Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation

Alissa L Russ,[1,2,3] Alan J Zillich,[1,2,3] Brittany L Melton,[4] Scott A Russell,[1] Siying Chen,[1] Jeffrey R Spina,[5,6] Michael Weiner,[1,2,7] Elizabette G Johnson,[8] Joanne K Daggy,[1,9] M Sue McManus,[10] Jason M Hawsey,[11] Anthony G Puleo,[12] Bradley N Doebbeling,[1,2,13] Jason J Saleem[14]

ABSTRACT

Objective   To apply human factors engineering principles to improve alert interface design. We hypothesized that incorporating human factors principles into alerts would improve usability, reduce workload for prescribers, and reduce prescribing errors.

Materials and methods   We performed a scenario-based simulation study using a counterbalanced, crossover design with 20 Veterans Affairs prescribers to compare original versus redesigned alerts. We redesigned drug–allergy, drug–drug interaction, and drug–disease alerts based upon human factors principles. We assessed usability (learnability of redesign, efficiency, satisfaction, and usability errors), perceived workload, and prescribing errors.

Results   Although prescribers received no training on the design changes, prescribers were able to resolve redesigned alerts more efficiently (median (IQR): 56 (47) s) compared to the original alerts (85 (71) s; p=0.015). In addition, prescribers rated redesigned alerts significantly higher than original alerts across several dimensions of satisfaction. Redesigned alerts led to a modest but significant reduction in workload (p=0.042) and significantly reduced the number of prescribing errors per prescriber (median (range): 2 (1–5) compared to original alerts: 4 (1–7); p=0.024).

Discussion   Aspects of the redesigned alerts that likely contributed to better prescribing include design modifications that reduced usability-related errors, providing clinical data closer to the point of decision, and displaying alert text in a tabular format. Displaying alert text in a tabular format may help prescribers extract information quickly and thereby increase responsiveness to alerts.

Conclusions   This simulation study provides evidence that applying human factors design principles to medication alerts can improve usability and prescribing outcomes.

BACKGROUND AND SIGNIFICANCE

Computerized medication alerts can reduce medication errors, but there are still challenges to fully realize their benefits.[1][2] Seidling et al[3] found that alert acceptance was most strongly predicted by the quality of the alert interface display. However, 'serious gaps'[1] remain in understanding how to effectively display alert information to prescribers.[1][2][4][5] To this end, Horsky et al[6] outlined interface design principles for prescribing decision support, but derived principles from implementation reports because of a lack of studies on interface design. Recognizing that alerts are a type of warning, Phansalkar et al[7] summarized human factors

guidelines for hazard signs and warning labels to inform the design of medication alerts. Four content components are recommended for text-based warnings: (1) a *signal word* to capture attention and convey the probability and severity of injury; (2) *hazard information* that identifies the danger; (3) *instructions* to avoid the hazard; and (4) information on potential *consequences* if the hazard is not averted.[8] One human factors handbook on warnings presents over 800 pages of evidence-based research findings and standards for warning design across various industries.[9] Many principles apply to the medication alert interface but have not been consistently implemented.[7]

A few studies have modified and compared the interface design features of alerts; most of these have focused on alert intrusiveness,[10–13] which generally refers to the extent to which alerts interrupt prescribers' work or require acknowledgment.[14] One simulation study compared modal alerts against non-modal alerts, as well as no alert. A modal design forces the prescriber to interact with the alert before interacting with other components of the interface, whereas a non-modal alert does not. Both alert types were associated with reduced prescribing error rates compared to no alerts, but the error rate for non-modal alerts was 3.6 times higher than that for modal alerts.[11] Similarly, two randomized controlled trials of laboratory monitoring alerts showed that non-modal alerts did not significantly improve laboratory test ordering compared to no alert.[12][13]

From the communication-human information processing (C-HIP) model in the human factors literature, one can deduce that alerts are more likely to capture prescribers' attention and be effective if the alert is more intrusive. This model[15] posits that for a warning (eg, alert) to be effective, it must first capture the attention of the intended recipient, a step called 'attention switch.'[16] While intrusiveness is desirable for attention switch, one investigation found that a drug interaction alert—which could not be easily overridden—led to unintended consequences of delayed treatment.[4] This illustrates the challenge of designing effective and safe alerts. In addition to intrusiveness, the layout, text format, and interface display of alerts likely influence alert effectiveness and medication safety.[3][7][14][17] We found no studies that incorporated human factors principles into medication alerts and evaluated the impact of the resulting design changes. Our

objective was to do this in a simulation environment with prescribers. We hypothesized that applying human factors principles to the design of the medication alert interface would improve usability, reduce prescribers' workload, and reduce prescribing errors.

MATERIALS AND METHODS
Organizational setting and details of the alert system

A simulation study was conducted at the Veterans Affairs (VA) Health Services Research and Development (HSR&D) Human-Computer Interaction and Simulation Laboratory in a VA medical center (VAMC).[18] The original alert interface in this study (figure 1) mirrors the VA alert system used for patient care. The VA system incorporates clinical information from VA or First DataBank, depending on the type of alert, and is implemented at over 150 VAMCs and 1400 clinics.[14 19] This study focused on drug–allergy, drug–drug interaction, and drug–disease alerts, since these alerts represent different design challenges, commonly occur during prescribing, and are important for patient safety. In this article, 'alert dialog box' refers to the pop-up box presented to prescribers, and 'alerts' refers to one or more associated warning messages within the box. This article was prepared according to STARE-HI guidelines.[20]

Alert redesigns

Redesigned features (table 1) were selected based upon available evidence as well as field observations and interviews with prescribers.[14] The design team was guided by two human factors engineers and consisted of physicians, pharmacists, a nurse practitioner, and informatics experts. The team iteratively redesigned

alerts using paper-based prototyping over 8 months. Redesigns were shared with an advisory panel of national VA informatics and medication safety leaders for additional input. Alerts were redesigned (figure 2) to appear in one alert dialog box at a single time point, immediately after the prescriber entered prescription information. A limited-function, high-fidelity mockup of the VA's electronic health record (EHR) system was created, which was *not* integrated into patient care, to safely evaluate the alert designs. For this study, we created two copies of the same mock EHR system, one with the original alerts designs and one with the redesigned alerts.

