

# A framework for evaluating the appropriateness of clinical decision support alerts and responses

Allison B McCoy,<sup>1</sup> Lemuel R Waitman,<sup>2</sup> Julia B Lewis,<sup>3</sup> Julie A Wright,<sup>3</sup>  
David P Choma,<sup>3</sup> Randolph A Miller,<sup>1</sup> Josh F Peterson<sup>1,4,5</sup>

<sup>1</sup>Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA

<sup>2</sup>Department of Biostatistics, University of Kansas Medical Center, Kansas City, Kansas, USA

<sup>3</sup>Division of Nephrology, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>4</sup>Center for Health Services Research, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>5</sup>Geriatric Research Education Clinical Center (GRECC), VA Tennessee Valley Healthcare System, Nashville, Tennessee, USA

## Correspondence to

Dr Allison B McCoy, School of Biomedical Informatics, The University of Texas Health Science Center at Houston (UTHealth), The University of Texas at Houston-Memorial Hermann Center for Healthcare Quality and Safety, 6410 Fannin St, UTPB 1100, Houston, TX 77030, USA; allison.b.mccoy@uth.tmc.edu

Received 14 February 2011

Accepted 6 July 2011

Published Online First

17 August 2011

## ABSTRACT

**Objective** Alerting systems, a type of clinical decision support, are increasingly prevalent in healthcare, yet few studies have concurrently measured the appropriateness of alerts with provider responses to alerts. Recent reports of suboptimal alert system design and implementation highlight the need for better evaluation to inform future designs. The authors present a comprehensive framework for evaluating the clinical appropriateness of synchronous, interruptive medication safety alerts.

**Methods** Through literature review and iterative testing, metrics were developed that describe successes, justifiable overrides, provider non-adherence, and unintended adverse consequences of clinical decision support alerts. The framework was validated by applying it to a medication alerting system for patients with acute kidney injury (AKI).

**Results** Through expert review, the framework assesses each alert episode for appropriateness of the alert display and the necessity and urgency of a clinical response. Primary outcomes of the framework include the false positive alert rate, alert override rate, provider non-adherence rate, and rate of provider response appropriateness. Application of the framework to evaluate an existing AKI medication alerting system provided a more complete understanding of the process outcomes measured in the AKI medication alerting system. The authors confirmed that previous alerts and provider responses were most often appropriate.

**Conclusion** The new evaluation model offers a potentially effective method for assessing the clinical appropriateness of synchronous interruptive medication alerts prior to evaluating patient outcomes in a comparative trial. More work can determine the generalizability of the framework for use in other settings and other alert types.

## INTRODUCTION

Many electronic medical record (EMR) systems and computerized provider order entry (CPOE) systems incorporate clinical decision support (CDS) features that attempt to improve patient safety. For example, to prevent adverse drug events, CDS alerts associated with electronic prescribing and CPOE address errors such as wrong patient, wrong drug, or wrong dose.<sup>1</sup> Previous studies have examined the usability, rates of responses and overrides, and patient outcomes associated with CDS alerts.<sup>2–19</sup> However, few studies have concurrently examined the appropriateness of CDS alert generation and the corresponding appropriateness of providers' clinical responses to each alert. Ensuring

long-term CDS sustainability and patient safety critically depends on one's ability to evaluate both alerts and subsequent clinical actions in the context of an actual patient care episode.<sup>20</sup>

We propose a conceptual framework for evaluating synchronous, interruptive computerized medication safety alerts. Our framework utilizes expert reviewers to judge both the appropriateness of an alert (ie, whether it should trigger in a given specific clinical context and whether its advice is clinically sound) and the appropriateness of the provider's response (ie, whether the resulting actions of the provider would potentially have clinical benefit). The framework defines alert system success or failure using metrics from the entire alert episode, from display to both immediate and delayed user response. Importantly, the framework complements comparative evaluations of CDS that might determine whether the rate of process and patient outcomes are attributable to the alert.

To demonstrate its utility, we applied the new CDS evaluation framework to a previously described system that provided medication safety alerts for patients with potential acute kidney injury (AKI) within a tertiary care academic medical center.<sup>21</sup> We demonstrated that the earlier system improved the rate and timeliness of providers' responses to clinical events; however, high rates of alert overrides (78%) persisted. As a result, uncertainty remained about whether the alerts were excessive and whether observed providers' responses were appropriate. We conceived the new evaluation framework to better analyze the successes and limitations of the AKI alerts.

