n = 6)
A15	Are alert numbers, override rates, and override reasons reviewed by the organization? Alert information captured by the system are reviewed by the hospital.	<p><u>Understandability:</u> Participants had difficulty understanding the sentence: "the sentence is badly turned", "the wording of the question is not clear". Specifically, some participants questioned what "override rate" is: "is this the right term?", "I don't understand "override", "override" of drugs?", and they felt that "organization" could be more precise: the organization is the hospital? it should be written like that" (n = 9)</p> <p><u>Relevance:</u> Some participants found this question irrelevant to them because "[there is] no data extraction module [in their CDSS]" (n = 1)</p>
A17	Does the alerting system maintain existing customization following an upgrade? Drug-drug interaction alerts continue to be switched off after a system upgrade.	<u>Organization:</u> Participants felt that this question should be directly linked to A16 to fully understand it: "difficult example to understand, need to link to previous question" (n = 1)
B1	Does the alerting system allow multiple team members (i.e. doctors, nurses and pharmacists) to determine what alerts have been triggered? The system allows doctors, nurses and pharmacists to see triggered alerts for a patient.	<p><u>Relevance:</u> Some participants, because their alert system is only for pharmacists, found this question irrelevant to them: "In our case, the software is only for pharmacists" (n = 2)</p> <p><u>Understandability:</u> Some participants felt that this question is the same as A10: "Wait, this is the same question as before, isn't it? They would like clarification on the purpose of the question: "What is that? What is the difference with the question before?" (n = 3)</p>
B2	Does the alerting system allow multiple team members (i.e. doctors, nurses and pharmacists) to identify which alerts have been responded to? The system allows doctors, nurses and pharmacists to see what response has been made to an alert.	<u>Understandability:</u> some participants focused more on the response to the alert rather than the fact that multiple professionals can identify that response: "What response was made to an alert, right?" (n = 3)
B3	Does the alerting system allow multiple team members (i.e. doctors, nurses and pharmacists) to view responses to alerts? The system allows doctors, nurses and	<u>Relevance:</u> When the alert system is independent of the rest of the information system and is only intended for one type of professional, participants question the value of this item: "The alert system allows the

