Evaluation of medication alerts in electronic health records for compliance with human factors principles

Shobha Phansalkar,[1,2,3,4] Marianne Zachariah,[1] Hanna M Seidling,[5] Chantal Mendes,[1] Lynn Volk,[1] David W Bates[1,2,3]

[1]Partners Healthcare Systems, Inc., Wellesley, Massachusetts, USA
[2]Division of General Internal Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA
[3]Harvard Medical School, Boston, Massachusetts, USA
[4]Wolters Kluwer Health, Indianapolis, Indiana, USA
[5]Department of Clinical Pharmacology and Pharmacoepidemiology, University of Heidelberg, Heidelberg, Germany
[6]Cooperation Unit Clinical Pharmacy, University of Heidelberg, Heidelberg, Germany

Correspondence to
Dr Shobha Phansalkar, Division of General Medicine and Primary Care, 1620 Tremont Street, OBC-3, Boston, MA 02120, USA; sphansalkar@partners.org

Received 15 August 2013
Revised 7 March 2014
Accepted 11 March 2014
Published Online First 29 April 2014

To cite: Phansalkar S, Zachariah M, Seidling HM, *et al. J Am Med Inform Assoc* 2014;**21**:e332–e340.

ABSTRACT

Introduction Increasing the adoption of electronic health records (EHRs) with integrated clinical decision support (CDS) is a key initiative of the current US healthcare administration. High over-ride rates of CDS alerts strongly limit these potential benefits. As a result, EHR designers aspire to improve alert design to achieve better acceptance rates. In this study, we evaluated drug–drug interaction (DDI) alerts generated in EHRs and compared them for compliance with human factors principles.

Methods We utilized a previously validated questionnaire, the I-MeDeSA, to assess compliance with nine human factors principles of DDI alerts generated in 14 EHRs. Two reviewers independently assigned scores evaluating the human factors characteristics of each EHR. Rankings were assigned based on these scores and recommendations for appropriate alert design were derived.

Results The 14 EHRs evaluated in this study received scores ranging from 8 to 18.33, with a maximum possible score of 26. Cohen's κ (κ=0.86) reflected excellent agreement among reviewers. The six vendor products tied for second and third place rankings, while the top system and bottom five systems were home-grown products. The most common weaknesses included the absence of characteristics such as alert prioritization, clear and concise alert messages indicating interacting drugs, actions for clinical management, and a statement indicating the consequences of over-riding the alert.

Conclusions We provided detailed analyses of the human factors principles which were assessed and described our recommendations for effective alert design. Future studies should assess whether adherence to these recommendations can improve alert acceptance.

INTRODUCTION

Clinical decision support (CDS) when implemented in electronic health records (EHRs) has the potential to prevent medication errors and decrease adverse patient outcomes.[1 2] Despite their importance, the vast majority of CDS alerts are over-ridden,[3] leaving much of their potential benefit untapped.[4] There are several reasons for high alert over-ride rates, but the most common include incorrect alert content[3] or inappropriate presentation of the alert within the context of prescribing.[5 6] The knowledge base determining which alerts are presented and the actual display of the alert are both important in determining alert acceptance.[7 8] Previous research in this area has shown that consideration of human factors principles can play a prominent role in alert acceptance.[8]

Previously, we developed and validated an analytical tool called the Instrument for Evaluating Human-Factors Principles in Medication-Related Decision Support Alerts (I-MeDeSA) in order to assess compliance with human factors principles (table 1).[9] This instrument was designed specifically for the evaluation of drug–drug interaction (DDI) alerts in EHRs. Our aim in this study was to draw a comparison across EHRs of DDI alerts and their compliance with human factors principles using the I-MeDeSA instrument. A secondary aim of this study was to provide recommendations for appropriate alert design based on this evaluation.

METHODS

The I-MeDeSA instrument

The I-MeDeSA instrument was developed and validated to allow EHR designers to examine the compliance of alerts with human factors principles. The instrument measures alerts on the following nine human factors principles: alarm philosophy, placement, visibility, prioritization, color learnability and confusability, text-based information, proximity of task components being displayed, and corrective actions. Each principle exists as a construct of individual questions which are scored. There are a total of 26 questions (or items) for the nine constructs. Each item receives a score of '1' if the item characteristic is present and a score of '0' if it is absent. Details of the development and validation of the I-MeDeSA instrument were previously reported and the instrument is available upon request.[9] The maximum score a system could achieve in this evaluation was 26. The design principles of high-scoring systems were assessed and these characteristics were highlighted as recommendations for appropriate alert design. Alternatively, low-scoring systems were analyzed to expose undesirable characteristics.

Participating institutions

EHRs with CDS functionalities, specifically DDI alerting, were selected for inclusion in this study. In order to gain a broad understanding of alert design, we did not limit the sample of EHRs by setting. The sample consisted of EHRs developed in-house at academic medical centers and commercially offered by EHR vendors. Nine institutions agreed to participate in the study: seven academic medical centers and two EHR vendors. No financial incentive was offered for participation. A total of 14 EHRs—eight developed in-house and six

Table 1 Items in the nine constructs of the Instrument for Evaluating Human-Factors Principles in Medication-Related Decision Support Alerts (I-MeDeSA) used for evaluating electronic health records

