

Defeating information overload in health surveillance using a metacognitive aid innovation from military combat systems

Journal of Defense Modeling and Simulation: Applications, Methodology, Technology



2017, Vol. 14(4) 371–388

© The Author(s) 2016

DOI: 10.1177/1548512916667246

journals.sagepub.com/home/dms



Daniel C McFarlane¹, Alexa K Doig², James A Agutter³, Jonathan L Mercurio⁴, Ranjeev Mittu⁵, Lara M Brewer⁶, and Noah D Syroid⁷

Abstract

Modern sensors for health surveillance generate high volumes and rates of data that currently overwhelm operational decision-makers. These data are collected with the intention of enabling front-line clinicians to make effective clinical judgments. Ironically, prior human–systems integration (HSI) studies show that the flood of data degrades rather than aids decision-making performance. Health surveillance operations can focus on aggregate changes to population health or on the status of individual people. In the case of clinical monitoring, medical device alarms currently create an information overload situation for front-line clinical workers, such as hospital nurses. Consequently, alarms are often missed or ignored, and an impending patient adverse event may not be recognized in time to prevent crisis. One innovation used to improve decision making in areas of data-rich environments is the Human Alerting and Interruption Logistics (HAIL) technology, which was originally sponsored by the US Office of Naval Research. HAIL delivers metacognitive HSI services that empower end-users to quickly triage interruptions and dynamically manage their multitasking. HAIL informed our development of an experimental prototype that provides a set of context-enabled alarm notification services (without automated alarm filtering) to support users' metacognition for information triage. This application is called HAIL Clinical Alarm Triage (HAIL-CAT) and was designed and implemented on a smartwatch to support the mobile multitasking of hospital nurses. An empirical study was conducted in a 20-bed virtual hospital with high-fidelity patient simulators. Four teams of four registered nurses (16 in total) participated in a 180-minute simulated patient care scenario. Each nurse was assigned responsibility to care for five simulated patients and high rates of simulated health surveillance data were available from patient monitors, infusion pumps, and a call light system. Thirty alarms per nurse were generated in each 90-minute segment of the data collection sessions, only three of which were clinically important alarms. The within-subjects experimental design included a treatment condition where the nurses used HAIL-CAT on a smartwatch to triage and manage alarms and a control condition without the smartwatch. The results show that, when using the smartwatch, nurses responded three times faster to clinically important and actionable alarms. An analysis of nurse performance also shows no negative effects on their other duties. Subjective results show favorable opinions about utility, usability, training requirement, and adoptability. These positive findings suggest the potential for the HAIL HSI system to be transferrable to the domain of health surveillance to achieve the currently unrealized potential utility of high-volume data.

Keywords

Nurse cognition, multitasking, metacognition, situation awareness, human interruption, task management, wearable, attention aid, alarm fatigue, alarm safety, context, alarm notification, information overload, human–systems integration, patient simulation, clinical simulation, smartwatch, decision aid

¹Royal Philips, Patient Care & Monitoring Solutions, USA

²University of Utah, College of Nursing, USA

³University of Utah, College of Architecture & Planning, USA

⁴Lockheed Martin Advanced Technology Laboratories, USA

⁵US Naval Research Laboratory, Information Technology Division, USA

⁶University of Utah Health Sciences Center, Department of Anesthesiology, USA

⁷University of Utah, Anesthesiology Center for Patient Simulation, USA

Corresponding author:

Daniel C McFarlane, 3000 Minuteman Road, Andover, MA, 01810, USA.

Email: mcfarlane@acm.org

I. Background and problem

Field experience shows that developing human–systems integration (HSI) solutions for health surveillance operations is an extremely difficult problem. This challenge is at the intersection of high-stakes decision making, complex high-volume high-rate data, and complex human operator roles. Front-line operators are responsible for performing three very different types of multitasks simultaneously. They must: (1) *track* the progress of activities, people, or groups previously marked as “interesting”; (2) *scan* high-volume multi-source data for new things of interest; and (3) *manage* the dynamic allocation of limited surveillance resources, including the operator’s own cognitive attention. High-volume and high-rate data may describe a complex and rapidly changing situation, therefore a surveillance operator must maintain his/her focus of attention on what is “important”, as this changes over time. The things that seem “important” now can change. In addition, some things that are currently unimportant may become “interesting.” Operators have human cognitive limitations for attention, and it is easy for them to become overwhelmed with the track + scan + manage triple activity of surveillance operations.

1.1 Generalizing human–systems integration innovations from other surveillance domains

Examples of domains that require track + scan + manage surveillance by human operators include the following: commercial aviation control; commercial aviation cockpit control; nuclear power plant control; military intelligence operations; cyber defense operations; insider threat monitoring; maritime domain awareness; military anti-air warfare (AA); etc. In each of these domains, operators must organize their own cognitive multitask balancing to accomplish their track + scan + manage responsibilities. The data and knowledge requirements are different across domains, but the metacognitive work of balancing limited cognitive attention resources is similar.

Despite early enthusiasm for the potential to use fully automated methods to pre-process large internet-derived data sets for health surveillance, the problem is somewhat difficult, due in large part to noisy input data.¹ A review of studies that exploited internet use and search trends to monitor influenza and dengue shows that internet-based surveillance systems have not demonstrated the capacity to replace traditional surveillance systems and are best viewed as an extension to traditional systems.² These results suggest that a useful health surveillance system will likely need to integrate with healthcare providers and their existing tools.

Scientific literature highlights successful HSI technical innovations that have improved performance in

surveillance operations for other (non-health) domains. HSI improvements that support operators’ track + scan + manage multitask workflow could potentially be generalizable to health surveillance operations. The US Navy’s Human Alerting and Interruption Logistics (HAIL) technology is a candidate for useful generalization to health surveillance. The HAIL project started as domain-agnostic basic research on metacognitive aiding to support interruption of operators with alarms/alerts during human-computer interaction (HCI).

Supporting operator metacognition for triaging interruptions was recognized as a technology shortfall for military surveillance operations such as AA.^{3–5} The HAIL basic research produced a foundational interdisciplinary theory of human interruption that describes the metacognitive processes of human interruption during surveillance operations.⁶ Laboratory research with human subjects discovered a breakthrough technology for metacognitive aiding called “negotiation-based coordination.”^{7–9} HAIL was first produced as a mature domain-independent technology innovation, and then applied to AA surveillance operations to improve operator performance under frequent alarm/alert interruptions. This AA application of HAIL was transitioned into the US Navy’s Aegis Weapon System (HAIL-AEGIS) and is now supporting surveillance operations on the majority of US Navy cruisers and destroyers.^{10,11}

Many cognitive activities are similar for operators between healthcare surveillance and AA surveillance operations. Operators from both domains must maintain situation awareness (SA) while performing rapid life-critical multitasking in dense information environments with frequent interruptions. Hospital clinicians performing health surveillance for individual patients have some differences to AA operations. The complexity of multitasking is somewhat higher in health surveillance, while the rate of interruption is somewhat higher in AA surveillance. Another key difference is that, unlike AA operators who mostly work seated at fixed consoles, hospital clinicians’ work is highly mobile. A design effort was conducted to address these differences and expand a HAIL application for healthcare surveillance.

1.1.1 “Track-while-scan”.

Automation can support the track + scan + manage HSI workflow. For example, military radar sensors support AA surveillance operations by automatically balancing the use of limited radar resources to collect data for both scan and track. This radar mode, called “track-while-scan,” facilitates collecting data about both known aircraft (“track”) and discovering new aircraft (“scan”).¹² A dynamic resource allocation algorithm runs as a separate third process (the management, or “while” work). It considers the changing priorities and changing

situation and re-plans a scheduled use of the shared resource. Performance of the management (or “while”) function is especially important to mission success. Poor management of limited radar assets for “track-while-scan” can result in loss of previously seen objects and/or failure to detect new objects.

This automated “track-while-scan” function for data collection is the same pattern as the human track + scan + manage workflow for using the data. Both provide a dynamically balanced use of a limited resource (radar energy or human cognitive attention) for accomplishing multiple simultaneous activities. Both can usefully be called “track-while-scan”. In human military command and control (C2) team workflow, inadequate management of cognitive attention resources in human “track-while-scan” activities can result in failed human performance of scheduled tasks and/or failure to notice and respond to important change. In this terminology usage, HAIL provides a set of “track-while-scan” HSI metacognition aiding services that improve human operators’ multitask balancing performance. Like radar automation, HAIL helps human surveillance operators “track,” “scan,” and manage (“while”).

Field studies from many domains report that workers’ attentional resources during “track-while-scan” activities can be easily overwhelmed.⁵ For example, answering a cell phone call while driving can overcommit drivers’ cognitive attentional resources and negatively impact their performance on “track-while-scan” responsibilities. Drivers can make “track” errors (e.g., miss a turn) and/or “scan” errors (e.g., fail to notice the car in front stop unexpectedly). In field studies, practitioners from multiple different domains complain about alarm designs that indiscriminately grab users’ attention resources to make her/him attend to an alarm (the “scan” function). This “grab users’ attention” design approach would only be appropriate in the rare case where the occurrence of a detected event is obviously more important than anything else the human practitioner may be doing. In hospital nursing, this could perhaps be a valid “code” (a confirmed life-threatening patient crisis). In the military C2 example, this might be a confirmed launch of an inbound enemy missile.