Scenario development

Three standardized patient scenarios were constructed over 12 months (see online supplementary file, appendix I). Scenarios were identical across the two alert designs, except for the patient's name, and were clinically relevant for outpatient primary care. Altogether, scenarios included 19 possible alerts: 5 drug–allergy, 11 drug–drug interactions, and 3 drug–disease alerts for creatinine clearance. We included alerts that were likely to be familiar to prescribers as well as alerts that were likely to be unfamiliar. Some alert dialog boxes contained multiple alerts. Pharmacological information was verified by consulting Micromedex.[21]

We deliberately developed scenarios where correct and incorrect responses to alerts could be clearly defined, enabling us to evaluate how the alert design influenced prescribing errors. Correct actions encompassed a range of anticipated responses: changing the medication, changing the dose, canceling a new medication order, discontinuing a current medication, and

Figure 1 Screen shots of the original alert dialog boxes. The screen shots are from our mock-up of the Veterans Affairs (VA) alert system and are nearly identical to the alerts actually used in clinical care. (A) A drug–disease alert for creatinine clearance. This alert only appears when the prescriber initiates the prescribing process but before the prescriber has selected a specific medication. (B) Drug–allergy and drug–drug interaction alerts triggered by an ibuprofen order. This alert dialog box appears immediately after a prescriber enters information for the prescription, and is formally known as an 'acceptance order check.' (C) The session alert dialog box, which appears when the prescriber attempts to sign the order and is intended to help coordinate care across prescribers (eg, medical students who initiate the order, and physicians who sign it). If one or more alert(s) in figure 1B are bypassed earlier during the ordering session, then the ordering process continues and the overridden alerts appear again in this session alert dialog box. In addition, when two new medications that interact with each other are ordered, the drug–drug interaction alert is only displayed via this session alert. This session alert is formally known as the 'session order check.'
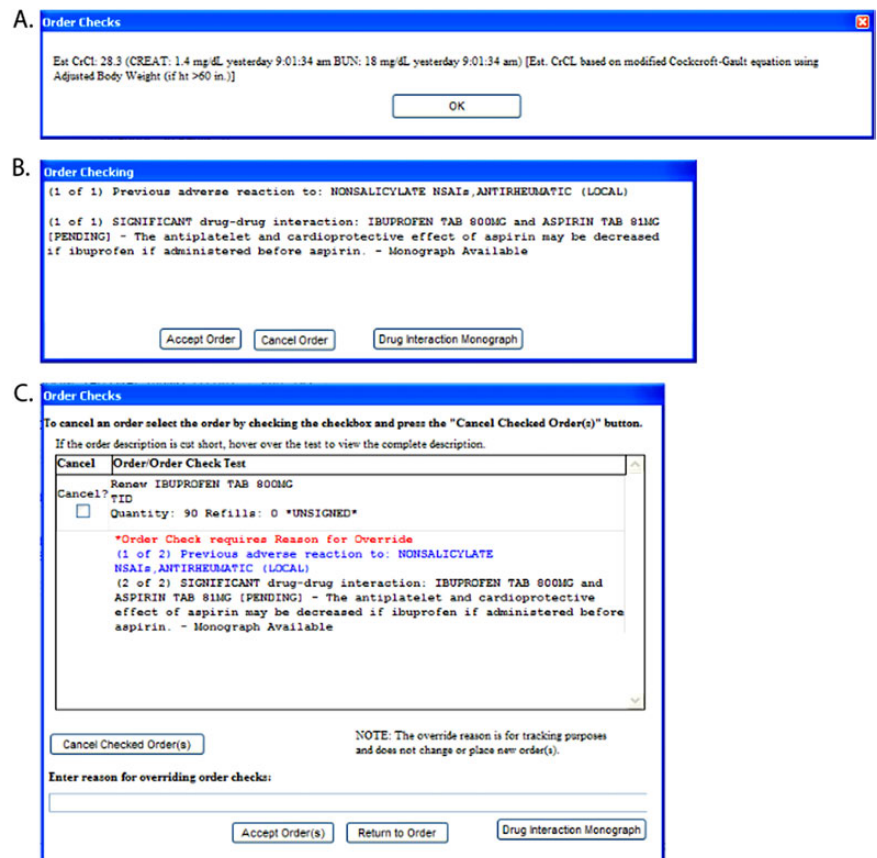
Table 1 Summary of key changes to alert design

Redesign decision	Human factors principle(s)	Rationale or example
A. Alerts presented in a tabular format	Chunking[23]	Grouping information can reduce cognitive effort.
B. Similar information is presented in the same column (ie, see first and second columns of figure 2) with consistent tabs and spacing	Proximity compatibility principle[24 25]	For the second column in figure 2, 'symptoms' was used for the adverse reaction alerts only since, unlike the other alerts, this refers to the patient's medical history. Previous design iterations listed 'severity' in the second column for drug interactions, but the order of this column was changed so that 'risk' information was presented in the same location as for the other alert types.
C. Order of alert types within one alert dialog box; if alerts related to any of the following categories, then they were always presented in this order: ▶ Adverse reaction ▶ Drug–drug interactions ▶ Low creatinine clearance	Visual alerts should be located in the visual field in order of importance[7]	We used a similar order as the original alert design, but intentionally placed alerts about low creatinine clearance last (figure 2). Since the actual severity of an individual warning message can depend on patient-related variables, we assigned an order based on the alert type (adverse reaction, drug–drug interaction, or creatinine clearance) and the general propensity of that alert type to result in an adverse drug event. A standardized order of alert types was decided based upon: 1. Type of risk to the patient, with most severe appearing first 2. Potential for immediate harm, with an allergic response being the most immediate.
D. Medication name added to alert header	Situation awareness[26]	Header indicates what drug is being ordered to help prescribers accurately perceive and understand the prescribing task in the event of disruptions. (An example header is shown in figure 2, 'ordering ibuprofen tab.')
E. Use of signal words: ▶ Adverse reaction ▶ Interactions ▶ Low creatinine clearance	Warning design guidelines recommend including a signal word that conveys level or degree of hazard[7 16]	The American National Standards Institute broadly recommends 'DANGER,' 'WARNING,' and 'CAUTION' for industrial hazards.[16] We considered these terms but ultimately decided against them because: 1. The level of danger associated with an individual alert often depends on patient-related variables that the computer system can neither account for nor detect; thus, these terms may frequently be inaccurate or misleading for a given alert 2. Concerns that use of these terms would decrease prescribers' ability to quickly distinguish between various types of medication hazards and alerts 3. The selected terms were believed to be more clinically meaningful, more closely match end-users' terms, and provide basic information on type of hazard.
F. Scrolling eliminated	Function allocation[27] and hazard control hierarchy[8 28]	With the original alerts, there is a risk that warning messages may be hidden beneath the visual field because the design includes a scrolling mechanism. In this case, manipulating the viewing area is more appropriately allocated to the computer rather than the person. The redesign eliminates the need to scroll; instead, the alert dialog box expands to present text. Thus, we 'designed out' the risk of alerts remaining hidden beneath the viewing area. Designing out the risk is a strong approach for hazard control.
G. Links to additional information embedded within individual alerts	Proximity compatibility principle[24 25] and principle of consistency[29]	Related information should be shown together on interface displays. The coloring and style of links used were intended to be similar to those used in other applications (eg, web sites) and thus aligned with end-users' expectations for hyperlinks.
H. Action buttons separated by greater distance, with 'Cancel' still on the right	Hazard control hierarchy[8 28]	Added space between action buttons 'Accept order' and 'Cancel order' in an attempt to safeguard against accidental clicks of an unintended action. Safeguards generally provide moderate hazard control.

