BMJ Open

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

Romaric Marcilly [1,2] Wu Yi Zheng,[3] Regis Beuscart,[1] Melissa T Baysari [4]

¹Univ. Lille, CHU Lille, ULR 2694 - METRICS: Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France
²Inserm, CIC-IT 1403, F-59000 Lille, France
³Black Dog Institute, Randwick, New South Wales, Australia
⁴The University of Sydney, Faculty of Medicine and Health, School of Medical Sciences, Biomedical Informatics and Digital Health, Sydney, New South Wales, Australia

Correspondence to
Dr Romaric Marcilly;
romaric.marcilly@univ-lille.fr

ABSTRACT

Introduction Research has shown that improvements to the usability of medication alert systems are needed. For designers and decisions-makers to assess usability of their alert systems, two paper-based tools are currently available: 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 aims to compare the validity, usability and usefulness of both tools to identify their strengths and limitations and assist designers and decision-makers in making an informed decision about which tool is most suitable for assessing their current or prospective system.

Methods and analysis First, TEMAS and I-MeDeSA will be translated into French. This translation will be validated by three experts in human factors. Then, in 12 French hospitals with a medication alert system in place, staff with expertise in the system will evaluate their alert system using the two tools successively. After the use of each tool, participants will be asked to fill in the System Usability Scale (SUS) and complete a survey on the understandability and perceived usefulness of each tool. Following the completion of both assessments, participants will be asked to nominate their preferred tool and relay their opinions on the tools. The design philosophy of TEMAS and I-MeDeSA differs on the calculation of a score, impacting the way the comparison between the tools can be performed. Convergent validity will be evaluated by matching the items of the two tools with respect to the usability dimensions they assess. SUS scores and answers to the survey will be statistically compared for I-MeDeSA and TEMAS to identify differences. Free-text responses in surveys will be analysed using an inductive approach.

Ethics and dissemination Ethical approval is not required in France for a study of this nature. The results will be published in a peer-reviewed journal.

INTRODUCTION

Evaluating the usability of medication alert systems

Medication alerts aim to provide relevant information and recommendations to clinicians to help them prevent and manage iatrogenic risks, like drug–drug interactions, or drug overdoses. Yet, the recommendations

Strengths and limitations of this study

► This study is novel, as the two tools for assessing the usability of medication alert systems, the instrument for evaluating human-factors principles in medication-related decision support alerts and the tool for evaluating medication alerting systems, have not been compared before.

► The protocol is rigorous, including both quantitative and qualitative outcomes.

► The study is limited by differences in the tools' design and philosophy, making direct comparison of some outcomes (eg, reliability) impossible.

proposed by alerts are typically not accepted or followed by prescribers.[1 2] Worse, inappropriate alerts can lead to automatic override behaviour and clinicians' exhaustion due to alert fatigue,[3] resulting in errors in interpretation, or outright rejection of the alert system.[4] Alert systems have been the subject of many studies to understand how and why alerts are not effective in changing prescribing decisions. One explanation lies in the usability of these systems: alerts with good usability are associated with better user satisfaction,[5] faster work, fewer prescription errors and less workload for clinicians compared with alerts with poor usability.[6] Usability design principles must therefore be taken into account when designing, evaluating or choosing alert systems.

Usability knowledge must be made available and accessible to support the development of usable alert systems or to assist decision makers during a procurement process. Even if usability design principles based on experts' consensus[7 8] or on evidence[9] are available in the literature, they may not be accessible, understandable or useful for designers and decision makers with little or no knowledge of human factors. To overcome this problem, tools for verifying compliance with human

factors principles have been developed. To date, two paper-based tools have been developed and are available in the literature. These are presented below.