1.1.2 Leveraging HAIL theory and analysis methods. The “track-while-scan” HSI analogy can be used as a framework to analyze HSI problems for health surveillance operations. The HAIL theory base frames surveillance operations HSI problems in terms of the many different multitasks that human operators have to manage concurrently. It also highlights the importance and complexity of the operators’ meta-level task of continuously revising their cognitive multitask plan. Only through constant effort are people able to maintain their focus of attention on what

is important as this changes. In addition, this meta-level work itself consumes a portion of the person’s available cognitive attention resources. An analysis of the HSI design challenge for surveillance operation from this perspective emphasizes the complexities of dynamic resource allocation of human cognitive resources. The design must also consider the scope of individual differences and learning effects for cognitive multitasking skills across the user population.

HAIL methods for analysis of a HSI problem enable researchers to build a model of the operators’ cognitive and metacognitive workload and workflows. For example, the occurrence of a new automatically generated alert may not at first glance seem like a complicated HSI problem. A naive device-centric HSI perspective may mistakenly put requirements on the user to change his/her workflow when receiving alerts –“The operator needs to immediately respond to the alert and investigate.” In practice, however, this situation can be extremely complex and simplistic approaches to delivery of alerts often fail.⁵ A better approach is to specify requirements of the system design to support operators’ “track-while-scan” responsibilities.

Analysis of fieldwork has shown that a model of operators’ mental state at the time of an alert can reveal that people’s cognition is fully engaged. At any point in time, operators are likely involved in a complex multitask that is taxing their cognitive limits for tracking, scanning, and meta-level managing. Their cognitive working memories would be stretched to capacity trying to remember the details of each of the many multitasks on their cognitive stack as they work sequentially through a multitask plan. Within the context of the user’s current cognitive resource allocation, a new alert is received as an interruption. It breaks into this complex cognitive situation and demands use of the person’s cognitive attention resources. He/she may have only a tiny amount of this resource that is not already committed. From the operator’s point of view, he/she needs to accomplish a quick triage of the alert using minimal cognitive attention resources. He/she needs to judge the alert’s importance relative to everything else he/she is cognitively holding onto, and then decide how/whether to fit it into his/her multitasking plan. It’s almost never a simple prospect for an operator to instantly drop everything else and focus 100% of his/her attention on the topic of the alert.

The HAIL interdisciplinary theory base shows that operators use their metacognitive resources to perform interruption triage. Some interruptions signal important change that must be addressed. The majority, however, are typically not actionable and do not merit any change in multitask scheduling. HAIL delivers metacognitive HSI services that empower users to quickly triage interruptions and dynamically manage their multitasking. This helps teams remain flexible to unexpected changes while

protecting their primary multitask performance from possible distraction and error. HAIL is founded on basic HSI research that models how human operators use metacognition to triage cognitive multitasking. HAIL services aid operator metacognition in the following ways: (1) minimizing the cost (effort, time, risk) of triaging interruptions (alarms/alerts, etc.) as they happen; and (2) maximizing the potential value of alerts with flexible alarm generation automation.

An analysis of these costs and values of the data available for surveillance operations can reveal problems that negatively affect operators' ability to manage their attention resources. The costs of using data may be overwhelmingly high, and/or low average value of data streams (including alerts) may be unusably low. Together with a model of operators' cognitive resource allocation for their "track-while-scan" workflow, an analysis of the value/cost ratio of data can inform the design of HSI solutions and HCI decision aids.

1.1.3 Leveraging HAIL technologies. The HAIL negotiation-based coordination technology is a set of interactive HCI services that support surveillance operators' "track-while-scan" HSI needs. A person's internal dynamic resource allocation of their cognitive attention resources must consider the domain situation changes represented in the data stream. The HAIL basic research discovered a breakthrough that operators have the cognitive capacity and valid need to be involved in triaging change.

Artificial intelligence (AI) algorithms have shown proficiency in pattern-matching for known patterns of interest. They have also been applied successfully to recognize novel situations. In practice, however, there is usually a very high occurrence of "strange" situations that were not foreseen when the AI was developed, and only a tiny fraction of these are actually important to the surveillance mission. In contrast, people have much superior capabilities for assigning the meaning of previously unknown "strange" patterns and determining which are of interest. This capability allows people to deliver superior prioritization of multitask handling of changing situations compared to full automation. Existing HSI designs often do not realize this potential because either (1) AI pattern-matching is not exploited well and people are overwhelmed with responsibilities to process high-volume raw data that could better be done by automation or (2) AI is inappropriately engaged as full automation to apply meaning and prioritization to the host of unexpected "strange" occurrences that happen during operations.

Negotiation-based coordination supports users' needs to participate in the triage of change and prioritize their multitask plan. Designing a computer-based aid to deliver negotiation-based coordination is challenging because

users must work within a highly dynamic environment with constantly changing task loading. HAIL provides services for the following: (a) announcing alerts with context packaging; (b) flexible control of AI pattern-matching according to known policy; (c) user-controlled triage of alerts and other interruptions; and (d) multitask context recovery.

1.1.4 Leveraging HAIL evaluation methods. The HAIL focus on operators' metacognitive dynamic attention allocation is useful for identifying the key evaluation metrics. HSI metrics need to be at a systems level and not focused on niche concerns about individual components.¹³ However, since human cognition and metacognition are impractically difficult to measure directly, it is important to find a key observable external behavior for measuring performance. If an operator is performing well on their "track-while-scan" responsibilities, they will quickly notice "important" changes in the situation and environment and thus can prioritize their response. Under realistic heavy workload and complex high-volume data with wildly mixed relevancy, operators should still be able to quickly recognize "important" changes and respond. An ineffective HSI system will not shelter users from becoming overwhelmed by a high percentage of irrelevant data. HSI system design must result in a favorable "value/cost" ratio, where the value of data is improved through preprocessing, and the cost of triaging is minimized.

1.2 Health surveillance operations by hospital nurses

Health surveillance operations focus on individual patients in hospital and attempt to track their health status over time. Hospital nurses working in acute care settings (e.g., medical, surgical, orthopedic, cardiac, and other non-intensive care unit areas) are responsible for maintaining global SA while attending to the needs of multiple patients.¹⁴ SA is the activity of maintaining a continual cognitive awareness of all dynamically changing aspects of the work environment that are relevant to making progress toward work objectives. In the USA, acute care nurses typically care for between four and six patients during the course of an eight or 12-hour shift, with the panel changing as patients are discharged and admitted to the unit or ward. Nurses perform mobile work and multitask among the different patient rooms and other work areas, although physically attending to only one patient at a time.^{15–18} Ebright and colleagues¹⁹ identify this organizational work strategy as "stacking," whereby nurses may leave one task incomplete due to an interruption or other interference (e.g., waiting for missing information) and move on to the next one in order to optimize time and to increase the likelihood of completing other tasks during the shift.

The cognitive work of the nurse is demanding and often unpredictable.^{19–22} At the start of a shift, nurses may plan the sequence of care for their patients. However, in practice there are a myriad of unplanned events and interruptions that must be attended to and incorporated into their multitasking. Most critical are acute changes in a patient's status, including physiological deterioration that can lead to the onset of life-threatening crises. Therefore, as nurses execute their planned work with individual patients or perform focused tasks such as medication administration, they must also stay aware and apprised of all clinically "important" changes in patient status (i.e., continuously tracking-while-scanning) and be ready to dynamically revise their multitask scheduling across all patients (i.e., manage their surveillance resources).^{19,23} Because of the ubiquitous nature of nursing work, interruptions, and the sheer number of short-duration nursing tasks, nursing workflow is rarely a liner process.²⁴ A nurse must be able to perform a sequence of non-sequential mobile tasks and maintain vigilance while simultaneously performing the meta-level work of dynamically revising their multitasking schedule – a "track-while-scan" workflow.

