

Understanding Advice Sharing among Physicians: Towards Trust-Based Clinical Alerts

DEBALEENA CHATTOPADHYAY¹, ROMISA ROHANI GHAHARI¹, JON DUKE² AND
DAVIDE BOLCHINI^{1*}

¹*School of Informatics and Computing, Indiana University, Indianapolis, IN, USA*

²*Regenstrief Institute, Indianapolis, IN, USA*

*Corresponding author: dbolchin@iupui.edu

Safe prescribing of medications relies on drug safety alerts, but up to 96% of such warnings are ignored by physicians. Prior research has proposed improvements to the design of alerts, but with limited increase in adherence. We propose a different perspective: before re-designing alerts, we focus on improving the trust between physicians and computerized advice by examining why physicians trust their medical colleagues. To understand trusted advice among physicians, we conducted three contextual inquiries in a hospital setting (22 participants), and corroborated our findings with a survey (37 participants). Drivers that guide physicians in trusting peer advice include: timeliness of the advice, collaborative language, empathy, level of specialization and medical hierarchy. Based on these findings, we introduce seven design directions for trust-based alerts: endorsement, transparency, team sensing, collaborative, empathic, conflict mitigating and agency laden. Our work contributes to novel alert design strategies to improve the effectiveness of drug safety advice.

RESEARCH HIGHLIGHTS

- In three contextual inquiries, we examined why clinical advice is trusted among physicians.
- We uncovered eight key drivers that are at play when physicians trust peer advice.
- These include medical specialization, hierarchy, empathy and inclusive language.
- A survey with 37 participants validated the qualitative findings.
- We introduce seven design directions for trust-based alerts that improve physicians' compliance.
- These include endorsed, team-sensing, agency-laden and empathic alerts.

Keywords: health care information systems; health informatics; contextual design; clinical alerts; trust in socio-technical systems; contextual inquiry

Editorial Board Member: Timothy Bickmore

Received 27 February 2015; revised 8 June 2015; accepted 24 July 2015

1. INTRODUCTION

The safe prescribing of patient medications via Computerized Physician Order Entry (CPOE) routinely relies on drug safety alerts. The most common type of such alerts, drug–drug interaction (DDI) alerts, are a primary form of clinical decision support (CDS), but their effectiveness remains surprisingly low: up to 96% of such alerts are ignored by physicians on a daily basis (Payne *et al.*, 2002; Schedlbauer *et al.*,

2009). Noncompliance to DDI alerts leads to increased risk of prescribing unsafe medications, which may cause severe health complications and even death. The primary reason for this poor adherence is ‘alert fatigue’; a state in which physicians, feeling bombarded by numerous alerts of questionable clinical importance, become desensitized. Physicians report that in order to ‘get through their daily work’ they are forced to override several of the DDI alerts, including critical ones

(Ash *et al.*, 2007; Glassman *et al.*, 2002; Sittig *et al.*, 2006). Numerous strategies have attempted to reduce alert fatigue—from decreasing the intrusiveness of alerts to simplifying their visual design to improving the specificity of underlying data (Duke *et al.*, 2013; Luna *et al.*, 2007; Paterno *et al.*, 2009; Shah *et al.*, 2006). Still, physician adherence remains low, as does satisfaction and trust in clinical alert systems (Hayward *et al.*, 2013; Jung *et al.*, 2013).

To address this long-standing problem, this paper proposes a different perspective: before looking at how to improve alerts, we look at how to improve trust between physicians and computerized advice. A starting point to address this issue is looking at whom physicians do trust—their medical colleagues and mentors. While the medical field has sought to move away from anecdotal decision-making to evidence-based medicine (Guyatt *et al.*, 1992; Sackett *et al.*, 1996), the narratives and discourse that physicians share continue to influence the medical practice strongly (Enkin and Jadad, 1998; Greenhalgh, 1999).

This influence of anecdotal evidence on medical practice continues because medicine lacks rules that can be generally and unambiguously applied to every case at hand (Hunter, 1996). This characteristic of the medical practice tends to increase physicians' belief in personal or trusted experiences above impersonal data—even when those data are scientifically rigorous. In such instances of uncertainty, attachment to anecdotal evidence reflects the well-studied cognitive phenomenon of heuristic processing (Chaiken *et al.*, 1989; Newell and Simon, 1972). During heuristic processing, individuals rely on rules of thumb for judgement, such as outcomes of prior similar events (representative heuristic) or outcomes of events that are easy to imagine (availability heuristic, Kahneman and Tversky, 1972). Failure to apply relevant heuristics may lead to erroneous judgements (e.g. mistrust an alert) due to cognitive biases, such as biasing systemic processing of source credibility when alert content is ambiguous (Chaiken and Maheswaran, 1994) or ignoring prior probabilities of adverse drug-related events during inductive reasoning (nonstatistical reasoning, Kunda, 1999; Nisbett *et al.*, 1983).

Thus, to improve physician responsiveness to drug safety warnings, we must recognize that providing accurate, evidence-based data is not enough. We must delve into how and why physicians come to trust the information that they are provided, which in turn actively influences their heuristic processing. What are the characteristics of the information source (whether a mentor, a colleague, or the medical literature) that physicians trust in their clinical work? How and when is this information delivered? How do physicians make the determination of whether this information is reliable? Under which conditions physicians trust and/or follow the advice of peers? Exploring these questions will provide the necessary foundational knowledge to ideate alert designs, which we believe will have the potential to increase physician's adherence.

To that aim, our contribution is 2-fold:

- We report an exploratory sequential mixed methods study (a field observation followed by a survey) aimed at capturing the characteristics of collegial discussions around prescribing medications, looking at how information is communicated and whether it is accepted or not.
- Informed by the results of our empirical observations and survey, we introduce and elaborate an initial set of foundational principles regarding adherence to clinical advice. These principles, aiming to garner physicians' trust, can drive novel design strategies for computerized CDS.

The results of our work uncovered novel principles for an effective alert design that are based on what physicians consider important when taking advice from peers in the context of their clinical activities. Thus, by understanding what makes physicians' advice trustworthy in the ecosystem of clinical activities, we contribute to the intellectual basis for designing substantially better DDI alerts for a broad variety of CPOE systems.

2. RELATED WORK

A substantial body of research studied the nature of DDI alerts and their effectiveness (for an example workflow, see Figure 1). DDI alerts are considered a kind of CDS because they guide physicians in avoiding, monitoring, or modifying potentially interacting medications (Kuperman *et al.*, 2007). In spite of such crucial importance, DDI alerts are routinely dismissed by physicians, with between 49% and 96% of such alerts overridden (Nightingale *et al.*, 2000; Payne *et al.*, 2002; Schedlbauer *et al.*, 2009; Van Der Sijs *et al.*, 2006; Weingart *et al.*, 2003). These overrides have important clinical consequences: 5–14% of inpatients and 1–7% of outpatients who receive potentially interacting medications will experience an adverse drug event (Magro *et al.*, 2012; Weingart *et al.*, 2003). The elderly are particularly at risk, with 20–60% of patients over 65 being exposed to a potential DDI, and 5–15% experiencing an actual DDI-induced adverse event (Becker *et al.*, 2008; Doucet *et al.*, 1996; Magro *et al.*, 2012; Obreli-Neto *et al.*, 2012).

To improve DDI alerts, prior studies have looked at different design aspects, such as organizational (Lo *et al.*, 2009; Shah *et al.*, 2006), presentational (Luna *et al.*, 2007; Phansalkar *et al.*, 2010) or contextual (Duke and Bolchini, 2011; Duke *et al.*, 2013). For example, Duke and Bolchini (2011) introduced *Contextual Drug–Drug Interaction* (CADDI) alerts which demonstrated a successful model for integrating relevant, patient-specific data into alerts. In a preliminary laboratory evaluation, physicians perceived CADDI alerts to be efficient, time-saving and easy to understand. However,

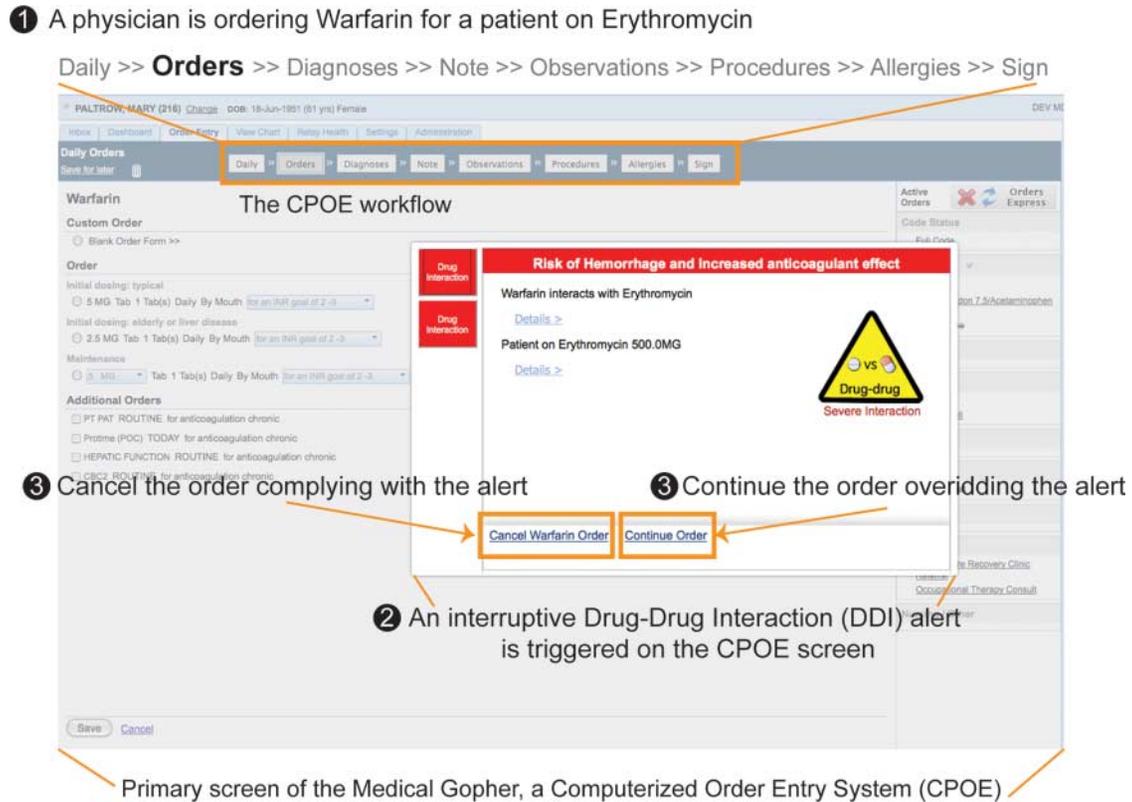


Figure 1. An example workflow of DDI alerts: a CPOE, the Medical Gopher, triggers a DDI alert as (1) a physician attempts to order Warfarin for a patient on erythromycin (Duke *et al.*, 2014). (2) This severe DDI alert is displayed interruptively on the CPOE screen expressing the severity of the interaction and (3) providing the physician with choices to either cancel the order or continue by overriding the alert. Alert overrides are tracked by the Gopher and may be used to modify the alert delivery. For example, alerts could also be displayed as noninterruptive reminders on the sidebar.

a follow-up, randomized and controlled trial with CADDI found surprisingly different results (Duke *et al.*, 2013). The study found that despite incorporating relevant contextual data indicating increased patient risk of an interaction, physician adherence rates remained low (~15%). And these high-risk patients did indeed go on to experience adverse drug events at significantly higher rates than other patients. Overall, these findings underscore the contention that only improving content, contextual cues, or layouts are insufficient to improve alert adherence.

