International Journal of Medical Informatics 133 (2020) 104027

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

An experimental investigation of the impact of alert frequency and relevance on alert dwell time

Melissa T. Baysari[a,b,*], Wu Yi Zheng[a,b], Amina Tariq[c], Maureen Heywood[d], Grace Scott[b], Ling Li[b], Bethany A. Van Dort[a,b], Kasun Rathnayake[b], Richard Day[e,f], Johanna I. Westbrook[b]

[a] The University of Sydney, Faculty of Health Sciences, Sydney Australia
[b] Centre for Health Systems & Safety Research, Australian Institute of Health Innovation, Macquarie University, Sydney Australia
[c] School of Public Health and Social Work, Faculty of Health, Queensland University of Technology, Australia
[d] Department of Pharmacy, St Vincent's Hospital, Sydney Australia
[e] Department of Clinical Pharmacology & Toxicology, St Vincent's Hospital, Sydney Australia
[f] St Vincent's Clinical School, UNSW Medicine, UNSW, Sydney Australia

ARTICLE INFO

Keywords:
Alert
Alert dwell time
CPOE

ABSTRACT

Aim: This study aimed to determine the impact of alert frequency and relevance on alert dwell time.
Method: A 2 × 3 design was used where 127 university students completed 60 prescribing tasks and were presented with a variable frequency of computerized alerts (low, medium and high) with variable relevance (low and high). Participants were instructed to override an alert if it was not relevant to their prescription, and to cancel the order if the alert signalled an error in their order.
Results: Participants presented with a small number of alerts spent more time attending to alert content than participants presented with a medium or high number of alerts (respectively median 15.6 s vs 10.8 vs 10.2 s). Alert relevance had no impact on alert dwell time. Alerts requiring an override response were 4.5 times more likely to be correctly actioned than alerts requiring the order to be cancelled.
Discussion: Dwell time was influenced by alert frequency, with greater exposure to alerts associated with shorter dwell times. We hypothesize that this was because participants came to learn that spending time on alert information was unnecessary. We propose that when users experience no consequences or feedback from overriding alerts they quickly learn that this action is more efficient and so more rewarding than taking any other action.

1. Introduction

Alerts in computerized provider order entry (CPOE) systems can have a positive impact on prescribing behaviors [1–3]. However, there is also a large body of work demonstrating that providers override alerts, up to 95% of the time [4–6]. A recent systematic review of the effectiveness of medication alerts in hospital CPOE systems revealed that approximately half of the studies examining the impact of an alert set (e.g. drug-drug interaction alerts, dose-range alerts) on prescribing found a positive effect; the remainder found no impact or a detrimental impact of alerts on prescribing behaviors [7].

Research evaluating alert effectiveness has largely comprised assessments of alert overrides, with override rate viewed as a surrogate inverse indicator for alert effectiveness. If an alert is frequently overridden, then it naturally follows that the alert is not providing prescribers with useful or relevant information. More recently, researchers have cautioned against using override rates as a means of assessing alert effectiveness [8,9]. This is primarily because an override does not tell the full story about an alert's impact on prescribing (e.g. changes that are made to prescriptions long after the alert is triggered and clicked past). In a study that used field observations of prescribers to explore prescriber-alert interactions, it was discovered that some alerts that were overridden were still useful as they prompted prescribers to discuss information with patients [10]. This positive effect would not have been captured if alert override had been used as the only indicator of effectiveness.

Relying on override rate to assess computerized alerts also assumes that alerts are being read and determined to be irrelevant by users. Our research suggests that this is unlikely to be the case when users are experiencing alert overload [11]. We shadowed teams of doctors as

* Corresponding author at: Charles Perkins Centre D17, The University of Sydney, NSW 2006, Australia.
E-mail address: Melissa.baysari@sydney.edu.au (M.T. Baysari).

https://doi.org/10.1016/j.ijmedinf.2019.104027
Received 2 October 2018; Received in revised form 10 October 2019; Accepted 28 October 2019

they prescribed medications on ward-rounds using a CPOE and observed a very large number of alerts being triggered (approximately half the medication orders triggered one or more alerts). We noticed that system users not only overrode most of the alerts triggered, but rarely read the alert content [11].