## BACKGROUND

Early CDS research generated substantial optimism that medication-related CDS, including alerting systems, could prevent adverse drug events and improve patient safety.<sup>1 22 23</sup> Indeed, alerts are now important components of most point-of-care clinical information systems, including those for CPOE, bar-coded medication administration, and pharmacy-based medication management. However, demonstrating the clinical benefits of alerting systems often requires extensive evaluations. Evaluations of CDS system implementations have not always demonstrated improved patient outcomes as expected.<sup>24–27</sup> User non-adherence, through overriding of alerts (ie, failure to perform the recommended action), has often been cited as the basis for suboptimal CDS efficacy. A systematic review of medication-related CDS reported that

overrides occurred after 49–96% of displayed CDS alerts.<sup>2</sup> While several projects have attempted to reduce alert override rates, their methods lacked a generalizable approach to evaluating key trade-offs between avoiding false positive alerts, where clinician time and attention is diverted, and avoiding false negative alerts, which silently leave patients at risk.<sup>28–31</sup> Studies examining alert override rates have used chart review or user feedback to conclude that many alert overrides are clinically justifiable. Such justifications include the clinical insignificance of an overridden alert, known patient tolerance for a drug or drug dose, and documented clinician intention to monitor the patient.<sup>2–12</sup> However, studies have often not examined the extent to which override rates reflect justifiable non-adherence, such as non-response to clinically irrelevant alerts. Also, evaluations often lacked third-party clinical expert adjudications that could determine whether provider actions subsequent to an alert had potentially beneficial or potentially harmful consequences for the patient.

**MODEL DEVELOPMENT PROCESS**

To develop the conceptual model and corresponding CDS alert evaluation framework, we first reviewed the CDS-related health information technology literature. We sought out key unifying concepts and studies that attempted to measure the accuracy (ie, binary classification measures) and clinical appropriateness of medication alerting systems. While developing the framework, we asked expert clinicians to test iteratively our proposed alert metrics. During each iteration, using EMR log files, reviewers independently evaluated 10–15 patients for which CDS medication alerts had occurred. Experts later discussed all cases as a group to achieve consensus about the utility of the evaluation metrics. Each iterative evaluation updated the framework so that it clarified data element definitions and better ensured that framework application procedures were effective.

Our review identified two sets of alert-related measurements: ‘alert appropriateness’ and ‘provider response appropriateness’. Alert appropriateness evaluations identify false positive alerts, which result from errant or absent underlying knowledge or patient data, contain clinically irrelevant advice, or display to the wrong person. Evaluations of provider response appropriateness determine whether a provider acted in a potentially clinically beneficial manner after viewing an alert. A growing literature suggests that provider response appropriateness depends on the appropriateness of the original alert. Accurate judgments regarding appropriateness of providers’ responses requires detailed reviews of the scenarios in which inappropriate responses occurred, including ability to detect non-justifiable overrides and incorrect responses to false positive alerts.

**MODEL DESCRIPTION**

Table 1 depicts the general framework model for evaluating CDS alerts and responses. We describe the framework below in greater detail, giving examples applied to different CDS alert implementations.

**Guidelines for alert review**

The framework can apply to various alert scenarios and analyses. Accordingly, review criteria will depend on individual investigator decisions, although some general principles apply to most alert reviews. Whenever possible, evaluators examining alert appropriateness should blind the expert clinician reviewers to subsequent clinician actions and to patients’ clinical sequelae. Such reviewer blinding allows determinations of alert and response appropriateness to be made as if the review had occurred at the time that the alert initially displayed to the provider and ensures that knowledge of subsequent responses to the alert does not influence the reviewer’s characterizations.

Factors for review may be ‘explicit’, involving well-defined, rule-based criteria, or ‘implicit’, depending on expert clinical judgment. When using explicit criteria, investigators conduct pilot reviews with non-study datasets to develop the evaluation criteria before beginning the review process. Explicit review criteria best fit clinical situations where extensive evidence-based guidelines exist, such as determining if the medications recommended at the time of discharge after myocardial infarction were prescribed. By contrast, for clinical alert scenarios lacking standardized correct responses, expert-based reviews using implicit criteria provide a flexible means of judging the appropriateness of actual provider responses. Previous adverse drug event studies have often described and applied methods for implicit review.<sup>32 33</sup>

**Alert appropriateness**

The appropriateness of alerts measure depends on expert review of alerts within CDS and EMR system log files. Adjudication of clinical relevance for the alert requires an understanding, gained through EMR review, of the patient conditions at the time the alert first displayed.