Instrument for evaluating human factors principles in medication-related decision support alerts (I-MeDeSA)

I-MeDeSA was developed to assess the compliance of drug–drug interaction alerts in electronic health records (EHRs) using human factors design principles identified by Phansalkar et al[8] and enhanced with feedback from users.[10] After validation, it was used to evaluate and compare the adherence of alerts to human factors design principles in several EHRs in the USA,[11] before being translated and evaluated in the Korean context,[12] and then in the Australian context.[13 14] Recently, I-MeDeSA was also extended to enable assessment of a broader range of alerts.[15]

In I-MeDeSA, alert systems are assessed against nine human factors principles via the completion of 26 items: alert philosophy, placement, visibility, prioritisation, colour, learnability and confusability, text-based information, proximity of task components being displayed, and corrective actions. An item is scored '1' if the characteristic it describes is present in the alert system, and '0' if absent. Therefore, the maximum possible score for an alert system is 26, representing a high level of adherence to human factors principles.

The intended audience of I-MeDeSA is not specified in the original papers. However, if intended to be used by designers with little or no knowledge of human factors as well as for institutional decision-makers when selecting and purchasing an alert system, the I-MeDeSA may be difficult to use, as no instructions on how to complete the assessment are included. In the literature, the instrument has been used individually by evaluators mainly with a background or expertise in human factors or usability.[10 11 13–15] In one paper, I-MeDeSA's users were all medical informatics experts.[12]

Several problems have also been highlighted with the tool: definitions are not precise enough, rationale for the items is not clear, insufficient examples are included which have led to differences in the interpretation of the items,[12 13] and the use of conditional items also over-penalises some systems.[13 14] These shortcomings make it difficult to take advantage of the design principles I-MeDeSA incorporates.

Tool for evaluating medication alert systems (TEMAS)

To address the shortcomings identified in I-MeDeSA, Zheng et al[16] developed a tool to assess the usability of medication alerts using a list of evidence-based usability principles specific to medication alert systems: the TEMAS. The designers' objective was to develop a tool that could be used by hospital staff with extensive knowledge of the prescribing software and the alert system (eg, computerised physician order entry (CPOE) pharmacist) to evaluate their system. The tool can also be used during the procurement process to guide the purchase of a usable alerting system.

TEMAS is composed of 66 items, adapted from 60 evidence-based design principles,[9] organised into six sections: optimise the signal-to-noise ratio, support collaborative work, fit the clinicians' workflow and mental model, display relevant data within the alert, ensure the system rules are transparent to the user, and include actionable tools within the alert.

In contrast to I-MeDeSA, TEMAS was not developed with the intention of generating a usability score. Indeed, depending on the context of use and the system in use, the importance of each item could change; thus, weighting items and calculating a score would be arbitrary. Instead, TEMAS is intended to be used as a checklist to identify the usability-related strengths and weaknesses of an alert system. Since knowledge about the alert system is often shared between several experts, TEMAS is explicitly intended to be completed by a group of local experts of the alert system and not by a single person.

A preliminary evaluation of TEMAS was recently performed[16] and this led to the identification of a number of confusing items. As a result, 22 items were revised to reduce ambiguity.

Objectives

In the hospital setting, compliance of an alert system with human factors design principles can be evaluated at different points in the life of the alert system: for example, when comparing candidates in a procurement process, or following an upgrade of the alert system to assess changes. Both I-MeDeSA and TEMAS are available tools for assessing compliance of alert systems with human factors design principles but they have been designed for different purposes: by using a score, I-MeDeSA is intended to be normative, whereas TEMAS is more of a formative approach, enabling the strengths and weaknesses, and areas for improvement of alert systems to be identified. Thus, the usefulness of these tools may depend on the context in which compliance of the alert system with the human factors design principles is being sought.

This study aims to compare I-MeDeSA and TEMAS in order to identify their strengths and limitations and help end-users identify the most suitable tool for their current or prospective alert system assessments. The study will be carried out in France. Therefore, both tools need to be translated and validated in French prior to performing a comparison.

► Aim 1: validate the translation of TEMAS and I-MeDeSA.
► Aim 2: compare TEMAS and I-MeDeSA in their ability to evaluate the usability of an alert system. For this objective, we aim to estimate the convergent validity of both tools. In parallel, we will investigate the perceived usability and usefulness of both tools, their perceived strengths as well as users' preferences for the tools.