We applied human factors principles to alert interface design, relying heavily on established warning design guidelines[9 15] and Nielsen's principles for interface design.[22]

overriding the alert. A pharmacist and physician from the research team reviewed the scenarios and outlined a pre-defined list of correct and incorrect actions for each task. Scenarios were further refined based on pilot tests with three prescribers who were not included in the main study.

Study design

A scenario-based simulation study was conducted with prescribers using a counterbalanced, crossover design for a within-subject, repeated measures (two sessions: time 1 and time 2) comparison of original versus redesigned alerts. Each prescriber completed two 30 min sessions. In the first session, half of the prescribers received the original alerts and half received the redesigned alerts. Three scenarios were presented in the same order during both sessions. A washout period (2 weeks minimum) was used for each prescriber to reduce the likelihood that they would remember the scenario and previous responses to alerts.

Figure 2 Screen shots of the redesigned alerts. The task that could lead to the alert shown read, '[The patient] has chronic pain due to osteoarthritis and has been managing it with ibuprofen since 2006. [The patient] asks you for a new prescription for his ibuprofen since he is about out and has no more refills. Begin renewing ibuprofen.'

Prescriber selection and recruitment

All VAMC outpatient prescribers, including physicians, nurse practitioners, and clinical pharmacists who worked in primary care and had at least 1 year of VA computerized provider order entry (CPOE) experience were invited to participate. Students and residents were excluded from the study.

Simulation procedure

One researcher facilitated all sessions and read a standardized, scripted introduction to each prescriber. Prescribers were informed that the goal was to compare two different alert designs; they did not receive any training on the alert designs as part of the simulation procedure. Prescribers were asked to complete scenarios as though the patients were real and were informed that no medication orders would be sent to any pharmacy or patient. A physical barrier separated the facilitator from prescribers to reduce the potential for bias related to the facilitator's proximity. Prescribers did not have access to reference materials. The facilitator refrained from offering guidance unless the prescriber experienced a technical difficulty or was unable to proceed without assistance.

Prescribers completed tasks in the same sequence. Clinical information, such as patient's age, diagnoses, laboratory results, and medication lists, was provided on an introductory sheet and in the mock EHR system. Prescribers were asked to complete each of the first two scenarios within 7 min and the researcher gave verbal warnings when 3 and 1 min remained, to simulate the time pressure in a primary care environment. If prescribers exceeded the time limit, they were asked to complete the tasks; thus, prescribing errors were not due to prematurely stopping the scenario. For the first two scenarios, prescribers were asked to 'think aloud'[30] by verbalizing their thought process as well as positive and negative reactions to alerts. If prescribers remained quiet when encountering an alert, the researcher reminded them to think aloud. The evidence is mixed on whether verbalizations may inappropriately slow or speed time measurements.[31] Therefore, taking a conservative approach, we asked prescribers to refrain from the think aloud protocol for the third scenario, where we measured efficiency. Since prescribers had at least 1 year of experience with the original alert design (modeled after the VA alert system), but none with the redesigned alerts, the third scenario was selected to more accurately compare efficiency across the two alert designs by mitigating potential learning effects. We did not specify a time limit for the third scenario.

Data collection and outcome measures

Usability

Video data were recorded via a webcam and Morae software to capture prescribers' verbal and non-verbal behavior along with computer screen, keyboard, and mouse activities.[18] We assessed four usability attributes that are widely accepted[32]: (1) *learnability*: how easy is it for end-users to accomplish fundamental tasks the first time they encounter the design? (2) *efficiency*: after learning the technology, how quickly can end-users complete tasks? (3) *satisfaction*: how well do end-users like using the design? and (4) *usability errors*: does the technology support a low error rate so that few usability issues/critical errors occur while the end-user is using the technology?[32 33]

Learnability and usability errors were evaluated by reviewing video recordings. Efficiency was measured from time-stamped video recordings from the third scenario via: (1) scenario completion time, measured from the appearance of the first alert dialog box until the disappearance of the last alert dialog box; and (2) time spent on alerts, measured by adding the time spent on each alert dialog box that occurred during the scenario. Time was measured from the appearance to the disappearance of each alert dialog box and included any repeat alerts. Satisfaction was measured via the validated, 19-item Computer System Usability Questionnaire (CSUQ),[34] which prescribers were asked to complete after finishing all scenarios. Debrief interviews were conducted as time permitted.

Perceived workload

Workload was measured after each scenario when prescribers completed an electronic tool based on the validated, paper-based NASA Task Load Index (NASA TLX) instrument.[35] The NASA TLX measures perceived mental workload across six subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration. It also yields a global workload score. Higher scores indicate higher perceived workload.

Prescribing errors

To assess prescribing errors, one pharmacist reviewed each video and categorized responses to alerts against the team's predefined list of correct and incorrect actions. A second pharmacist double-checked these categorizations against the predefined criteria. If the correct categorization was unclear, a third team member was consulted.[36] For example, this included cases where a prescriber did not receive an alert due to a technical issue and instances where a prescriber vocalized intent to respond to an alert but did not complete the action. We followed the predefined criteria throughout the analysis. Prescribing errors included errors of

omission (no prescription) and errors of commission (inappropriate prescriptions). In other words, if a prescriber took action based upon an inappropriate alert, or did not take action in response to a valid alert, both situations were counted as a prescribing error. Since multiple alerts can be presented in the same alert dialog box, a maximum of 11 prescribing errors were possible per session.