**Alert display appropriateness**

The first component of alert appropriateness determination involves experts deciding whether the triggering circumstances and the alert recommendations are clinically relevant and indicated. Each clinical subspecialty domain or topic may have different technical factors that contribute to and determine alert appropriateness, and these evolve over time. Causes of inappropriate alerts may include lack of granularity in CDS

**Table 1** Medication alert assessment and measures framework

	Alert evaluation measure	Review assessment	Calculation of measure
Alert appropriateness measures	False positive alert rate	Was alert display appropriate?	$\frac{\text{Inappropriate alerts}}{\text{Total alerts}}$
	Rate of alert urgency	If alert is appropriate, is urgent response to alert expected?	$\frac{\text{Urgent alerts}}{\text{Appropriate alerts}}$
Provider response appropriateness measures	Alert override rate	Did provider override alert?	$\frac{\text{Overridden alerts}}{\text{Total alerts}}$
	Alert adherence rate	Did provider make alert-indicated change?	$\frac{\text{Alert-indicated responses}}{\text{Total alerts}}$
	Provider response time	If provider made alert-indicated change, when did provider respond to alert?	Summary measure with variance
	Rate of provider response inappropriateness	Was provider response to alert appropriate?	$\frac{\text{Inappropriate responses}}{\text{Total alerts}}$

knowledge or failing to account for demographic, medication, laboratory, or other EMR data elements that may countermand otherwise appropriate reasons for an alert. For example, a medication alert that indicates recommended safe doses for pediatric patients should not display for adult patients. Data elements recorded in unstructured text notes typically cannot influence CDS system recommendations, even when such information is critical. For example, an inaccessible text comment in an infectious disease consultant's note indicating that a patient's infection is life-threatening should warrant an alert about monitoring renal function carefully rather than an alert to discontinue a nephrotoxic antibiotic that is the best therapy for the infection. Experts may also require access to structured, explicit knowledge in EMRs. For example, on the basis of a history of previous anaphylaxis, an alert may suggest that clinicians either undertake a desensitization procedure or prescribe a different medication. While such an alert is almost always relevant, an expert reviewer should determine the authenticity of the historical basis for such an alert from the EMR. From the experts' ratings, one can calculate the 'false positive alert rate' by dividing the number of 'inappropriate alerts' by 'total alerts'. For an individual patient, multiple instances of the same alert may occur that are independent events. Evaluators must adjust alert-related analyses accordingly by sampling only one alert per unique patient—medication pair or by applying statistical methods to control for non-independence of events.

#### Alert response urgency

For those alerts determined to be appropriate, expert evaluators should also determine the urgency of provider response to the alert—that is, how quickly a change in therapy should occur to avert potential patient harm. Judgments regarding alert response urgency depend on each patient's clinical situation and vary significantly among patients and settings. A given evaluation may code urgency determinations in a binary manner (eg, an urgent alert requires a response within 48 h) or in a categorical manner (eg, emergent—within 2 h; urgent—within hours; or routine—within the next few days). One can calculate the 'rate of alert response urgency' from experts' ratings of 'urgent alerts' divided by 'appropriate alerts'.

Expert reviewers with access to patient records can often identify the causative factors that determine alert display appropriateness and alert response urgency. These comprise a secondary evaluation measure, 'contributing factors to false positive and non-urgent alerts'. This measure allows system developers to carefully determine how best to present a given alert in terms of its disruptiveness and its method of display.<sup>1</sup> Improving alerting algorithms to increase specificity can potentially enhance alert-related end-user acceptance and corresponding clinical outcomes.

#### Provider response appropriateness

Evaluation of provider responses should involve setting an appropriate time window for each type of alert before the onset of any formal evaluation. The time window duration derives in part from alert response urgency. Responses can occur within the context of the user session that generated the alert or at a later time. In teaching hospitals, junior clinicians such as medical students and interns may consult more senior clinicians, such as residents and attending physicians, before taking actions. Similarly, in private non-teaching hospitals, a nurse may choose to consult the attending physician before responding to an alert. Evaluation of user response appropriateness can utilize

electronic methods, such as a computerized script to extract information from CPOE logs, and manual expert review of printed or electronic alert logs and patient charts.

#### Alert overrides

Through CPOE transaction database queries or via parsing of usage log files, evaluators should first identify each individual alert and determine how providers interacted with the actual alert displayed. Overrides occur when providers do not carry out the change in care recommended by the alert. We define an alert override as an alert followed by either no change in care when one was recommended, or a change that occurs that differs substantially from what the alert recommended. This definition differs from many previous definitions of alert overrides, which often include only initial non-responsiveness to an alert. Calculation of the override rate involves dividing 'overridden alerts' by 'total alerts'.