METHOD AND ANALYSIS

Patient and public involvement

No patients or members of the public were involved in the design of this protocol. No patients or members of the public will be involved in conducting this study nor in the reporting and dissemination plans.

Recruitment of study sites

Hospitals (public and private) will be included in the study without size criterion. The only inclusion criterion is that hospitals must have a medication alert system either as part of their EHR/CPOE system or a standalone system. If an organisation has several alert systems in use, they will be the subject of separate evaluations.

Study sites will be contacted by emailing and/or calling persons identified as responsible for the alert system. There will be no financial reimbursement for participation, but the hospital will be free to use the evaluation results of their alert system.

Materials for the study (instructions, tools, etc) will be sent by mail and email.

A minimum of 12 study sites will be recruited (see below for explanation).

Study design

The translation of the tools into French will be done prior to and independently to the evaluation of TEMAS and I-MeDeSA. The original version of I-MeDeSA will be used.[10]

For the evaluation phase, at each site, a group of hospital staff members identified as 'experts in the alert system' will be invited to participate. First, the name and version of the alert system(s) in use (or CPOE/EHR) will be collected along with the background and role of each participant. Then, participants will use both tools to assess their alert system. The running order of the tools will be counterbalanced to prevent an order effect. Both TEMAS and I-MeDeSA will be completed as a group. At each site, the same participants will use both I-MeDeSA and TEMAS to evaluate their alert system.

After the completion of each assessment, participants will individually complete the System Usability Scale (SUS)[17] along with a set of questions about the understandability and usefulness of the tool. After the completion of both assessments, another set of questions will ask participants about their preferred tool and added value of each tool. This survey will also be answered individually.

Aim 1: validate the translation of TEMAS and I-MeDeSA

The translation will be performed by a human factors researcher with knowledge and experience in medication alert systems. The validity of the translation will be assessed by asking three human factors specialists fluent in French and in English, and with expertise in alert systems to express their level of agreement with the translation of each item ('I disagree', 'rather disagree', 'rather agree', 'I agree').

For each tool, an inter-rater agreement score will be calculated (Krippendorf's α,[18]) grouping the negative categories on the one hand and the positive ones on the other. Items for which 'I disagree' 'rather disagree' or 'rather agree' categories have been selected will be discussed in a group to improve their translation until an agreement is reached.

Aim 2: compare TEMAS and I-MeDeSA in their ability to evaluate the usability of an alert system

Convergent validity: do TEMAS and I-MeDeSA measure the same usability dimensions?

To assess whether both tools evaluate the same dimensions of usability, the convergent validity of both tools will be estimated. Convergent validity is usually estimated through a correlation coefficient.[19] However, TEMAS' design philosophy does not include allocating a score to an alert system, preventing a correlation coefficient from being calculated.

Instead, we will determine whether the usability concepts captured by both tools are the same. First, an expert in the field (first author) will match TEMAS' items (n=66) to I-MeDeSA's items (n=26) with respect to the dimension of usability they assess. The level of detail included in the items across tools may differ, resulting in a single item in one tool being mapped to multiple items in another. A second expert (last author) will then perform this task independently, blind to the results of the first expert's matching, and the mapping between experts will be compared. An agreement score will be calculated (Krippendorf's α), and the disagreements will be discussed until a consensus is reached. Any remaining disagreements will be resolved through discussion with a third expert (second author). This mapping will make it possible to identify the dimensions of usability that are common to both tools and those that are specific to each.

Perceived understandability, usability and usefulness of the tools and user preference for a tool

While using the tools (I-MeDeSA and TEMAS), participants will have an opportunity to record any potential difficulties they faced in understanding each item through free-text fields. These comments will be reviewed and analysed to identify potential difficulties and discrepancies in the interpretation of items.

After completing each assessment with a tool, participants will individually complete a short questionnaire on the perceived usability of the tool.

After the use of each tool

After the use of each tool, an adaptation of Brookes' SUS[17] will be individually completed by the participants. The original version of the SUS was designed to provide a quick assessment of perceived usability of interactive systems. TEMAS and I-MeDeSA are not interactive tools but rather static documents. Therefore, the wording of the items has been adapted from the rewording proposed by Perrier et al[20] and Grudniewicz et al[21] (cf. table 1).