Alert dwell time is a relatively under-explored outcome measure that provides useful information about an alert's impact on prescribing. Alert dwell time is the time elapsed from the generation of an alert (i.e. alert presentation) to dismissal of the alert window (i.e. cancel medication order or override alert) [12]. Thus, in instances where alerts are not being read or read fully, one would expect alert dwell time to be shorter than in cases where alerts are being attended to completely.

To date, few studies of alert dwell time have been undertaken. In a US study, dwell time for over 100,000 computerized alerts generated over a three-year period in a children's hospital was calculated [12]. Median dwell time was determined to be 8 s with more frequently occurring alerts acted on faster than those that were presented only once (7 s vs. 11 s). In experimental studies, alert dwell time (representing 'efficiency with which alerts are dealt with') has been used to examine the impact of alert interface design on prescribing behavior [13,14]. These studies have shown that improving the design of alerts (by applying human factors principles) results in faster responses to alerts [13,14].

In this experimental study, we set out to explore the impact of two alert characteristics, frequency and relevance, on alert dwell time. These two constructs were selected because they have been reported to influence users' attitudes and behaviour towards computerized alerts [6,11,15,16]. That is, when users experience a large number of alerts and alerts that are not clinically relevant to their patients, this has been reported to lead to frustration and annoyance, and to users dismissing alerts without fully attending to alert content.

2. Method

2.1. Design

This study employed a 3 × 2 design, including three levels of alert frequency (low, medium, and high), and 2 levels of alert relevance (low and high) as shown in Fig. 1.

We defined a relevant alert as one that conveyed an error in a medication order, based on the patient's characteristics (e.g. age, allergies, etc) and medications. For example, if a patient was allergic to penicillins, prescribing a medication containing a penicillin would trigger a relevant alert.

2.2. Testing environment

Participants completed the study using the training module of the commercial CPOE system, MedChart® (http://www.dxc.technology/providers/offerings/139499/140202-medchart_electronic_medication_management). MedChart® is an end-to-end medication management system currently in use in several leading tertiary care hospitals in Australia, New Zealand and the United Kingdom. The study was conducted in a simulated environment, with fictional patient information. Scenarios were developed by a clinical pharmacist highly experienced in using the CPOE system and the alerts. During the study, participants' interactions with the system (i.e. their screens) were video recorded using Morae® software (http://www.techsmith.com/morae.asp).

2.3. Participants

127 university undergraduate students, including medicine, medical science, clinical science, psychology and science students (87 females, 40 males; average age 20.8, range 17–49 years) were recruited via advertisements in student newsletters and mailing lists, and through posters around university campuses. Students were intentionally recruited for the study because they had no prior experience prescribing using a CPOE system and no previous exposure to medication alerts. We pilot tested all scenarios with non-clinical participants and sought student feedback to ensure that scenarios and alert content were understandable. An example alert appears in the Appendix (Figure A1).

To minimise participants focusing on alerts during the experiment, students were blinded to the true purpose of the study (i.e. to examine alert dwell time). Instead, participants were informed that the study aimed to evaluate the usability of the CPOE system. To simulate a real-life, time-pressured prescribing environment, participants were also instructed to complete the prescribing tasks as quickly as possible.

2.4. Study procedure

Students received a 15-minute demonstration on how to prescribe using the CPOE system and how to respond to alerts. The same investigator (WYZ) delivered this one-on-one training to all participants, facilitated by a series of short training videos. During training, participants were instructed to override an alert and continue with their order if they believed the alert was not relevant to their prescription, and to cancel or remove the order (i.e. not proceed with the prescription) if they believed the alert signalled an error in their order. Thus, overriding alerts resulted in orders being prescribed, and adhering to alert recommendations resulted in orders being cancelled. Alert overrides were used to determine the proportion of correct responses to alerts. All information required to make an assessment of alert relevance (e.g. patient allergy status) was made available to participants on the prescribing screen and on the task sheets for each of the six scenarios (see Figures A2a and A2b in the appendix).