#### Alert adherence

Independent of the provider's interaction with the alerts, evaluators must determine whether providers performed an alert-related action at any point during the alert interaction or within the designated response time window following the alert display. Providers may have initially overridden an alert but later followed the alert recommendation. Multiple responses may occur within the time window. For example, following an alert warning of increasing serum creatinine levels, clinicians might modify an antibiotic dose as indicated, and also order therapeutic drug monitoring. Likewise, although a provider may have indicated intention to comply with alert advice within the context of the alert display, the provider may not always follow through with the response suggested, and no actual response occurs. Calculating 'alert-indicated responses' divided by 'total alerts' yields the 'provider alert-adherence rate', which is a primary evaluation measure that complements the 'alert override rate'.

#### Provider response time

Evaluators should record the relative time, after alert display, at which providers responded to the alert. If available, they can determine the exact response time electronically from usage log files. Responses may be immediate (eg, discontinuation of a medication as soon as an alert is displayed) or delayed, possibly occurring in a separate, later order entry session involving a different clinician on the patient's care team. The delay between an alert-triggering event and the time at which a provider uses the clinical information system also affects the response time. For example, a patient may have an elevated serum creatinine level reported shortly after an afternoon blood draw, but no provider accesses the CPOE system to receive an alert until the following day. 'Provider response time' is an evaluation measure that should be measured and analyzed in continuous form. Monitoring response time also informs system developers and clinical leaders by indicating whether alert-related patient care decisions occur within an appropriate time frame, which is especially important for 'urgent alerts' or delays resulting from latent system utilization. Knowledge gained from this measure can motivate implementation of different escalation strategies or alternative alert deployment mechanisms.

#### Expected alert responses

To determine whether the provider response to an alert was appropriate, experts must first determine the expected responses

to an alert. In many alerting scenarios, prior evidence-based guidelines or widely accepted practices dictate explicit, specific criteria for appropriate alert responses. For example, an alert to discontinue ongoing oral potassium supplements and to monitor serum potassium levels for a patient whose potassium level had gradually increased to 6.0 mEq/l should rarely be overridden. However, in some clinical scenarios, multiple clinically beneficial responses may qualify as appropriate. For example, in the setting of progressive renal dysfunction, a provider might decrease the dose of a renally excreted drug, discontinue the drug outright, or order serial drug levels to monitor whether the dose should change. In the case of inappropriate alerts, one appropriate response might involve continuing therapy unchanged, while another would involve adjusting medication-related variables in a clinically beneficial manner contrary to what the alerts suggested. Thus, reviewers could identify multiple actions or taking no action as appropriate provider responses to alerts. The secondary evaluation measure ‘expected responses to alerts’ provides evaluators with a gold standard against which they can measure observed responses following alerts to determine alert response appropriateness. The heterogeneity of clinical scenarios makes gold standard references different for many alerting systems, and prior work has highlighted the need for a reference standard for determining appropriate responses in evaluation frameworks.<sup>34</sup>

**Alert response appropriateness**

As a final determination, expert reviewers should categorize the appropriateness of the provider’s response to each alert. Experts must consider the individual patient’s conditions at the time the alert was displayed in addition to the expected responses to alerts discussed above. Appropriate responses occur when an expert reviewer judges that one or more actions taken by the alert recipient were potentially clinically beneficial for the patient. Experts categorize an inappropriate alert response as actions that pose potential harm to the patient. Inappropriate responses may take several forms, such as failure to implement an important, clinically beneficial action following an appropriate alert that recommended it, following the incorrect advice of a clinically inappropriate alert, or undertaking a potentially harmful action triggered by the alert but not related to the alert

content. Evaluators can use metrics for response appropriateness that are binary (eg, appropriate or inappropriate) or scaled (eg, appropriate, acceptable, or inappropriate). Responses that differ from the expected, appropriate response that are judged unlikely to cause patient harm would be deemed acceptable rather than inappropriate. One calculates the ‘rate of provider response appropriateness’ primary evaluation measure by dividing ‘appropriate responses’ by ‘total alerts’. This measure helps evaluators and institutional staff to understand the true effect of an alert intervention. Further expert analyses can help to ascertain ‘contributing factors to provider response appropriateness’.