Besides organizational, presentational and contextual factors, individual factors such as time, trust/mistrust and motivation play a crucial role in physicians' adherence to alerts (Zheng *et al.*, 2011). For example, research shows that physicians' nonadherence to alerts is often due to a *distrust of CDS systems* (Van Der Sijs *et al.*, 2006). This distrust stems from complaints that most DDI alerts are irrelevant to the individual patient or of minor clinical significance (Ash *et al.*, 2007; Weingart *et al.*, 2009). Physicians also feel that current alerts do not present persuasive scientific evidence (Weingart *et al.*,

2009). Whereas clinical advice is constantly shared among physicians based on personal clinical experience or reference to existing literature, computerized alerts do not elicit the same level of trust from physicians.

This dichotomy in responding to clinical advice shared by peers vs. computers may also be rooted in how individuals make social judgements—cues available during physician–physician interactions may be unavailable during physician–computer interactions. For example, in a social milieu, from a social-cognition perspective, decisions are determined by a multitude of factors (Kunda, 1999), such as how concepts are formed and new instances are categorized (attitude priming, Higgins *et al.*, 1976); whether systematic or heuristic processing is at play (Chaiken *et al.*, 1989); what kind of biases are affecting heuristic processing (e.g. illusory correlations, Chapman and Chapman, 1969); or how moods are influencing the choice of reasoning strategies (Kunda, 1999; Taylor, 1991). When among peers, physicians' reasoning strategies about clinical judgements may be affected by their social surroundings, which in turn may dictate whether or

not they trust and follow shared advice. For example, a supervisor's reminiscence of a salient event could serve as a crucial heuristic to make a decision. A group's discussion of professional ethics could affect a physician's mood leading to a choice of careful, systematic processing of all facts at hand. In contrast, the lack of similar cueing in computerized clinical alerts might be contributing toward physicians' noncompliance.

Additional factors that contribute to mistrust in alerts include the fact that such warnings are often poorly timed in the clinical workflow (Krall and Sittig, 2002). DDI alerts trigger when physicians have already initiated the ordering of a medication—when they are in an implemental mindset, which entails a focus on achieving the chosen option and gives rise to directional goals. Directional goals motivate individuals to draw a particular conclusion and selectively consider beliefs and rules that support their desired conclusions (Kunda, 1990, 1999). By arriving late in the decision-making process and interrupting the interaction flow, current alerts risk motivated reasoning for already-formed decisions (Kunda, 1990), thus failing to elicit physicians' trust.

It is thus clear that even with improvements in alert saliency, accuracy and frequency, there remains a critical lack of trust between physicians and computerized decision support systems. It follows that, before looking at how to improve alerts, we must look at how to improve the trust between physicians and computerized advice. If the issue of trust is not properly addressed, distrust in the advice given threatens the foundations of the design of alert systems, thus making all other improvements marginal or irrelevant.

Whom do physicians trust when exchanging advice about medication prescribing? How clinical judgements are reached? Exploring this type of questions will provide the necessary foundational knowledge to ideate alert designs that have the potential to garner physician's trust, and thus increase the level of adherence. Our approach, detailed in the remainder of the paper, is first to examine the dynamics of exchanging medical advice among physicians during inpatient team meetings. We then use this knowledge to inform novel principles for effective alert design that are based on what physicians consider important when taking advice from peers in the context of their clinical activities.

3. METHODS

To understand how physicians share medical advice during inpatient team meetings, we employed an exploratory sequential mixed methods study approach (Creswell and Clark, 2007). First, we conducted a modified version of the contextual inquiry (CI) technique (Beyer and Holtzblatt, 1998), and then administered a survey to empirically evaluate our qualitative findings. Both CIs and surveys were conducted in the Eskenazi Health Network (Eskenazi Health, 2014).

3.1. Contextual inquiry

3.1.1. Procedure and participants

CI—an ethnographic technique—aims at understanding user requirements by actively observing users (partnership) in their work environment (context) while engaged in work pertinent to the scope of investigation (focus). Our objective was to unearth key factors that drive trust and compliance among physicians in an inpatient environment while they prescribe medications. However, inpatient environment is a complex domain where extensively trained physicians engage in expert work. This domain complexity poses challenges to the data interpretation process because human-computer interaction (HCI) researchers often lack clinical expertise (Chilana *et al.*, 2010). Hence, instead of following the original partnership model of CI, we collaborated with an expert physician (third author) and adopted a *persistent partnership model* (Chilana *et al.*, 2010). The domain expert took part in all observation sessions. After each observation, an interpretation session was organized where the domain expert answered researchers' questions regarding the session and confirmed (or corrected) their interpretations.

Three CIs—totaling 255 min (over 4 h)—were conducted in an Eskenazi Health Hospital in Indianapolis between December 2013 and January 2015. Overall, the CI sessions observed and engaged 22 health care professionals during inpatient team meetings. Each team meeting comprised interns (1–2), medical students (2–3), pharmacists (1–2), residents (1–3) and a team attending (Figure 2). Across all sessions, the inpatient team meeting aimed at reviewing admitted patients' data and taking necessary medical decisions (such as ordering tests, consulting specialists or prescribing medications). The focus of our CI comprised: inherent properties of the context (e.g. calm, secluded, or interruptive, noisy), roles and relationships (e.g. who are the mentors, mentees, and peers), frequency of advice sharing, manner of seeking advice (e.g. reactive, proactive or solicited), manner in which advice was shared and phrased (e.g. communication style) and dynamics of compliance (e.g. when advice was followed or when and whose advice was overruled). Researchers took pictures and recorded notes during the observation.

3.2. Survey

3.2.1. Survey design

To further our understanding of the themes discovered in the CIs (see Results), we conducted a 19-item online survey with a greater number of physicians involved in clinical work (Section 1, Supplementary Material). Themes that emerged from the inquiry were pruned and selected based on their relevance to our focus: dynamics of trusted advice among physicians in their daily clinical work. Each of the chosen themes was then operationalized in one or more questions.

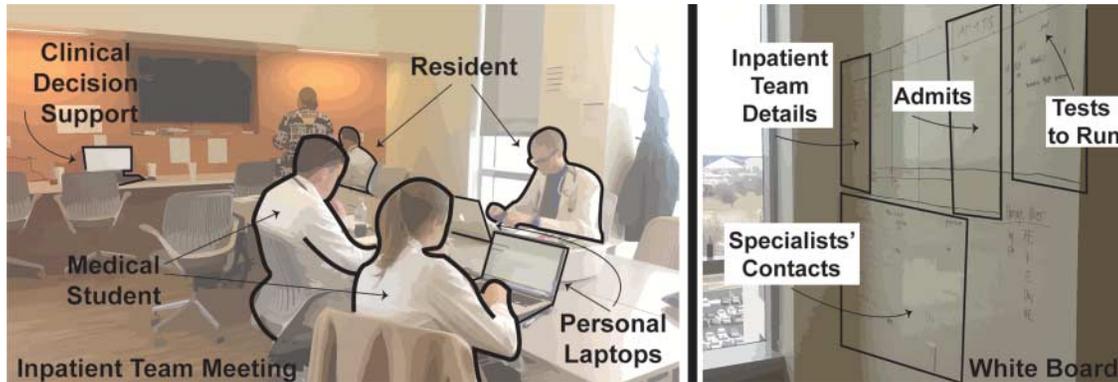


Figure 2. Snapshots from the second observation session: in the inpatient team meeting, most physicians were situated at a central table (left), engaged with either patient data or their personal digital devices (e.g. laptops, tablets). A white board (right) showed snippets of information that required discussion (e.g. admits), future action (e.g. tests) or potential consultation (e.g. specialists' contacts). Notably, no CDS systems were part of this ecosystem.

Overall the survey was structured as follows: we provided a medical scenario where physicians played the part of a second-year medical resident. The scenario included a drug recommendation from different colleagues that differed from their decision. Physicians used a 10-item Likert-type scale to rate which colleague's advice they were most likely to trust or follow.

We chose to orient the scenario to a second-year resident's mindset, because a second-year resident sits at the middle of the inpatient team hierarchy—below attending physicians but above interns and medical students. So our intent was to put the respondents in a frame of mind where they will be susceptible to hierarchical influences but not at the bottom of the medical hierarchy. Medical training carries a strong imprint on the minds of those going through it. We were confident that current residents or attending physicians (our target group for this study) could place themselves in the frame of mind of being a resident, particularly as this study is less about explicit medical knowledge than about the relationships in an inpatient environment.

In part A of the survey, participants considered the same recommendation that was coming from different colleagues. In part B of the survey, participants considered two conflicting recommendations that were coming from two different colleagues. In part C, participants reported (in free text) their verbal communication style to express their differences with a recommendation coming from a more senior colleague (having a higher role in the medical hierarchy, e.g. attending). Questions about user demographics included gender, age, current position, years of clinical experience and the percentage of time spent in an inpatient environment.

3.2.2. Survey administration

The online survey was developed and administered using Google Forms (Google, 2014), and was approved by the

Indiana University Institutional Review Board (Protocol 1409022835). The three parts of the survey were not randomized across participants because their ordering was crucial to the medical scenario presented in the survey. However, the different responses (corresponding to different colleagues' advice) were counterbalanced using a balanced Latin Square. Survey responses were anonymous. As a token of appreciation for their participation (around 10 min), participants had a choice of receiving a \$10 Amazon gift card via email. The survey was randomly emailed to 87 residents, fellows and attending physicians recruited through the Eskenazi Health Network's mailing lists between September 2014 and January 2015.

3.2.3. Statistical analysis

The 10-item Likert-type measurements were analyzed at interval level. We used R for our statistical analysis that was conducted at two levels: *Question level*, where each question was an independent factor; and *Theme Level*—where multiple questions corresponding to certain themes (extracted during the CI) were combined as independent factors (see Table 1). During the combination of multiple questions, each question was given equal weightage.

Overall, our objective was to explore how well the survey responses validated the themes that emerged from the CI. Each survey question that asked participants' trust in (or compliance with) a single colleague (Part A, Q1–Q10) was mapped exclusively to one of the themes that represented the corresponding colleague's role; questions where participants considered two conflicting recommendations coming from two different colleagues (Part B, Q11–Q18) were also exclusively mapped to either of the two themes that represented the two colleagues' role. For example, Q13 (that asked whether to trust a medical student with an *NEJM* study or the team attending) was pertinent to two themes—*Demonstrated*

Knowledge of Evidence from the Literature and Role in the Medical Hierarchy. Hence, a response (x out of 10) was categorized as either of the following: (a) left polarized (rating of $\leq 5/10$) and mapped to the left theme (i.e. evidence from literature) with the response value of $10 - x + 1$; or (b) right polarized (rating of $> 5/10$) and mapped to the right theme (i.e. medical hierarchy) with the response value of x .