Training was directly followed by Phase 1, during which participants were required to order 60 medications using the CPOE system and to respond to any computerized alerts that were triggered. As shown in Fig. 1, students were exposed to a variable rate of alerts and a variable rate of relevant alerts during Phase 1, depending on their allocated group. For orders that triggered an alert, only one computerized alert was displayed per order.

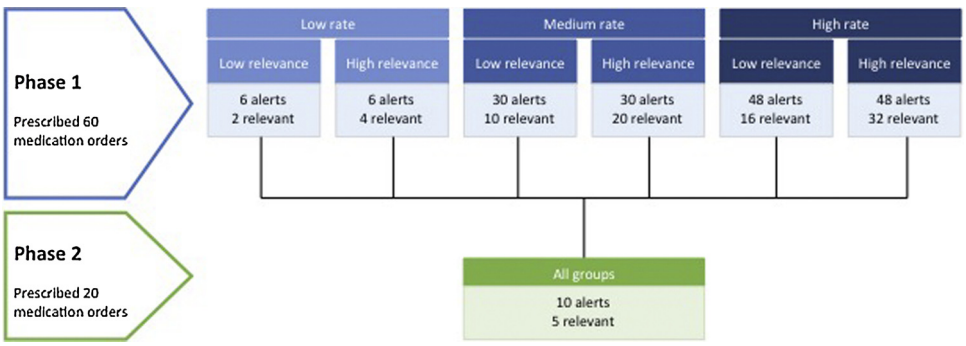

Fig. 1. Study procedure. Participants received a variable number of alerts during Phase 1 but the same alerts during Phase 2.

To assess the impact of alert frequency and relevance on alert dwell time, all participants then completed Phase 2, where they were required to prescribe 20 medications and experienced an alert rate and relevance rate of 50%. Participants completed the same scenarios and were exposed to the same computerized alerts during Phase 2. Phase 2 allowed a comparison between participants' behaviors in the six different groups (who were exposed to variable levels of alerts in Phase 1) when subsequently exposed to the same level and type of alerts (see Fig. 1).

Following completion of the study, participants were informed of the true purpose of the study and any questions they had were answered by the researcher.

This study was approved by Macquarie University and University of New South Wales Human Research Ethics Committee.

2.5. Statistical analysis

Only data from Phase 2, where all participants were exposed to the same computerized alerts, were analysed. Our primary outcome measure was alert dwell time for participants in each of the six groups exposed to different alert frequency and relevance. Distribution of alert dwell time data was checked and non-parametric tests were deemed to be appropriate. We examined whether there were differences between groups in alert dwell time using a Kruskal-Wallis Test. Our secondary outcome was accurate responses to alerts. We tested for differences between groups in correct responses to alerts using a logistic regression analysis. The relationship between alert dwell time and correct responses to alerts was examined using a logistic regression analysis, with alert dwell time as the independent variable and response to alerts (correct or incorrect) as the dependent variable. Descriptive statistics were presented for demographic data. All analyses were carried out using IBM SPSS Statistics 25 (IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp).

3. Results

We found no association between alert relevance and alert dwell time (all $p > 0.05$) or correct responses to alerts (Wald = 0.2, $p = 0.7$), thus we collapsed the relevance groups into one. All subsequent analyses were performed on the three groups which experienced different levels of alert frequency (high, medium and low).

3.1. Alert dwell time during Phase 2

Median dwell time during Phase 2 was 12 s (IQR: 6.6–19.8 s). We found statistically significant differences in dwell time by alert frequency ($X^2(2) = 72.6$, $p < 0.001$). As shown in Table 1, participants who were presented with a low rate of alerts during Phase 1 displayed longer dwell times in Phase 2 than those who were presented with a medium (mean rank = 745.6 vs. mean rank = 580.9, $p < 0.001$) or high rate of alerts in Phase 1 (mean rank = 745.6 vs. mean rank = 544.4, $p < 0.001$). There was no difference in alert dwell time between those in the medium frequency group and those in the high frequency group ($p = 0.418$).