Number	Item
Alarm philosophy	
1i	Does the system provide a general catalog of unsafe events, correlating the priority level of the alert with the severity of the consequences?
Placement	
2i	Are different types of alerts meaningfully grouped? (ie, by the severity of the alert, where all level 1 alerts are placed together, or by medication order, where alerts related to a specific medication order are grouped together)
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?
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)
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.)
Visibility	
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.
3ii	Is the background contrast sufficient to allow the user to easily read the alert message? (ie, dark text on a light background is easier to read than light text on a dark background)
3iii	Is the font used to display the textual message appropriate for the user to read the alert easily? (ie, a mixture of upper and lower case lettering is easier to read than upper case only)
Prioritization	
4i	Is the prioritization of alerts indicated appropriately by color? (ie, colors such as red and orange imply a high priority compared to colors such as green, blue, and white)
4ii	Does the alert use prioritization with colors other than green and red, to take into consideration users who may be color blind?
4iii	Are signal words appropriately assigned to each existing level of alert? For example, 'Warning' would appropriately be assigned to a level 1 alert and not a level 3 alert. 'Note' would appropriately be assigned to a level 3 alert and not a level 1 alert
4iv	Does the alert utilize shapes or icons in order to indicate the priority of the alert? (ie, angular and unstable shapes such as inverted triangles indicate higher levels of priority than regular shapes such as circles)
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.
Color	
5i	Does the alert utilize color-coding to indicate the type of unsafe event? (ie, drug–drug interaction (DDI) vs allergy alert)
5ii	Is color minimally used to focus the attention of the user? As excessive coloring used on the screen can create noise and distract the user, there should be less than 10 colors.
Learnability and confusability	
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.
Text-based information: Does the alert possess the following four information components?	
7i	A signal word to indicate the priority of the alert (ie, 'note,' 'warning,' or 'danger')
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.
7iia	If yes, are the specific interacting drugs explicitly indicated?
7iii	An instruction statement (telling the user how to avoid the danger or the desired action)
7iiia	If yes, does the order of recommended tasks reflect the order of required actions?
7iv	A consequence statement telling the user what might happen, for example, the reaction that may occur if the instruction information is ignored.
Proximity of task components being displayed	
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,' or a link to a medical reference website providing additional information?
Corrective actions	
9i	Does the system allow corrective actions that serve as an acknowledgement of having seen the alert? For example, 'Accept' and 'Cancel' are corrective actions, while 'OK' is an acknowledgment.
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.
9ii	Is the system able to monitor and alert the user to follow through with corrective actions? Referring to the previous example, if the user tells the system that he/she will reduce the warfarin dose but fails to follow through on that promise, does the system alert the user?

commercial products—were analyzed. Details of the EHRs evaluated and their host institutions along with software version numbers are provided in table 2. The protocol was approved by the Partners Healthcare Research Committee. Additionally, we sought approval from individual organization IRBs and EHR vendors as required.

Table 2 Description of institutions, electronic health records (EHRs), version numbers, and inpatient/outpatient use

Institution	EHR	Version number	Inpatient/outpatient
Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA	WebOMR	WebOMR 2009	Outpatient
GE Healthcare (vendor), Waukesha, Wisconsin, USA	Centricity	2005 6.0	Outpatient
Harvard Vanguard Medical Associates, Newton, Massachusetts, USA	EpicCare (Epic)	2007 IU3	Outpatient
NextGen Healthcare (vendor), Horsham, Pennsylvania, USA	NextGen Ambulatory	5.5.28	Outpatient
Northwestern Memorial Faculty Foundation, Chicago, Illinois, USA	EpicCare (Epic)	Spring 2008 IU7	Outpatient
Northwestern Memorial Hospital, Chicago, Illinois, USA	PowerChart (Cerner)	2007.19	Inpatient
Northwestern Memorial Hospital, Chicago, Illinois, USA	PowerChart Office (Cerner)	2007.19	Outpatient
Partners Healthcare, Boston, Massachusetts, USA	Longitudinal Medical Record	8.2	Outpatient
Partners Healthcare, Boston, Massachusetts, USA	BICS (tiered)	November 2010	Inpatient
Partners Healthcare, Boston, Massachusetts, USA	BICS (un-tiered)	November 2010	Inpatient
Regenstrief Institute, Indianapolis, Indiana, USA	Gopher	5.25	Outpatient
US Department of Veterans Affairs (VA), Washington, DC, USA	CPRS	1.0.27.90	Inpatient and outpatient
Vanderbilt University, Nashville, Tennessee, USA	RxStar	Not available	Outpatient
Vanderbilt University, Nashville, Tennessee, USA	WizOrder	Not available	Inpatient

Screenshot collection

Box 1 details the instructions that were provided to the site coordinator at each participating institution for providing the screenshots from the EHRs. We sought permission to publish screenshots and obtained authorization from individuals at participating healthcare and/or vendor institutions. In order to preserve the anonymity of the systems, we redacted screenshots when necessary, as seen in the figures displayed in this article. Some vendors refused permission to publish their screenshots despite an Institute of Medicine report in 2011 that specifically recommended disallowing this practice in the interest of safety improvement. We decided to include these EHRs in the analysis but have not presented their screenshots for publication in this article. System numbers henceforth referred to in the article are in no particular order to those described in table 2 to preserve the anonymity of the EHRs.

Evaluation of screenshots

Two reviewers employed the I-MeDeSA to independently evaluate the screenshots of DDI alerts provided by the participating institutions detailed in table 2. Both individuals had experience in evaluating the usability of CDS: one had a background in medical informatics and pharmacy and the second had expertise in clinical information systems research. For EHRs with tiered levels of DDI alerts, the reviewers analyzed the alert levels individually by applying every item of the I-MeDeSA instrument to each level. The resulting scores for each level were then averaged to determine final scores for the system overall. If the information was incomplete or the workflow sequence/details were unclear from the screenshots alone, the reviewers requested a walk-through of the medication ordering and alerting processes via a web conference, followed by independent scoring of the items as above. After completing their independent evaluations, the reviewers met to compare their assessments of each EHR. If there was a disagreement among scores, then a third reviewer with expertise in medical informatics, human factors, and qualitative research methodologies, helped arrive at consensus. If reviewers were unable to come to an agreement on an item score, a third reviewer was enlisted to determine the final score. Cohen's κ was calculated to measure inter-rater reliability between reviewers.