1.3 The alarm safety problem for healthcare

Many types of medical devices, including physiological monitors and infusion pumps, are designed to generate alarms that purposefully "require" immediate nurse awareness and/or intervention at the device's bedside location.^{25,26} These clinical alarms, which primarily involve an audible signal, are important for detecting acute changes in physiological state (e.g., hypoxemia detected by pulse oximetry), but can be a source of interruptions to nurses' workflow.^{25–27} The rate of alarm signals in a hospital (including all medical devices) is, on average, 350 alarms per patient per day – equivalent to one alarm approximately every 4 minutes per patient.^{27,28} At this rate, a nurse with four patients will experience an alarm approximately every minute. This interferes with their multitasking workflow, breaks SA, and can cause harmful adverse events for patients.^{29,30} Healthcare workers (and researchers) call this problem "alarm fatigue."³¹

1.3.1 Information overload due to alarms. Information overload due to the high frequency of clinical alarms is compounded by the fact that only about 5–15% are clinically significant.^{25,31–33} The other 85–95% of alarms are false signals, signal noise, and artifact, or are clinically unimportant. The nurse must first hear the alarm (which means they must be in the vicinity), enter the patient room, and then because many alarms are not self-correcting, silence or turn off the alarm. At this point they are required to spend time in the room acquiring information to determine

whether it was clinically important. If the vast majority of alarms are not important or clinically significant, then the nurse must either expend considerable time going to the patient's bedside to determine whether or not each alarm is actionable, or as several observational studies have shown, may end up ignoring most alarms, some of which may be signaling a life-threatening event, such as respiratory depression or hypotension. Due to the lack of auditory differentiation among different devices and patients, a nurse in the hallway or at the nursing station often cannot tell what alarm is sounding,³⁴ or whether an alarm is for one of their patients or not, and may bet on the probability that it is not theirs. Between 2005 and 2010 the Food & Drug Administration (FDA) Manufacturer and User Facility Device Experience (MAUDE) database identified 216 deaths related to medical alarms, one third of which were related to nursing staff either not hearing the alarm, inappropriately silencing or turning off the alarm, or not responding for other reasons.³⁵

The time and effort needed to check on a new alarm leads to slower task performance³⁶ and increases cognitive disruption, which is associated with an increase in the rate of medication errors.^{37,38} The constant distraction and interruption of nursing workflow by alarms reduces the amount of focused time nurses need for critical thinking work (e.g., identifying subtle changes in patient vital signs and laboratory results, then concluding that a patient is developing sepsis).²⁴ The cost of checking individual alarms with high frequency makes it near impossible for nurses to use audible alarm signals as a basis for clinical decision making and task prioritization. This situation conspires to make the current value/cost ratio for alarms unusually low.

In addition to interrupting workflow, alarms generate auditory sounds (beeps) emanating from the device at the bedside where the patient and family members are forced to listen. In many instances, the nurse is not immediately in a location to hear it. As a result, patients and family members may be bothered and/or distressed by the beeping, although there is little research to quantify this negative effects of alarms^{39,40}

1.3.2 Solutions to reduce information overload due to alarms. A range of solutions to the problem of information overload due to clinical alarms have been proposed by the Association for the Advancement of Medical Instrumentation (AAMI), including improving device standards, institutional policies, and the national regulatory environment.²⁷ System-level solutions for improving alarm management include efforts to decrease the overall frequency of alarms (especially false alarms), the development of "smart alarms" that have delayed audible signals and self-correcting abilities so that the audible signal is

silenced when the patient state returns to normal, and tailoring alarm threshold values not customized for the patient.^{27,41} An additional and complimentary solution involves developing a remote surveillance system where relevant alarm signals are delivered directly to the nurse using a mobile device (pager, wireless phone, mobile phones, tablet, or smartwatch) so that the nurse: (1) is aware that an alarm in another location is sounding; (2) knows which device is alarming; and (3) does not have to expend the cognitive, physical, or time costs of going to the patient's room if the alarm is non-actionable. In order for this system to work optimally, several functions must be in place, including providing enough contextual data so that the nurse can triage alarms correctly (i.e., differentiate between clinically important and clinically unimportant or false signals) and the ability for the nurse to silence or reset the alarm from the mobile device.

2. Aim and hypothesis

The goal of this research was to determine if HAIL methods, theory, and technologies could be used to improve operator performance in health surveillance operations, especially under high volumes of data and high rates of alert interruption. HAIL services aid operator metacognition by (1) minimizing the cost (effort, time, risk) of triaging interruptions (alerts and notifications) as they happen and (2) maximizing the potential value of alerts with flexible alarm generation automation. This study investigates whether this HSI innovation could be leveraged for analyzing and explaining the HSI challenges for operators' "track-while-scan" workflow, developing interactive services to support operator multitasking and articulating key evaluation metrics.

A wearable prototype was designed and created as an experimental platform that delivers HSI services to nurses. This prototype leverages HAIL methods and is called the "HAIL Clinical Alarm Triage" application (HAIL-CAT). It enables tracking the status of between four and six patients simultaneously, as well as triaging clinical patient alarms (and other alerts) with the contextual information required to recognize important change amid a high volume of non-actionable data. The hypothesis was that the introduction of this HAIL-inspired wearable metacognitive aid for health surveillance would increase nurse performance on their "scanning" responsibilities, while not negatively impacting nurse performance on their "tracking" multitasking responsibilities. This hypothesis focuses experimental design primarily on the central "scanning" concerns for defeating alarm fatigue and improving health surveillance. It is possible that the HAIL-CAT prototype could also improve nurses' performance on their critically important "tracking" responsibilities. However, a single

focus for this study was selected to maximize internal validity and the statistical power of the results and to simplify the experiment design. More detailed exploration of the HAIL-CAT effects on "tracking" performance is indicated for a future study.

An empirical investigation was conducted on 20–21 January 2015 to assess the potential for leveraging the HAIL HSI innovation to improve nurse performance for health surveillance operations under heavy workload conditions. The newly designed HAIL-CAT wearable prototype (informed by HAIL theory, methods, and technology) delivered a set of mobile, context-enabled alarm notification services to support nurses' metacognition for clinical alarm triage.

3. Design of the Human Alerting and Interruption Logistics Clinical Alarm Triage user interface

The goal for the design of the user interface (UI) for the HAIL-CAT was to present each nurse end-user with two types of interactive services relative to each patient: (1) an alerting service for delivery of integrated alarms from bedside monitors, infusion pumps (including a pump for patient-controlled analgesia), and call light system events; and (2) contextual information about the patient. Additional literature shows the potential for intelligent mobile routing of alerts to reduce the costs of checking⁴² and integrated context information in the display to improve nurse awareness of patients' changing status.^{15,43}

Following a human-centered design protocol, the interdisciplinary research team utilized a design and development process to develop and evaluate a new interface for clinical alarms presented on a mobile and optimally wearable device. This five-step process, which is described in more detail below, consisted of the following: (1) systematic research to gain insight into the problem of information overload due to the high volume and rate of clinical alarms; (2) development of ideas to map interface components with results of these insights; (3) refinement of these concepts based on upon direct input from end-users; (4) implementation of concepts into a workable prototype; and (5) a rigorous evaluation with end-users.

Firstly, the team conducted a systematic literature review to understand the past and current work being conducted in wearable technology development and the current literature on workflow for nurses and alarm management strategies in a hospital environment. In addition, the team conducted semi-structured interviews with three nurses at a large urban hospital to understand their wants and needs, and the workflow of nurses surrounding response to clinical alarms. It was during this part of the design process that the newly available smartwatch was

chosen as the platform for the alarm triage system, since the majority of nurses interviewed said that a smartphone in their pocket would be difficult to check with gloves on and that attaching a smartphone to their forearm would be heavy and not visible when wearing a medical gown.

Information from the literature review and interviews was used to develop a series of interface design options that could deliver alarm information on the small screen of a smartwatch, and would be intuitive and useful for nurses. Four graphic concepts were developed and then shown on a computer screen to five different clinical experts. These experts provided feedback about the utility and desirability of these concepts and gave input on what additional features would be needed. This information was then compiled and utilized to guide the next step of refinement.

This user feedback and our expert understanding of UI requirements informed selection and refinement of one of the design concept options. A process of iterative design with rapid user feedback was used to refine this final concept and develop a design prototype. In this process, the prototype was presented on paper to nurses for feedback. Based on this input, the various screen prototypes went through another round of revision. In each revision, the screen prototypes were shown to two or three experts. This prototype was revised four times until the design feedback reached convergence.

The focus of the design work was on creating a research prototype to investigate nurse “scanning” performance, not to develop a fully deployable commercial product. The iterative design process dealt with multiple important practical issues related to patient safety and the technical feasibility of deploying a reliable solution. One safety topic of special concern for the design was the need to avoid

spreading infection between patients. The form factor of a wrist watch raised the question of potential infection control issues. Iterative design cycles and pilot testing with registered nurses (RNs), however, confirmed that most nurses currently wear watches. Introducing a smartwatch platform for use by nurses in most types of hospital units was judged to probably not introduce any additional safety risk to infection control.

The final design for the smartwatch-based HAIL-CAT (Figure 1) presents alarm information in four screens that are easily navigable by nurses. If there are no active alarms, the nurse is presented with a home screen with a list of that shift’s patients. In our experiment, nurses were given five patients. From this home screen nurses can select a patient and be presented with a list of alarms and other notifications that are current for that particular patient. Selecting one alarm presents the nurse with the details of that particular alarm and contextual vital sign information about the patient. This provides nurses with relevant information about the broader context of the alarm in order to help prioritize their attention and efforts. If one or more alarms are active, the home screen becomes a list of alarms (for all patients) that the nurse can quickly click on and review with the aforementioned contextual information.