Data analyses

Summed responses to CSUQ subscales were each analyzed with a paired t test with a Sidak multiple comparison adjustment to control the overall confidence at 95% using SAS V.9.2 (Cary, North Carolina, USA). As outcomes for efficiency and prescribing errors were not normally distributed, the Wilcoxon signed-rank test was performed using SPSS V.20 (IBM). All of these analyses account for within-subject, repeated measures (time 1 and time 2). CSUQ, efficiency, and prescribing errors were analyzed irrespective of the order of alert design since this was a counterbalanced study with a 2-week washout period to mitigate potential learning effects from time 1 to time 2.

For the NASA TLX analysis, we used raw TLX scores,[37] [38] since the weighting procedure has been shown to have limited benefit.[39] [40] NASA TLX scores for workload were analyzed using repeated measures analysis of variance (ANOVA).[41] [42] The linear mixed effects model included fixed effects for scenario (1, 2, 3), alert design (original, redesign), presentation order (redesign first, original first), subscale (each of the six described in the 'Perceived workload' subsection), and the interaction between alert design and subscale. The model also included a random effect for prescriber to account for the correlation between measurements from the same prescriber (ie, clustered data). The intra-provider agreement for NASA TLX was assessed using intra-class correlation coefficients.[43] This determines the proportion of total variability due to variability between prescribers. Comparisons were made between designs for each subscale and p values were adjusted using the Sidak multiple comparison method. NASA TLX analysis was completed using SAS V.9.2.

For all analyses, a statistically significant difference existed for p values <0.05; where data were not normally distributed, median values are presented instead of means.

RESULTS

Prescribers' characteristics

Twenty prescribers (six men and 14 women; 43% response rate) from five different primary care clinics completed the study. Prescribers included 14 physicians, two nurse practitioners, and four clinical pharmacists. The mean duration of experience with the VA CPOE system was 7.5 years (range: 1–13.5 years). The mean age of prescribers was 41 years (range: 29–56 years). Sessions were conducted between October 2011 and March 2012.

Usability attributes

1. Learnability of redesigned alerts. In general, prescribers were able to respond to the redesigned alerts without training or guidance. Three prescribers, however, had difficulty overriding the first redesigned alert that they encountered and required prompting from the facilitator to proceed past the alert dialog box (see also: 4. Usability errors). One prescriber experienced this again when encountering a second alert but was then able to override the redesigned alerts thereafter.
2. Efficiency. Prescribers completed tasks more quickly with the redesigned alerts (median (IQR): 195 (54) s vs original alerts: 246 (108) s; p=0.010). Similarly, prescribers spent

Table 2 Satisfaction scores from the Computer System Usability Questionnaire (CSUQ), as rated by prescribers (N=20)

Score	Original alerts	Redesigned alerts	p Value
Overall satisfaction (items 1–19)	4.3 (0.3)	**5.2 (0.3)**	**0.033**
System usefulness (items 1–8)	4.5 (0.3)	5.2 (0.3)	0.372*
Information quality (items 9–15)	4.2 (0.3)	**5.2 (0.3)**	**0.039***
Interface quality (items 16–18)	4.3 (0.3)	**5.2 (0.3)**	**0.013***

Ratings are derived from 7-point Likert-type scales ranging from 1=strongly disagree to 7=strongly agree. Results are shown as mean (SEM). Statistically significant findings are shown in bold.
*p Values from paired t test adjusted using the Sidak method to control the overall confidence at 95%.

significantly less time on redesigned alerts (56 (47) s vs original alerts: 85 (71) s; p=0.015).

3. Satisfaction. Overall, prescribers were more satisfied with the usability of the redesigned alerts compared to the original alerts (table 2).
4. Usability errors. The redesign reduced or eliminated some types of errors (table 3) but also introduced new usability errors (table 4).

Perceived workload

Based on the statistical model, there was a significant difference in the NASA TLX global workload score between the two alert designs (mean±SEM for original design: 38.6±5.5; redesign: 36.0±5.0; p=0.042). However, there was no significant difference between the original and redesigned alerts for any individual subscale (figure 3). The intra-class correlation coefficient was estimated to be 0.389 and represents the fraction of variability in the NASA TLX that is due to variability between prescribers. See online supplementary file, appendix II for detailed findings.

Prescribing errors

Compared to the original alerts, the redesigned alerts significantly reduced prescribing errors. The median (range) per prescriber across all scenarios was 4 (1–7) errors with the original alerts and 2 (1–5) errors with the redesigned alerts (p=0.024); a maximum of 11 errors were possible per prescriber. For this prescribing error analysis, N=18 prescribers, since two prescribers did not receive one of the alerts due to technical difficulties. Figure 4A–C shows the distribution of error frequency across prescribers, across patient scenarios, and by error type (omission vs commission), respectively. With the original alerts, 60% of prescribers made four or more errors, but with the redesigned alerts, 61% of prescribers made only one or two errors (figure 4A). Most of the original alerts are displayed twice (figure 1B,C), while each redesigned alert was presented once. Thus, we examined whether showing the original alerts a second time substantially changed prescribing actions. In response to the second appearance of an alert, four (20%) prescribers canceled a medication, switching to a correct action; however, three (15%) prescribers stopped a medication order, switching to an incorrect action.