RESULTS

We evaluated 14 EHRs on their display of DDI alerts and found considerable variability in their compliance with human factors principles. In table 3, we have given the scores attained by the EHRs on each of the human factors constructs measured using the I-MeDeSA instrument. EHR systems achieved scores ranging from 8 to 18.4 out of a total score of 26 points, with the average total score being 13.6±2.7 (52.6%). On average, systems scored best on visibility items (2.8/3, 94.3%) and almost equally well on the constructs of Proximity of task components (0.71/1, 71.4%) and Placement (2.8/4, 70.8%), and poorest on providing an Alarm philosophy (0.14/1, 14%). The other two low scoring constructs were Prioritization (1.3/5, 25.7%) and Learnability and confusability (0.29/1, 28.6%). Inter-rater reliability was high (Cohen's κ=0.86).

DISCUSSION

A significant focus of the domain of DDI alerting has been on the content of the alerts; however, little attention has been paid to how these alerts are actually presented in the EHR and how appearance may impact alert acceptance. This comparison across EHRs allowed us to provide recommendations for appropriate DDI alert design from the human factors perspective.

System 10 was the highest scoring system (18.4/26) and scored the highest points on the constructs of Learnability and confusability, Text-based information, Proximity of task components, and Corrective actions. The highest scoring systems (systems 10, 6, and 7) all received perfect scores on the construct of Proximity of task components. Systems 6 and 7 scored equally on every construct and hence received the same total score of 16/26. These systems also received perfect scores on the four constructs of Placement, Visibility, Color, and Proximity of task components. It is interesting to note that systems 6 and 7 scored no points on the construct of Prioritization. They also performed poorly on the construct of Learnability and confusability. This tells us that even high-scoring systems differ considerably in their design aspects. While overall scores maybe a good indicator of human factors compliance in general, specific design principles are not always equally adhered to. When comparing systems, designers should take into account not just overall scores but also specific construct scores to identify principles that can be improved to

Box 1 Instructions provided to participating institutions for capturing screenshots of drug–drug interaction (DDI) alerts from their electronic health records

Please provide a screenshot for each level of DDI if you have multiple severity levels.

Within these screenshots we are looking for the following characteristics:

I. Visual distinctions based on severity of alert:
- (i) Symbols/icons to indicate severity
- (ii) Words to indicate severity
- (iii) Colors to indicate/prioritize severity
- (iv) Size of font to indicate/prioritize severity

II. Response to the alert:
- (i) What are the possible actions that the provider can take to over-ride an alert for each severity level (continue order/cancel/discontinue pre-existing drug, etc.)
- (ii) Reasons for over-riding the alert: if this is a drop down list, please provide a screenshot with the entire list visible
- (iii) Is there a place for a free-text entry of a reason for over-riding the alert?

III. Summary screen:
- (i) Are interactions ordered by severity or by sequence in which the medications were ordered?
- (ii) Any symbols/icons to alert the provider of possible interactions on this screen?
- (iii) What actions can the provider take from the summary screen to modify the order or respond to the alert?

IV. Alert message:Please provide a screenshot of
- (i) The information the provider sees regarding why the alert was generated
- (ii) The reaction
- (iii) Indication of clinical significance
- (iv) Any additional links that the provider can access to obtain additional information, for example, an infobutton, drug monographs, or web links
- (v) Whether an alternative medication is suggested

V. Types of medication-related alerts:
- (i) Any other types of alerts besides DDIs and therapeutic duplications (eg, medication alerts for allergies, renewals, etc.)
- (ii) Any other types of informational medication-related alerts that are shown to the provider

match the clinical context in which they are used. For example, an alert provided to a medical student needs to have sufficient text-based information to convey details on understanding the consequences and mechanism of action of an interaction. In contrast, it is probably more important to focus on aspects of Prioritization in a busy setting like an emergency department, where providers need to process disparate pieces of information in a very short period of time.

The lowest scoring system was system 13 (8/26), a home-grown system at an academic medical center that performed worst among all systems on the following seven out of nine constructs: Alarm philosophy, Placement, Prioritization, Color, Learnability and confusability, Proximity of task components, and Corrective actions. Systems 11 (9/26), 1 (12/26), and 14 (12/26) also received overall scores below 50%. System 13 was the only system to obtain no points on the construct of Color because the system lacked the capability to distinguish between types of interactions based on color and also employed many different colors on the screen. In addition, system 13 scored no points on the construct of Corrective actions because it did not provide users with the capability to acknowledge any of the alerts. Systems 11, 13, and 14 all scored 0/5 on the Prioritization construct. Interestingly, two of the highest scoring systems (systems 6 and 7) also received 0/5 on this construct.

Most of the evaluated systems employed adequate measures for incorporating Visibility principles (overall score=94.3%) by making the alerts easily distinguishable from the rest of the screen and applying a font that was easy to read. Most systems also performed well on the constructs Proximity of task components and Placement. The overall score on Proximity of task components was 71.4% and most systems offered links to outside sources of information (drug monographs, medical information websites, or electronic physician desk references) within spatial and/or temporal proximity to the alert. The construct of Placement was also well employed (overall score=70.8%) through use of meaningful grouping of alerts and taking into account appropriate timing for the appearance of the alert. However, the majority of systems assessed did *not* provide an alert philosophy statement or any other variant of well-defined guidelines for classifying the prioritization of alerts.