**Classification of alert episodes**

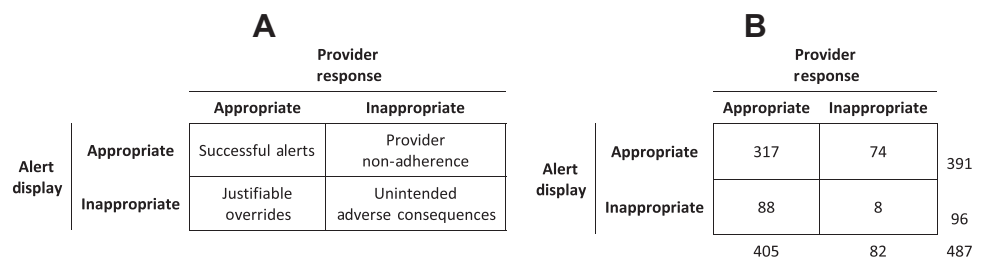
Through experts determining whether each displayed alert and provider response was appropriate, one can classify each alert episode into one of four categories: ‘successful alerts’, ‘provider non-adherence’, ‘justifiable overrides’, and ‘unintended adverse consequences’ (figure 1A). Appropriate provider responses to appropriate alerts make up successful alerts. Inappropriate provider responses to appropriate alerts, previously labeled active failures,<sup>35</sup> comprise provider non-adherence. This classification of alert episodes is the most common category of non-compliance described in prior literature and includes lack of response or an incorrect response to a displayed appropriate alert. Justifiable overrides occur when providers respond in a clinically beneficial manner to inappropriate alerts. Inappropriate, potentially harmful provider responses to inappropriate alerts are classified as unintended adverse consequences.<sup>35</sup> The relationships among these classifications and the primary outcomes in table 1 are depicted in figure 1C. We note that a more comprehensive review of all patient records could identify those patients whose clinical situations should have triggered an alert, yet no alert occurred. This comprises an additional ‘failure to alert’ category. However, this metric is beyond the scope of our current model, since practical limitations constrain reviews to instances where alerts did occur.

**VALIDATION THROUGH EXAMPLE**

**Study setting and methods**

Vanderbilt University Hospital (VUH) is an academic, tertiary care facility with over 500 adult beds and 50 000 admissions

**Figure 1** Evaluation framework applied to the retrospective evaluation of acute kidney injury medication alerts. (A) Classification of alert episodes through the evaluation framework. (B) Results of the retrospective evaluation of acute kidney injury medication alerts. (C) Calculation of primary outcomes in the alert evaluation framework related to the alert episode classifications.



**C**

$$\text{False positive alert rate} = \frac{\text{Inappropriate alerts}}{\text{Total alerts}} = \frac{96}{487} = 20\%$$

$$\text{Alert override rate} = \frac{\text{Provider non-adherence} + \text{Justifiable overrides}}{\text{Total alerts}} = \frac{74 + 88}{487} = 33\%$$

$$\text{Rate of provider response inappropriateness} = \frac{\text{Inappropriate provider responses}}{\text{Total alerts}} = \frac{82}{487} = 17\%$$

annually. At VUH, care providers have for more than 15 years used locally developed and maintained inpatient CPOE and inpatient/outpatient EMR systems. Those systems integrate extensive CDS capabilities, including: advisors to select recommended medications; medication dosing advice; and alerts about drug–allergy, drug–laboratory, and drug–drug interactions.<sup>1 36 37</sup> At VUH, a previous intervention utilized interruptive, CPOE-based alerts to warn providers about potential AKI. Evaluation of that intervention indicated a significant increase in the rate and timeliness of provider modification or discontinuation of targeted nephrotoxic or renally cleared medications.<sup>21</sup> However, the evaluation noted that initial override rates were high.

We retrospectively evaluated the AKI alerts using our new framework. We reviewed 300 randomly selected inpatient admissions, or cases, from 697 who were admitted as inpatients to VUH between November 2007 and October 2008 and received at least one AKI medication alert. For patients receiving multiple alerts, experts only performed reviews on the first distinct patient–medication pairs encountered; subsequent alerts involving the same patient–medication pair were ignored. Two study nephrologists (DC, JW) used the new alert evaluation framework to perform a retrospective electronic chart review of the selected patients. A web-based collection tool facilitated reviews by displaying patient orders, laboratory results, and alerts, and to record experts' alert and provider response appropriateness categorizations according to the new framework. The tool also facilitated reviewer blinding to subsequent clinical events through selective data display. Each nephrologist reviewed 200 cases, 100 of which overlapped with the other reviewer (300 cases total). Whenever the reviewers disagreed, a third nephrologist (JL) reviewed the case for consensus. The Vanderbilt Institutional Review Board approved this study.

## RESULTS

Figure 1 depicts the evaluation results in the context of the new framework. Reviewers evaluated 487 total alerts. Inter-rater reliability for alert appropriateness was moderate; reviewers agreed for 85% of alerts ( $\kappa=0.46$ ). After resolving disagreements and reaching consensus, the reviewers selected 391 alerts as appropriate to display; the false positive alert rate was 20%. All appropriate alerts were determined to be urgent. Contributing factors to alert display inappropriateness included no AKI actually present because of laboratory error, medications or conditions interfering with creatinine assays, and insufficient change in glomerular filtration rate for 54% of inappropriate alerts. Additional contributing factors to alert display inappropriateness included: the pre-alert medication dose was acceptable because of a previous adjustment; drug doses not subject to adjustment because of short-duration prophylaxis; presence of clinician monitoring of therapeutic drug levels for 33% of inappropriate alerts.