DISCUSSION

To our knowledge, this is the first scenario-based simulation to systematically apply human factors principles to medication alert design and examine the effect on usability, prescribers'

Table 3 Summary of usability errors that were reduced or eliminated with the redesigned alerts

Original alerts		Redesigned alerts		
Design feature	**Associated usability errors**	**Redesign feature**	**Associated usability errors**	**Potential safety implications**
Prescribers are expected to use a scrolling mechanism when the alert text exceeds the visual field.	19 medication alerts* were missed across 6 (30%) prescribers because alerts were hidden by a scrolling mechanism. Two were clinically appropriate alerts for 'critical' drug interactions that remained unnoticed during prescribing.	The scrolling mechanism was eliminated. Instead, the alert dialog box resizes itself to accommodate the amount of alert text.	None	Alerts, including critical alerts, may be inadvertently missed by prescribers when using an alert interface with a scrolling mechanism. A design that eliminates the need to scroll eliminates the risk of alerts that are hidden below the visual field.
'Cancel' checkbox can be selected on the session alert to cancel medication orders (figure 1C).	Two (10%) prescribers interpreted the checkbox as a way to denote what medications were being *ordered* with a corresponding override justification.	The 'Cancel' checkbox design was eliminated.	None, but one limitation of the redesign was that only the medication order being processed could be changed directly from an alert.	With the checkbox design in figure 1C, there is a risk that some medications may be unintentionally cancelled.
To cancel a medication from the session alert (figure 1C), the prescriber must select 'Cancel' and then click 'Cancel Checked Order(s).'	12 (60%) of prescribers vocalized an intent to cancel or discontinue the medication, but did not click the 'Cancel Checked Order(s)' button; thus, they did not complete the cancelation and the medication was ordered.†	The 'Cancel Checked Order(s)' button design was eliminated.	None	The canceling process in figure 1C does not adequately prevent usability errors.† At a minimum, many prescribers are likely to encounter error boxes that disrupt prescriber workflow.
Most alerts that appear at an earlier stage of the ordering process are shown a second time in the session alert (see figure 1B vs C).	Five (25%) prescribers expressed confusion or frustration that the session alert dialog box repeated alerts that appeared earlier in the ordering process.	Session alert removed. Each alert was presented only once.	None, but for the VA system, other design modifications would be needed to coordinate care across prescribers and account for unsigned medications.	Repeated alerts within the same ordering session cause confusion for some prescribers and may promote alert fatigue.
Medications that are in the process of being discontinued in response to an alert still trigger alerts later in the ordering process (figure 1C stage) since the discontinuation has not yet been signed.‡	Five (25%) prescribers expressed confusion when this alert dialog box (figure 1C) included alerts for medications that were in the process of being discontinued. Many of these prescribers expressed uncertainty about what would happen if they selected the 'Cancel' check box for these medications.	If a prescriber starts to discontinue a medication, this does not trigger alerts.	None	Alerts triggered by medications being discontinued cause confusion for prescribers and may promote alert fatigue.

The results presented are applicable across multiple types of alerts (eg, drug–drug interactions, etc).
*Alerts included drug–drug interactions for simvastatin/amiodarone, simvastatin/diltiazem, warfarin/ciprofloxacin, phenelzine/fluoxetine, and selegiline/dextromethorphan.
†An error box would appear in the live system, which should mitigate the risk of not completing the cancelation process, but we were unable to incorporate this error box in our prototype system; we did not count these instances as errors in our analysis of prescribing errors.
‡Medication changes do not become final in the system until they are signed by a prescriber, and therefore, medications undergoing changes (including discontinuation) are still reviewed by the alert system.
VA, Veterans Affairs.

workload, and prescribing errors. Our hypotheses were supported by study results: overall, incorporating human factors principles into the interface design of medication alerts: (1) improved usability for prescribers; (2) reduced perceived workload; and (3) reduced prescribing errors.

Usability

Redesigned alerts supported usability attributes by fostering learnability, reducing some usability errors, increasing prescribers' satisfaction ratings for alerts, and increasing prescribing efficiency. Overall, prescribers were able to learn how to use the redesigned alerts on their own, and design changes reduced critical usability errors. Findings that prescribers sometimes canceled a medication when they thought they were ordering it—or vice versa depending on the alert design (tables 3 and 4)—have not been previously reported for medication alerts. For the redesign (table 1, row H), the 'cancel order' button may have been inadvertently selected as a way to proceed, since buttons on the lower right of a screen display are often used for 'next'

on other software interfaces, rather than stop or 'cancel' as shown in figure 2, but we cannot confirm this interpretation.

Prescribing efficiency increased with the redesigned alerts, even though prescribers had experience with the original alerts but received no training on design modifications. Efficiency data demonstrate that time savings were due to redesigned alerts, rather than other parts of the ordering process. Improved efficiency was likely the result of: (1) modifying the design to present each alert only once, which reduced the number of alerts and alert dialog boxes; and (2) displaying information in a tabular format. Only the original alerts presented creatinine clearance alerts separately (see figure 1 and 2), but this does not explain the differences in efficiency since there were no creatinine clearance alerts during the third scenario where efficiency was measured. Both good and poor alert designs have the potential to shorten the time spent on warnings[44]: a good design may provide better cognitive support and help prescribers extract information quickly, while a poor design may cause prescribers to prematurely dismiss alerts.[44] Thus, time data need to be interpreted in light of other findings: redesigned alerts

Table 4 Summary of new usability errors that were introduced by the redesigned alerts.

Feature of redesigned alert	Usability error	Potential reasons why errors occurred	Potential safety implications
Increased spacing between 'Cancel' and 'Accept' action buttons (figure 2 vs figure 1B)	Two (10%) prescribers vocalized an intention to proceed with an order, but selected a 'Cancel' button on the alert that stopped the ordering process (figure 2).	Lower right portions of a screen display are often used for 'next' on other software interfaces, rather than stop or 'cancel.'	Some orders may be unintentionally canceled, depending on the layout of action buttons on alerts. Increased action button spacing in redesign is NOT recommended (see table 1, row H).
'Accept' action button is disabled until the prescriber enters required override justification(s).	Three (15%) prescribers had difficulty overriding the first redesigned alert that they encountered and required assistance from the facilitator.	A disable button design is not used elsewhere in the EHR. Prescribers may have interpreted the button to mean that overriding the alert was not an option.	A disable button design (figure 2) may impede the processing of medication orders. This design should not be used on alerts, especially when it conflicts with the conventions of the EHR interface.

EHR, electronic health record.

were associated with improved efficiency *and* significant reductions in prescribing errors. This indicates that improved efficiency was *not* due to prescribers prematurely dismissing the redesigned alerts.

Satisfaction ratings were generally higher for the redesigned alerts and were likely due to efforts to reduce text but add more essential data closer to the point-of-decision. Satisfaction findings for 'interface quality' are aligned with literature reports on hazard signs and warning labels that individuals prefer an outline format (eg, tabular layout) rather than prose, and that formatting can help individuals easily find information.[16] Although design changes were favorably received, the ideal amount of information to present is still unknown. In addition, ratings for the usefulness of alerts were higher with the redesign but were not statistically significant. With the original alerts, prescribers perceived the design as still 'useful' and rated this aspect higher than the other three satisfaction scores.