Performance of systems by human factors principle

Only two (systems 3 and 5) of 14 systems offered documentation of guidelines on alarm philosophy and provided information to the user specifying the algorithms used to assign priority levels to DDIs. This is clearly an area where systems should make their criteria for alerting more transparent to the user, allowing them to understand why certain interactions are deemed more severe than others.

Systems 6 and 7 obtained a perfect score on the Placement construct. The alert presented by these systems clearly indicated the offending drug combination, the nature of the interaction, and the recommended care management, as shown in figure 1. The DDI alert appeared after the user put in an order for a drug that interacted with another drug on the patient's medication profile. The type of interaction, in this case a DDI, was indicated by using an icon to cue the user, indicating meaningful grouping among alerts. In the lower-left quadrant, the user was provided with response options, such as documentation of override reasons via a drop-down list or in free text and the option to cancel the order. These options were placed in close proximity to the information on the interacting drug pair.

Systems (systems 2, 11, 13, and 14) that performed poorly on this construct failed to display drug interaction information in an appropriate manner. These systems did not require a response to the alert and/or presented information from a drug monograph that included superfluous information which hampered users from easily identifying critical information.

Eleven EHRs scored 3 out of 3 on the Visibility construct. Systems employed good visibility principles by placing the alert so that it would occupy either the computer screen or appear in the user's direct line of sight. Systems that obtained high scores utilized alerts with colors and fonts that were easy to read, such as dark text on a light background and a mixture of upper and lower case text. Light text on a dark background is harder to read than dark text on a light background. System 11, which received the lowest score on this construct, provided insufficient information and also employed a white font on a dark blue background, making it difficult for reviewers to read the alert. See figure 2 for details.

Table 3 Electronic health record (EHR) system scores by human factors principle and overall ranking system numbers

EHR system	Alarm philosophy	Placement	Visibility	Prioritization	Color	Learnability and confusability	Text-based information	Proximity of task components being displayed	Corrective actions	Total score	Overall system ranking
Sys 1	0/1	3/4	3/3	2/5	1/2	1/1	0/6	1/1	1/3	12/26	6
Sys 2	0/1	2/4	3/3	3/5	1/2	1/1	3/6	1/1	1/3	15/26	2
Sys 3	1/1	3/4	3/3	2/5	1/2	0/1	3/6	1/1	1/3	15/26	3
Sys 4	0/1	3/4	3/3	2/5	1/2	0/1	3/6	1/1	1/3	14/26	4
Sys 5	1/1	3/4	3/3	2/5	1/2	0/1	2/6	1/1	1/3	14/26	4
Sys 6	0/1	4/4	3/3	0/5	2/2	0/1	5/6	1/1	1/3	16/26	2
Sys 7	0/1	4/4	3/3	0/5	2/2	0/1	5/6	1/1	1/3	16/26	2
Sys 8	0/1	2.7/4	2.3/3	1.3/5	1.7/2	1/1	3.7/6	0/1	1.3/3	14/26	4
Sys 9	0/1	3/4	3/3	1/5	2/2	0/1	4/6	0/1	1/3	14/26	4
Sys 10	0/1	3/4	2.3/3	2.7/5	1/2	1/1	5.7/6	1/1	1.7/3	18.4/26	1
Sys 11	0/1	2/4	2/3	0/5	1/2	0/1	2/6	1/1	1/3	9/26	7
Sys 12	0/1	3/4	3/3	2/5	1/2	0/1	3/6	0/1	1/3	13/26	5
Sys 13	0/1	2/4	3/3	0/5	0/2	0/1	3/6	0/1	0/3	8/26	8
Sys 14	0/1	2/4	3/3	0/5	1/2	0/1	4/6	1/1	1/3	12/26	6
Overall score %	0.14/1 14%	2.8/4 70.8%	2.8/3 94.3%	1.3/5 25.7%	1.2/2 59.6%	0.29/1 28.6%	3.3/6 55.2%	0.71/1 71.4%	1/3 33.3%	13.6/26 52.6%	

The systems are in no particular order to those described in table 2 to preserve the anonymity of the EHRs.

System 2 received a score of 3 out of 5, the highest score of all systems, in the Prioritization construct. This system demonstrated appropriate use of color-coding to indicate the priority of the alert. What was exceptional about the system was the fact that the user could choose a color scheme for each level of severity of the interaction. However, while this gives a lot of flexibility to the user in terms of their preferences, there is no guarantee that users will always make appropriate choices in terms of the colors that are indicative of the severity of the interaction (item 4i in table 1). The system tested utilized the color scheme of red, yellow, and white to indicate major, moderate, and minor interaction severity, respectively (figure 3A). Utilization of red and green colors precluded system 2 from achieving a perfect score since the system failed to account for color-blind users who cannot distinguish between red/green and yellow/orange/gray combinations. The lack of accommodation for color-blind users was a common failing in 12 out of the 14 EHRs indicating prioritization with the use of the red/green combination. Appropriate use of signal words and/or symbols for indicating priority is considered good practice to