	pharmacists to see the override reasons documented for alerts.	validation loop by the pharmacist but is not accessible to other professionals" (n = 2)
B4	Does the alerting system trigger alerts to the appropriate team member? The system triggers medication administration alerts for nurses.	<u>Relevance:</u> When the alert system is independent of the rest of the information system and is only intended for one type of professional, participants question the value of this item: "the tool can only target pharmacists during pharmaceutical validation, not nurses or physicians" (n = 2)
B5	Does the alerting system display alert content differently for different users while maintaining core patient and alert information? The system presents more pharmacological data in alerts to pharmacists.	<u>Understandability:</u> Some participants had to reread the item to understand it: "wait I didn't understand anything there", "I always look at the example it's clearer" (n = 3)
C2	Does the alerting system display alerts instantly (i.e. no lag time)? The system triggers an alert during order entry.	<u>Understandability:</u> Some participants do not see the difference between this item and C1: "Ambiguous question and very similar to the previous one" "Almost similar question with C01" "The example is the same as the previous question" "The nuance is fine" "Questions C01 and C02 do not differ much and the nuance is difficult to understand" "The nuance is fine" "Questions C01 and C02 do not differ much and the nuance is difficult to understand", "the example reduces the field of reflection" (n = 6)
C4	Do alerts appear in a central location, over the CPOE/EMR screen? Alerts pop up in the middle of the computer screen.	<u>Relevance:</u> When the alert system is standalone, participants found the item not applicable: "it is not in the EHR", "It's stand-alone" (n = 6)
C5	Does the alerting system allow quick and easy responses to alerts? A single mouse click is required to respond to an alert.	<u>Relevance:</u> When the alert system does not allow for a response to the alert: "Not applicable because you can't respond to it: It's an informational system, you can't issue an action from it" (n = 1)
C6	Are users promptly returned to the appropriate stage of workflow after making a response to the alert? After overriding an alert, the user is allowed to continue with the order.	<u>Understandability:</u> Some participants found that "the item and the example don't go together" and that the example changes their understanding of the item: "I don't understand the item the same way anymore". The term "workflow" is not easily understood by everyone: "it is difficult to understand" (n = 3) <u>Relevance:</u> When the alert system is standalone, the item is not applicable: "Not concerned, not integrated in the prescription" "the question does not arise because we do not respond to the alert" (n = 3)
C7	Are alerts of different severity distinguishable from one another? Alerts of different severity are presented using different colours and shapes.	<u>Understandability:</u> Some participants felt that this item was a bit like A11: "I feel like it's been said before", "The examples direct and reduce the field of answer [...] it helps but it directs" (n = 2)
C9	Does the alerting system use interruptive alerts only for warnings of high severity? A pop up alert is triggered for duplication of controlled medications, but not for allergies that result in a minor rash.	<u>Relevance:</u> Some alert systems do not have interruptive alerts, participants found this item inappropriate: "it's a separate software, no alerts are interruptive" (n = 3) <u>Understandability:</u> Some participants found that the example does not fit well with the item: "the example does not agree with the item: which of the two should be answered?" (n = 1)
C14	Does the alerting system present less critical information on demand? Mechanism of interaction in a drug-drug interaction alert is accessed via a hyperlink.	<u>Understandability:</u> Participants had different understandings of the item: "info or critical alerts?" They had difficulty seeing the link between the example and the item: "the example is poorly done, is not related to the item", "the ex. is not good, I don't understand the ex." (n = 4)
C15	Does the alerting system use consistent terms, phrases, classifications, colours and definitions in all alerts? An allergy alert is always classified as severe, major, moderate or minor, and the warning message is always in red.	<u>Understandability:</u> Not all users understood the item in the same way at first reading. Above all, the fact that several elements are listed in the item made them wonder how to answer when all the elements are not homogeneous in their alerts: "Too many elements in the question so we put "partial"" (n = 2)
D1	Does the alert include information on why the alert was triggered? Medication names, dosages, and severity of interactions are included in drug-drug interaction alerts.	<u>Understandability:</u> Some participants found that the example does not correspond to the item: "I don't see the link between the example and the item", "Bad example, it misleads" (n = 4)
D2	Does the alert include information on the nature of the unsafe event and its likelihood? The drug-drug interaction alert explains that the interaction is highly likely to cause respiratory depression.	<u>Validity:</u> Participants explained that it is not possible to put the probability of the event occurring because of its dependence on the clinical context: "too difficult to establish the probability" (n = 1) <u>Understandability:</u> Some participants found that the example does not correspond to the item: "the example does not correspond" (n = 1)
D3	Does the alert use colour and a signal word to indicate severity of the unsafe event? High severity alerts are always in red and include the word 'warning'.	<u>Understandability:</u> The participants were embarrassed to answer with "and": "it is written "and" not "or", we don't have a color" (n = 4)
D4	Does the alert include information on the mechanism of the unsafe event? This should be presented on demand.	<u>Understandability:</u> Some participants found that the example does not correspond to the item: "the example does not fit" (n = 1)