Workload

Redesigned alerts resulted in a modest but significant reduction in perceived workload. Neither alert design was particularly demanding, since global workload scores were below the midpoint of 50. Prescribers perceived their performance favorably for both designs, since lower TLX scores indicate more successful performance. Perceptions of performance were not aligned with results for prescribing errors, since errors were high for the original alerts but significantly reduced with the redesigned alerts.
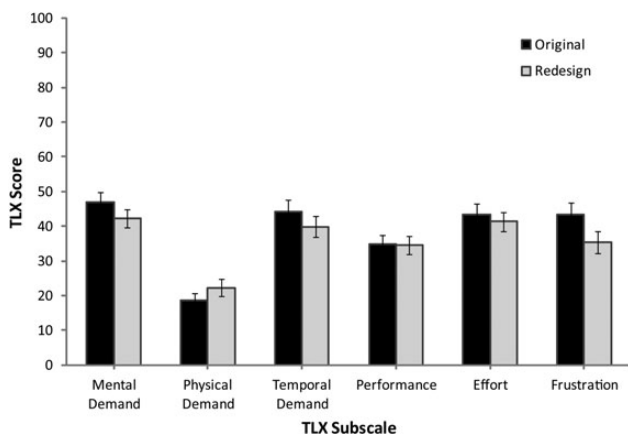

Figure 3 Prescribers' (N=20) perceived workload for the NASA Task Load Index (TLX) subscales. Results are shown as mean±SEM. For all subscales, higher scores indicate greater perceived workload. There was no statistically significant difference between the original and redesigned alerts for any individual subscale.

Prescribing errors

Three aspects of the redesign likely improved prescribing. First, redesigns followed human factors principles. This reduced some prescribing errors, such as those caused by 'hidden alerts' (table 3). Second, organizing information into groups can help maintain an individual's attention and facilitate information search and acquisition, a key step that must occur for warning effectiveness.[16 45] The tabular format for redesigned alerts may have helped prescribers read and encode information, thereby reducing prescribing errors. Third, redesigned alerts provided some clinical data closer to the point-of-decision. For example, details on patients' previous adverse reactions were readily available in the mock EHR system for both alert designs, but only the redesigned alerts displayed adverse reaction details. Moreover, the mock EHR system did not provide information on the risks of low creatinine clearance or what specific medications to avoid; these details were provided by the redesigned alerts only. These redesign features are aligned with literature that underscores the importance of decision support tools that provide information at the right time and place.[46 47]

Overall, prescribing errors for both alert designs were higher than we anticipated. We did not observe any cases where prescribers appeared to click past an alert dialog box without examining it, and prescribers were asked to 'think aloud' when they encountered alerts; thus, a lack of 'attention switch'[16] (ie, shifting cognitive focus) to the alert dialog box is unlikely to explain the number of prescribing errors. Even though redesigned alerts significantly reduced prescribing errors, there was no significant difference in how prescribers rated the alert usefulness (table 2) or their performance (figure 3) across the two alert designs. These findings indicate that subjective feedback from prescribers is insufficient to evaluate the safety and utility of alert systems in an accurate manner.

This study has limitations that should be considered when interpreting results. Prescribers were informed that the patients were fictitious and may not have always responded to alerts in the same manner as they would for an actual patient. In addition, our list of pre-defined correct and incorrect actions may be more conservative than actual clinical practice. Prescribing errors were lower for the third scenario, and this may be due to a learning effect and/or this scenario may have been clinically easier to address. Moreover, in clinical practice, orders are double-checked by a pharmacist during dispensing to help prevent errors from reaching the patient, and not all prescribing errors lead to an adverse event. This study focused on outpatient care, and findings may or may not apply to inpatient care. Prescribers were recruited from a single VAMC, and although we evaluated alert designs that are used nationally across the VA, prescribers in other geographic regions may have responded differently. Nevertheless, we have no evidence
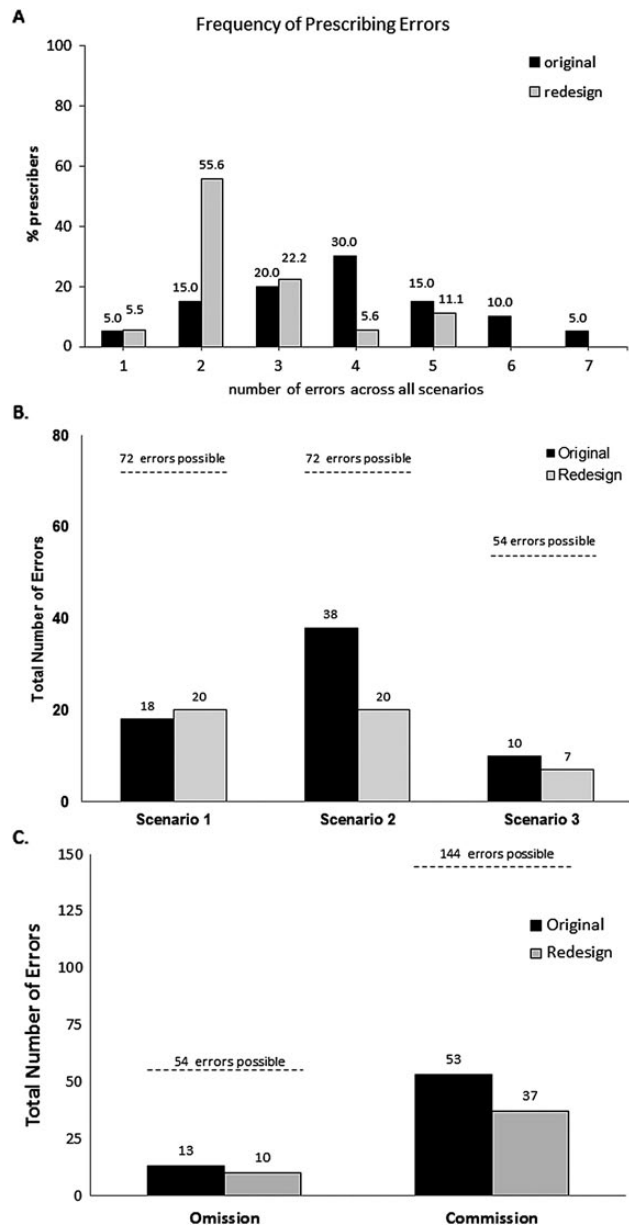
Figure 4 (A–C) Frequency of prescribing errors. These include both errors of omission (no prescription) and errors of commission (inappropriate prescriptions). (A) Distribution of error frequency across all scenarios when prescribers used the original versus redesigned alerts. Prescribers made fewer errors with the redesigned alerts, shifting the distribution to the left. N=20 prescribers for the original design. N=18 prescribers for the redesign, since two prescribers did not receive one of the drug–drug interaction alerts due to technical difficulties. (B) Frequency of prescribing errors across the three patient scenarios (N=18 prescribers for each alert design). (C) Frequency of omission and commission errors (N=18 prescribers for each alert design). The majority of errors were errors of commission (and more commission errors were possible). For (B) and (C), the maximum number of possible errors for each scenario and error type, respectively, is noted on the graph. For example, for scenario 1 in figure 4B, four prescribing errors were possible with the alerts (4×18 prescribers=72 errors possible). To allow for direct comparison of the absolute number of errors for (B) and (C), we excluded all data from two prescribers who did not receive one of the alerts due to technical difficulties; thus, the total possible errors for these two graphs is 11×18 prescribers=198.