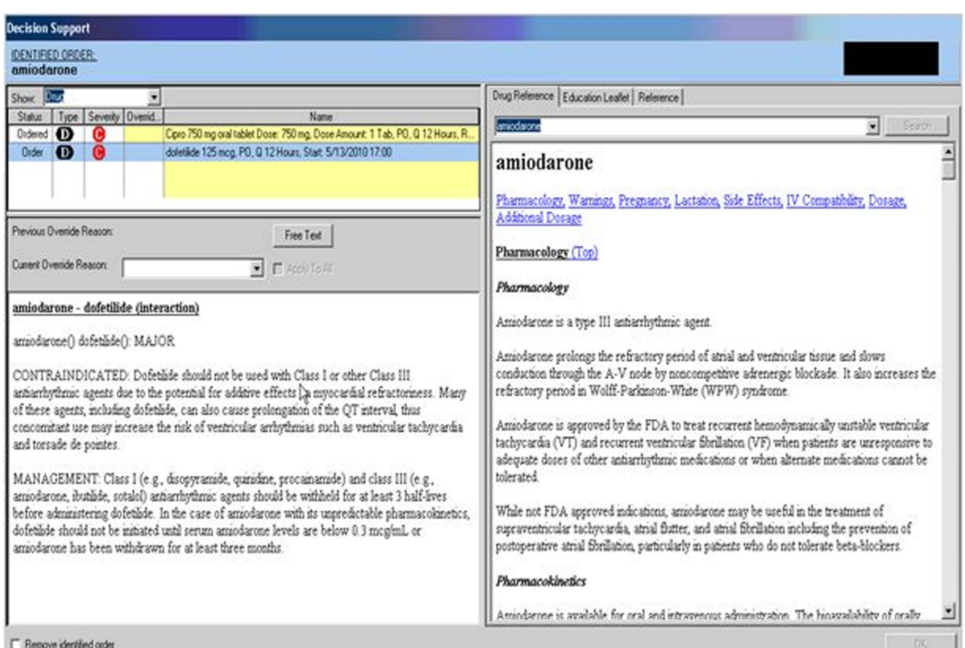

Figure 1 Example of a system that scored highly on the construct of Placement by identifying the type of interaction, allowing the user to easily enter in their response to the alert, linking the alert to the medication order by appropriate timing, and providing the critical information needed to act on the drug-drug interaction alert.

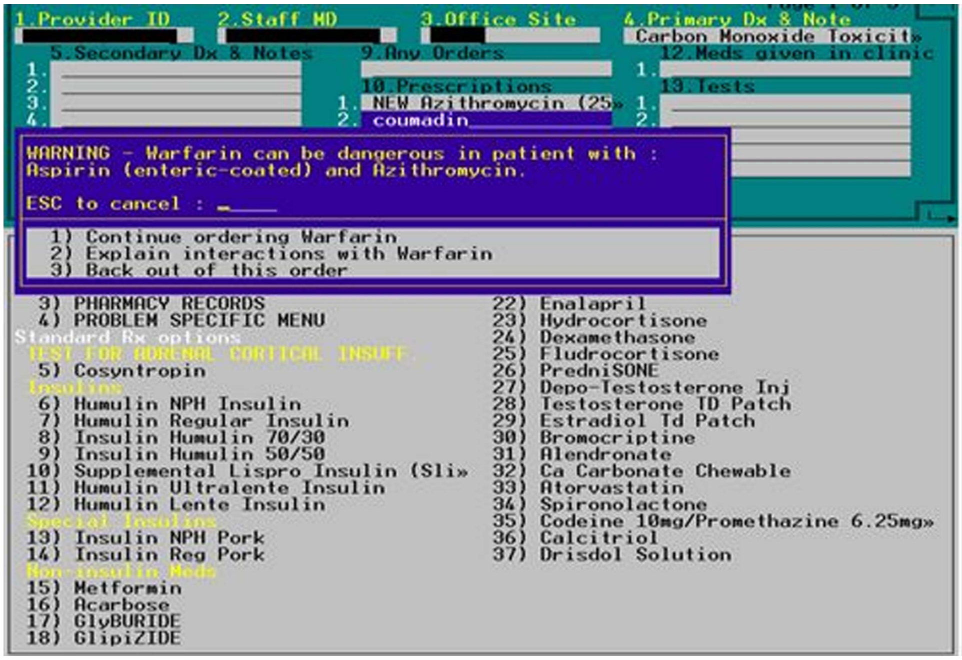

Figure 2 The drug–drug interaction alert presented here shows insufficient information for the user to act on the interaction between warfarin and the interacting drugs on the patient's medication profile. In addition, reviewers found it difficult to read the statement indicating the interacting drugs in bright yellow font on a dark blue background.

accommodate color-blind users. Seven systems chose to utilize signal words. Only system 2 employed the appropriate use symbols to indicate the priority of an interaction. In this system, the highest severity alerts used a red exclamation point (figure 3B), while the lower severity alerts were indicated by an exclamation point within an inverted yellow triangle (figure 3C). While we did not evaluate this under the construct of Prioritization, an important consideration is the way in which figures and icons are used in the alert box and also in the context of the entire EHR. The figures and icons utilized within the context of one type of alert should be homogeneous with other types of alerting employed in the EHR. So, if a particular icon, say a stop sign, represents a high level of an alert in one context (eg, DDIs), then it should be consistently employed to represent similar severity in another context (eg, drug–allergy interactions).