The design of the UI allows nurses to attend to their planned multitasking (i.e., assessing patients, administering medications, coordinating care, etc.) while efficiently triaging alarms and mitigating the potentially negative effects of the alarm’s “interruption.” When an alarm occurs, the nurse is alerted via a short, non-obtrusive vibration on their wrist. The nurses are able to see which patient and which alarm is activated, the patient’s room, and vital signs, including heart rate (HR), systolic blood pressure

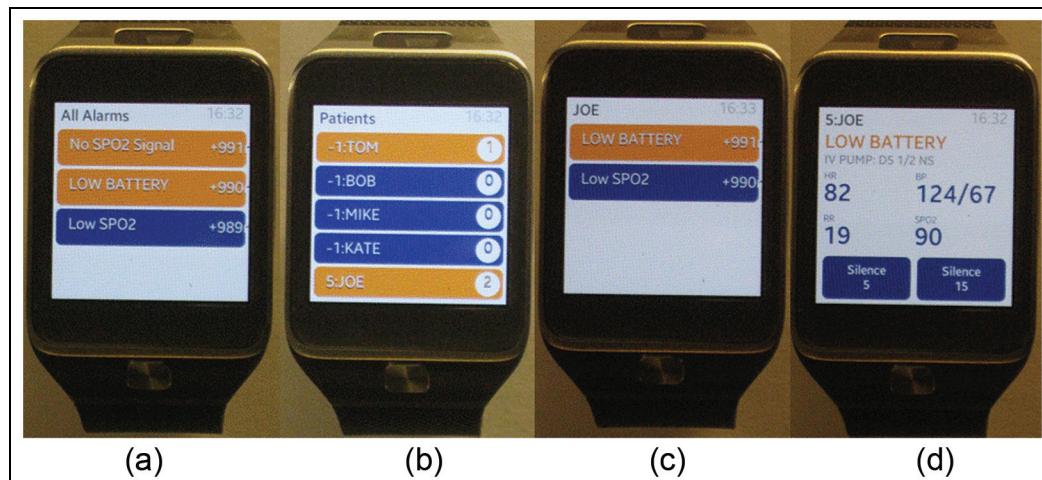


Figure 1. Final user interface design of the Human Alerting and Interruption Logistics – Clinical Alarm Triage (HAIL-CAT) for a smartwatch: (a) “All Alarms List”; (b) “Home – Patient List”; (c) “Individual-Patient-List”; (d) “Individual-Alarm.”

(SBP), diastolic blood pressure (DBP), respiratory rate (RR), and SpO₂ (oxygen saturation of hemoglobin measured by pulse oximetry). The nurse is also presented with the option to silence the alarm for a short period of time or dismiss the alarm.

4. Research methods

4.1 Research design

This investigation uses a variant of a “clinical trial” evaluation design in which an experimentally controlled intervention is introduced into a clinical work setting with a clinician participant, but does not expose human patients to any risk. The method is a “clinical trial in patient simulation.” Participating clinicians perform highly realistic multitasking within a full-scale simulated clinical unit, but with patient mannequins instead of real people. The experiment for this study was conducted in a 20-bed patient simulation laboratory with teams of RN volunteers. This “clinical trial in patient simulation” method provides the benefits of experimental control and external validity like in a traditional clinical trial. It allows within-subjects repeated measures experimental designs that are extremely useful for clinical study because there is a very great diversity of workflow pattern and experience among nurses. Also, like an observational study, it has the additional benefit of not exposing human patients to risk. Further, it has the benefit of controlled repeatability that dramatically increases the statistical power of the analyses of results – something that neither a traditional clinical trial method nor observational study support.

A within-subjects clinical trial in patient simulation experiment was conducted in a 20-bed simulation laboratory that is a replica of an acute care unit or ward in a modern hospital. Four teams of four nurses each (16 total nurse participants) were enrolled in this experiment. Each nurse was assigned responsibility to care for five simulated patients located at random across the 20 beds in the simulated hospital unit.

There were two treatment conditions: an experimental treatment that included participants wearing and using the HAIL-CAT on a smartwatch and a control condition with no wearable aid. The four nurses within each session team received conditions in the same order, but the order of conditions across teams was randomized and balanced. The 180-minute clinical patient care scenario used to evaluate the effects of the new HAIL-CAT prototype on nurse performance had two 90-minute sessions to simulate two sections of a single day. The first part of the scenario simulated the start of a traditional nursing day shift (clock time 8–9:30 a.m.). The second part of the scenario simulated these same 20 patients toward the end of the shift on the same day (4:30–6 p.m.).

4.2 Participants

Sixteen licensed RNs (15 female, 1 male) actively working at one of seven hospitals within the Salt Lake City region of Utah, USA, participated in the study. All nurse participants reported currently working in acute care, critical care, or emergency department settings, and had experience caring for between four and six patients. The median years of experience working as a RN in a hospital setting was 6.5 years (range from 0.75 to 16 years). All 16 RNs reported their attitude toward technology as either “positive” or “very positive.”

4.3 Outcomes and measures

The “scan” metric of nurse performance was “how well nurses could dynamically use the information streams to recognize and respond to important change in their patients’ statuses.” This was operationalized for the experiment as the response time for nurses to arrive in the affected patient’s room after an auto-generated alert/alarm about an important change in the patient’s status. For the purposes of this experiment, the nurse had to go to the bedside in order to cancel the alarm (rather than having the ability to remotely dismiss the alarm), so that the response time for important and non-important alarms could be compared.

The “track” metric of nurse performance was “how well nurses performed their heavy assignment of pre-planned shift tasks.” This was operationalized by observation from the two conditions of the overall time that nurse participants took to complete their assigned tasks (assessment, medication administration, clinical procedures, and healthcare coordination) and noise level from active alarms. The experiment tests whether use of the HAIL-CAT prototype improves nurse “scan” performance without negatively impacting their “track” performance.

4.4 Design of the simulation scenario for the experiment

A two-part 180-minute scenario was designed with realistic patient characteristics and simulated clinical data. As described previously, nurses cared for their assigned patients during the first 90 minutes of the simulation, which represented 8–9:30 a.m., then paused and switched experimental conditions (HAIL-CAT versus control) and continued care for their same assignment patients, although the timeframe moved ahead to 4–6:30 p.m. The morning session involved administering morning medications and a predefined set of nurse tasks and the afternoon session included a late afternoon set of medications and another predefined set of nurse tasks.

Table 1. Five simulated patient archetypes with types of actionable and non-actionable alarms.

Patient summary and simulated risk	Simulated actionable changes trigger alarms	Simulated non-actionable changes trigger alarms
Contagious disease and recent surgery (clinically inferable risk of sepsis infection)	Simulated onset of sepsis causes multiple physiological changes, including triggering alarm for low blood pressure	Unimportant status changes to infusion pump device, and others
Heart disease (clinically inferable risk of heart failure)	None	Unimportant alarms related to recurring low blood pressure, and others
Recent surgery; patient-controlled delivery of pain narcotic (clinically inferable risk of respiratory depression)	Simulated onset of respiratory depression causes multiple physiological changes, including triggering alarm for low respiration rate; Simulated occlusion of infusion pump for pain narcotic	Unimportant status changes to infusion pump device, and others
Blood clot in leg (clinically inferable risk of pulmonary embolism)	Simulated occlusion of infusion pump of anti-clotting drug; Simulated onset of pulmonary embolism causing multiple physiological changes, including triggering alarm for low SpO ₂ (blood oxygenation)	Unimportant status changes to infusion pump device, and others
Pneumonia (clinically inferable risk of respiratory distress)	Simulated onset of unknown distress causing multiple physiological changes, including alarm for low heart rate	Unimportant alarms related to recurring low SpO ₂ , and others

Each of the four nurses was assigned responsibility for the care of five patients, which represented a realistic yet heavy nursing shift assignment. So that each nurse was exposed to similar types and acuity of patients, we developed five different patient archetypes (Table 1) and replicated them in four sets, one set for each nurse on a team. Patients (mannequins with simulated vital signs and device behavior) had different names and were randomly distributed across the 20 beds in the simulation laboratory. The vital signs for every patient were individually generated using a custom autoregressive-moving-average (ARMA) time series algorithm that was parameterized using a validated model (based on actual clinical data previously recorded) of the five different types of patients. As a result, no two patients had the same vital sign trajectory. Patients of the same type, however, had vital sign simulations with the same median values for each vital sign type. Subjective observations during the experiment showed that nurses were unaware that their teammates were caring for sets of patients that were similar to their own.

Before the start of the scenario, nurses received a change of shift report consisting of paper medical chart documentation summaries of each patient's diagnosis, history, and orders for medical and nursing care. A list with pre-scheduled tasks was also provided for every patient and included the following:

- full standard assessment of each patient;
- charting all observations and care delivery;

- administer multiple prescribed medications for each patient;
- perform prescribed procedures and treatments;
- carry out infection control precautions, including full personal protective equipment (PPE) whenever visiting the contagious patient (one of five for each nurse);
- involve nursing leadership for anything serious;
- coordinate patients' care with multiple other hospital workers.

Nurses were also asked to use their normal clinical judgment in deciding how and when to respond to alarms and call button events. This task loading was designed with variation in workload over time to maximize realism. Nurses were heavily loaded with pre-assigned multitasking responsibilities for the first 60 minutes of both of the 90-minute sessions. The remaining 30 minutes of each session were loaded more lightly with pre-assigned tasks. The PPE used to care for the infectious patient was a translucent blue plastic, and allowed nurses to read the watch through the PPE gown sleeve. High rates of health surveillance data were available from patient monitors and infusion pumps. Thirty alarms per nurse were generated in each 90-minute segment of the data collection sessions; only three were clinically important alarms. This represents the average rate (as reported in the literature) of relevant and non-relevant alarms (i.e., false alarms, noise,



Figure 2. Simulation laboratory facility showing five patient “rooms” divided by curtains.