Box 1   Summary of key study recommendations.

▶ To reduce prescribing errors, human factors principles and warning design guidelines should be systematically integrated into alert interface designs
▶ A tabular format for alert text may improve efficiency, increase prescribers' perception of information quality, and promote better prescribing decisions.
▶ Text on the main alert dialog should not be hidden by scrolling mechanism(s), since this can be a safety risk. (See **Table 3**)
▶ Alerts should undergo formal, systematic usability testing[45] to assess the potential for critical usability errors, including design features that may lead prescribers to cancel or order a medication unintentionally.
▶ Repeating alerts in the same ordering session for a given patient case did not substantially reduce prescribing errors. Design strategies should be implemented to reduce this type of alert repetition and decrease the risk of alert fatigue.
▶ Mechanisms intended to coordinate care across prescribers with the original design resulted in repeated alerts and several usability issues. The redesigned alerts have the ability to support coordination in some, but not all, cases. For example, more advanced alerting mechanisms need to be developed for medications that are "on hold" or in progress to coordinate care among multiple prescribers. This is especially important as medication lists become shared across different healthcare institutions.

suggesting that the findings would be limited to only one type of clinical setting or geographic location. Box 1 summarizes key recommendations based upon study findings. Finally, although studies support the use of the tabular format for presenting warnings,[16] this study did not separate the impact of the tabular format from that of other modifications in the redesign. This study was not conducted to specifically examine the effect of a singular design feature, rather, the redesigned alerts encompassed a set of design changes based upon established human factors principles (table 1) which, used together, produced the findings.

CONCLUSION

Incorporating human factors principles into alert design significantly improved usability for prescribers and reduced prescribing errors. Study findings suggest that a tabular format for presenting multiple alerts and grouping similar information together may aid prescribing decisions. This research also identified some features, such as scrolling, that pose high patient safety risks. Many of our findings are consistent with evidence from warnings literature, but this study provides some of the first experimental evidence about the presentation of information on computerized medication alerts. Results indicate that even in an environment where prescribers are likely to shift their cognitive focus from the ordering system to alerts, prescribing errors remained high. This finding underscores the need to improve alert interfaces. Ultimately, alert system

designers should design alert interfaces based on research evidence to promote safety.

Author affiliations
[1]Center for Health Information and Communication, Health Services Research and Development Service CIN 13-416, Department of Veterans Affairs, Veterans Health Administration, Richard L. Roudebush VA Medical Center, Indianapolis, Indiana, USA
[2]Indiana University Center for Health Services and Outcomes Research and Regenstrief Institute, Inc., Indianapolis, Indiana, USA
[3]College of Pharmacy, Purdue University, West Lafayette, Indiana, USA
[4]School of Pharmacy, University of Kansas, Lawrence, Kansas, USA
[5]Department of Veterans Affairs, VA Greater Los Angeles Healthcare System, Los Angeles, California, USA
[6]David Geffen School of Medicine, University of California, Los Angeles, California, USA
[7]Department of Medicine, Indiana University School of Medicine, Indianapolis, Indiana, USA
[8]Department of Psychology, Purdue University, West Lafayette, Indiana, USA
[9]Department of Biostatistics, Indiana University School of Medicine, Indianapolis, Indiana, USA
[10]Department of Nephrology Services, Department of Veterans Affairs, Temple, Texas, USA
[11]Department of Veterans Affairs, Office of Information, Bay Pines, Florida, USA
[12]Department of Veterans Affairs, Office of Information, Salt Lake City, Utah, USA
[13]Department of BioHealth Informatics, School of Informatics and Computing, Indiana University-Purdue University, Indianapolis, Indiana, USA
[14]Human Factors Engineering, Office of Informatics and Analytics, Veterans Health Administration, Louisville, Kentucky, USA

REFERENCES

1 Schedlbauer A, Prasad V, Mulvaney C, et al. What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior? *J Am Med Inform Assoc* 2009;16:531–8.

2 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 decision-support systems—I-MeDeSA. *J Am Med Inform Assoc* 2011;18(Suppl 1):i62–72.

3 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.

4 Strom BL, Schinnar R, Aberra F, et al. Unintended effects of a computerized physician order entry nearly hard-stop alert to prevent a drug interaction: a randomized controlled trial. *Arch Intern Med* 2010;170:1578–83.

5 Khajouei R, Jaspers MW. The impact of CPOE medication systems' design aspects on usability, workflow and medication orders: a systematic review. *Methods Inf Med* 2010;49:3–19.

6 Horsky J, Schiff GD, Johnston D, et al. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. *J Biomed Inform* 2012;45:1202–16.

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 Wogalter MS, Laughery KR. Ch 32: Warnings and hazard communication. In: Salvendy G. ed *Handbook of human factors and ergonomics*. 3rd edn. Hoboken, NJ: John Wiley and Sons, 2006:889–911.

9 Wogalter MS. Handbook of warnings (human factors and ergonomics). Salvendy G, ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2006, pp 1–841.

10 Paterno MD, Maviglia SM, Gorman PN, et al. Tiering drug-drug interaction alerts by severity increases compliance rates. *J Am Med Inform Assoc* 2009;16:40–6.

11 Scott GP, Shah P, Wyatt JC, et al. Making electronic prescribing alerts more effective: scenario-based experimental study in junior doctors. *J Am Med Inform Assoc* 2011;18:789–98.