	Mechanism of interactions in a drug-drug interaction alert is accessed via a hyperlink.	
D5	Does the alert include relevant patient information and provide a link for users to obtain further patient information? Patient lab results can be accessed by clicking on a link in the alert.	<u>Relevance</u> : Participants pointed out that the item is not adapted to all alert systems: "we are stand-alone, we have to go into the patient record to get the information" (n = 1)
D6	Does the alert provide clinically appropriate recommendations and suggest alternatives (i.e. drug, dose and frequency)? For two medications that may interact and lead to hypokalaemia, the alert recommends close monitoring of potassium concentrations.	<u>Understandability</u> : The "recommendations AND alternatives" seems to make the co-occurrence of the two items mandatory to answer yes: ""and" seems to make "reco" and "alternative" mandatory", "it implies that there must be both the recommendation and the alternative, if only one of the two info is present then you would have to answer no. This is a shame" (n = 1)
D7	In the case of multiple suggestions, does the alerting system display these in the order of importance? In an alert warning of a potential interaction between tacrolimus and diltiazem, the recommendation is to monitor the concentrations and effects (e.g. on renal function) of tacrolimus more frequently if diltiazem is started or stopped, and adjust the tacrolimus dose as necessary.	<u>Organization</u> : Some participants explained that if the answer to D6 is no, then D7 should not have to be answered: "if D06 = no, it's a shame to have to answer D07" "there should be something more ergonomics" (n = 2) <u>Understandability</u> : Some participants felt that "the example is too complicated and too long to be useful" (n = 1)
D10	Are users notified if alert recommendations are not followed? The system notifies the doctor if a dose reduction has not been actioned.	<u>Organization</u> : Some participants explained that if the answer to D9 is no, then D10 should not be answered: "it is not applicable because we answered no to the previous one" "Not applicable because we said no to D9; they are strongly linked" (n = 2)
E3	Does the alerting system inform users of the severity levels in use? Clicking on a 'more information' link in the help page informs the users that 'contraindicated' is the highest severity level of drug-drug interaction alerts triggered by the system.	<u>Understandability</u> : Some participants had difficulty understanding the example: "The wording of the question is clear but the example is not", "Wait, I don't understand". Some think that the question has already been asked, referring to A11 and C7: "It feels like it's repeating itself", "We've been asked this before", "The difference is why was it coded like that?" (n = 5)
E4	Does the alerting system provide an explanation to users on how severity is classified? Clicking on a 'more information' link in the help page explains why the interaction between allopurinol and azathioprine is classified as 'contraindicated'.	<u>Understandability</u> : Participants found this question similar to E3: "is it the same question as before?", "E03 and E04 the difference is not obvious" and some found it hard to understand: "Wait but I don't understand the question actually", "it becomes philosophical". The example lost some participants a little more: "It's crap their examples", "The example is not clear", "it loses me", "No on the item but yes on the example" (n = 8)
E5	Does the alerting system inform users of what data are used to trigger alerts? Clicking on a 'more information' link in the help page informs the user that patient information and the Multum drug database are used to trigger drug-drug interaction alerts.	<u>Understandability</u> : Some participants felt that this question is the same as A3", "Yes on the item, No on the example", "The example influences the comprehension of the item. Without reading it, I understand the item a little differently" (n = 3)
E6	Does the alerting system inform users of the types of orders that will trigger alerts? Clicking on a 'more information' link in the help page informs the user that both order sentences and free text orders can trigger alerts.	<u>Understandability</u> : Some participants did not understand the question: "I don't understand the question" "What does the question mean?" "You have to read the example to understand", "The question is twisted, you have to read the example to understand it". The expression "types of orders" is problematic: "a type of prescription? it doesn't make sense" "the question doesn't say what a type of prescription is, it's vague", "E06: The definition of 'type of prescription' is missing" (n = 7)
F1	Does the alert provide a function for the user to modify an order from within the alert? The user can modify the dose of a medication by clicking the 'modify' button within the alert.	<u>Relevance</u> : For alert systems that are not linked to a prescription software, participants found this question irrelevant "we only read the alert, we do nothing from the alert", "our system is independent, it is not linked to a prescription software, we cannot do anything" (n = 2)
F5	Does the alerting system allow users to delay an alert so that it can be actioned at a later time? The user is able to delay the triggering of an allergy alert for 24 hours so that the patient can be monitored before making a decision on whether to continue with the medication.	<u>Validity</u> : Participants discussed the value of the "delay alert" feature: "I don't see the purpose of such a feature" (n = 1)
F9	Does the alerting system provide users with a list of override reasons to select from? The user selects 'patient tolerates drug combination' from a drop-down list after overriding a drug-drug interaction alert.	<u>Understandability</u> : Participants found this question very close to F8: "same question as on the reasons just before" (n = 1)

Appendix 7. Perceived usefulness of I-MeDeSA and TEMAS (complete).