4. RESULTS

4.1. Contextual inquiry

4.1.1. Data analysis

Following each CI, we conducted an interpretation session with the domain expert and analyzed our data to produce five work models (Beyer and Holtzblatt, 1998). The *Sequence Model* identified the sequence of activities, triggers that activated those sequences and team members’ intent at the time of action. The *Flow Model* documented individual team members’ responsibilities, their communication styles (informal or formal, with or without artifacts) and their communication topic with peers. The *Artifact Model* documented the physical objects used in an inpatient team’s decision-making, and the *Physical Model* identified the constraints posed by the workspace. Finally, the *Cultural Model* identified the underlying influences on decision-making, the invisible expectations, values, policies and mindset. After the completion of all CI sessions, we consolidated the individual work models. The consolidated sequence, flow, and cultural model instrumented our follow-up *theme extraction*. The consolidated sequence model defined alternative steps (abstracted from the individual observations) that the actors followed to accomplish a corresponding intent (Table 2). The consolidated flow model identified the different roles shared among the team members and their responsibilities (Figure 3). The consolidated cultural model revealed whose decisions were being influenced by whom in an inpatient meeting (Figure 4). Using these consolidated work models, we identified and iteratively refined a

set of emerging themes representing the dynamics of sharing medical advice among physicians during inpatient team meetings.

4.1.2. Sequence model

The consolidated *Sequence Model* confirmed that the primary intent of the inpatient meeting was concurring on treatment decisions about admitted patients—led by the primary attending and residents (Table 1). We also uncovered three secondary intents: training (medical and pharmacist students), deciding whether a specialist consult is required and finding out the rationale behind a previous decision. For example, during the team meeting, the attending or the resident(s) would use a patient’s condition to train medical students about ordering required tests (e.g. ‘Did you get ECHO? How could you not get that? [...] Everyone with new Afib needs ECHO’), recognizing common pitfalls (e.g. ‘You should be really careful if you are going around a black box warning [exceeding dosage]’), or identifying scenarios when a specialist’s consult is deemed necessary (e.g. ‘I’m not the best to ask, would see what renal says’). When failing to recall the exact reason for an earlier course of treatment (e.g. ‘Because we talked about it last time, and decided not to go aggressive on the dose’), a resident used Gopher (a CPOE system, Duke et al., 2014) to track down his earlier notes and recognize the rationale (‘Now reading my note again, I remember [...]’). While the resident looked for earlier notes to update the team, the team continued discussion about the next patient—a breakdown that was necessary to maintain the timeliness of the advice.

Our model also indicated two sub-sequences during the inpatient meeting. Medical students provided additional patient information on resident’s request (trigger) to arrive at a decision (intent); and the attending or the residents temporarily left the meeting to address additional issues requiring immediate attention (intention) when receiving a call/pager alert (trigger). Considering specialists’ consultation followed one of the following sequences: the attending endorsed the advice (e.g. ‘As far as duration of abx [antibiotics], that is a GI

Table 1. Connecting themes mined from CI to survey questions. Each theme is characterized either as a global (-G) or a situated factor (-S).

Themes mined from contextual inquiry	Survey questions ^a
1 Specialization/Expertise (SP-G)	Q1, Q6, Q11L, Q12R, Q15L, Q14L, Q16R, Q18L
2 Role in the Medical Hierarchy (MH-G)	Q2, Q7, Q11R, Q15R, Q13R, Q17R
3 Evidence of Understanding the Patient Situation (UP-S)	Q3, Q8, Q12L, Q16L
4 Demonstrated Experience (DE-S)	Q4, Q9, Q14R, Q18R
5 Demonstrated Knowledge of Evidence from the Literature (EL-G)	Q5, Q10, Q13L, Q17L
6 Collaborative and Inclusive Language (CL-G)	Q19

^aL indicates the mapping when the response provided to the corresponding question was left polarized (rating of $\leq 5/10$) and R indicates the mapping when the response was right polarized (rating of $> 5/10$). For survey questions, see Section 1, Supplementary Material.

Table 2. The consolidated *sequence model* for an attending (or a resident): an inpatient team's meeting followed the sequence of finding a patient's current condition, discussing treatment options and deciding next steps. Typically, 5–10 patients were reviewed in 1–2 h. We identified three secondary intents: training medical students, finding out the rationale behind a previous decision and reviewing specialists' advice.

Activity	Intent	Abstract step
<i>Find</i> out about a patient's current condition	Find out if the admitted patient requires immediate attention, such as medication, tests, or a consult from a specialist	<i>Trigger:</i> find out about an admitted patient
		From an intern From a medical student
<i>Discuss</i> next steps for the patient's treatment	Arrive on a decision about patient's treatment (primary)	Ask medical students/interns about potential course of treatment for the patient
	Train medical students and interns (secondary)	Discuss with residents about treatment decisions
	Review specialist's advice about the patient (secondary)	Gather more patient data from the medical students
	Find out the rationale behind a previous decision regarding a patient's treatment plan (secondary)	Consider whether to follow a specialist's advice or not
		Ask pharmacist for advice on prescribing medicines (e.g. dose or alternatives)
		Endorse a decision from a resident or a student
		Correct a student's response about potential course of action or disagree with a resident
		When failing to recall why a certain treatment was prescribed previously, browse one's own notes in a computerized system (e.g. CPOE) to recognize the rationale.
		<i>Breakdown:</i> When the attending/resident(s) leave the meeting room, discussions would either stop, or other members would reiterate them as they rejoined.
		<i>Breakdown:</i> Browsing previous patient notes always overlaps with the rest of the team continuing discussion about other patients. Hence, the team member engaged in browsing is virtually absent from the team's discussion during that time.
<i>Decide</i> next steps for the patient's treatment	Decide the tests to run	Train medical students about what (paper) notes to send to specialists about a consult in progress.
	Prescribe medications with a certain dose	Train students about what tests are absolutely necessary to order for the patient
	Ask medical students to consult specialists	Train about common pitfalls that students might encounter while using a computerized ordering system after the meeting
	Respond to specialists about a consult in progress	

call'); as new information became available, the team decided to have another round of consulting (e.g. 'I will sleep better at night with a neurological consultant'); or the attending overruled the consultation (e.g. 'Why heparin dip? Why not

lovenox?'). The key finding of the sequence model was the *interruptive nature* of the meeting environment that caused inefficiencies in the advice-sharing process. An example of this breakdown occurred when the primary attending temporarily

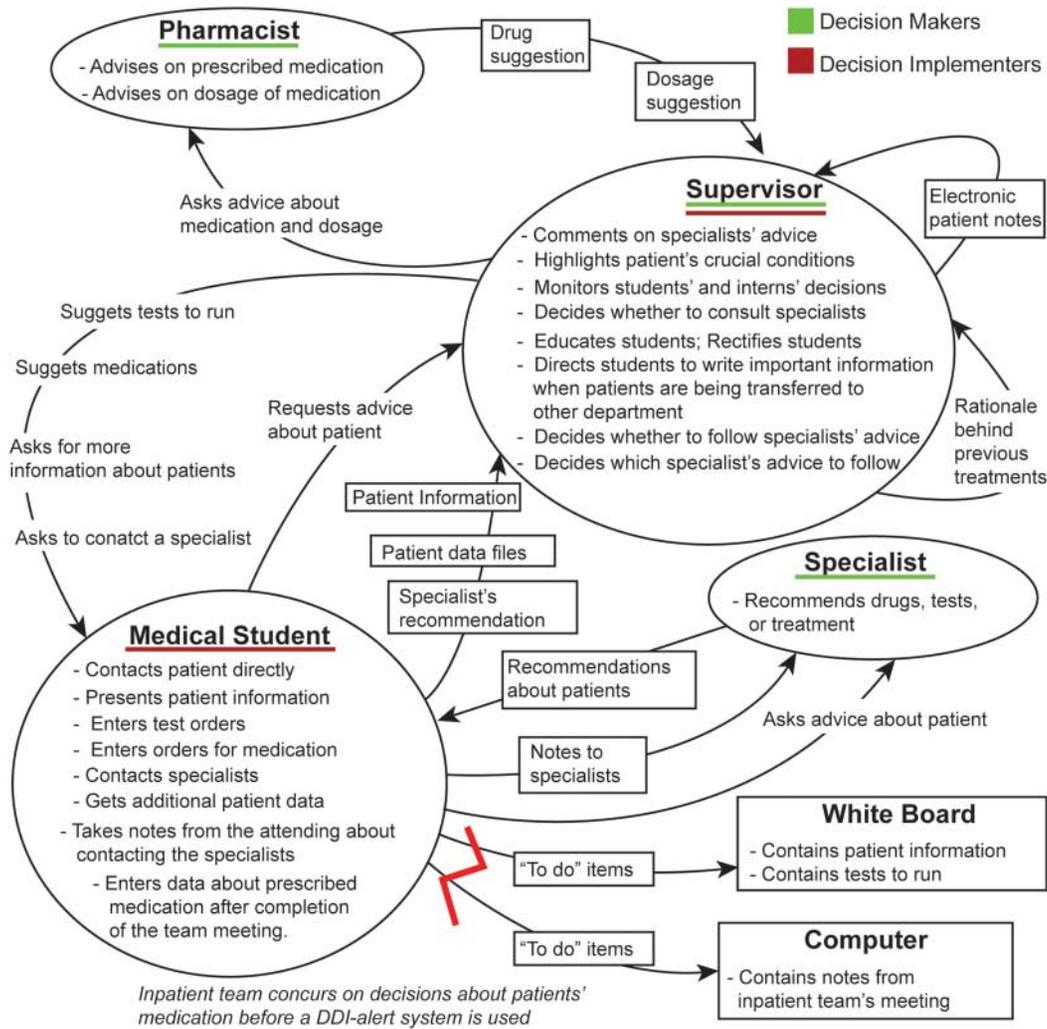


Figure 3. The consolidated flow model: two primary functional roles—decision-makers and decision-implementers—shared across different individuals emerged during inpatient team’s decision-making. The ‘Supervisor’ role represents attending and residents. The ‘Medical Student’ role represents students and interns. The lightning bolt represents a breakdown: final decisions are concurred upon by the team before orders are entered in a computerized system, thus making clinical alerts displaced in time with respect to the decision-making flow.

left the meeting and rejoined, and other members had to reiterate the team’s discussions (Table 2).

4.1.3. Flow model

The consolidated Flow Model (Figure 3) unpacked the different functional roles in the inpatient team meeting, categorized how individual physicians shared the responsibilities of these roles, and revealed how communication flowed among these roles (and hence the corresponding individuals). Two primary functional roles emerged from the model—decision-makers and decision implementers. Attending and residents primarily served as decision-makers (e.g. ‘I think we should go ahead and do it.’). But when specialists or pharmacists provided the necessary advice, they acted as decision implementers (e.g.