12 Lo HG, Matheny ME, Seger DL, et al. Impact of non-interruptive medication laboratory monitoring alerts in ambulatory care. *J Am Med Inform Assoc* 2009;16:66–71.

13 Palen TE, Raebel M, Lyons E, et al. Evaluation of laboratory monitoring alerts within a computerized physician order entry system for medication orders. *Am J Manag Care* 2006;12:389–95.

14 Russ AL, Zillich AJ, McManus MS, et al. Prescribers' interactions with medication alerts at the point of prescribing: a multi-method, in situ investigation of the human-computer interaction. *Int J Med Inform* 2012;81:232–43.

15 Wogalter MS, Conzola VC, Smith-Jackson TL. Research-based guidelines for warning design and evaluation. *Appl Ergon* 2002;33:219–30.

16 Wogalter MS. Ch 5: Communication-human information processing (C-HIP) model. In: Wogalter MS. ed *Handbook of warnings*. Mahwah, NJ: Lawrence Erlbaum Associates, 2006:51–61.

17 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.

18 Russ AL, Weiner M, Russell SA, et al. Design and implementation of a hospital-based usability laboratory: insights from a Department of Veterans Affairs laboratory for health information technology. *Jt Comm J Qual Patient Saf* 2012;38:531–40.

19 About VHA. [updated Oct 5, 2011; cited Jan 4, 2013]; http://www.va.gov/health/aboutVHA.asp (accessed 28 Feb 2014).

20 Talmon J, Ammenwerth E, Brender J, et al. STARE-HI—Statement on reporting of evaluation studies in health informatics. *Int J Med Inform* 2009;78:1–9.

21 Micromedex Solutions [Internet database]. Truven Health Analytics. Greenwood Village, CO [cited 2011]; http://www.micromedexsolutions.com/micromedex2/librarian (accessed 28 Feb 2014).

22 Nielsen J. Enhancing the Explanatory Power of Usability Heuristics. Human Factors in Computing Systems. CHI. 1994:152–8.

23 Sanders M, McCormick E. *Human factors in engineering and design*. 7th edn. McGraw-Hill, 1976.

24 Marino CJ, Mahan RR. Configural displays can improve nutrition-related decisions: an application of the proximity compatibility principle. *Hum Factors* 2005;47:121–30.

25 Wickens CD. *Engineering psychology and human performance*. 2nd edn. Harper Collins, 1992.

26 Wright MC, Taekman JM, Endsley MR. Objective measures of situation awareness in a simulated medical environment. *Qual Saf Health Care* 2004;13(Suppl 1):i65–71.

27 Czaja SJ, Nair SN. Human factors engineering and systems design. In: Salvendy G. ed *Handbook of human factors and ergonomics*. 3rd edn. Hoboken, NJ: John Wiley and Sons, 2006:32–49.

28 Scanlon MC, Karsh BT. Value of human factors to medication and patient safety in the intensive care unit. *Crit Care Med* 2010;38(6 Suppl):S90–6.

29 Wickens CD, Lee JD, Liu Y, et al. An introduction to human factors engineering. 2nd edn. Upper Saddle River, NJ: Pearson Prentice Hall, 2004.

30 Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *Int J Med Inform* 2009;78:340–53.

31 Lewis JR. Section 8: Human-computer interaction, Ch 49: Usability testing. In: Salvendy G. ed *Handbook of human factors*. 3rd edn. Hoboken, NJ: John Wiley and Sons, 2006:1282.

32 Nielsen J. Ch 2: What is usability? In: *Usability engineering*. San Francisco, CA: Morgan Kaufmann, 1993:23–48.

33 Nielsen J. Usability 101: Introduction to Usability. Jan 4, 2012 [cited Feb 5, 2013]. http://www.nngroup.com/articles/usability-101-introduction-to-usability/ (accessed 28 Feb 2014).

34 Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int J Hum-Comput Interact* 1995;7:57–78.

35 Hart SG, Staveland LE. Development of the NASA TLX (Task Load Index): results from empirical and theoretical research. In: Hancock PA, Meshkati N, eds; Human Mental Workload. North-Holland: Elsevier Science, 1988;239–50.

36 Miller AM, Boro MS, Korman NE, *et al*. Provider and pharmacist responses to warfarin drug-drug interaction alerts: a study of healthcare downstream of CPOE alerts. *J Am Med Inform Assoc* 2011;18(Suppl 1):i45–50.

37 Hart S. NASA-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Egronomics Society 50th Annual Meeting*. 2006:904–8.

38 Sawin DA, Scerbo MW. Effects of instruction type and boredom proneness in vigilance: implications for boredom and workload. *Hum Factors* 1995;37:752–65.

39 Nygren TE. Psychometric properties of subjective workload measurement techniques: implications for their use in the assessment of perceived mental workload. *Hum Factors* 1991;33:17–33.

40 Hendy KC, Hamilton KM, Landry LN. Measuring subjective workload: when is one scale better than many?. *Hum Factors* 1993;35:579–601.

41 Satterfield K, Ramirez R, Shaw T, *et al*. Measuring workload during a dynamic supervisory control task using cerebral blood flow velocity and the NASA-TLX. *Proceedings of the Human Factors and Egronomics Society 56th Annual Meeting*. 2012:163–7.

42 Temple JG, Warm JS, Dember WN, *et al*. The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Hum Factors* 2000;42:183–94.

43 Nanji KC, Slight SP, Seger DL, *et al*. Overrides of medication-related clinical decision support alerts in outpatients. *J Am Med Inform Assoc* 2014;21:487–491.

44 Smith-Jackson TL, Wogalter MS. Methods and procedures in warnings research. In: Wogalter MS. ed *Handbook of warnings*. Mahwah, NJ: Lawrence Erlbaum Associates, 2006:23–33.

45 Wogalter MS, Vigilante WJ. Ch 18: Attention switch and maintenance. In: Wogalter MS. ed *Handbook of warnings*. Mahwah, NJ: Lawrence Erlbaum Associates, 2006: 245–65.

46 Hayward J, Thomson F, Milne H, *et al*. 'Too much, too late': mixed methods multi-channel video recording study of computerized decision support systems and GP prescribing. *J Am Med Inform Assoc* 2013;20:e76–84.

47 Bates DW, Kuperman GJ, Wang S, *et al*. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003;10:523–30.