Providers initially (ie, within the alert-containing order entry session) overrode 400 (82%) displayed alerts, although 228 (47%) of the alerted medications were modified or discontinued within 24 h. For medications modified or discontinued before patient discharge, the median time to response was 13 h. Inter-rater reliability for response appropriateness was fair; reviewers agreed for 78% of alerts ( $\kappa=0.37$ ). After resolving disagreements and reaching consensus, reviewers adjudicated provider responses to alerts as inappropriate for 82 (17%) of 487 alerts. Expected responses to alerts included modification of the order, discontinuation of the medication, monitoring of drug levels, documentation of the indication, monitoring of creatinine, and no

change in therapy. Of the alert responses determined to be inappropriate, only 8 (10%) resulted from an alert adjudicated as inappropriate. Thus unintended adverse consequences were rare compared with provider non-adherence (figure 1B).

## DISCUSSION

Medication-related CDS is widely integrated into current clinical information systems. Yet, the appropriateness of displayed alerts or subsequent end-user decision making is not well understood. Consequently, when a new alerting system fails to improve a targeted process or patient outcome, the causative factors may not be apparent. Speculation then often focuses on alert fatigue or clinically irrelevant alert content in the absence of a deeper analysis of objective data. Application of our proposed framework can produce a more complete picture of the effectiveness of implementing clinical alerts. A key attribute of the framework is that it determines appropriateness by capturing relevant clinical context information at the time of a triggered alert and by applying expert knowledge. Our framework approach determines appropriateness of alerts and responses independently. We anticipate that the framework, once verified and evolved for other settings, can serve as a means to evaluate, post facto, underperforming existing systems. It can also serve to evaluate a CDS system in preparation for a comparative trial that measures its patient-level clinical impact.

Prior research in this area has typically described the prevalence of alerts and reported overall override rates, usually limited to the context of the alert-related ordering session. Few previous studies have determined independently whether clinicians should have ignored clinically irrelevant or improper alerts, and how and when users' non-adherence to alerts might be justifiable. Some studies have evaluated the clinical context of alerts when overrides occurred, but did not report alert and response appropriateness in settings where providers accepted the alert advice.<sup>4 5</sup> Weingart *et al*<sup>5</sup> examined a subset of all displayed alerts to determine alert validity and expert agreement with overrides, although no measures of unintended adverse consequences were reported. To prevent alert fatigue, CDS implementers must monitor and identify situations that frequently trigger inappropriate alerts and take well-informed steps to improve alert specificity. Our validation study suggests that formal, model-based evaluation of implemented CDS systems can describe and quantify categories of alert appropriateness and corresponding categories of user responses.

Our framework includes in its analyses downstream responses to alerts that occurred outside of the initial alert context. Previous CDS evaluations have often classified the provider response as an override of a displayed alert when the provider takes no action within a short time (eg, by the end of the ordering session, or within a few hours). Such studies do not account for alert-related changes to care made at a later time. This is important, as we previously found that providers overrode 67% of alerts when initially displayed but still later changed the order for 80% of alerts.<sup>21</sup> We hypothesize that providers in our study delayed their responses to alerts in order to discuss therapy with other members of the patient's care-providing team, to look up reference material, or to consult a specialist (eg, nephrologist or clinical pharmacists). Another scenario important to CDS evaluation occurs when providers respond to an alert and indicate that they plan to comply by modifying or discontinuing an existing order, but fail to complete the action after leaving the alerting context. Previous studies may have classified such responses as appropriate for the alert if they did

not examine post-alert behaviors. Correct classification of provider responses to alerts, whether by sophisticated automated log file analyses or by expert review, must include appropriate-length follow-ups of clinicians' post-alert actions for accurate reporting and evaluation of CDS alert effects.

While qualitative evaluations are useful in describing the types of errors that can occur from use of CDS, gaps exist in the literature about best approaches to formal quantitative evaluation of CDS effectiveness and potential associated CDS unintended adverse consequences. A recent study examined nurses' responses to CDS alerts about documentation of patient weights.<sup>38</sup> The study determined that overrides occurred more often for false positive alerts, and that false positive alerts rarely (2.7%) resulted in inappropriate entries. A separate study of provider responses to simulated drug safety alerts found that incorrect alerts more often resulted in errors than correct alerts.<sup>13</sup> Such analyses help investigators to understand the beneficial and detrimental roles of CDS in helping to improve outcomes as well as facilitating errors. We believe that such analyses should occur for all types of CDS.