‘Has she received IV iron? Renal [says] they recommend IV iron.’). However, advice from different specialists was also vetted based on associated risks to the patient if identified by one of the experts. We observed specialty boundaries: In a certain instance, the ID’s (infectious disease) recommendation was overturned based on the Renal’s recommendation, because according to the Renal, the former advice endangered patient’s condition. Interns, medical students (and sometimes pharmacist students) primarily acted as decision implementers. Although they actively took part in the decision-making process, their decision was almost always endorsed or corrected by the supervisors (attending or residents). Pharmacists and specialists always served as decision-makers—they provided advice to the inpatient team when solicited. As indicated by

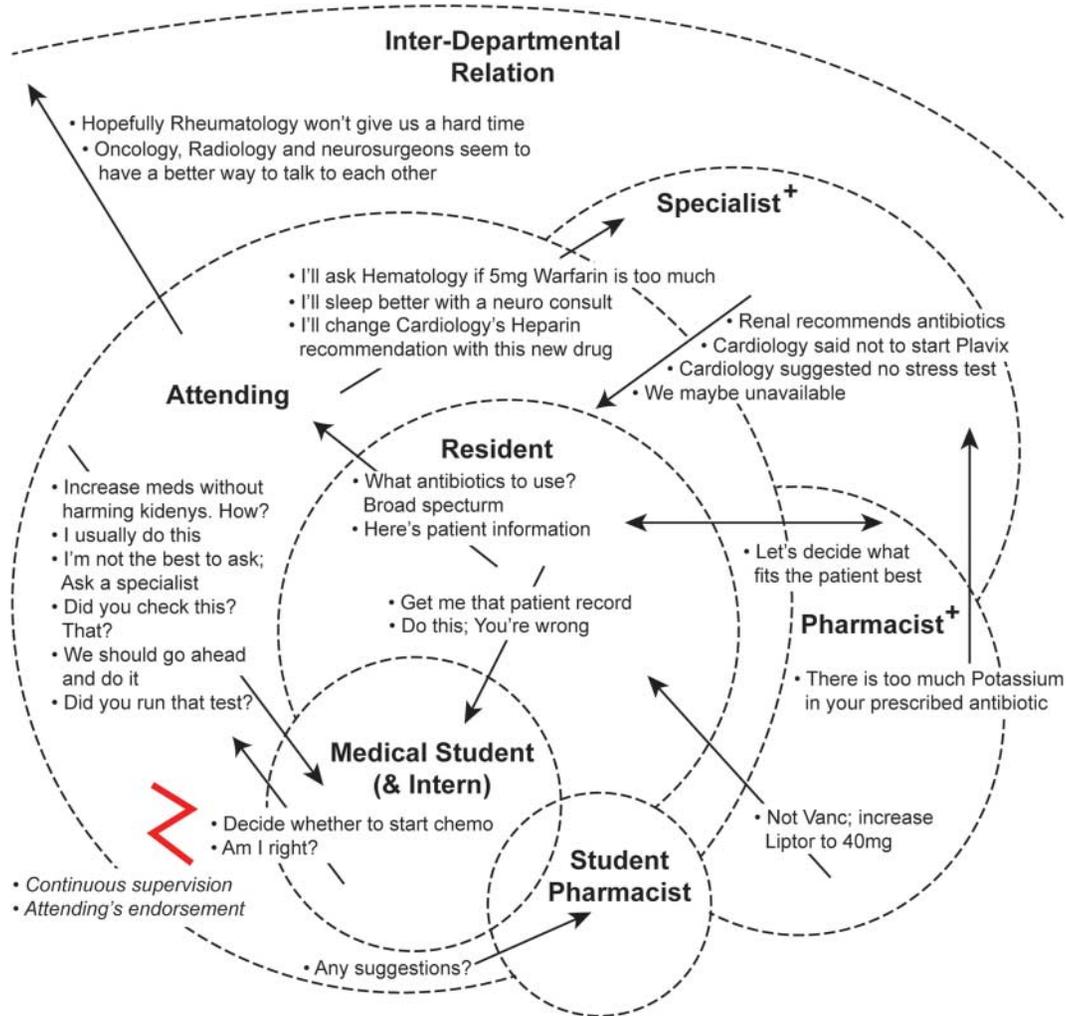


Figure 4. The consolidated *cultural model*: when sharing advice about medication prescribing, inpatient teams enact a complex network of influences that drives decision-making. ‘+’ indicates one or more individuals belonging to a group that exhibits the same kind of cultural influence. The lightning bolt represents a breakdown: medical students’ decisions are continually supervised—rectified or endorsed by the attending and the resident(s). The lack of this close supervision while individually interacting with the CPOEs may lead medical students to override important alerts.

the flow model, supervisors endorsed the final decision for any patient treatment, and medical students executed that decision (often using the computerized system to order tests or prescribe medications).

The model also identified important breakdowns during the activities. A crucial breakdown in the communication flow was the fact that physicians used the computerized system as a terminal ‘documenting’ step in their decision-making process. By then, alternative treatment suggestions had already been presented, dosage and tests discussed, future consults agreed upon and pharmacists’ opinion noted. Hence, current computerized systems did not play any active role in the inpatient team’s decision-making process—they were displaced in time with respect to the decision flow.

4.1.4. *Cultural model*

The consolidated *Cultural Model* (Figure 4) identified the influencers affecting the team’s advice sharing, the styles in which they influence the process and the extent of their influence. Crucially, the cultural model identified a pervasive mindset of an inpatient team: a strict adherence to medical hierarchy and a strong preference for experts’ opinions. Overall, our model revealed three distinct traits of the underlying culture of an inpatient team’s advice-sharing process.

First, the inter-departmental relations influenced the team’s decisions to engage in soliciting consults from domain experts (e.g. ‘N-surg [neurosurgeons] didn’t put in clear recs [recommendations] for what they wanted to do. [...] Maybe

rad-onc [radiology oncology], and n-surg have a better way to talk to each other because we don't get calls from them.')

Second, attending and residents (supervisors in the flow model) completely influenced the activity of the medical students and interns (primary decision implementers in the flow model). However, supervisors rarely provided students simple instructions to carry out. Instead, the attending (or residents) would ask students to present their course of action, and either endorse or correct their decisions, train them about required protocols, and inquire them about other pertinent concerns (e.g. 'If you were to increase his meds, but not harm his kidneys, what would you do?'; 'I usually do this. '; 'Have you checked her Thyroid?'; 'What does EPO do?'; 'Did you write for platelets also?'). Most importantly, all decisions that the team concurred upon were presented as a collaborative inference (e.g. 'We'll do this. '; 'Should we treat for a full course?'). Overall, whereas the flow model uncovered a supervisor-supervisee relationship evident in an inpatient meeting, the cultural model surfaced an undertone—identifying it more closely to a mentor-mentee relationship.

Third, we observed a strong influence of expert opinions on the team's decisions. However, although the primary attending would often follow specialists' (or pharmacists') suggestions about treatment options, there were instances where the attending would decide to disregard their suggestions. Our observations suggested that these decisions were based on team members' (especially supervisors') prior experience, thus emphasizing experience-based practices oriented towards ethical clinical judgement. However, further investigation is required to confirm our observations, and infer about the relative importance between *medical expertise* and *medical hierarchy*.

4.1.5. Other work models

Although the artifact model and the physical model did not play a key role in uncovering elements of trusted advice among peers, they identified the following. First, we observed a distinction between the two different ways artifacts were used—*exclusive-use* and *shared-use*. Exclusive-use included artifacts such as patient notes from morning rounds, personal laptops, tablets or use of a computer terminal providing CDS. Team members individually inputted information into these artifacts and checked the output. Two artifacts were primarily used in a shared manner: the patient history (either on paper or in a computer terminal, such as Gopher) and a centrally-placed whiteboard (Figure 2) that contained frequently-accessed information, such as admitted patient's details, specialists' contacts, team members' names, and things to do after the conclusion of the meeting (e.g. tests to order, or medications to prescribe). Second, the physical model identified two separate spaces: the *brainstorming* space and the *implementation* space. The brainstorming space was organized as a central collaborative space with team members facing each other, such as in a roundtable. The implementation space was modeled as individual stations, where a team member would

use a computer terminal to order tests or prescribe medications. Such a spatial organization reflected an inherent assumption that CPOEs are used as read-only during the team's discussions while orders were inputted or modified individually prior or after the meeting. This serial ordering crucially emphasizes how the current paradigm of clinical alerts is not a part of the vital ecosystem that encapsulates the clinical decision-making process.

4.2. Elements of trusted advice among physicians in clinical settings: emerging themes

Informed by our consolidated work models, we identified a set of themes. These themes were developed and iteratively refined to demonstrate the elements of trusted advice among physicians in clinical settings. They crystallized our findings from the CI and were used to design our follow-up survey.

4.2.1. (SP-G) Specialization

Individuals with professionally recognized specialization influenced inpatient team's decision-making. Specialists were trusted because they entailed better knowledge about the discipline, greater accountability and higher standards of care. We observed *Specialty Boundaries*: conflicting advice from different specialists were examined based on the priority of associated risks to the patient identified by one expert, but not another.

4.2.2. (MH-G) Role in the medical hierarchy

The professional roles of individuals in an inpatient team (e.g. attending, intern or medical student) distinctly influenced whose advice was trusted more during the team discussions. The higher the individuals belonged in the medical hierarchy, the more their peers were to trust their advice.

4.2.3. (DE-S) Demonstrated experience

Individuals with extensive experience about a patient, a medical condition, an environment, or a medication, garnered their peers' trust. They delivered crucial knowledge—specific to a situation-at-hand—that was unavailable to or unnoticed by any of their peers.

4.2.4. (UP-S) Evidence of understanding the patient's situation

Team members earned trust by demonstrating a thorough understanding of the patient's condition. Typically, they articulated accurate patient history, such as rationale for previous course of treatment(s), updated the team on recent developments of patient's symptoms or proposed a convincing future plan of action.

4.2.5. (E-S) Empathy

Individuals showing empathy toward the patient were actively trusted by the team members. The overall empathy among the

inpatient team members to care for the patient was the key factor that helped them to finally concur with a decision.

4.2.6. (EL-G) Demonstrated knowledge of evidence from the literature

By referencing relevant studies in the medical literature pertinent to a patient's condition, team members might be able to gain credibility and overall trust.

4.2.7. (CL-G) Collaborative and inclusive language

In spite of the medical hierarchy evident in an inpatient team meeting, team conversations demonstrated an inclusive tone (e.g. 'Let's do this.'). This style was effective in driving the team's decisions as it allowed individuals to identify themselves with the team and recognize their responsibility.

4.2.8. (TA-S) Timeliness of the advice

The inpatient team reviewed patients in a sequential manner. This temporal feature of the inpatient team meetings operated as a double-edged sword. On the one hand, the rigid linearity of the meeting structure made it an inclusive process: all team members used their time efficiently and participated and contributed to all aspects of the decision-making. On the other hand, unforeseen events (e.g. a physician who had to leave temporarily the room; or a physician who could not recall patient details at a given moment) put the meeting at risk of missing crucial input from the virtually absent team members, easily breaking the participatory nature of the decision-making process. For these reasons, timeliness of the shared advice played a crucial role in the decision-making process.

Taken together, the eight themes that emerged from our CI may also be classified at a higher level in two groups: global factors (themes marked with -G) such as specialization or role in medical hierarchy, and situated factors (themes marked with -S), such as demonstrated experience or evidence of understanding the patient situation.