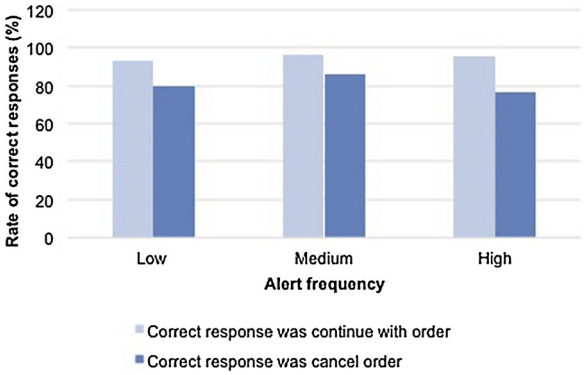

Fig. 2. Percentage of correct responses to alerts during Phase 2 by alert frequency group and type of response required.

3.2. Accuracy of responses to alerts during Phase 2

We found no significant differences between alert frequency groups in accurate responses to alerts (Wald = 0.08, $p = 0.8$). Alert dwell time was also not associated with correct responses to alerts (Wald = 0.02, $p = 0.9$).

Fig. 2 shows the rate of correct responses to alerts for participants in each group, by type of response required to alerts. As shown in Fig. 2, when participants were required to cancel an order in response to an alert, they made fewer correct responses than when they were required to override the alert and continue with their order. Alerts requiring an override were 4.5 times more likely to be correctly responded to than alerts requiring the order to be cancelled (Wald = 51.0, $p < 0.001$).

4. Discussion

This experiment showed that varying the alert frequency experienced by participants resulted in significant differences in alert dwell time. Participants presented with a large number of alerts spent less time attending to alert content than participants presented with a small number of alerts. However, alert relevance appeared to have little impact on alert dwell time or correct responses to alerts.

Previous studies have shown that an individual's attitude towards automation and subsequent use of a system is influenced by their trust in automation, which is impacted by the reliability or accuracy of the automation [17]. Interestingly, we found that alert relevance (i.e. accuracy of automation) had no impact on alert dwell time or on correct responses to alerts. We hypothesize that this was because participants received no feedback on whether their responses to alerts were correct or incorrect. As a result, they were unable to ascertain how reliable the automation was, that is, how accurate the alerting system was in identifying errors in patients' orders. This is not inconsistent with what occurs in practice. Feedback is rarely provided to hospital doctors on prescribing in general [18–20], and even less so with respect to their responses to computerized alerts. Equipped with limited information on the consequences of their alert overrides, doctors quickly learn that the override response is an action which allows them to proceed with their

Table 1
Comparison of dwell times and correct responses to alerts in the Low, Medium and High alert frequency groups during Phase 2. All participants prescribed 20 medications and were presented with the same 10 computerized alerts.

Alert frequency group	Median dwell time (seconds)	Interquartile range[c] (seconds)	Number of alerts responded to correctly (%)
Low alert rate[a,b] (10%)	15.6	9.0 – 25.8	346 (86.7)
Medium alert rate[a] (50%)	10.8	6.0 – 18.6	387 (91.1)
High alert rate[b] (80%)	10.2	5.4 – 16.8	361 (86.2)

[a] Significant difference in median dwell time between low group and medium group.
[b] Significant difference in median dwell time between low group and high group.
[c] The interquartile range represents the 25th percentile and the 75th percentile.

order rapidly and efficiently.

Supporting this hypothesis is our finding that participants, in response to alerts, were more likely to correctly override an alert compared to correctly cancelling an order. This could suggest that the act of overriding an alert produced a more favourable outcome (i.e. a complete list of medications as per the task sheet) for participants than the alternative. Again, this aligns with research which has shown that alerts are often viewed as a barrier to doctors quickly prescribing an appropriate medication for their patients [6,21].

Importantly, the notion of participants learning the value of overriding alerts also explains our key study result, that participants in the low alert rate group displayed longer dwell times than those in the medium and high alert rate groups. Following greater exposure to alerts, participants could have come to learn that spending time on alert information was unnecessary to complete the task.