The proposed framework is currently limited by the validation, which occurred in an interruptive, synchronous alert system. Evaluation of other alert types, such as passive, informational alerts, will likely require modification or supplementation of the existing metrics. The framework is also not designed to demonstrate that process or patient outcomes are attributable to the alert; such evaluation requires a comparative or controlled trial. However, we anticipate that the framework will aid preparation for comparative trials by providing a preliminary measure of alert and provider response appropriateness that determines the true anticipated effectiveness of an alert system.

## CONCLUSION

High rates of alert overrides and low rates of alert adherence can hinder the success of otherwise well-designed CDS alerting systems. We developed and tested a framework that utilizes expert review for evaluating the clinical appropriateness of CDS alerts and providers' responses to the alerts that extends previous work by others. Further utilization of such CDS evaluation frameworks can lead to more standardized, comparable evaluations on a more widespread basis. Such frameworks could potentially enhance the field's ability to evaluate CDS performance and impact on clinical care delivery, and provide feedback to developers about optimizing CDS implementations that improve patient safety.

**Acknowledgments** We thank Ioana Danciu for her technical assistance and Dean Sittig for his critique of the manuscript. We were funded in part by National Library of Medicine grants T15 LM007450 (ABM) and R01 LM009965 (to ABM, LRW, JBL, JAW, DPC, RAM, JFP). Some data collection was supported by NCR/NIH grant UL1 RR024975.

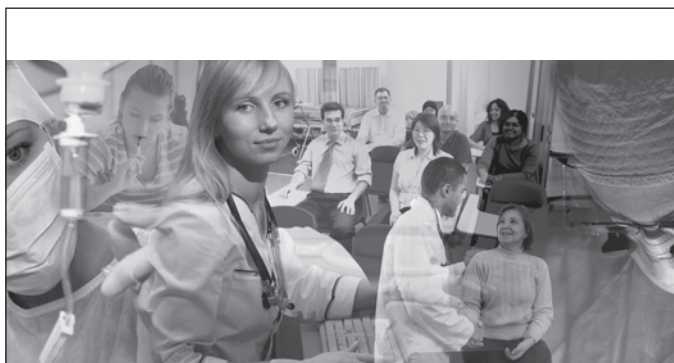
**Ethics approval** Vanderbilt University Institutional Review Board.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