Figure 3. Patient simulator with identification and allergy bracelet, nasal cannula for oxygen delivery, peripheral intravenous catheter, and pulse oximeter probe on a finger.

and/or clinically unimportant).²⁷ A call light system generated five additional notifications per nurse.

The vital signs for every patient were individually generated using a custom ARMA time series algorithm that was parameterized using a model of the five different types of patients (Equation (1)). Each scenario session (a.m. and p.m.) included a set of changes from the normal simulation baselines for each patient.

Equation (1) shows the ARMA time series used to generate simulated patient vital signs:

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

4.5 Simulation facility

The 20-bed Intermountain Healthcare Simulation Learning Center at the University of Utah College of Nursing (Salt Lake City, UT) replicates a full-scale acute care hospital unit or ward. Architecturally, it is a rectangular space with

a central nursing station in the middle and 10 hospital beds on each side. Figure 2 shows one five-bed area with the curtains opened. For the experiment, the curtains were closed except for a door-sized opening in each. A call light indicator is located near the ceiling in front of each patient “room.” The patient mannequins used in this study were SimMan® Essential patient simulators (Laerdal, Wappinger Falls, NY, USA) situated in Hill-Rom-1000 hospital beds. The simulators reproduce breathing, heart sounds, and peripheral pulses, and have realistic skin and surface tissues with veins to receive intravenous (IV) catheters and injections. Each patient “room” also contained a full complement of hospital furniture, standard gas interfaces (oxygen, medical air, Medivac suction), pulse oximetry, communications systems (phone, call light control, code button, and bathroom call button), a bedside computer for documentation, a sphygmomanometer, and a sharps disposal container. Every patient “room” was also equipped with an android tablet that simulated an integrated bedside monitor and up to two infusion pumps.

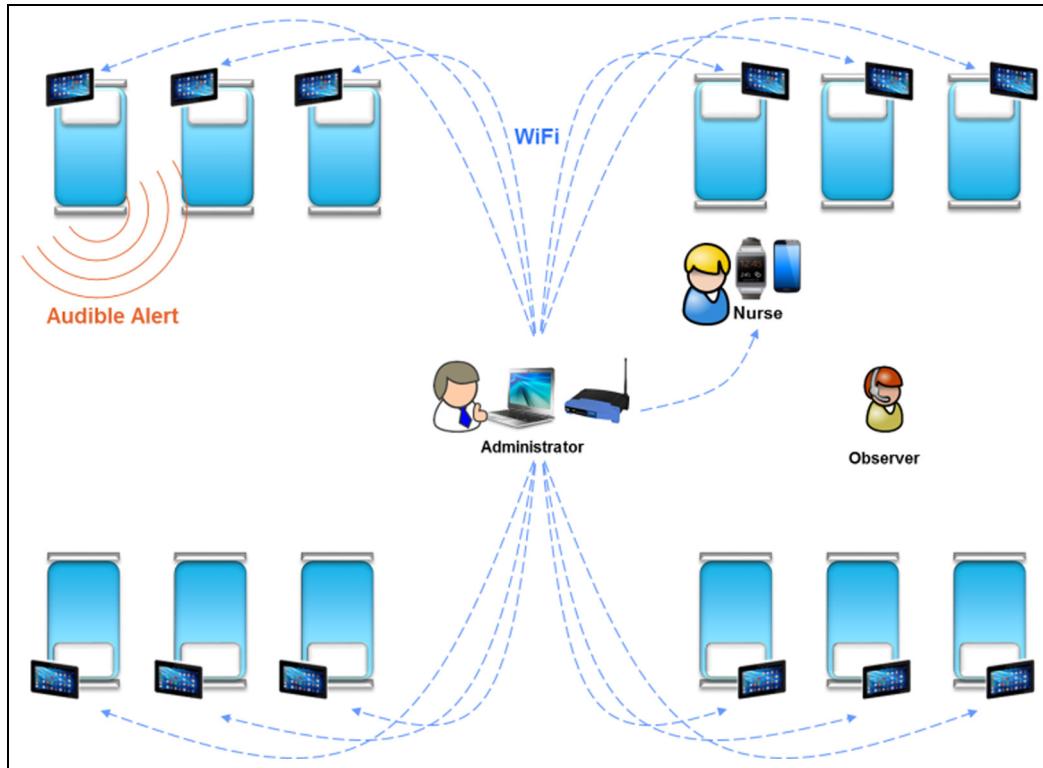


Figure 4. Components for the alarm delivery platform include an administrator with control software wireless connected to an android tablet at each patient bedside (only 12 of the 20 are shown), plus an android phone and smartwatch for each nurse (only one of four are shown).

Figure 3 shows one of the mannequins and its right arm with patient ID bracelet, IV catheter, and pulse oximeter finger probe.

At the central nursing station were two medication preparation areas, each with an automated medication dispensing unit (Oncocell, Mountain View, CA, USA; Figure 4). Nurses in the study had to retrieve all medications from these dispensing units and used medication preparation supplies stocked in this central area. The central nursing station also included two wall phones that nurses used to call and speak with other hospital areas (e.g., pharmacy) and the patients' physicians to coordinate care.

4.6 Development of the alarm delivery system for the experiment

HAIL informed the development of the HAIL-CAT experimental platform to support a clinical trial in patient simulation. Components for the alarm delivery system are illustrated in Figure 4. The prototype is implemented on Samsung Gear 2 watches using Java software technologies and common-off-the-shelf (COTS) WiFi networking. It provides a set of wearable context-enabled alarm

notification services to support users' metacognition for interruption triage. The smartwatch function was supported by a server-side integrated data environment and analytics engine. A central experimental simulation server provided patient vital simulation updates every second and simulated scenario changes or "deviations" from patients' normal statuses and it integrated all data, generated alarms, supported integrated metacognitive-aiding alarm mediation, and delivered user-directed information query services.

A scenario is defined by a set of patient vitals for the duration of the experiment and a file that contains vital sign deviations for the patients at various times throughout the simulation. The patient vitals files represent the nominal vital signs for the patient given that there is no health event taking place. Vital information is sent to the responsible nurse, where they are displayed on the watch and also sent to the patient's assigned bedside tablet. The deviations file defines time when the vitals for a patient should deviate from the normal behavior into a new mode. These deviations have the potential to trigger alarms that are activated when vital signs cross specified thresholds. If an alarm is triggered, an alert message is sent to the nurse and to the patient's bedside tablet. On the tablet, the alarm

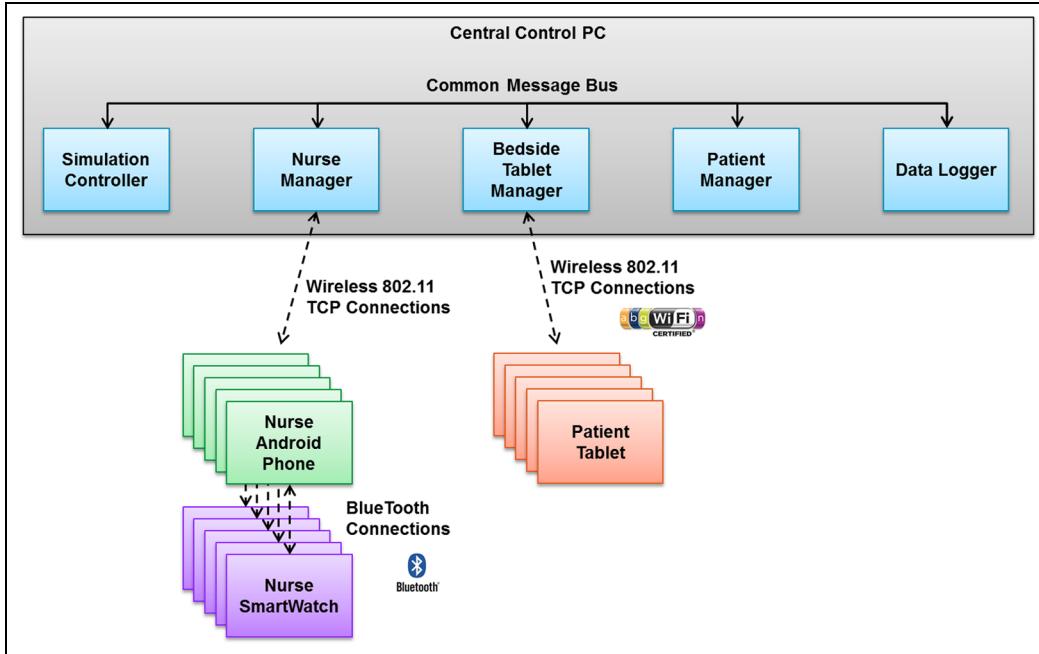


Figure 5. Human Alerting and Interruption Logistics – Clinical Alarm Triage architecture, including the integration of software components.

information is displayed and an audible alarm begins to transmit from the tablet. Along with the alert information, the tablet also has buttons that can be used to temporarily silence the alarm or to mark the alert as resolved. The alert message is also sent to the nurse's watch, where the information is displayed as well as buttons to silence the alert for either five or 15 minutes. The nurses are also able to use the smart watch to view a list of all their patients, the active and silenced alarms for their patients, and the current vital signs of their patients.