As a first step of validating with a larger sample of physicians the crucial themes emerging from our formative study, we designed and administered an online survey (see Table 1). This survey enabled us to address six crucial themes emerging from our contextual inquiries. In the following section, we discuss the results of our survey.

4.3. Survey validation

Of the 87 questionnaires sent, 37 were returned (22 females). Respondents were mostly less than 30 (17) or less than 40 (11) years old and were mostly either resident (19) or attending physicians (17). Seventeen physicians worked less than 5 years while 4 worked for more than 25 years. Except four respondents, all physicians currently worked in an inpatient environment, and 20 physicians spent more than 50% of their time in an inpatient environment.

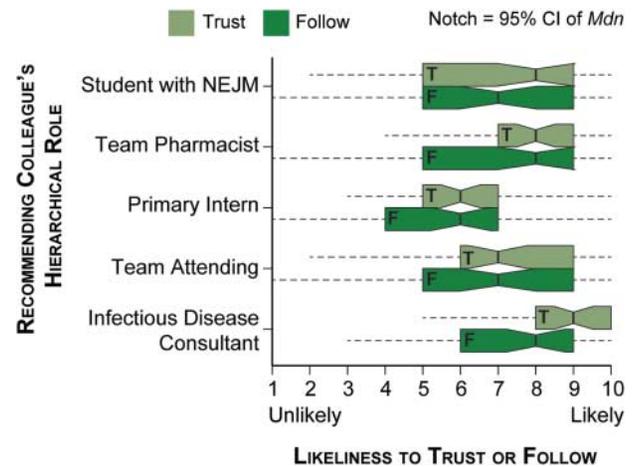


Figure 5. Recommending colleagues' hierarchical role significantly affected how much a second-year resident would trust or follow their recommendation. ID (infectious disease) consultant would be significantly more trusted than the primary intern.

Because Shapiro–Wilk tests were significant, $P < 0.001$, and Q-Q plots were nonlinear, we analyzed performance data using nonparametric tests for within-subject experimental design. In our univariate data analysis (survey question as the unit of analysis), we used box plots to detect outliers. Box plots are suitable for exploring nonparametric, quantitative data with a heuristic of $1.5 \times$ inter-quartile range beyond the upper and lower extremes for outliers (Hodge and Austin, 2004; Laurikkala *et al.*, 2000; Seigel, 1988). Box plots detected three outliers (P6, P7 and P23) for questions 14 and 18 (Part B of the survey). For both of these questions, participants 6, 7 and 23 had responded '1'. We believe that this was a human error—participants inadvertently clicked on the first option of the scale while scrolling through a list of 18 questions. Outliers were removed for all analysis including Part B of the survey.

A Friedman's ANOVA showed that recommending colleagues' hierarchical role significantly affected how much a second-year medical resident would *trust* their recommendation, $n = 37$, $\chi^2(4) = 33.44$, $P < 0.001$ (Figure 5). Post-hoc Wilcoxon Signed-rank tests (Bonferroni correction for 10 comparisons, $\alpha = 0.005$) showed the infectious disease consultant (Mdn = 9, IQR = 2) is significantly more likely to be trusted than the primary intern with a medium effect size, (Mdn = 6, IQR = 2), $p < 0.001$, $r_{\text{TRUST}} = 0.43$, and also the team pharmacist (Mdn = 8, IQR = 2) is significantly more likely to be trusted than the primary intern with a medium effect size, $P < 0.001$, $r_{\text{TRUST}} = 0.41$.

Similarly, recommending colleagues' hierarchical role significantly affected how much a recommendation is likely to be *followed*, $\chi^2(4) = 24.76$, $P < 0.001$ (Figure 5). Post-hoc Wilcoxon signed-rank tests ($\alpha = 0.005$) showed the infectious disease consultant (Mdn = 8, IQR = 3) is significantly more likely to be followed than the primary intern with a medium

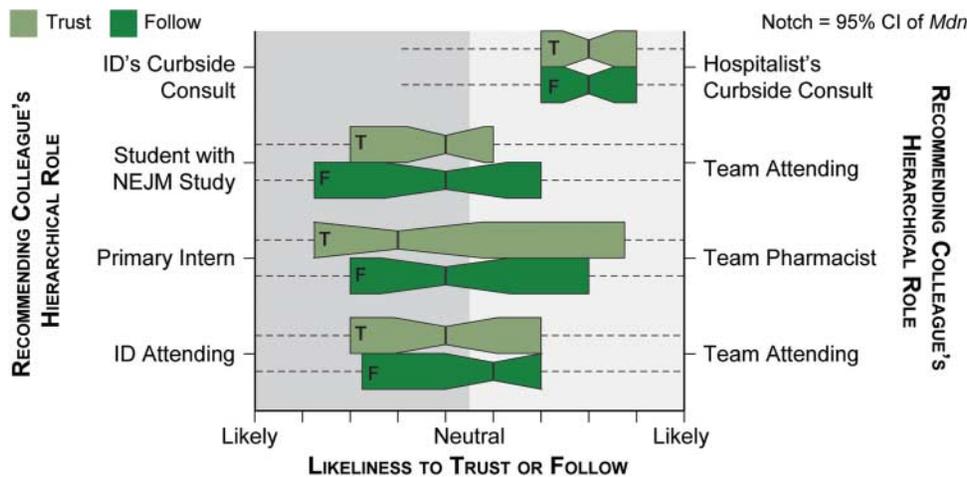


Figure 6. When faced with conflicting recommendations coming from peers, a second-year resident would trust and follow a curbside consult from Hospitalist than one from the new ID fellow.

effect size (Mdn = 6, IQR = 3), $p = 0.002$, $r_{\text{FOLLOW}} = 0.37$, and also the team pharmacist (Mdn = 8, IQR = 4) is significantly more likely to be followed than the primary intern with a medium effect size, $P = 0.001$, $r_{\text{FOLLOW}} = 0.37$.

Planned comparisons ($\alpha = 0.05$, 10 comparisons) showed that respondents are likely to trust significantly more than follow the recommendation of an infectious disease consultant with a medium effect size, $P < 0.001$, $r_{\text{TRUST VS FOLLOW}} = 0.44$; and likely to trust significantly more than follow the team pharmacist, $p = 0.004$, $r_{\text{TRUST VS FOLLOW}} = 0.33$.

Part B of the survey ($n = 34$, after removing outliers), where conflicting recommendations were provided, showed that respondents are equally likely to trust and follow either of the recommending colleagues (center of the scale fell within 95% CI of the median), except when the conflict was between the curbside consults from new ID Fellow and Hospitalist: respondents were more likely to both trust and follow the consult from Hospitalist (Mdn = 8, IQR = 2). Planned comparisons using Wilcoxon signed-rank tests showed no significant differences between the likeliness to trust and the likeliness to follow any of the conflicting pairs of colleagues, $P_s > 0.05$ (Figure 6).

When survey questions were compiled into themes (see Section 3), themes significantly affected the likeliness of a medical resident to trust a recommendation, $n = 34$, $\chi^2(4) = 23.80$, $P < 0.001$, and to follow a recommendation, $\chi^2(4) = 22.96$, $P < 0.001$ (Figure 7). Post-hoc Wilcoxon signed-rank tests ($\alpha = 0.005$, 10 comparisons) showed that both likeliness to trust and follow a recommendation is driven by specialization significantly more than medical hierarchy, $P_s < 0.005$, $r_{\text{TRUST}} = 0.42$, $r_{\text{FOLLOW}} = 0.35$. Specialization is also likely to be trusted and followed more significantly than evidence of understanding patient situation, $P_s < 0.001$, $r_{\text{TRUST}} = 0.45$, $r_{\text{FOLLOW}} = 0.43$. Likeliness to trust and follow

is also driven by experience significantly more than evidence of understanding patient situation, $P_s < 0.005$, $r_{\text{TRUST}} = 0.38$, $r_{\text{FOLLOW}} = 0.45$.

Planned comparisons ($\alpha = 0.05$, 5 comparisons) also showed that specialization is significantly more likely to be trusted than followed, $P = 0.008$, $r_{\text{TRUST VS FOLLOW}} = 0.32$, and similarly evidence of understanding patient situation is significantly more likely to be trusted than followed, $P = 0.001$, $r_{\text{TRUST VS FOLLOW}} = 0.40$.

N-Vivo V.10 (Richards, 1999) was used to analyze the open-ended Q19 (see Section 1, Supplementary Material), which asked how second-year medical residents would communicate their recommendation when in conflict with the team attending. We followed the traditional three-stage analysis (Winkelman *et al.*, 2005): responses were first coded into component phrases or keywords (open coding), then grouped into a framework of ideas (axial coding), and finally organized into more structured higher level themes (selective coding).

Two distinct themes regarding communication style emerged: (*linguistic*) *frame of reference* and the *tone of communication*. Respondents either chose an egocentric (first-person perspective) or an allocentric (third-person perspective) frame of reference. In an egocentric style, they referred to themselves (e.g. 'I believe it is beneficial to try'), to the attending (e.g. 'Do you think it is wise to go with amlodipine at this time?'), or to themselves and the attending together (e.g. 'How about we try X instead?'). During allocentric framing, respondents were implicit (e.g. '[...] it would be better to try another medication'), evidence-based (e.g. 'Based on known side effects [...]') or referring to the patient (e.g. '[...] patient puts high value on not having lower extremity edema').

Respondents' tone of communication was (1) reflective or prompting a discussion (e.g. 'what do you think about trying a [...]'), (2) politely prescriptive (e.g. 'I think there may

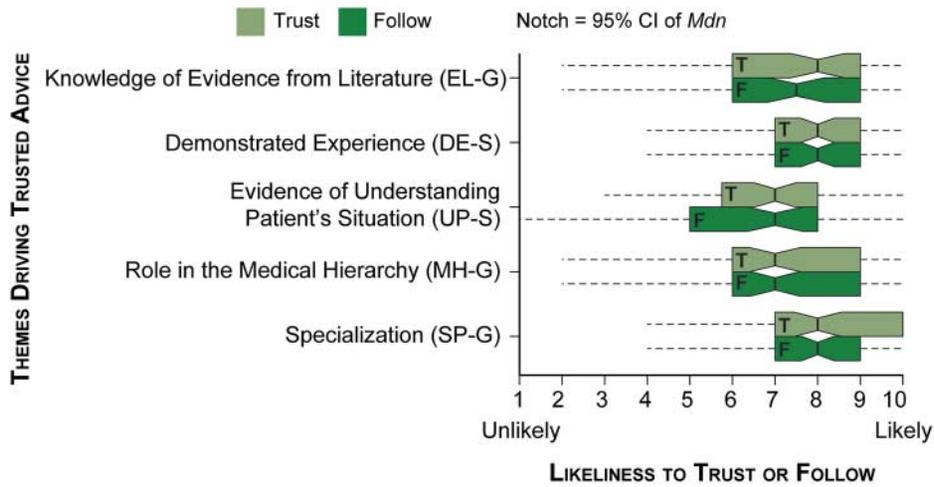


Figure 7. Themes emerging from our CI significantly affected the likeliness of a second-year resident to trust or follow a recommendation. Specialization would be trusted and followed significantly more than role in medical hierarchy.