	I-MeDeSA	TEMAS
--	----------	-------

Identification of shortcomings and areas for improvement	+	"Detailed description of what would be expected", "The tool allows, despite its shortcomings, to identify the major shortcomings/defects and to draw up areas for improvement", "Highlighting of the tool's shortcomings and highlighting of the improvements to be made", "Useful for identifying ways of improving the CDSS, identifying what the software does or does not do", "Ideas for development/improvement", "Allows us to ask ourselves new questions that will allow us to improve the tool", "Highlights the tool's weaknesses" (n = 7)	"Identify areas for improvement", "Identify strengths and weaknesses (what the software does not do and what we would like it to do)", "However, it helps to improve the tool in the end", "Final synthesis allows us to identify the strengths and weaknesses of the evaluated tool", "Highlights the failures of our CDSS", "Realize the points of improvement to be considered", "Reflection on the missing functionalities", "Questioning on the possibilities of improvement of our CDSS, especially the items and balance sheet", "Easy identification of the items to be improved", "Gives ideas of evolutions to propose" (n = 10)
Raising awareness of good practices	+		"The questions make us aware of what we can expect from the software", "More educational (ideal idea)" (n = 2)
Exploring the alert system	+	"Useful because it allows you to explore different aspects of the alert system", "Allows you to evaluate the CDSS as a whole", "Investigates the different functionalities of a CDSS" (n = 3)	"Useful because it reviews all aspects of CDSS" (n = 1)
Not suitable for all alert systems	-	"Several questions not applicable to our CDSS" (n = 1)	"Several items do not correspond to the functionalities of our CDSS", "Several items only apply to the CDSS integrated in other EHR/CPOE", "Be clearer in asking the question: is the decision support system integrated with the Electronic Patient Record or is it an independent application?" (n = 3)

Appendix 8. Perceived added value for I-MeDeSA and TEMAS (complete)

I-MeDeSA	
Shorter and simpler (n = 21)	"Even though I-MeDeSA questions are clearer", "[I-MeDeSA] Simpler, easier to use", "[I-MeDeSA] Simpler", "[I-MeDeSA] Simpler, quicker", "[I-MeDeSA] Simpler", "[I-MeDeSA] Quicker to fill out", "[I-MeDeSA] Quicker to use", "[I-MeDeSA] Simpler/quicker tool", "[I-MeDeSA] It's shorter", "[I-MeDeSA] Quicker", "[I-MeDeSA] Quicker to fill out", "[I-MeDeSA] Shorter; allows for more recurrent use", "[I-MeDeSA] Easier to use than TEMAS", "I-MeDeSA, I found it simpler" "At first glance [TEMAS] is longer so it can be discouraging", "TEMAS system heavier", "[TEMAS] is longer", "[TEMAS] Heavier and longer", "Despite its complexity, TEMAS...", "[TEMAS] longer", "TEMAS too complex, too long"
More understandable (n = 8)	"[I-MeDeSA] Easier to access and understand," "[I-MeDeSA] easier to understand questions," "[I-MeDeSA] easier to understand", "[I-MeDeSA] More understandable," "[I-MeDeSA] Simplicity and lack of ambiguity in wording." "TEMAS has items that are sometimes not very explicit despite the examples", "Items and examples should be less ambiguous", "Item categories are more meaningful to me [in TEMAS] than those of I-MeDeSA"
Theme more precise (n = 8)	"More precise on the "interface", but maybe too much? ", "Presence of items on the front or the rules of legibility (mix of lower / upper case) that we do not find on TEMAS", "Tools (...) that highlights certain usability items", "Goes into much more detail regarding the display of alerts (shapes, colors)", "It allows a better approach to the user experience and the ergonomics of a system", "Very focused on the ergonomics of using a CDSS", "Focused on the hierarchy of alerts and the readability of this hierarchy", "Questions about the visual aspect and not only about the content"
Conditional items (n = 1)	"You see, there are questions that should not be answered: it depends on the previous question"
Score (n = 1)	"the score"
TEMAS	
More complete and relevant themes (n = 44)/ too precise I-MeDeSA theme (n = 12)	TEMAS: "Takes into account a greater number of aspects", "More comprehensive", "More complete and logical", "More items explored by TEMAS and more relevant questions", "More complete to review the critical points of a CDSS", "More exhaustive in its functions", "Because it is more complete and allows to evaluate the CDSS in its entirety and to target points of improvement more easily", "The global approach of the tool makes it possible to identify points of improvement on several aspects of the CDSS", "Despite its complexity, the TEMAS questionnaire addresses many points and focuses less on ergonomics", "TEMAS is