4.6.1 Experiment software architecture. The software prototype was developed as a set of integrated programs and applications on Android operating system (OS) mobile COTS hardware and a Linux server on a PC laptop. The solution consisted of the following: android tablets at each patient bedside to simulate bedside monitors, infusion pumps, and patient-controlled analgesia (PCA); paired android smartphones (Samsung Galaxy S) and Samsung Gear 2 smartwatch for each nurse; and central control software to manage wireless connections and execute the simulation scenario for the experiment. Figure 5 shows the HAIL-CAT platform components and their wireless communications.

The system architecture design manages most processing functions on the server side. This minimizes the amount of mobile distributed code that has to be maintained and simplifies system operation. A *Simulation*

Controller reads the configuration files and distributes a time-step message that indicates the progression of time throughout the simulation. It also starts and stops all simulations. A software component called the *Nurse Manager* manages connections to the four nurse Android phones. It also manages patient-to-nurse assignment mapping. It uses this information to listen to messages coming across the common message bus to determine which messages need to be routed to which nurses. Another software component, the *Bedside Tablet Manager*, maintains wireless network data connections to the bedside tablets. It also maintains the mapping of patient-to-bed assignment so that messages can be routed accordingly. The *Nurse Manager* and *Bedside Tablet Manager* also receive messages for their wireless clients and forward messages onto the common message bus so that they can be received by other components of the software. The *Patient Manager* software component simulates patient status. It sends out current simulated vitals, monitors for alarm conditions, and generates alarms according to a configuration policy. A *Data Logger* listens to all the messages being sent over the message bus and logs all the information necessary for analysis.

4.7 Execution of the simulated scenarios in the simulation laboratory

The study was approved by the University of Utah's Institutional Review Board and determined to pose "less

than minimal risk” to nurse participants. Nurses were compensated for participating. Instructions emphasized that they were free to stop participation at any time.

After the nurse participants arrived at the simulation laboratory, they read and signed a consent form, and received orientation to the study procedures. This orientation consisted of an explanation of their responsibilities in caring for five patients and an introduction to the facility and training on the patient simulators. They received 10 minutes of training on the experimental platform that included an orientation to the HAIL-CAT on the smartwatch and smart phone and bedside tablets. (This training was conducted prior to the start of the afternoon session if nurses were using the smartwatch for that session.) Prior to starting the morning session, nurses reviewed a printed change-of-shift summary report and abbreviated medical/nursing charts for their five patients, which including diagnosis, history, and medical orders.

During the entire scenario, each nurse was shadowed by an observer wearing a hat-mounted video-camera for hands-free high-definition (HD) recording. The observer recorded the times (on paper with clipboards) that nurses entered and exited each patient room and noted field observations with unstructured text. Nurses were also followed by a confederate “family member” (role played by a research assistant) who used a script to support simulated conversations between the nurse and patients and their visitors. Whenever a nurse would approach a patient’s room, this “family member” would quickly enter the room and sit down on a chair near the patient. The “family member” would open a magazine with the concealed script inside in order to respond to all questions as if continually present in the room, speaking with the patient frequently. The “family member” would also respond verbally to any question the nurse asked directly to the patient providing information that would otherwise be accessible to the nurse in a real hospital. “Family member” confederates were well practiced in this process, and nurses were able to have rapid realistically natural conversations with every simulated patient and/or their family. Other confederates in the scenario included two nurse practitioners who have primary responsibility for the 20 patients and could answer nurses’ queries, two certified nursing assistants who assisted with patient care and activation of the call light system, and an individual in a remote room answering the phone and assuming the role of individuals in other hospital areas (for example, pharmacy, radiology, etc.) or the patient’s primary physician. Observation and exit interviews confirm that nurses felt that the interactions with the confederates were highly realistic.

After both parts of the scenario were completed, nurses were asked to complete a survey that included questions about demographics and work experience.

They also completed a 17-question usability survey about the smartwatch-based HAIL-CAT and a series of nine questions that compared using the HAIL-CAT on the smartwatch to the control (i.e., not using the smartwatch). Lastly, nurses participated in a semi-structured interview conducted by the nurse’s observer experimenter that focused on strengths of and issues with the HAIL-CAT.

4.8 Data collection procedures

The server for the experimental platform collected four types of comprehensive data logs. The first log type was a second-by-second log of the vital values for every one of the 20 simulated patients. These were the vital signs data that were available to the nurses (HR, SBP/DBP, RR, SpO₂) and were included as part of the contextual information provided with the alarms on the bedside tablets, and on the smartwatches and smart phones under the treatment condition. The second type of log was an alarm event log that recorded state change events for every alarm. Each alarm had a unique ID, and the alarm log captured the from-state and the to-state. For example, when an alarm was generated (because of a deviation above or below threshold), the log shows it transitioning from its scheduled state to its active state.

The third type of log was a deviation log that recorded simulation events related to device status deviation from patients’ normal. Onset and change in deviation event states were hidden from both observers and nurse participants (double blind). A deviation could drive a specified patient’s vital sign either above or below the alarm threshold. The system alarm generation mechanism would then recognize this as an alarm state and initiate an alarm signal. Some deviations were designed not to cause alarms, but to show some planned physiological change in association with other deviations that would cause alarms. The final type of log was a recording of screen transitions UI events for use of the smartwatch. These were all of the actions taken by nurse participants using the smartwatch during the treatment condition of the experiment.

5. Results

The HAIL-CAT wearable prototype delivers HSI services to support nurses’ complex “track-while-scan” workflow. The experiment tests the hypothesis that this novel wearable prototype system will improve operators’ “scan” performance while not negatively affecting their “track” performance. HSI support services for “track” multitasking includes a mobile HCI function to enable a nurse to check at-will on all of the patients’ statuses anytime and anywhere. Nurses’ “scan” multitask responsibilities are

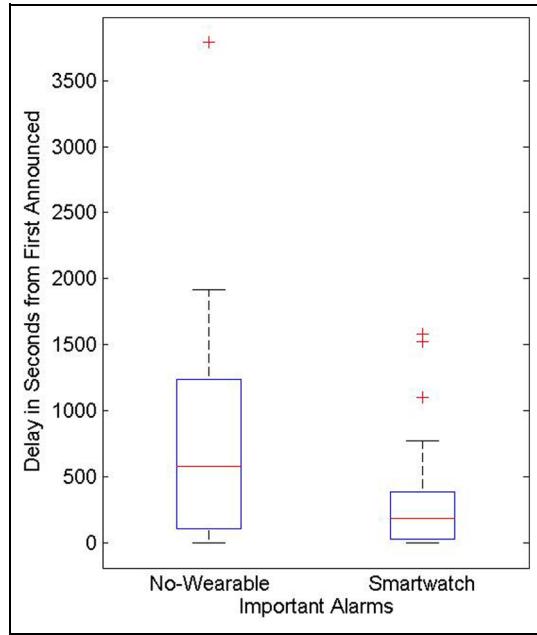


Figure 6. Nurse time to respond to clinically important alarms (seconds).

supported by mobile HCI with hands-free alerting with context information to minimize the cost to nurses' cognitive resources for triaging alert interruptions. The results described here are organized by observations related to support of nurses "scan" and "track" multitasks.

The HAIL-CAT also delivers HCI services to support nurses' meta-level responsibility to manage the dynamic allocation of their own cognitive attention resources. Observing this directly, however, is not practicable. Instead, experimental effects of aiding nurse internal meta-cognition performance are inferred through objective observations of nurses' external behavior on "track" and "scan" multitasks. A subjective exit questionnaire also reports nurse participants' opinions about the services and comparison with baseline.

5.1 Effect on operator "scan" performance

The primary metric for vigilance or "scan" performance was, "How quickly did a patient's nurse arrive at the patient's bedside after a 'clinically important' (i.e., actionable) alarm condition occurred (amid high 'track' multi-task workload and high-volume high-rate non-actionable information)?" The data are time delays until nurse response (in seconds) for each of three important events per nurse for each of the two conditions (a total of six). When using the HAIL-CAT on the smartwatch, nurses arrived at the patient's bedside three times faster (median

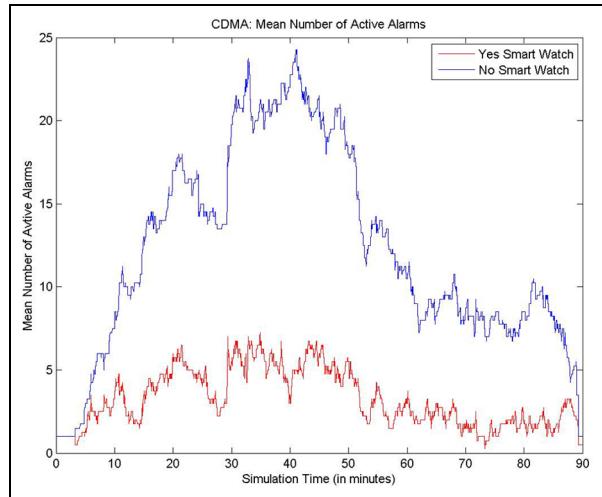


Figure 7. Distribution of the number of active sounding alarms.

time) when responding to an actionable alarm condition (Figure 6).