1. Miller RA, Waitman LR, Chen S, et al. The anatomy of decision support during inpatient care provider order entry (CPOE): empirical observations from a decade of CPOE experience at Vanderbilt. *J Biomed Inform* 2005;**38**:469–85.
2. van der Sijs H, Aarts J, Vulto A, et al. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc* 2006;**13**:138–47.
3. Weingart SN, Toth M, Sands DZ, et al. Physicians' decisions to override computerized drug alerts in primary care. *Arch Intern Med* 2003;**163**:2625–31.
4. Hsieh TC, Kuperman GJ, Jaggi T, et al. Characteristics and consequences of drug allergy alert overrides in a computerized physician order entry system. *J Am Med Inform Assoc* 2004;**11**:482–91.
5. Persell SD, Dolan NC, Friesema EM, et al. Frequency of inappropriate medical exceptions to quality measures. *Ann Intern Med* 2010;**152**:225–31.
6. Grizzle AJ, Mahmood MH, Ko Y, et al. Reasons provided by prescribers when overriding drug-drug interaction alerts. *Am J Manag Care* 2007;**13**:573–8.
7. van der Sijs H, Mulder A, van Gelder T, et al. Drug safety alert generation and overriding in a large Dutch university medical centre. *Pharmacoepidemiol Drug Saf* 2009;**18**:941–7.
8. Ko Y, Abarca J, Malone DC, et al. Practitioners' views on computerized drug-drug interaction alerts in the VA system. *J Am Med Inform Assoc* 2007;**14**:56–64.
9. Isaac T, Weissman JS, Davis RB, et al. Overrides of medication alerts in ambulatory care. *Arch Intern Med* 2009;**169**:305–11.
10. Shah NR, Seger AC, Seger DL, et al. Improving acceptance of computerized prescribing alerts in ambulatory care. *J Am Med Inform Assoc* 2006;**13**:5–11.
11. Swiderski SM, Pedersen CA, Schneider PJ, et al. A study of the frequency and rationale for overriding allergy warnings in a computerized prescriber order entry system. *J Patient Saf* 2007;**3**:91–6.
12. Litzelman DK, Tierney WM. Physicians' reasons for failing to comply with computerized preventive care guidelines. *J Gen Intern Med* 1996;**11**:497–9.
13. van der Sijs H, van Gelder T, Vulto A, et al. Understanding handling of drug safety alerts: a simulation study. *Int J Med Inform* 2010;**79**:361–9.
14. Tamblyn R, Huang A, Taylor L, et al. A randomized trial of the effectiveness of on-demand versus computer-triggered drug decision support in primary care. *J Am Med Inform Assoc* 2008;**15**:430–8.
15. 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.
16. 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.
17. van der Sijs H, Bouamar R, van Gelder T, et al. Functionality test for drug safety alerting in computerized physician order entry systems. *Int J Med Inform* 2010;**79**:243–51.
18. Koppel R, Metlay JP, Cohen A, et al. Role of computerized physician order entry systems in facilitating medication errors. *JAMA* 2005;**293**:1197–203.
19. Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc* 2004;**11**:104–12.
20. Bloomrosen M, Starren J, Lorenzi NM, et al. Anticipating and addressing the unintended consequences of health IT and policy: a report from the AMIA 2009 Health Policy Meeting. *J Am Med Inform Assoc* 2011;**18**:82–90.
21. McCoy AB, Waitman LR, Gadd CS, et al. A computerized provider order entry intervention for medication safety during acute kidney injury: a quality improvement report. *Am J Kidney Dis* 2010;**56**:832–41.
22. Kuperman GJ, Bobb A, Payne TH, et al. Medication-related clinical decision support in computerized provider order entry systems: a review. *J Am Med Inform Assoc* 2007;**14**:29–40.
23. 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.
24. Gurwitz JH, Field TS, Rochon P, et al. Effect of computerized provider order entry with clinical decision support on adverse drug events in the long-term care setting. *J Am Geriatr Soc* 2008;**56**:2225–33.
25. Strom BL, Schinnar R, Bilker W, et al. Randomized clinical trial of a customized electronic alert requiring an affirmative response compared to a control group receiving a commercial passive CPOE alert: NSAID-warfarin co-prescribing as a test case. *J Am Med Inform Assoc* 2010;**17**:411–15.
26. Tamblyn R, Reidel K, Huang A, et al. Increasing the detection and response to adherence problems with cardiovascular medication in primary care through computerized drug management systems: a randomized controlled trial. *Med Decis Making* 2010;**30**:176–88.
27. Strom BL, Schinnar R, Abera 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.
28. Lee EK, Mejia AF, Senior T, et al. Improving patient safety through medical alert management: an automated decision tool to reduce alert fatigue. *AMIA Annu Symp Proc* 2010;**2010**:417–21.
29. van der Sijs H, Aarts J, van Gelder T, et al. Turning off frequently overridden drug alerts: limited opportunities for doing it safely. *J Am Med Inform Assoc* 2008;**15**:439–48.
30. van der Sijs H, Kowlesar R, Aarts J, et al. Unintended consequences of reducing QT-alert overload in a computerized physician order entry system. *Eur J Clin Pharmacol* 2009;**65**:919–25.
31. Del Beccaro MA, Villanueva R, Knudson KM, et al. Decision support alerts for medication ordering in a computerized provider order entry (CPOE) system. *Appl Clin Inf* 2010;**1**:346–62.
32. Bates D, Leape L, Petrycki S. Incidence and preventability of adverse drug events in hospitalized adults. *J Gen Intern Med* 1993;**8**:289–94.

33. **Leape LL**, Bates DW, Cullen DJ, *et al.* Systems analysis of adverse drug events. *JAMA* 1995;**274**:35–43.
34. **Classen DC**, Avery AJ, Bates DW. Evaluation and certification of computerized provider order entry systems. *J Am Med Inform Assoc* 2007;**14**:48–55.
35. **Reason J**. Human error: models and management. *BMJ* 2000;**320**:768–70.
36. **Giuse DA**. Supporting communication in an integrated patient record system. *AMIA Annu Symp Proc* 2003:1065.
37. **Geissbühler A**, Miller RA. A new approach to the implementation of direct care-provider order entry. *Proc AMIA Annu Fall Symp* 1996:689–93.
38. **Cimino JJ**, Farnum L, Cochran K, *et al.* Interpreting nurses' responses to clinical documentation alerts. *AMIA Annu Symp Proc* 2010;**2010**:116–20.



## BMJ Group, supporting you throughout your career...

**At BMJ Group we have resources available to you at every stage of your career.**

Whether you are a medical student or doctor in training looking to keep up with the latest news and prepare for exams, or a qualified doctor who wants the latest medical information, to attend conferences, or looking for your next job, BMJ Group has something to offer. For the latest information on all of our products and services register to receive email updates at

**[group.bmj.com/registration](http://group.bmj.com/registration)**

---

BMJ  BMJ Journals BMJ Careers BMJ Evidence Centre BMJ Learning 