Our study was limited in that we explored alert dwell time in a small number of participants in a simulated environment. The practicalities of running a study of this nature (i.e. reasonable length of time for each session and adopting real-life alerts that were not too difficult) prevented us from exposing students to a very large task burden. We did not monitor whether participants read all relevant patient details on screen and in the study packs, preventing us from knowing whether alert relevance was accurately determined by users. Recruitment of students was intentional to control previous exposure to alerts, but in doing so, our results, especially those related to clinical relevance of alerts, are likely to have been influenced by their limited clinical knowledge and understanding.

Overall, this study showed that exposure to a low frequency of alerts resulted in users spending more time attending to alert information than exposure to a high frequency of alerts. We hypothesize that this is because participants exposed to frequent alerts learnt the value of not attending to alert content so they could proceed with their orders. When users receive no consequences or feedback from overriding alerts they quickly learn that this action is more efficient and so more rewarding than taking any other action. Thus, providing feedback to prescribers with respect to the consequences of their alert responses is likely to lead to greater attention directed to alert content. In designing effective alerting systems, organizations should consider reviewing and removing alerts where an override response from the user consistently has no consequence.

Authors' contributions

MB, AT and LL contributed to conception and design of the study, WYZ undertook recruitment, data collection and analysis, MH assisted with design of scenarios and alerts, GS, LL and KR assisted with data analysis, all authors contributed to data interpretation, writing of the manuscript and approved the final manuscript for submission.

Summary Points

What was already known on the topic

- Computerised alerts have the potential to impact on prescribing behavior, but most alerts are overridden by users
- When presented with too many irrelevant alerts, users report that they stop reading the alert information

What this study has added to our knowledge

- Users presented with a large number of alerts spend less time reading alerts than users presented with a small number of alerts
- In this simulated setting, alert relevance had no impact on time spent attending to alert content
- Participants made more correct responses to alerts when they were required to override the alert compared to when they were required to cancel the order

Declaration of Competing Interest

The authors have no conflicts of interests.

Acknowledgement

This research was supported by NHMRC Program Grant 1054146.