Systems 6, 7, and 9 received a perfect score on the construct of Color. These systems used appropriate colors to distinguish between the different types of alerts (drug–drug, drug–allergy, drug duplicate). Using less than 10 colors in an alert is recommended in order to avoid confusion. Thirteen out of the 14 systems used less than 10 colors within their alerts but failed to make a distinction between the different alert types. Systems 6 and 7 both received high scores because they used color-coded letters to mark DDIs and drug–allergy interactions (DAIs). DDIs were indicated by a blue 'D' and DAIs were marked by a red 'A.' Nine systems failed to communicate alert type with the use of color. System 13 received the lowest score for using over 10 colors but not having appropriate color-coding for alert types.

Four systems (systems 1, 2, 8, and 10) received a 1 out of 1 on this construct. All four systems had tiered alerts that were marked by unique visual characteristics such as distinct colors and shapes (figure 3A). Like Prioritization, the Learnability and confusability principle applies best to systems with a tiered alerting philosophy. Systems failing to satisfy this principle either lacked a tiered alerting system completely or displayed tiered alerts that used signal words rather than distinct visual features such as a specific color and font. This is because the use of a signal word to identify the severity of the alert is not considered a visual characteristic (figure 4).

Text-based information was evaluated by reviewing the information displayed when an alert was first generated. Any information accessed through additional clicks or displayed within a monograph that was outside the actual alert box was disregarded. System 10 scored the highest in this construct with 5.7 out of 6 (figure 5), followed by systems 6 and 7 with 5 out of 6. These three systems specifically identified the interacting drugs, clearly indicated management steps in the appropriate order, and outlined the potential consequences if the alert were over-ridden. However, signal words did not accompany level-three alerts in system 10, and were not used at all in systems 6 and 7, which precluded them from obtaining a perfect score. The use of signal words allows the user to distinguish between alert severities. Lower scoring systems (systems 1, 5, and 11) required users to actively seek interaction information by clicking on additional links. Systems 1 and 11 performed worst because they did not automatically present alert information, thus requiring users to take additional steps to access management instructions and consequence statements (figure 2).

The Proximity of task components construct aims to evaluate systems that provide the option for users to access additional sources of information, such as drug monographs, electronic physician desk references, or knowledge links to websites. Providing a link to such sources of information directly from the alert, or close by, caters to the needs of the user and increases usability. Most EHRs (10 out of 14 EHRs) adequately employed features that allowed the user to access informational components for decision making via drug monographs and/or links to websites.

System 10 scored the highest on this principle with 1.7 out of 3, and was closely followed by system 8 with 1.3 out of 3. These two systems stood out from the others because they used

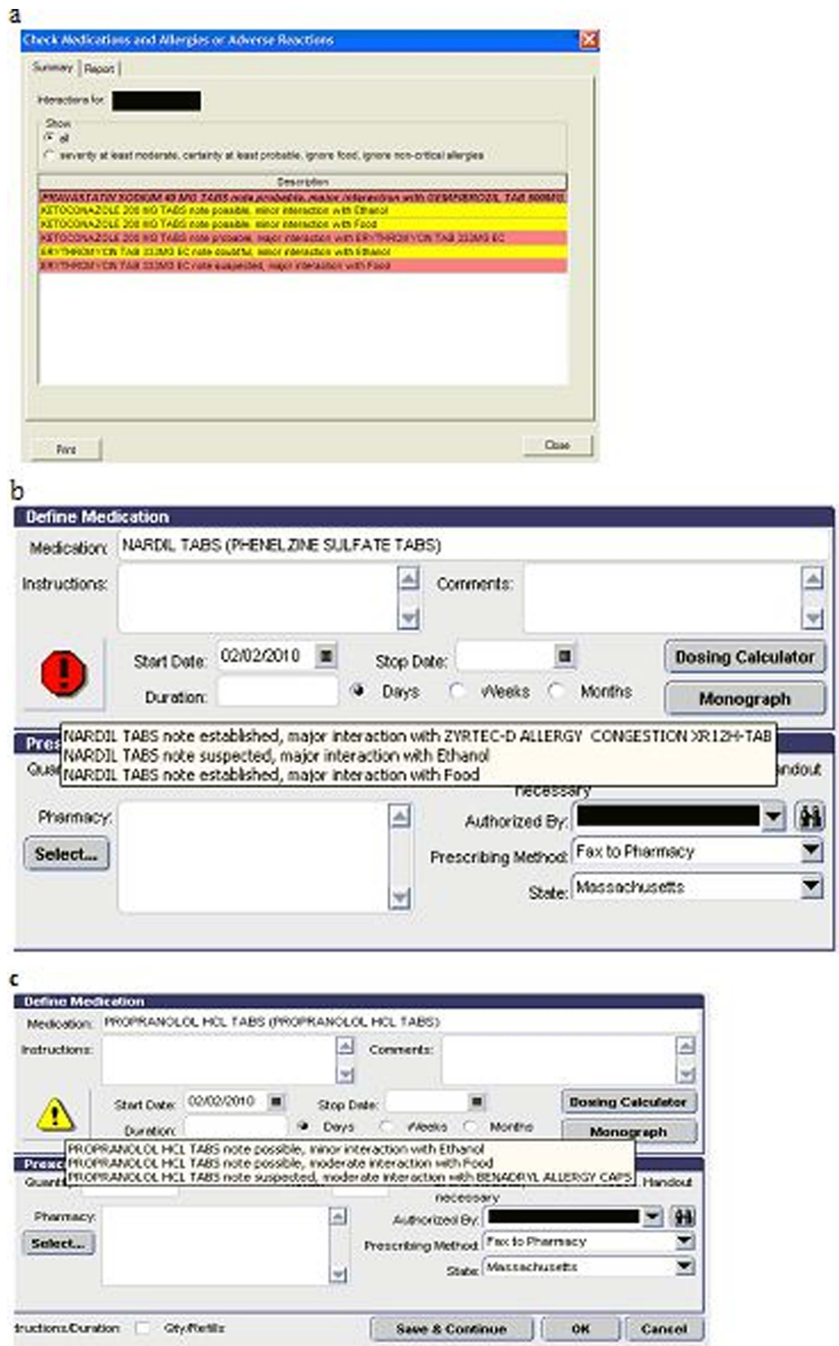

Figure 3 (A–C) Illustration of the Prioritization principle. (A) Use of color-coding for distinction between alert severities (B) and (C). Use of symbols to indicate appropriate severity levels.