These results support an assertion that HAIL-inspired HAIL-CAT improves operator "scan" performance in health surveillance operations. Times are truncated by the end of the 90-minute scenario parts; therefore, any alarm not responded to before the end is marked as completed at that time. Considering this limitation and the fact that the smartwatch only provides information delivery services (no filtering), the three-fold observed improvement in response time represents a very large and positive impact on nurse clinical performance. The HAIL-CAT approach shows strong potential to dramatically improve nurse team performance with real patients.

Although the HAIL-CAT condition showed a striking improvement in response time (median, standard deviation, and max), nurse performance on the "scanning" task was not perfect. Two outliers were observed with response times of approximately 1500 seconds. This suggests the importance of future work to further explore the limits of optimizing nurse "scanning" performance. Observations showed that nurses typically looked at each screen for only a second or two. They also frequently shifted through two or three screens rapidly in less than a second while navigating the UI. Total accumulated UI interaction time for nurses equaled only a few minutes.

5.2 Effect on operator "track" performance

The overall time that nurse participants took to complete their assigned tasks (assessment, medication administration, clinical procedures, and healthcare coordination)

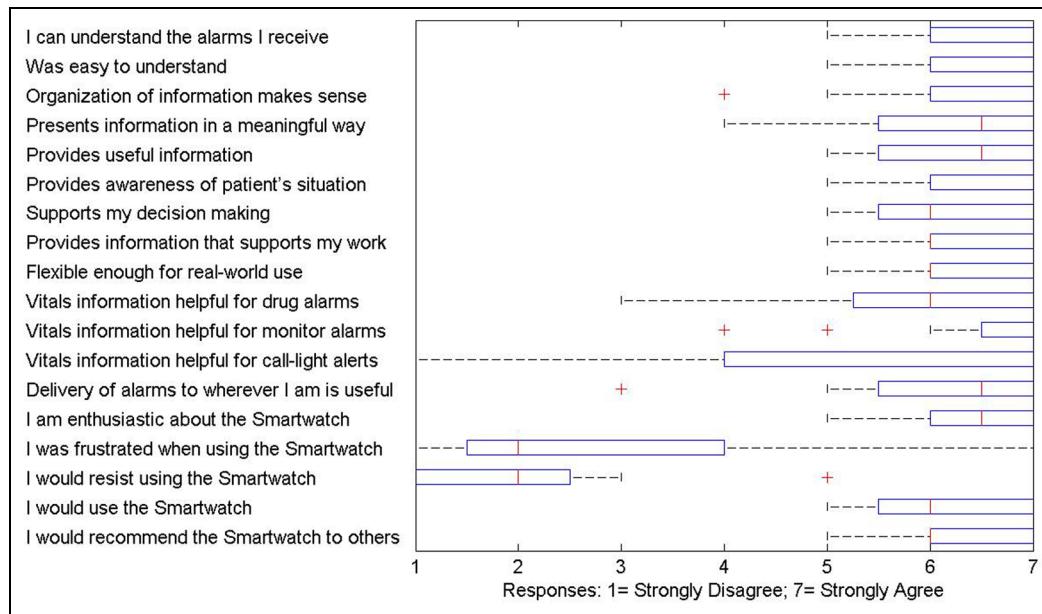


Figure 8. Feedback regarding usability of Human Alerting and Interruption Logistics Clinical Alarm Triage for the smartwatch platform on a Likert scale of 1 (strongly disagree) to 7 (strongly agree).

was equivalent with and without the smartwatch ($t = 0.28$, $p = .78$). This finding demonstrates that use of the smartwatch did not interfere with the normal nursing workflow.

The count of active sounding alarms in the simulation laboratory was also recorded. The smartwatch allowed nurses to silence alarm bed-side audio at will from the wearable prototype. The overall audio noise level for the 20-bed simulation laboratory from active, sounding alarms was much lower under the experimental smartwatch condition (see Figure 7). This is not a negative impact on the unit's work environment for nurses performing "track" tasks and therefore supports the hypothesis.

5.3 Observations and subjective evaluations of the HAIL-CAT prototype delivered to the smartwatch

During the experiment we observed the nurses utilizing the watch frequently to prioritize their tasks and triage patient alarms. In addition, we observed nurses using the watch to check vital signs of patients at will. The qualitative findings from the post-simulation interview indicate that nurses perceived this technology as valuable, useful, and easy to learn. The utility and usability of the smartwatch prototype compared very favorably relative to baseline (no wearable). Nurses also responded favorably about the utility, usability, and adoption of the prototype wearable services (Figure 8). All nurses said they would use the HAIL-CAT prototype for real work. Only one nurse indicated some

indifference. All 16 participants stated that the 10-minute training session was adequate.

Exit interviews did not discover any concern for infection control issues for use of a wrist watch form factor. Nurses said that the smartwatch would need to be able to not take damage from being rubbed with hand sanitizer after leaving each patient. Note that the smartwatch platform adopted for this experimental prototype did not support being rubbed with hand sanitizer. To conduct a future clinical trial with real patients would require a more clinically mature prototype that could deliver on all patient safety and device reliability concerns, including easy sanitization.

6. Discussion and future work

The results indicate the potential value of future research to explore whether HAIL-CAT can positively affect nurses' performance on "tracking" activities during heavy multitasking and frequent interruptions. It would be worthwhile to explore the limits of utility of HAIL-CAT for increasingly more intense situations and/or spikes of high workload, information overload, and frequency of interruption of health surveillance operators. A high degree of individual differences was observed for nurses. It would be useful to explore the diversity of strategies adopted for managing multitasking during interruption. The type and amount of contextual information included in

announcements could be explored to try to optimize performance. Future work could also explore the benefits of noise reduction across the unit on patient and clinician experience.

This article focuses on the HAIL metacognition methods and technology and their leverage to address health surveillance for hospital acute care patients. Modeling operator metacognition for interruption triage events informs useful aiding services. Design and implementation of the prototype are discussed in some detail. Technical pragmatics of running the simulation experiment and data collection are also included. Clinical description about the scenario, experimental procedures, and results is summarized at a high level to aid accessibility for a broad defense science audience. Additional highly clinical details are reported separately in a medical journal – interested medical professionals are invited to contact the authors directly for more information.

7. Conclusions

Information overload is a debilitating problem that negatively affects the performance of human operators working health surveillance operations. Modern sensors for health surveillance generate high volumes and rates of data that currently overwhelm operational decision-makers. These data are collected with the intention of enabling health operations decision-makers to achieve SA and make effective decisions amid a changing health situation. Ironically, HSI analysis shows that the flood of data instead causes information overload in operators and actually degrades their effectiveness. Information overload from clinical alarms can blind nurses to important changes in patients' statuses and hence delay their response to crises.

Results from this clinical trial in patient simulation show that the US Navy's HAIL HSI innovation from combat systems can be leveraged to improve nurse performance in hospital health surveillance operations. No automated filtering was introduced. Instead, nurses were provided with a wearable metacognition aid in the form of the HAIL-CAT delivered to the nurse on a smartwatch to support their need to prioritize their own cognitive attention resources. Nurses responded to clinically important, actionable alarms three times faster on average when they were using the HAIL-inspired wearable prototype. This is a dramatic increase in nurse performance on their "scan" multitask responsibility to recognize important change in patients' statuses. No negative consequences were observed affecting nurse performance on their pre-scheduled multitasks – their "track" responsibilities.

Empirical observation, together with the domain-independent foundation of the HAIL basic research, suggests that the HAIL HSI theory, methods, and technologies

could generalize. It may be possible to create HSI solutions that defeat data overload across multiple different types of health surveillance operations.

Acknowledgments

Special thanks for contributions from George Mason University, Ohio State University, Carnegie Mellon University, and the US Naval Academy. The NATO Science & Technology HFM-254 panel on health surveillance (2015) provided insightful feedback on this work. Research and development was performed by teammates from the US Naval Research Laboratory, Lockheed Martin Advanced Technology Laboratories, the University of Utah, and the Royal Philips, Innovation Office.

Funding

This work was supported by the US Office of the Assistant Secretary of Defense for Research and Engineering (ASD(R&E)) to transform "data to decisions" and defeat information overload for health surveillance. The leveraged AEGIS innovation is called the Human Alerting and Interruption Logistics – AEGIS (HAIL-AEGIS) technology, and was originally sponsored by the US Office of Naval Research (ONR).