References

[1] A. Schedlbauer, V. Prasad, C. Mulvaney, S. Phansalkar, W. Stanton, D.W. Bates, A.J. Avery, What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior? J. Am. Med. Inform. Assoc. 16 (2009) 531–538.
[2] A.X. Garg, N.K.J. Adhikari, H. McDonald, M.P. Rosas-Arellano, P.J. Devereaux, J. Beyene, J. Sam, B. Haynes, Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review, J. Am. Med. Assoc. 293 (2005) 1223–1238.
[3] W.L. Galanter, R.J. Didomenico, A. Polikaitis, A trial of automated decision support alerts for contraindicated medications using computerized physician order entry, J. Am. Med. Inform. Assoc. 12 (2005) 269–274.
[4] T.C. Hsieh, G.J. Kuperman, T. Jaggi, P. Hojnowski-Diaz, J. Fiskio, D.H. Williams, D.W. Bates, T.K. Gandhi, Characteristics and consequences of allergy alert overrides in a computerized physician order entry system, J. Am. Med. Inform. Assoc. 11 (2004) 482–491.
[5] S.N. Weingart, M. Toth, D.Z. Sands, M.D. Aronson, R.B. Davis, R.S. Phillips, Physicians' decisions to override computerised drug alerts in primary care, Arch. Intern. Med. 163 (2003) 2625–2631.
[6] H. van der Sijs, J. Aarts, A. Vulto, M. Berg, Overriding of drug safety alerts in computerized physician order entry, J. Am. Med. Inform. Assoc. 13 (2006) 138–147.
[7] N. Page, M.T. Baysari, J.I. Westbrook, A systematic review of the effectiveness of interruptive medication prescribing alerts in hospital CPOE systems to change prescriber behavior and improve patient safety, Int. J. Med. Inform. 105 (2017) 22–30.
[8] T.H. Payne, L.E. Hines, R.C. Chan, S. Hartman, J. Kapusnik-Uner, A.L. Russ, B.W. Chaffee, C. Hartman, V. Tamis, B. Galbreth, P.A. Glassman, S. Phansalkar, H. van der Sijs, S.M. Gephart, G. Mann, H.R. Strasberg, A.J. Grizzle, M. Brown, G.J. Kuperman, C. Steiner, A. Sullins, H. Ryan, M.A. Wittie, D.C. Malone, Recommendations to improve the usability of drug-drug interaction clinical decision support alerts, J. Am. Med. Inform. Assoc. (2015).
[9] A.B. McCoy, L.R. Waitman, J.B. Lewis, J.A. Wright, D.P. Choma, R.A. Miller, J.F. Peterson, A framework for evaluating the appropriateness of clinical decision support alerts and responses, J. Am. Med. Inform. Assoc. 19 (2012) 346–352.
[10] A.L. Russ, A.J. Zillich, M.S. McManus, B.N. Doebbeling, J.J. Saleem, Prescribers' interactions with medication alerts at the point of prescribing: a multi-method, in situ investigation of the human-computer interaction, Int. J. Med. Inform. 81 (2012) 232–243.
[11] M.T. Baysari, J.I. Westbrook, K.L. Richardson, R.O. Day, The influence of computerized decision support on prescribing during ward-rounds: are the decision-makers targeted? JAMIA 18 (2011) 754–759.
[12] R.B. McDaniel, J.D. Burlison, D.K. Baker, M. Hasan, J. Robertson, C. Hartford, S.C. Howard, A. Sablauer, J.M. Hoffman, Alert dwell time: introduction of a measure to evaluate interruptive clinical decision support alerts, J. Am. Med. Inform. Assoc. (2015).
[13] A.L. Russ, S. Chen, B.L. Melton, E.G. Johnson, J.R. Spina, M. Weiner, A.J. Zillich, A novel design for drug-drug interaction alerts improves prescribing efficiency, Comm. J. Qual. Patient Saf. 41 (2015) 396–405.
[14] A.L. Russ, A.J. Zillich, B.L. Melton, S.A. Russell, S. Chen, J.R. Spina, M. Weiner, E.G. Johnson, J.K. Daggy, M.S. McManus, J.M. Hawsey, A.G. Puleo, B.N. Doebbeling, J.J. Saleem, Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation, J. Am. Med. Inform. Assoc. (2014).
[15] J.E. van Doormaal, P.G. Mol, R.J. Zaal, P.M. van den Bemt, J.G. Kosterink, K.M. Vermeulen, F.M. Haaijer-Ruskamp, Computerized physician order entry (CPOE) system: expectations and experiences of users, J. Eval. Clin. Pract. 16 (2010) 738–743.
[16] K. Cresswell, J. Coleman, A. Slee, R. Williams, A. Sheikh, Investigating and learning lessons from early experiences of implementing ePrescribing systems into NHS hospitals: a questionnaire study, PLoS One 8 (2013) e53369.
[17] R. Parasuraman, V. Riley, Humans and automation: use, misuse, disuse, abuse, Hum. Factors 39 (1997) 230–253.
[18] M.T. Baysari, J.I. Westbrook, R.O. Day, Understanding doctors' perceptions of their prescribing competency and the value they ascribe to an electronic prescribing system, Stud. Health Technol. Inform. 178 (2012) 1–6.
[19] J. Bertels, A.M. Almoudaris, P.J. Cortoos, A. Jacklin, B.D. Franklin, Feedback on prescribing errors to junior doctors: exploring views, problems and preferred methods, Int. J. Clin. Pharm. 35 (2013) 332–338.
[20] M. Reynolds, S. Jheeta, J. Benn, I. Sanghera, A. Jacklin, D. Ingle, B.D. Franklin, Improving feedback on junior doctors' prescribing errors: mixed-methods evaluation of a quality improvement project, BMJ Qual. Saf. 26 (2017) 240.
[21] M.T. Baysari, J. Del Gigante, M. Moran, I. Sandaradura, L. Li, K.L. Richardson, A. Sandhu, E.C. Lehnbom, J.I. Westbrook, R.O. Day, Redesign of computerized decision support to improve antimicrobial prescribing. A controlled before-and-after study, Appl. Clin. Inform. 8 (2017) 949–963.