Table 1 Adaptation of the System Usability Scale (SUS) to the context of the evaluation of the tool for evaluating medication alerting systems (TEMAS) and the instrument for evaluating human factors principles in medication-related decision support alerts (I-MeDeSA)

#	Original version	Adapted version
1	I think that I would like to use this system frequently	I think that I would like to use TEMAS/I-MeDeSA when I need to assess the usability of an alert system
2	I found the system unnecessarily complex	I found TEMAS/I-MeDeSA unnecessary complex
3	I thought the system was easy to use	I thought the TEMAS/I-MeDeSA was easy to use
4	I think I would need the support of a technical person to be able to use this system	I think that I would need support to be able to use TEMAS/I-MeDeSA
5	I found the various functions in the system were well integrated	I found the various features of TEMAS/I-MeDeSA (ex: titles of sections, items, response fields, etc) were well integrated
6	I thought there was too much inconsistency in this system	I thought there was too much inconsistency in TEMAS/I-MeDeSA
7	I would imagine that most people would learn to use this system very quickly	I would imagine that most people would learn to use TEMAS/I-MeDeSA very quickly
8	I found this system very awkward to use	I found TEMAS/I-MeDeSA very cumbersome to use
9	I felt very confident using the system	I felt very confident using TEMAS/I-MeDeSA
10	I needed to learn a lot of things before I could get going with this system	I needed to learn a lot of things before I could get going with TEMAS/I-MeDeSA

As I-MeDeSA and TEMAS are not intended to be used 'frequently', the wording of item #1 has been modified to read 'I think that I would like to use TEMAS/I-MeDeSA when I need to assess the usability of an alert system'. Minor changes in the wording of items have previously been shown to have no impact on the resulting scores of the SUS.[22]

The SUS does not need a large sample size to provide reliable information on perceived usability. It has been shown that samples as small as 12 or 14 participants yielded the same conclusion as much larger samples.[23] Therefore, it was decided to target a sample size of 12 study sites, with several respondents recruited at each site.

To gain a deeper understanding of the perceived usability and usefulness of both tools, four questions (including one optional) were added after the SUS:

1. 'The items in TEMAS/I-MeDeSA' are easy to understand': response on a 5-point Likert scale ranging from 'I disagree' to 'I agree'
2. 'I found TEMAS/I-MeDeSA useful to identify possible improvements to the alert system': response on a 5-point Likert scale ranging from 'I disagree' to 'I agree'
   What made TEMAS/I-MeDeSA useful/not useful?' (free text)
3. 'I would like to use TEMAS/I-MeDeSA in future projects': response on a 5-point Likert scale ranging from 'I disagree' to 'I agree'
   'In what kind of projects?' (free text)
4. Optional: 'Additional comments on the TEMAS/I-MeDeSA' (free text)

The SUS score will be calculated from all participants' responses for both tools and compared with the Bangor *et al*'s scale.[24] Then, inferential statistics (paired t-test) will be performed on each item (SUS and added items) and on the overall SUS score to determine the differences in perceived usability and usefulness of both tools. For each open-ended question and for the optional comment, meaningful semantic units, that is, sets of words representing a single idea that is sufficiently self-explanatory for analyses, will first be extracted from each response. Then, two experts with a background in human and social sciences and fluent in French (first and third author) will independently assign each unit a code that they will name and define. The resulting codes will be discussed by the two experts together until agreement is reached on a clear, unambiguous and exhaustive classification scheme whose codes are mutually exclusive and where each code represents a meaningful dimension. In case of persistent disagreement, a third expert will help with the decision-making. Finally, the results will be presented according to the frequency of themes raised by the participants. This analysis of free-text comments will allow us to explore the limitations and/or advantages of the tools with respect to the contexts (and systems) in which they are being used.

After the use of both tools

After the completion of both tools, participants will be asked four questions to determine their preference for a tool and their opinions on the tools.