	<p>better adapted to the diversity of situations, more relevant and clearer in its questions", "The field of items covered is broader", "Seems to be more complete, broader in the evaluation of the use, of the parameterization", "I find the grid more complete", "More complete tool", "More extensive", "TEMAS addresses a lot of points", "Finally more exhaustive than I-MeDeSA", "TEMAS because it is more generalist and scans more items/sections", "More complete, more different items", "Items have a greater cross-section", "Much more complete", "More generalist, more functionalities", "Because it provides a comprehensive approach to CDSS", "TEMAS because more comprehensive than I-MeDeSA", "More items, questions", "more comprehensive", "addresses more items and therefore better source for thinking about a rigorous procurement process", "More comprehensive", "More comprehensive tools", "more comprehensive (different items)", "Items are more comprehensive", "I feel like it covers a wider range of items to incorporate into a CDSS. I would say it allows for more relevant developments to be suggested that would not have been spontaneously thought of", "More complete tool targeting more points of a CDSS in a precise way", "It is more exhaustive and more systemic in its approach", "More exhaustive tool and therefore more useful to identify the improvement areas of the CDSS", "More questions therefore allows to dig deeper on the points to be improved", "encompasses more situations", "greater diversity in items", "more complete and therefore more accurate assessment of the software's characteristics", "goes further, particularly in the collaborative aspect and prescription" "TEMAS (...) is more complete than I-MeDeSA" "TEMAS is more complete. It questions several aspects of a CDSS. Even if the use of a CDSS in a hospital environment is complex, this tool does not focus on the user experience and addresses almost all the topics", "Projecting into the use in routine care in collaborative work mode and workflow", "Targets more items and goes deeper in the analysis of items"</p> <p>"[I-MeDeSA has] Less exhaustive questions, too focused on the same parameters of the tool : many items on the visual aspect", "Impression that I-MeDeSA is mainly based on the visual aspect, not always on the substance", "I-MeDeSA is essentially interested in the CDSS ergonomics on the user's side", "I-MeDeSA is too focused on ergonomics", "the tool [I-MeDeSA] is too much centered on ergonomics and not on the patient's care", "[I-MeDeSA] is (...) less detailed", "[I-MeDeSA] Improvements mostly in form", "[I-MeDeSA] nothing about prioritization apart from colors", "I-MeDeSA only talks about appearance, that's not enough", "[I-MeDeSA] functionality points not addressed", "[I-MeDeSA] doesn't cover all fields around the use of the system settings", "There are things that are not integrated in I-MeDeSA"</p>
<p>More precise, detailed (n = 28)</p>	<p>"More detailed items", "More detailed examples", "more specific example, "More detailed tool", "But it is more detailed", "Much more detailed", "More detailed", "More detailed items", "More detailed", "More detailed tool to see the improvements and expectations of a software", "In the end the items are more detailed and it is easier to see the improvements and expectations of a software". "detailed", "more detailed", "more precise items", "more details", "more precise tool allowing to see the improvements and the expectations of a software", "in the end the items are more detailed and that allows in my opinion a better evaluation of the software (one approaches more points and a little more in detail)", "more logical and more precise", "TEMAS presents more detailed items", "goes into more detail, especially on the content of the alerts", "a little more precise", "more detailed tool", "more precise, more detailed", "more thorough analysis", "With more precise questions, the points to be improved appear more clearly", "[usefulness of] points raised by the precision of the questions", "Detailed items", "Very detailed items", "it is more precise"</p>
<p>More understandable (n = 19)</p>	<p>TEMAS: "Clearer tool", "Questions are better turned", "Examples are more meaningful", "Items are more synthetic, I find them more understandable especially since they are accompanied by examples", "Items are easier to understand", "Examples are more relevant", "Items are easier to understand", "Examples help clarify items", "Items are more understandable", "Items are easier to understand", "Examples allowing the clarification of the items", "The items are more (...) understandable", "items more (...), easy to understand", "the questions are less open to interpretation", "with more telling examples", "Questions less long and easier to understand", "Items more (...) simple to understand"</p> <p>"[I-MeDeSA] is more complicated to understand. The sentences used are complex/more difficult to understand than those in TEMAS. The examples given by I-MedeSA are confusing: they stop the thought on a particular character", "The sentences given in I-MedeSA are complex, sometimes difficult to understand"</p>
<p>Better synthesis (n = 15)</p>	<p>"Synthetic", "The final summary", "Makes you think about the points of attention to be made", "Allows a final synthesis of the evaluation", "TEMAS, What saves it (a little) is "the points needing attention", "final synthesis allowing to prioritize the improvements to be made", "the final synthesis table can be used to identify the areas of improvement of a CDSS", "because of the points of attention, makes you think about the points of attention to be made", "Good synthesis of the evaluation of a system", "The points requiring particular attention +++", "the final synthesis with the prioritization", "Final summary table very relevant to highlight the strengths and weaknesses of the tool", "The final synthesis is useful for prioritizing needs", "the final table allows us to take stock and prioritize the items to be reworked", "allows us to take stock of the use of our CDSS"</p>