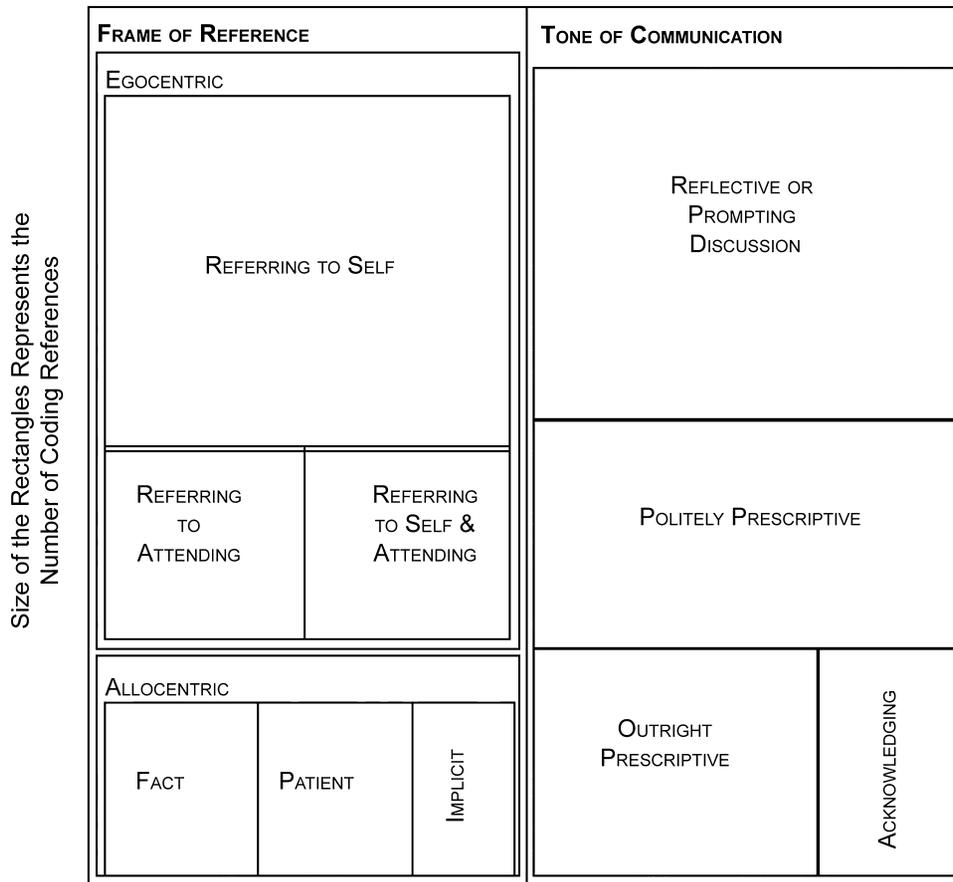


Figure 8. A treemap visualization showing the emerging nodes by number of coding references. The size of the rectangles represents the number of coding references. Self-reference was the most commonly coded frame of reference and reflective was the most commonly coded tone of communication.

Table 3. N-Vivo V.10 was used to analyze the open-ended Q19 (see Section 1, Supplementary Material), which asked how second-year medical residents would communicate their recommendation when in conflict with the team attending. Two overarching themes emerged about respondents' communication style when low-level ideas were compared and aggregated.

First order theme	Second order theme	Third order theme	Examples	
Frame of reference	Egocentric	Referring to Self	'I believe it is beneficial to try' 'I frequently have good results with amlodipine'	
		Referring to Attending	'I think that normally I would agree' 'are you concerned about what effect amlodipine' 'what do you think about trying a nitroglycerine drip' 'How likely do you think the dose we are choosing will exacerbate this patient's edema?'	
		Referring to Self and Attending	'[...] if we should try another medication first' 'Let's consider another medication as the patient has a h/o le edema' 'we should try another agent for BP control'	
		Allocentric	Referring to Fact	'Based on known side effects associated with amlodipine among patients [...]' 'because of the patient's prior history of lower extremity edema'
			Referring to Patient	'patient puts high value on not having lower extremity edema' 'patient has a history of LEE'
			Reference implicit	'it would be better to try another medication' 'What about the patient's with peripheral edema? Would amlodipine make it worse?'
	Tone of Communication	Reflective or Prompting Discussion		'What is your experience with it?' '[...] do you think it is wise to go with amlodipine at this time?' '[...] are you concerned about what effect amlodipine might have in the setting of her edema?'
			Politely Prescriptive	'I think there may be a better alternative that would' 'I would respectfully raise my concerns.'
		Outright Prescriptive	'I'd like to choose Lisinopril due [...]' 'we should try another agent for BP control'	
		Acknowledging	'I believe it is beneficial to try' 'I think that normally I would agree' 'I think amlodipine is good drug for lowering blood pressure in patients.'	

be a better alternative that would'), (3) outright prescriptive (e.g. 'we should try another agent for BP control. '), or (4) acknowledging (e.g. 'I think that normally I would agree [...]').

A treemap visualizing the emerging nodes by number of coding references showed that self-reference was the most commonly coded frame of reference and reflective was the most commonly coded tone of communication (Figure 8). Table 3 lists typical examples of how physicians' communication style was coded.

5. DISCUSSION

5.1. Emerging drivers of advice sharing among physicians

Specialization emerged as a key driver to elicit trust from physicians and to increase compliance with drug-related recommendations. This result was first identified during our CI and then supported by our survey (Figure 7). Thus, our findings provided empirical evidence to suggest that physicians tend to trust and follow the advice of peers whose area of specialization pertains to the case at hand—more than the advice coming from other colleagues in the team.

This emphasis on specialization seems to be tempered by another important factor: the demonstrated experience of the specialist. In fact, as emerged from our survey (Figure 6) respondents indicated that they would rather trust and follow the advice from a hospitalist (an experienced general practitioner) than a new ID fellow (a young specialist who just joined the team).

Other important drivers emerging from the CI observations include evidence-based knowledge (such as the latest literature on the topic) and the degree of understanding the patient's situation. These factors, however, were secondary to the role of medical hierarchy and specialization in guiding physicians' compliance—an observation that was further supported by our survey results. Furthermore, our results suggest that physicians would trust and follow specialists significantly more than colleagues who sit higher in the medical hierarchy (e.g. attending) with a medium effect size (Figure 7).

The importance of specialization as a driver of trusted advice may stem from two possible sources. First, physicians recognize that specialists have additional training and exposure to particular patient types and that they are up-to-date with the literature in their field and its application to individual patients. Second, the premium placed on specialist's advice is a reflection of the 'defensive' posture of medical practice, in which fears of medical liability (i.e. a lawsuit) drives much of the decision-making. In this context, the recommendations of a specialist are often considered the *standard of care*, and a physician who ignores these recommendations may do so at his peril should a lawsuit arise. In reality, the power of

specialists' advice is likely driven by both these factors in some combination.

In terms of language cues that are most common in sharing advice among physicians, our survey revealed a rich layout of strategies that physicians use to convey drug or treatment recommendations to colleagues. Crucially, our survey findings suggest that physicians mostly use a self-centered frame of reference to present their advice to peers (Figure 8). For example, when referring to themselves, physicians relied on leveraging their prior experience in similar cases or expressing strong conviction about a given course of action. Two much less common strategies concern the reference to the *attending* and the reference to *self and attending*. The latter directly connects to a theme emerged from a CI that centered on the use of *collaborative and inclusive language*. This style of communicating advice reflects a sense of teamness and collaborative decision-making that is at the heart of inpatient team dynamics. This finding is compounded by the emergence of another major theme that concerns the tone of the communication used in sharing advice: the use of a reflective language that prompts discussion, rather than a prescriptive one (Figure 8). For example, physicians strongly relied on asking questions to suggest potential problems in the recommended treatment or prompting a colleague to reflect on his experience in similar contexts before committing to a decision. This body of evidence suggests that arguments based on personal experience or expertise and the use of language that is inclusive of the team and mutually reflective are essential to shape the practice of sharing advice among physicians.

5.2. Design directions for trust-based alerts

The findings of this study have important implications for the design of future computerized alerts. Specifically they reveal many strategies for incorporating motivators for trust and following of clinical recommendations delivered by computerized decision support systems. In what follows, we propose several alerting strategies based on the results of this study.

5.2.1. Endorsed alerts

Presently, computerized alerts emerge from an anonymous source (i.e. the computer). While they are often accompanied by references from the literature, there is no sense of an actual person behind the advice. Based on the strong impact of specialization and medical hierarchy found in this study, we propose a 'brought to you by'-type sponsorship of clinical alerts. Specifically, an individual known and respected in the local environment—seen as an expert in the field of reference, and with a position of authority—could affix their name (or even image) to the alert. Such a connection to a person who triggers both the specialization and hierarchy drivers has the potential to increase adherence to an alert, as computing technology incorporating expertise has been suggested to

have increased powers of persuasion (principle of expertise, principle of reputed credibility, Fogg, 2003). Furthermore, endorsed alerts are also expected to influence physicians' mindset toward a careful consideration of all options and arrive at the best possible solution (accuracy goal, Kunda, 1990).

5.2.2. *Transparent alerts*

Currently, the decision to accept or reject an alert is private, in that it does not persist in the patient note or made visible to other team members. Based on the demonstrated influence of hierarchy and team on decision-making, we propose making physicians' decisions regarding clinical alerts visible in a patient's medical record. So, for example, if a DDI alert was overridden in the course of ordering a patient's medications, it would be stored along with the order documentation. This strategy would not be rendered as a judgment as to whether or not the decision was correct, but rather a form of 'sunlight' so that other physicians could follow the decision-making. We believe that this awareness—that others will be able to view physicians' choices—will lead to an increased consideration of recommended safety measures, similar to the principle of surveillance which suggests that such (overt) awareness of being monitored increases the likelihood of compliance towards a goal (Fogg, 2003).

5.2.3. *Team-sensing alerts*

In the inpatient environment, a time gap exists between the team's decision-making and the actual execution of an order by a physician in the CPOE system (typically by the intern or the resident). As a result of this gap, the recommendations made by the computer are delivered to the individual physician after the period of critical thinking by the team. This phenomenon has two costs. First, the team does not have the opportunity to consider the computerized advice during their decision-making process. Second, the physician entering the information must decide between circulating this information back to the team asynchronously or (far more likely) simply overriding the alert and moving on. An optimal alerting system would deliver guidance during the period of team discussion. Such solutions could be achieved via either high-technology (e.g. speech recognition) or low-technology (e.g. a scribe entering 'test' orders during the meeting). Alerts that consider the team dynamics will more likely be effective in influencing physicians while they are in a deliberative mindset—carefully considering all options to arrive at the best possible solution (accuracy goal, Kunda, 1990).

5.2.4. *Collaborative alerts*

Current computerized alerts are emotionally neutral. They provide relevant facts but are neither collaborative nor supportive in tone. Our findings suggest that the use of inclusive language that supports team-building (such as 'how about we? . . .') is effective in mitigating feelings of criticism while encouraging acceptance of advice. Computerized alerts

may be altered to provide a similar notion of collaboration. While this may be accomplished by the use of endorsed alerts (above), changing the language of alerts to reflect more joint ownership and responsibility for decisions may be effective at increasing adherence.