intelligent corrective actions. Intelligent corrective actions are superior to mere acknowledgements in response to an alert. The actions assist a user in appropriately completing a task through additional follow-up steps. For example, in system 10, when a user attempted to order amiodarone when ciprofloxacin was on the patient's current medication list, a level-two alert fired (figure 5). Two of the six response options were to discontinue the pre-existing drug or to adjust the dose of amiodarone. If the user responded that he/she would like to discontinue the pre-existing drug, the intelligent corrective action of the system automatically removed it from the patient's medication list and accepted the order for the amiodarone. No additional work was required by the user to discontinue the ciprofloxacin. The majority of the EHRs evaluated for this study lacked this feature. System 13 had the lowest score (0/3) on this construct. This system did not require any type of response or action by the user after an alert was fired (figure 6). These alerts functioned more like

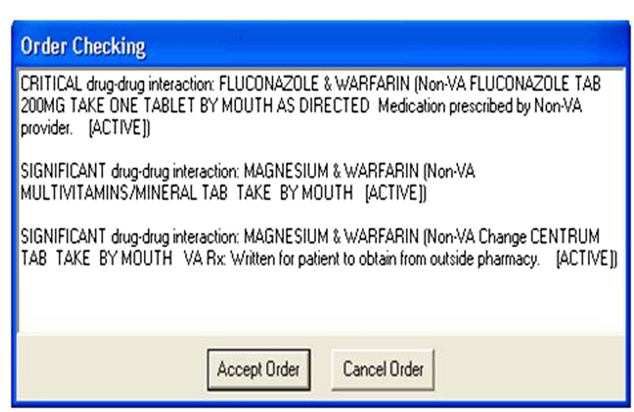

Figure 4 System 12 scored poorly on Learnability and confusability because it did not present unique visual characteristics for differentiating between alert severities.

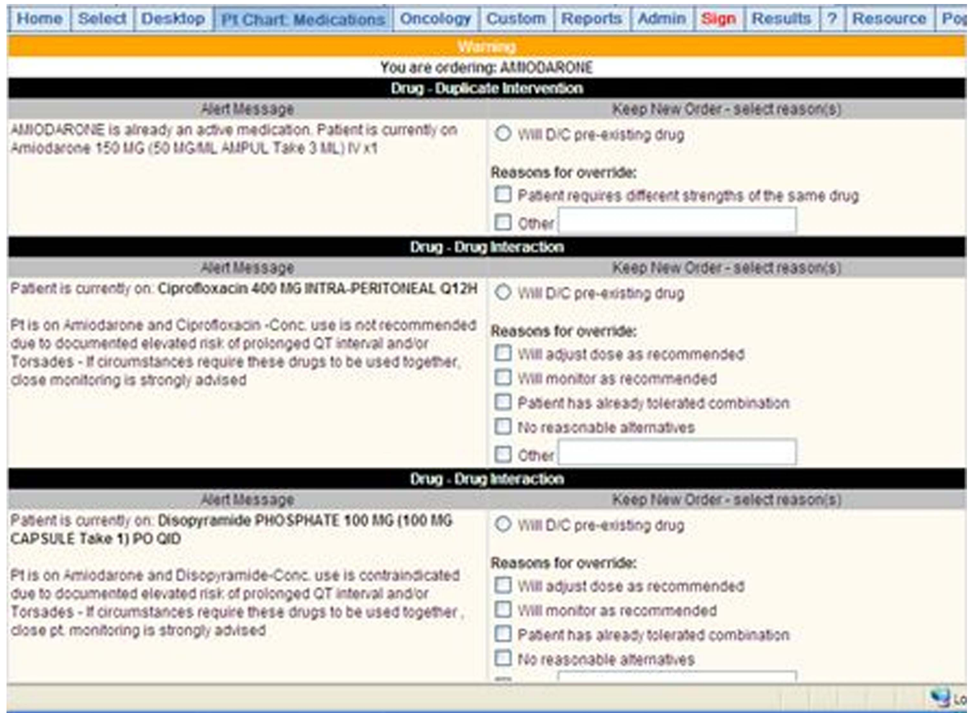

Figure 5 Interacting drugs, management steps, and potential consequence to the patient clearly presented when an alert is displayed.

notifications. To ease the user's interaction with the system, corrective actions—where appropriate—should clearly be included in the alert message and potentially be integrated into the user's alert response. For example, if the user states that he wants to stop one interacting drug, the respective drug might be deleted immediately from the drug list by the system itself. This interlinking of ordering, alerting, and modification of the order would represent best practice in terms of aligning multiple tasks in the ordering process. Further, an option should be available should the user choose to report a problem with the alert. Such an option does not compel action on the part of the user but might make it easier and more efficient to report problems and for system developers to retrospectively gather user feedback on the alerting capabilities. System developers could design the capability to capture exactly what the user is looking at when a problem is reported, to make it easier to understand the problem and try to address it.