References

1. Rubens M, Ramamoorthy V, Saxena A, et al. Public health in the twenty-first century: the role of advanced technologies. *Front Publ Health* 2014; 2: 1–4.
2. Milinovich GJ, Williams GM, Clements ACA, et al. Internet-based surveillance systems for monitoring emerging infectious disease. *Lancet Infect Dis* 2014; 14: 160–168.
3. Kushnir S, Heithecker CH, Ballas JA, et al. Situation assessment through collaborative human-computer interaction. *Nav Eng J* 1996; 108: 41–51.
4. Ballas JA, McFarlane DC, Achille LB, et al. Interfaces for intelligent control of data fusion processing. NRL, NRL/FR/5513–96–9806.
5. McFarlane DC and Latorella KA. The scope and importance of human interruption in human-computer interaction design. *Hum Comput Interact* 2002; 17: 1–61.
6. McFarlane DC. Interruption of people in human-computer interaction: a general unifying definition of human interruption and taxonomy. Naval Research Laboratory, Washington, DC, Peer-Reviewed Formal Report NRL/FR/5510–97–9870, 1997.
7. McFarlane DC. *Interruption of people in human-computer interaction*. Dissertation, The George Washington University, Washington, DC, 1998.
8. McFarlane DC. Coordinating the interruption of people in human-computer interaction. In: *human-computer interaction (INTERACT'99)*, Edinburgh, UK, 30 August–2 September 1999, pp.295–303. Amsterdam: IOS Press.

9. McFarlane D. Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Hum Comput Interact* 2002; 17: 63–139.
10. McFarlane DC. Engaging innate human cognitive capabilities to coordinate human interruption: the HAIL system. In: Forsythe C, Bernard ML and Goldsmith TE (eds) *Cognitive systems: human cognitive models in system design, part 3. Illustration of cognitive systems*. Mahwah, NJ: Laurence Erlbaum and Assoc. (LEA), 2006, pp. 37–152.
11. Lockheed Martin Advanced Technology Laboratories. The Human Alerting and Interruption Logistics-Surface Ship (HAIL-SS) system has successfully transitioned from laboratory to platform as an integrated component on the newly commissioned USS Sterett DDG-104. *Trans Brief* 2008; 2: 1–2.
12. Mahafza BR. *Introduction to radar analysis*. 1st ed. Boca Raton, FL: CRC Press, 1998.
13. McFarlane DC. A scare in the OR highlights value of systems engineering. *Biomed Instrumen Technol* 2014; 48: 44–50.
14. DeLucia PR, Ott TE and Palmieri PA. Performance in nursing. In: *Rev Hum Factors Ergon* 2005; 5: 1–40.
15. Koch SH, Weir C, Haar M, et al. Intensive care unit nurses' information needs and recommendations for integrated displays to improve nurses' situation awareness. *J Am Med Inform Assoc* 2012; 19: 583–590.
16. Westbrook JI, Duffield C, Li L, et al. How much time do nurses have for patients? A longitudinal study quantifying hospital nurses' patterns of task time distribution and interactions with health professionals. *BMC Health Serv Res* 2011; 11: 319.
17. Rajkomar A and Blandford A. A distributed cognition model for analysing interruption resumption during infusion administration. In: *Proceedings of the 30th European conference on cognitive ergonomics (ECCE'12)*, Edinburgh, UK, 28–31 August 2012, pp.108–111. New York: ACM.
18. Cheng M, Kanai-Pak M, Kuwahara N, et al. Dynamic scheduling based inpatient nursing support: applicability evaluation by laboratory experiments. In: *The international workshop on context-awareness for self-managing systems (CASEMANS 2009)*, Nara, Japan, 11 May 2009, pp.48–54. New York: ACM.
19. Ebright PR, Patterson ES, Chalko BA, et al. Understanding the complexity of registered nurse work in acute care settings. *J Nurs Admin* 2003; 33: 630–638.
20. Potter P, Wolf L, Boxerman S, et al. Understanding the cognitive work of nursing in the acute care environment. *J Nurs Admin* 2005; 35: 327–335.
21. Krichbaum K, Diemert C, Jacox L, et al. Complexity compression: nurses under fire. *Nurs Forum* 2007; 42: 86–94.
22. Patterson ES, Ebright PR and Saleem JJ. Investigating stacking: how do registered nurses prioritize their activities in real-time? *Int J Ind Ergon* 2011; 41: 389–393.
23. Tucker AL and Spear SJ. Operational Failures and interruptions in hospital nursing. *Health Serv Res* 2006; 41: 643–662.
24. Cornell P, Riordan M, Townsend-Gervis M, et al. Barriers to critical thinking: workflow interruptions and task switching among nurses. *J Nurs Admin* 2011; 41: 407–414.
25. Imhoff M and Kuhls S. Alarm algorithms in critical care monitoring. *Anesth Analg* 2006; 102: 1525–1537.
26. Schmid F, Goepfert MS and Reuter DA. Patient monitoring alarms in the ICU and in the operating room. *Crit Care* 2013; 17: 216.
27. Association for the Advancement of Medical Instrumentation. *A siren call for action: priority issues from the medical device alarms summit*. Arlington, VA: Association for the Advancement of Medical Instrumentation, 2011.
28. Gross B, Dahl D and Nielsen L. Physiologic monitoring alarm load on medical/surgical floors of a community hospital. *Biomed Instrumen Technol* 2011; 45: 29–36.
29. Clinical Alarms Task Force. ACCE healthcare technology foundation impact of clinical alarms on patient safety. *J Clin Eng* 2007; January/March: 22–33.
30. Classen DC, Resar R, Griffin F, et al. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff* 2011; 30: 581–589.
31. Drew BJ, Harris P, Zègre-Hemsey JK, et al. Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit patients. *PLoS One* 2014; 9: 1–23.
32. Siebig S, Kuhls S, Imhoff M, et al. Collection of annotated data in a clinical validation study for alarm algorithms in intensive care—a methodologic framework. *J Crit Care* 2010; 25: 128–135.
33. Nouira K and Trabelsi A. Intelligent monitoring system for intensive care units. *J Med Syst* 2012; 36: 2309–2318.
34. Sobieraj J, Ortega C, West I, et al. Audibility of patient clinical alarms to hospital nursing personnel. *Mil Med* 2006; 171: 306–310.
35. ECRI Institute. ECRI Institute releases top 10 health technology hazards for 2011, <http://psqh.com/ecri-institute-releases-top-10-health-technology-hazards-for-2011> (accessed 23 November 2015).
36. Garrett SK and Craig JB. Medication Administration and the complexity of nursing workflow. In: *Proceedings of the 21st Annual Society for Health Systems Conference and Expo*, Chicago, USA, 1–4 April 2009, pp. 1–6.
37. Westbrook JI, Woods A, Rob MI, et al. Association of interruptions with an increased risk and severity of medication administration errors. *Arch Intern Med* 2010; 170: 683–690.
38. Prakash V, Koczmara C, Savage P, et al. Mitigating errors caused by interruptions during medication verification and administration: interventions in a simulated ambulatory chemotherapy setting. *BMJ Qual Saf* 2014; 23: 884–892.
39. Cropp AJ, Woods LA, Raney D, et al. Name that tone. The proliferation of alarms in the intensive care unit. *Chest* 1994; 105: 1217–1220.
40. Ryherd EE, Okcu S, Ackerman J, et al. Noise pollution in hospitals: impacts on staff. *J Clin Outcome Manag* 2012; 19: 491–500.

41. Curry JP and Jungquist CR. A critical assessment of monitoring practices, patient deterioration, and alarm fatigue on inpatient wards: a review. *Patient Saf Surg* 2014; 8: 29.
42. Kafeza E, Chiu DKW, Cheung SC, et al. Alerts in mobile healthcare applications: requirements and pilot study. *IEEE Trans Inform Technol Biomed* 2004; 8: 173–181.
43. Koch SH, Weir C, Westenskow D, et al. Evaluation of the effect of information integration in displays for ICU nurses on situation awareness and task completion time: a prospective randomized controlled study. *Int J Med Inform* 2013; 82: 665–675.

Author biographies

Daniel C McFarlane, ScD, is a Senior Research Scientist at the Royal Philips, Patient Care & Monitoring Solutions, Innovation Office, Cambridge & Andover, Massachusetts, USA. He was previously a Fellow with Lockheed Martin Advanced Technology Laboratories, Informatics Lab, Arlington, VA, USA.

Alexa K Doig, PhD, RN, is an associate professor and chair of the Division of Acute and Chronic Care at the University of Utah, College of Nursing, Salt Lake City, Utah, USA. She is a Nurse Faculty Scholar alumna in the Robert Wood Johnson Foundation.

James A Agutter, MArch, is an assistant professor at the University of Utah, College of Architecture &

Planning, Salt Lake City, Utah, USA. He is also the owner of Applied Medical Visualizations (Medvis), Salt Lake City, Utah, USA.

Jonathan L Mercurio, MS, is a Senior Member of the Research Staff at Lockheed Martin Advanced Technology Laboratories, Spectrum Systems Laboratory, Cherry Hill, New Jersey, USA. His research background includes machine learning, RF systems, signal processing, and rapid prototyping.

Ranjeev Mittu, MS, is Branch Head at the US Naval Research Laboratory, Information Technology Division, Information Management and Decision Architectures Branch, Washington, District of Columbia, USA.

Lara M Brewer, PhD, is a research assistant professor at the University of Utah Health Sciences Center, Department of Anesthesiology, Salt Lake City, Utah, USA. Her research interests include improved patient monitoring and development of decision support systems.

Noah D Syroid, MS, is Technical Director at the University of Utah, Anesthesiology Center for Patient Simulation, Salt Lake City, Utah, USA. He is also Chief Technical Officer at Applied Medical Visualizations (Medvis), Salt Lake City, Utah, USA.