5.2.5. *Empathy-driven alerts*

Understanding the patient situation and other markers of empathy were influential in driving acceptance of advice. Drug safety alerts are currently generalized (i.e. not patient specific) and do not evoke a sense of empathy for the patient or the physician. An opportunity illuminated by our study is to embed more patient-specific information into alerts regarding their specific conditions and the impact of an adverse event. More subtly, we see the opportunity to embed physician-specific information regarding their previous decisions and what outcomes have resulted, thus readily catering representative instances toward making better probability judgements (Kahneman and Tversky, 1972). There is currently a gap between a physician's immediate response to an alert and their awareness of how previous patients have fared, thus increasing the risk of failure to detect an actual correlation between their prior decisions and adverse events (illusory correlations, Chapman and Chapman, 1969). Providing such information—tailored to patients and physicians—would increase both empathy and self-reflection on the part of the physician as well as increase the persuasive power of an alert (principle of tailoring, Fogg, 2003).

5.2.6. *Conflict-mitigating alerts*

Our results indicate that an advice emerges from a complex dialogue among physicians with different expertise and understanding of the patient situation. A specialist might provide an advice that is accurate in theory, but is not appropriate for the specific state of the patient. In these situations of conflicting advice, alerts should address this tension by, for example, prioritizing different courses of action based on the patient's condition.

5.2.7. *Agency-laden alerts*

Our findings highlight the importance for physicians to rely on their experience with the patient (e.g. documented through notes) to inform treatment decisions. Yet current alerts reason without considering the patient notes that an individual physician might have taken regarding possible courses of action. To loop in the 'agency' of a physician in making decisions, alerts can explicitly leverage individual notes in providing recommendations and thus garner an increased level of trust. This type of alerts enables physicians to track and become aware of their prior decisions, thus helping them to achieve a positive behavior (principle of self-monitoring, Fogg, 2003). Such easy access to prior decisions can also provide an accurate sense of prior probabilities, thus influencing the

physicians to apply statistical heuristic during decision-making (Kunda, 1999).

Overall, these requirements provide an empirically grounded basis to rethink current computerized clinical alerts and envision a new generation of computerized advice that can elicit more trust from physicians and thus potentially improve compliance to drug safety warnings.

5.3. Potential tradeoffs for trust-based alerts

Attending to clinical alerts halts a physician's workflow while prescribing medications. Our proposed design directions may run the risk of increasing overhead to the already taxed physicians' cognition and time. Thus, we advocate toward making the novel design elements either integrated or peripheral to the current design of alerts—both visually and to the decision-making flow. For example, we envision incorporating empathic or collaborative articulation to the current alert language without any substantial cognitive overload; or add peripheral modules to an alert inducing endorsement, transparency, or a sense of agency—with these add ins not required to be processed to reach a decision, but available for review. We do acknowledge that asking physicians to attend to alerts will improve reflection but may also increase the time needed to make an informed decision.

To further understand the cognitive overload associated with our proposed designs, we plan to pursue a controlled evaluation. In our evaluation, we will vary the number of modules presented at a time, the contents of each module, and provide a secondary task to simulate additional cognitive load required in multi-tasking clinical scenarios. Thus, future empirical evidence of physicians' inclination to adherence versus their cognitive overload will be instrumental in providing potential tradeoffs involved in designing trusted alerts.

5.4. Study limitations

Findings from our CI studies are limited by its scope: we investigated only sharing of advice during the inpatient meetings. After the meeting, interns or residents were individually responsible to carry out the team's decisions using CPOE systems. However, we did not conduct any follow-up observations to explore whether decisions taken during the meeting were indeed followed. In our early observations, we discovered that it was infeasible to ascertain whether or not advice was followed through after the meeting. The final clinical action taken for the patients—discussed during the inpatient meetings—depended on a multitude of factors. For example, a possible update on a patient's status, or any conflict in the availability of resources, such as staff or medications, could have led to discard an earlier advice. Such contextual factors were out of the focus of our CI, which concentrated on

the dynamics of sharing advice during the inpatient meetings. Thus to conserve our findings from those confounds, we limited our observation to team meetings.

To further validate our CI findings, we opted for a survey—instead of a real-life simulation—because we wanted to evaluate several alternative scenarios with a greater number of physicians. Although real-life simulations of our survey scenarios (18 variants) would be feasible and more ecologically valid, such simulations would pose scalability challenges due to the limited availability of the clinical population. Furthermore, simulation of information encounters to study decision-making in complex systems, such as emergency operations, usually span over two hours or more, where several interdependent aspects of decision-making are studied (Holzman *et al.*, 1995; Militello *et al.*, 2007). We, however, were investigating features of trusted advice that occur during independent, brief encounters (often less than 5 min), but involve multiple physicians at the scene. For example, as reported in our CI results, physicians review more than 10 patients in an hour-long inpatient team meeting.

Overall, at the cost of ecological validity, our survey increased the internal validity of our findings and incentivized more participation from physicians than a simulation would. Finally, surveying physicians for multiple scenario variants was useful to prune our design space for important drivers and facilitate the future prototyping of trust-based clinical alerts.

Our survey is limited in two ways. First, we administered the survey only to physicians affiliated with Eskenazi Health. This can be considered as convenient sampling. However, given that Eskenazi Health is one of the leading providers of health care in a major mid-western state in the USA, we expected our sample to be representative of a typical physician population. Second, our survey validated only six out of the eight themes that emerged from our CI. Further work is required to validate all our emerged themes.

In our survey, we explored and found a dichotomy between physicians' likeliness to trust and likeliness to follow clinical advice from peers. For example, we found that respondents were significantly more likely to trust than to follow an infectious disease consultant. Other interesting trends were also observed, but did not reach significance—owing to a limited sample size of our survey respondents. Such a divergence between trust and compliance has significant implications for clinical practice and merits further exploration.

Finally, we did not conduct longitudinal studies to explore how physician–patient relationships change over time and in turn influence advice-sharing among peers. The amount of time for which a physician knew a patient's condition was not factored in our analysis. Furthermore, we did not consider the familiarity among the team members. For example, whether it was the first day for an intern in the team or second year for a resident was not factored in our analysis.

6. CONCLUSIONS

Drug safety alerts intend to assist physicians in the safe prescribing of patient medications via CPOE systems but are frequently ignored (almost 96% daily). Prior design improvements, such as presentational or contextual cues, have failed to mitigate sufficiently this problem. Thus, in this paper, we introduced an alternative approach to alert design: to improve trust between physicians and computerized advice by understanding why physicians trust clinical advice from their peers. To that aim, we conducted three CI studies in an inpatient setting and an online survey. The focus of our CI included the manner of soliciting advice, the manner in which advice was shared and phrased, and dynamics of compliance. Consolidated work models from the CIs uncovered eight major themes. These themes, such as medical hierarchy, empathy, collaborative and inclusive language or specialization, were key drivers that guided physicians in trusting the shared advice.

To further validate the emerging themes, we designed an online survey that operationalized six of our uncovered themes into one or more questions. Based on our findings from the survey and the CI sessions, we proposed seven key design directions for trust-based alerts: alerts that include an expert's endorsement (based on the importance of medical hierarchy and specialization); visibility of physicians' decisions about accepting or overriding alerts (to teach by example); alerts that are 'active' and sensing the team's discussions; use of collaborative and inclusive language in alerts (from critiquing to collaborating with physicians); alerts that would foreground empathy by providing more facts about the patient situation; alerts that could mitigate conflicts arising from different advices from different specialists; and alerts that would accompany individual notes to augment physicians' sense of agency in the decision-making process.

In addition to eliciting more trust in clinical alerts, our proposed design drivers could work in concert to minimize habituation. The trust-based components of an alert could be modified and alternated, thus not repeatedly exposing physicians with alerts that look all the same (Phansalkar *et al.*, 2010). For example, in addition to the generic warning message currently presented in clinical alerts, trust-based alerts will also present individual physicians' notes, more facts about a patient's situation, or endorsement by different specialists—thus helping alerts to be perceived as different stimuli over time and reducing the likelihood of habituation.

In this work, we contributed toward an in-depth understanding of what makes physicians' advice trustworthy in an ecosystem of clinical activities. Our findings can inform fundamental design decisions of clinical alerts to improve physicians' compliance. An important result of our study is the set of design directions for trust-based alerts.

Our future work will focus on exploring the efficacy of our design principles. First, we are using these principles to design iteratively and prototype new types of clinical alerts.

Second, we are planning to evaluate these new alerts in two stages. Controlled empirical studies with high-fidelity prototypes will first evaluate physicians' likelihood of alert compliance and measure their cognitive load. Informed by these preliminary studies, we will refine our designs, and then deploy the trust-based alerts in current CPOE systems. Finally, longitudinal studies monitoring the use of the proposed clinical alerts in hospitals will be set up to validate the long-term effectiveness of trust-based alerts in terms of physicians' adherence, likelihood of habituation and avoidance of adverse drug-related events.

SUPPLEMENTARY MATERIAL

Supplementary material is available at www.iwc.oxfordjournals.org.

ACKNOWLEDGEMENTS

We thank our participants and the attending physicians of our observation sessions, Drs Paul Dexter and Justin More, for their time. We also thank Reecha Bharali for her assistance in data collection.

FUNDING

This research is based upon work supported by the National Science Foundation under Grant Number IIS-1343973. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- Ash, J.S., Sittig, D.F., Campbell, E.M., Guappone, K.P. and Dykstra, R.H. (2007) Some Unintended Consequences of Clinical Decision Support Systems. In AMIA Annu. Symp. Proc., 2007, pp. 26–30. <http://www.ncbi.nlm.nih.gov/pubmed/18693791> (accessed June 5, 2015).
- Becker, M.L., Visser, L.E., van Gelder, T., Hofman, A. and Stricker, B.H.C. (2008) Increasing exposure to drug–drug interactions between 1992 and 2005 in people aged ≥ 55 years. *Drugs Aging*, 25, 145–152. doi:10.2165/00002512-200825020-00006
- Beyer, H. and Holtzblatt, K. (1998) Contextual Design: Defining Customer-centered Systems. Morgan Kaufmann. CA: San Francisco.
- Chaiken, S. and Maheswaran, D. (1994) Heuristic processing can bias systematic processing: effects of source credibility, argument ambiguity, and task importance on attitude judgment. *J. Personal. Soc. Psychol.*, 66, 460–473. doi:10.1037/0022-3514.66.3.460

- Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and Systematic Information Processing within and Beyond the Persuasion Context. In Uleman, J.S. and Bargh, J.A. (eds) *Unintended Thought: Limits of Awareness, Attention, and Control*, pp. 212–252. Guilford Press, New York.
- Chapman, L.J. and Chapman, J.P. (1969) Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *J. Abnorm. Psychol.*, 74, 271–280. doi:10.1037/h0027592
- Chilana, P.K., Wobbrock, J.O. and Ko, A.J. (2010) Understanding Usability Practices in Complex Domains. In Proc. SIGCHI Conf. Human Factors Comput. Syst., pp. 2337–2346. ACM Press, New York, NY.
- Creswell, J.W. and Clark, V.L.P. (2007) *Designing and Conducting Mixed Methods Research*. Sage Publications, Thousand Oaks, CA.
- Doucet, J., Chassagne, P., Trivalle, C., Landrin, I., Pauty, M.D., Kadri, N., Ménard, J.F. and Bercoff, E. (1996) Drug–drug interactions related to hospital admissions in older adults: a prospective study of 1000 patients. *J. Am. Geriatrics Soc.*, 44, 944–948.
- Duke, J.D. and Bolchini, D. (2011) A Successful Model and Visual Design for Creating Context-aware Drug–drug Interaction Alerts. In AMIA Annu. Symp. Proc., 2011, pp. 339–348. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3243201/> (accessed June 5, 2015).
- Duke, J.D., Li, X. and Dexter, P. (2013) Adherence to drug–drug interaction alerts in high-risk patients: a trial of context-enhanced alerting. *J. Am. Med. Inform. Assoc.*, 20, 494–498. doi:10.1136/amiainjnl-2012-001073
- Duke, J.D., Morea, J., Mamlin, B., Martin, D.K., Simonaitis, L., Takesue, B.Y., Dixon, B.E. and Dexter, P.R. (2014) Regenstrief Institute’s Medical Gopher: a next-generation homegrown electronic medical record system. *Int. J. Med. Inform.*, 83, 170–179. doi:10.1016/j.ijmedinf.2013.11.004
- Enkin, M.W. and Jadad, A.R. (1998) Using anecdotal information in evidence-based health care: heresy or necessity? *Ann. Oncol.*, 9, 963–966.
- Eskenazi, Health. (2014) <http://www.eskenazihealth.edu/> (accessed June 5, 2015).
- Fogg, B.J. (2003) *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann, San Francisco, CA.
- Glassman, P.A., Simon, B., Belperio, P. and Lanto, A. (2002) Improving recognition of drug interactions: benefits and barriers to using automated drug alerts. *Medical Care*, 40, 1161–1171.
- Google Forms. (2014) <http://www.google.com/forms/about/> (accessed October 10, 2014).
- Greenhalgh, T. (1999) Narrative based medicine: narrative based medicine in an evidence based world. *Br. Med. J.*, 318, 323–325.
- Guyatt, G. et al. (1992) Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*, 268, 2420–2425.
- Hayward, J., Thomson, F., Milne, H., Buckingham, S., Sheikh, A., Fernando, B., Cresswell, K., Williams, R. and Pinnock, H. (2013) ‘Too much, too late’: mixed methods multi-channel video recording study of computerized decision support systems and GP prescribing. *J. Am. Med. Inform. Assoc.*, 20(e1), e76–e84. doi:10.1136/amiainjnl-2012-001484
- Higgins, E.T., Rholes, W.S. and Jones, C.R. (1976) Category accessibility and impression formation. *J. Exp. Soc. Psychol.*, 13, 141–154. doi:10.1016/S0022-1031(77)80007-3
- Hodge, V.J. and Austin, J. (2004) A survey of outlier detection methodologies. *Artif. Intell. Rev.*, 22, 85–126. doi:10.1007/s10462-004-4304-y
- Holzman, R.S., Cooper, J.B., Gaba, D.M., Philip, J.H., Small, S.D. and Feinstein, D. (1995) Anesthesia crisis resource management: real-life simulation training in operating room crises. *J. Clin. Anesth.*, 7, 675–687. doi:10.1016/0952-8180(95)00146-8
- Hunter, K. (1996) ‘Don’t think zebras’: uncertainty, interpretation, and the place of paradox in clinical education. *Theor. Med.*, 17, 225–241. doi:10.1007/BF00489447
- Jung, M. et al. (2013) Attitude of physicians towards automatic alerting in computerized physician order entry systems. *Methods Inf. Med.*, 52, 99–108.
- Kahneman, D. and Tversky, A. (1972) Subjective probability: a judgment of representativeness. *Cognit. Psychol.*, 3, 430–454. doi:10.1016/0010-0285(72)90016-3
- Krall, M.A. and Sittig, D.F. (2002) Clinician’s Assessments of Outpatient Electronic Medical Record Alert and Reminder Usability and Usefulness Requirements. In AMIA Annu. Symp. Proc., 2002, pp. 400–404. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2244227/> (accessed June 5, 2015).
- Kunda, Z. (1990) The case for motivated reasoning. *Psychol. Bull.*, 108, 480–498. doi:10.1037/0033-2909.108.3.480
- Kunda, Z. (1999) *Social Cognition: Making Sense of People*. MIT Press, Cambridge, MA.
- Kuperman, G.J., Bobb, A., Payne, T.H., Avery, A.J., Gandhi, T.K., Burns, G., Classen, D.C. and Bates, D.W. (2007) Medication-related clinical decision support in computerized provider order entry systems: a review. *J. Am. Med. Inform. Assoc.*, 14, 29–40. doi:10.1197/jamia.M2170
- Laurikkala, J., Juhola, M., Kentala, E., Lavrac, N., Miksch, S. and Kavsek, B. (2000) Informal Identification of Outliers in Medical Data. In Proc. Fifth Int. Workshop Intel. Data Anal. Med. Pharmacol., pp. 20–24. <http://citeseerx.ist.psu.edu/viewdoc/versions?doi=10.1.1.36.7407> (accessed June 5, 2015).

- Lo, H.G., Matheny, M.E., Seger, D.L., Bates, D.W. and Gandhi, T.K. (2009) Impact of non-interruptive medication laboratory monitoring alerts in ambulatory care. *J. Am. Med. Inform. Assoc.*, 16, 66–71. doi:10.1197/jamia.M2687
- Luna, D., Otero, V., Canosa, D., Montenegro, S., Otero, P. and de Quirós, F.G. (2007) Analysis and redesign of a knowledge database for a drug-drug interactions alert system. *Stud. Health Technol. Inform.*, 129, 885–889.
- Magro, L., Moretti, U. and Leone, R. (2012) Epidemiology and characteristics of adverse drug reactions caused by drug–drug interactions. *Expert Opin. Drug Saf.*, 11, 83–94. doi:10.1517/14740338.2012.631910
- Militello, L.G., Patterson, E.S., Bowman, L. and Wears, R. (2007) Information flow during crisis management: challenges to coordination in the emergency operations center. *Cogn. Technol. Work*, 9, 25–31. doi:10.1007/s10111-006-0059-3
- Newell, A. and Simon, H.A. (1972) *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ.
- Nightingale, P.G., Adu, D., Richards, N.T. and Peters, M. (2000) Implementation of rules based computerised bedside prescribing and administration: intervention study. *BMJ*, 320, 750–753. doi:10.1136/bmj.320.7237.750
- Nisbett, R.E., Krantz, D.H., Jepson, C. and Kunda, Z. (1983) The use of statistical heuristics in everyday inductive reasoning. *Psychol. Rev.*, 90, 339–363. doi:10.1037/0033-295X.90.4.339
- Obreli-Neto, P.R. et al. (2012) Incidence and predictors of adverse drug reactions caused by drug–drug interactions in elderly outpatients: a prospective cohort study. *J. Pharm. Pharmaceut. Sci.*, 15, 332–343.
- Paterno, M.D., Maviglia, S.M., Gorman, P.N., Seger, D.L., Yoshida, E., Seger, A.C., Bates, D.W. and Gandhi, T.K. (2009) Tiering drug–drug interaction alerts by severity increases compliance rates. *J. Am. Med. Inform. Assoc.*, 16, 40–46. doi:10.1197/jamia.M2808
- Payne, T.H., Nichol, W.P., Hoey, P. and Savarino, J. (2002) Characteristics and Override Rates of Order Checks in a Practitioner Order Entry System. In *AMIA Annu. Symp. Proc.*, 2002, 602–606. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2244252/> (accessed June 5, 2015).
- Phansalkar, S., Edworthy, J., Hellier, E., Seger, D.L., Schedlbauer, A., Avery, A.J. and Bates, D.W. (2010) A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. *J. Am. Med. Inform. Assoc.*, 17, 493–501. doi:10.1136/jamia.2010.005264
- Richards, L. (1999) Data alive! The thinking behind NVivo. *Qual. Health Res.*, 9, 412–428. doi:10.1177/104973239900900310
- Sackett, D.L., Rosenberg, W., Gray, J.A., Haynes, R.B. and Richardson, W.S. (1996) Evidence based medicine: what it is and what it isn't. *BMJ*, 312, 71–72. doi:10.1136/bmj.312.7023.71
- Schedlbauer, A., Prasad, V., Mulvaney, C., Phansalkar, S., Stanton, W., Bates, D.W. and Avery, A.J. (2009) What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior? *J. Am. Med. Inform. Assoc.*, 16, 531–538. doi:10.1197/jamia.M2910
- Seigel, A.F. (1988) *Statistics and Data Analysis: An Introduction*. John Wiley & Sons Inc., New York.
- Shah, N.R., Seger, A.C., Seger, D.L., Fiskio, J.M., Kuperman, G.J., Blumenfeld, B., Recklet, E.G., Bates, D.W. and Gandhi, T.K. (2006) Improving acceptance of computerized prescribing alerts in ambulatory care. *J. Am. Med. Inform. Assoc.*, 13, 5–11. doi:10.1197/jamia.M1868
- Sittig, D.F., Krall, M.A., Dykstra, R.H., Russell, A. and Chin, H.L. (2006) A survey of factors affecting clinician acceptance of clinical decision support. *BMC Med. Inform. Decis. Making*, 6, 6. doi:10.1186/1472-6947-6-6
- Taylor, S.E. (1991) Asymmetrical effects of positive and negative events: the mobilization–minimization hypothesis. *Psychol. Bull.*, 110, 67–85. doi:10.1037/0033-2909.110.1.67
- Van Der, Sijs, Vulto, A. and Berg, M. (2006) Overriding of drug safety alerts in computerized physician order entry. *J. Am. Med. Inform. Assoc.*, 13, 138–147. doi:10.1197/jamia.M1809
- Weingart, S.N., Toth, M., Sands, D.Z., Aronson, M.D., Davis, R.B. and Phillips, R.S. (2003) Physicians' decisions to override computerized drug alerts in primary care. *Archiv. Int. Med.*, 163, 2625–2631. doi:10.1001/archinte.163.21.2625
- Weingart, S.N., Massagli, M., Cyruлик, A., Isaac, T., Morway, L., Sands, D.Z. and Weissman, J.S. (2009) Assessing the value of electronic prescribing in ambulatory care: a focus group study. *Int. J. Med. Inform.*, 78, 571–578. doi:10.1016/j.ijmedinf.2009.03.007
- Winkelman, W.J., Leonard, K.J. and Rossos, P.G. (2005) Patient-perceived usefulness of online electronic medical records: employing grounded theory in the development of information and communication technologies for use by patients living with chronic illness. *J. Am. Med. Inform. Assoc.*, 12, 306–314. doi:10.1197/jamia.M1712
- Zheng, K., Fear, K., Chaffee, B.W., Zimmerman, C.R., Karls, E.M., Gatwood, J.D., Stevenson, J.G. and Pearlman, M.D. (2011) Development and validation of a survey instrument for assessing prescribers' perception of computerized drug–drug interaction alerts. *J. Am. Med. Inform. Assoc.*, 18(Suppl. 1), i51–i61. doi:10.1136/amiajnl-2010-000053