1. In a situation where I would have to compare alert systems (eg, to make a choice among several systems during the procurement process), I would prefer to use: I-MeDeSA or TEMAS (choose one option only)
   Why? (free text)
2. When identifying an alert system's weaknesses and areas for improvement, I would prefer to use: I-MeDeSA or TEMAS (choose one option only)
   Why? (free text)
3. What is the added value of TEMAS compared with I-MeDeSA? (free text)
4. What is the added value of I-MeDeSA compared with TEMAS? (free text)

Answers to the dichotomous questions will be analysed by inferential statistics (one-sample binomial test) to see if the proportion of preferences for TEMAS or I-MeDeSA differs significantly from 0.5. For each open-ended question, meaningful semantic units will be extracted from each participant's responses. Topic codes will be developed by two independent experts (first and last author) before being discussed to establish a clear and unambiguous classification scheme, with mutually exclusive codes and good internal consistency. A third expert (second author) will settle disagreements. The results will be presented according to the frequency of occurrence of the codes.

Ethics and dissemination
Compliance with ethical standards
This study is a human and social science study. The French law governing 'research involving the human person' exempts human and social science studies from requiring approval from an ethics committee.[25] For practical reasons, the data being collected are identifiable but will be deidentified for data analysis. This study is therefore subject to the European GDPR law[26]: a declaration of the protocol to the French National Commission on Informatics and Liberty has been made. Written informed consent will be obtained from each participant before they take part in the study. All participants will be recruited on a voluntary basis without any financial reimbursement provided.

Limitations and considerations
The purpose of this study is to compare the convergent validity, the usefulness and the usability of two tools available in the literature that are developed to assess the usability of medication alert systems. As the two tools have a different philosophy regarding their method of scoring (ie, calculation of a score for I-MeDeSA, not for TEMAS), a direct comparison of these tools on all measures is not possible (cf. table 2). However, whenever a proven method could not be applied (eg, convergent validity correlation), another suitable approach has been proposed (eg, coverage of usability domains).

Table 2 Difference in the evaluation of the instrument for evaluating human factors principles in medication-related decision support alerts (I-MeDeSA) and the tool for evaluating medication alert systems (TEMAS)

	I-MeDeSA	TEMAS
Translation validity	Krippendorf's α and consensus	
Convergent validity	Correlation of scores not applicable: matching the coverage of usability concepts	
Perceived understandability, usability, usefulness and preference	SUS and questions on the understandability and the usefulness of the tools, and on the preference and opinion of the participants	

SUS, System Usability Scale.

The three experts who performed the mapping between the TEMAS and I-MeDeSA items worked to different degrees on the development of TEMAS. To reduce this bias, the two experts performed the first step of the mapping independently, without consulting each other.

Dissemination
Usability assessment results of alert systems using I-MeDeSA and TEMAS will be made available to all participating sites to use as they see fit. In addition, results of the comparison between TEMAS and I-MeDeSA will be published in a peer-reviewed journal indexed in MEDLINE. This will make it possible for system vendors, researchers and health services to select and adopt an appropriate evaluation tool for their purpose and organisation, with the ultimate aim of improving the usability and the usefulness of medication alert systems.

Key milestones and timeframe
The first stage of the study, the translation of I-MeDeSA and TEMAS into French as well as its validation is anticipated to take 2 months beginning February 2021. The process of matching the items of the two tools with respect to the dimension of usability they assess will be carried out in parallel. Recruitment of hospitals and medication alert system experts is planned to start in April 2021. Instructions, tools, and response and consent forms will be sent to participants after they agree to take part. This data collection phase is expected to take approximately 9 months. Data analysis will be conducted as data are received from participants. The project is expected to conclude by January 2022.

ORCID iDs
Romaric Marcilly http://orcid.org/0000-0002-7077-7267
Melissa T Baysari http://orcid.org/0000-0003-1645-9126

REFERENCES
1 Edrees H, Amato MG, Wong A, *et al*. High-priority drug-drug interaction clinical decision support overrides in a newly implemented commercial computerized provider order-entry system:

override appropriateness and adverse drug events. *J Am Med Inform Assoc* 2020;27:893–900.