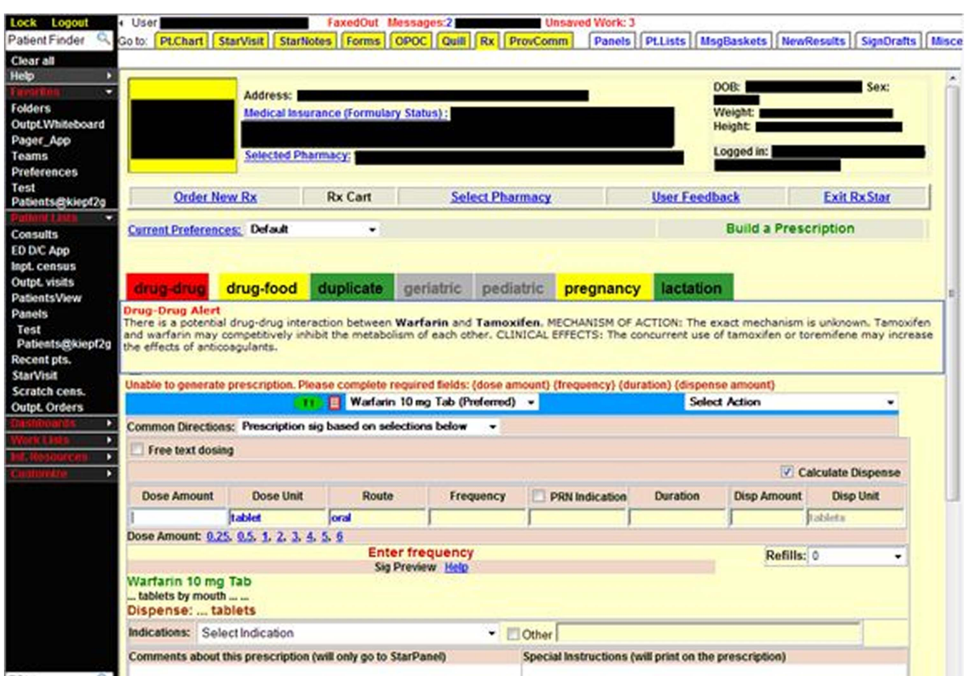

Figure 6 Poor corrective actions do not allow the user to provide a response to an alert. In addition, this system utilized more than 10 colors on the screen.

The discussion above provides detailed analyses on how the design of alerts may depart from well-known human factors principles that can be found in the patient safety literature. While rankings enable us to compare one system against another, given that the overall scores of the systems are so close, and the highest score is only 52.6%, the important message is that all systems fall short in meeting the principles of good alert design. Compliance with these design principles is particularly important when a large number of alerts are generated and lead to alert fatigue. This study is a first step in specifically evaluating DDIs which form a large part of the alerts fired in any EHR and are therefore a significant contributor to the alert fatigue experienced by users. Further empirical research is needed to validate whether compliance with these principles actually produces an effect on the rate at which users appropriately override alerts and experience alert fatigue. Over time, users become forgiving of a system that allows them to perform their work in a efficient manner, and users pay more attention to the efficiency of the system although design principles may be grossly overlooked. Actually assessing whether employing these design principles can have an impact on user's acceptance of decision support alerts is crucial in having EHR vendors and clinical information system designers pay close attention to these recommendations.

CONCLUSIONS

From the sample of EHRs evaluated, it was evident that systems are not consistently applying human factors principles to alert design. We have provided recommendations based on evaluation of these systems that designers and developers may want to consider. Future studies should focus on empirically evaluating whether consideration of these design principles actually impacts alert effectiveness or decreases alert over-ride rates. The findings of this study highlight elements of DDI alert design that can be improved by the application of human factors principles, consequently increasing usability and user acceptance of medication-related decision support alerts.

Acknowledgements The authors would like to thank the American Medical Informatics Association (AMIA) for their support on this project.

Contributors SP and DWB had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. SP and DWB: study concept and design, study supervision, and obtaining funding; MZ and HMS: acquisition of data; SP, MZ, HMS, and CM: analysis and interpretation of data and drafting of the manuscript; SP, DWB, HMS, CM, and LV: critical revision of the manuscript for important intellectual content; LV: administrative, technical, and material support.

Funding This study was sponsored by the Center for Education and Research on Therapeutics on Health Information Technology (CERT-HIT) grant (PI: David W Bates), Grant # U18HS016970 from the Agency for Healthcare Research and Quality (AHRQ).

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Weingart SN, Simchowitz B, Padolsky H, *et al*. An empirical model to estimate the potential impact of medication safety alerts on patient safety, health care utilization, and cost in ambulatory care. *Arch Intern Med* 2009;169:1465–73.
2 Roberts LL, Ward MM, Brokel JM, *et al*. Impact of health information technology on detection of potential adverse drug events at the ordering stage. *Am J Health Syst Pharm* 2010;67:1838–46.
3 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.
4 Magnus D, Rodgers S, Avery AJ. GPs' views on computerized drug interaction alerts: questionnaire survey. *J Clin Pharm Ther* 2002;27:377–82.
5 Saleem JJ, Russ AL, Sanderson P, *et al*. Current challenges and opportunities for better integration of human factors research with development of clinical information systems. *Yearb Med Inform* 2009;48–58.
6 Russ AL, Zillich AJ, McManus MS, *et al*. A human factors investigation of medication alerts: barriers to prescriber decision-making and clinical workflow. *AMIA Annu Symp Proc* 2009;2009:548–52.
7 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.
8 Seidling HM, Phansalkar S, Seger DL, *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.
9 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.