More precise response mode (n = 9)	"the answer is finer", "TEMAS allows a less clear-cut answer to the items", "I find the grid (...) more precise, especially with the examples", "Not binary", "The possibility of answering "partial" and commenting is great", "Possibility of nuance with partial answer", "The option "partial" helps", "Possibility of a "partial" response", "the big advantage over I-MeDeSA is when you have several types of alerts, you can nuance by partial"
Simpler and more concrete (n = 7)	"TEMAS system (...) less complex", "TEMAS tool is easier to use", "Easier to use because more concrete", "easier to use", "Simplicity of use", "Less complex questions", "more concrete"
More consistent (n = 1)	"Less inconsistencies"

Appendix 9. Contexts of use envisaged by the participants for I-MeDeSA and TEMAS.

		I-MeDeSA	TEMAS
Developing or improving a CDSS		"Development of CDSS", "To improve the graphic presentation of the software", "Improvement of prescription software".	"Software improvement roadmap", "CDSS development", "Development to improve the software".
Supporting the choice of CDSS	Choosing	"Choosing a CDSS", "Evaluation of prescription software, e.g., adaptation of questions for the implementation of any new software".	"Comparison of CDSS", "Comparison of two systems", "Choice of a prescription support software", "CDSS evaluation in the context of a change of EHR", "Purchase of a new CDSS", "CDSS choice"
	Writing specifications	"Specifications for a new CDSS ", "Purchase of a new CDSS ", "Constructing an ideal CDSS' specification", "In CDSS tenders"	"For CDSS tenders"
Evaluating a CDSS	Overall	"To evaluate aids, e.g., project on prescription assistance for obese patients",	"Evaluation of CDSS integrated within the DPI", "Evaluation of software", "Evaluation of CDSS", "For example to evaluate the use of the new chemotherapy software", "Analysis of information technology applied to care systems"
	Evaluating improvements after reengineering	"Evaluation of a redesign of the alert system", "to evaluate new versions of the CDSS software"	"Evaluation of a redesign of the alert system", "Comparison (...) of the versions of our CDSS", "Evaluation of future versions of the tool (upcoming development) as a milestone"