2   van der Sijs H, Aarts J, Vulto A, *et al*. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc* 2006;13:138–47.

3   Baysari MT, Tariq A, Day RO, *et al*. Alert override as a habitual behavior - a new perspective on a persistent problem. *J Am Med Inform Assoc* 2017;24:409–12.

4   Marcilly R, Ammenwerth E, Roehrer E, *et al*. Usability flaws in medication alerting systems: impact on usage and work system. *Yearb Med Inform* 2015;10:55–67.

5   Tsopra R, Jais J-P, Venot A, *et al*. 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, *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:e287–96.

7   Horsky J, Phansalkar S, Desai A, *et al*. Design of decision support interventions for medication prescribing. *Int J Med Inform* 2013;82:492–503.

8   Phansalkar S, Edworthy J, Hellier E, *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.

9   Marcilly R, Ammenwerth E, Roehrer E, *et al*. Evidence-Based usability design principles for medication alerting systems. *BMC Med Inform Decis Mak* 2018;18:69.

10 Zachariah M, Phansalkar S, Seidling HM, *et al*. Development and preliminary evidence for the validity of an instrument assessing implementation of human-factors principles in medication-related decision-support systems--I-MeDeSA. *J Am Med Inform Assoc* 2011;18 Suppl 1:i62–72.

11 Phansalkar S, Zachariah M, Seidling HM, *et al*. Evaluation of medication alerts in electronic health records for compliance with human factors principles. *J Am Med Inform Assoc* 2014;21:e332–40.

12 Cho I, Lee J, Han H, *et al*. 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.

13 Baysari MT, Lowenstein D, Zheng WY, *et al*. 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.

14 Lowenstein D, Zheng WY, Burke R, *et al*. Do user preferences align with human factors assessment scores of drug-drug interaction alerts? *Health Informatics J* 2020;26:563–75.

15 Snyder ME, Jaynes H, Gernant SA, *et al*. Alerts for community pharmacist-provided medication therapy management: recommendations from a heuristic evaluation. *BMC Med Inform Decis Mak* 2019;19:135.

16 Zheng WY, Van Dort B, Marcilly R. A tool for evaluating medication alerting systems (TEMAS): development and initial assessment. *JMIR Medical Informatics [Internet]* 2020;9 https://medinform.jmir.org/2021/7/e24022/

17 Brooke J. *SUS: a "quick and dirty" usability scale. Usability evaluation in industry*. London: CRC Press, 1996.

18 Krippendorff K. *Content analysis: an introduction to its methodology*. 4 Edn. Los Angeles London New Delhi Singapore Washington DC Melbourne: SAGE, 2019.

19 Furr RM. *Psychometrics - an introduction*. 3 edn. Los Angeles: SAGE, 2018.

20 Perrier L, Kealey MR, Straus SE. A usability study of two formats of a shortened systematic review for clinicians. *BMJ Open* 2014;4:e005919.

21 Grudniewicz A, Bhattacharyya O, McKibbon KA, *et al*. Redesigning printed educational materials for primary care physicians: design improvements increase usability. *Implement Sci* 2015;10:156.

22 Lewis JR, Sauro J. The factor structure of the system usability scale. In: *Human Centered Design [Internet]*. Heidelberg,Berlin: Springer Berlin Heidelberg, 2009: 94–103. http://link.springer.com/10.1007/978-3-642-02806-9_12

23 Tullis TS, Stetson JN. A comparison of questionnaires for assessing website usability. *usability professionals association (UPA) 2004 conference*, Minneapolis, Minnesota, 2004.

24 Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Hum Comput Interact* 2008;24:574–94.

25 Peute LW, Lichtner V, Baysari MT, *et al*. Challenges and best practices in ethical review of human and organizational factors studies in health technology: a synthesis of Testimonies. *Yearb Med Inform* 2020;29:58–70.

26 The general data protection regulation [Internet]. Available: http://www.consilium.europa.eu/fr/policies/data-protection-reform/data-protection-regulation/ [Accessed 1 Dec 2019].