

# Patient Safety, Potential Adverse Drug Events, and Medical Device Design: A Human Factors Engineering Approach

Laura Lin,<sup>\*,†,1</sup> Kim J. Vicente,<sup>\*,†,‡,§,2</sup> and D. John Doyle<sup>†,¶</sup>

<sup>\*</sup>*Cognitive Engineering Laboratory, Department of Mechanical and Industrial Engineering,*

<sup>†</sup>*Institute of Biomaterials and Biomedical Engineering, ‡Department of Computer Science,*

<sup>§</sup>*Department of Electrical and Computer Engineering, ¶Department of Anaesthesia, University of Toronto, Toronto, Ontario, Canada M5S 3G8*

Received November 27, 2001; published online February 22, 2002

Adverse drug events are the single leading threat to patient safety. Human factors engineering has been repeatedly proposed, but largely untested, as the key to improving patient safety. The value of this approach was investigated in the context of a commercially available patient-controlled analgesia device that has been linked with several alleged patient injuries and deaths. Several reports have stated that errors in programming drug concentration were made during these adverse drug events. A simulation of the commercially available interface was compared experimentally with a simulated prototype of a new interface designed according to a human factors process. Professional nurses, averaging over 5 years of clinical experience with the commercially available interface and only minimal experience with the new interface, programmed both interfaces. The new interface eliminated drug concentration errors, whereas the simulated commercially available interface did not. Also, the new interface led to significantly fewer total errors and faster performance. These findings may have broad implications for the design, regulation, and procurement of biomedical devices, products, or systems that improve patient safety in clinical settings. © 2001 Elsevier Science (USA)

**Key Words:** patient safety; adverse drug events; human factors engineering; medical device design; patient-controlled analgesia; medical error.

## INTRODUCTION

Adverse drug events are the single leading cause of medical injuries, accounting for 19.4% of all adverse events identified in the landmark Harvard Medical Practice Study [1]. Preventable adverse drug events lead to extended hospital stays that have been estimated to cost \$2 billion annually in the United States, not including costs of injuries to patients or malpractice costs [2]. In the follow-up Harvard Adverse Drug Event Study, analgesics (e.g., morphine, meperidine) were the most likely drugs to be associated with preventable medical injury [3]. Furthermore, 45% of the adverse drug events associated with analgesics involved misuse or malfunction of drug infusion devices of various types (e.g., patient-controlled analgesia). When preventable adverse drug events were analyzed as a function of stage of delivery, errors were most frequent in the drug ordering stage (49%), but 48% of these errors (i.e., 23.5% overall) were detected and intercepted. Errors in the administration stage were the second most frequent (26%), yet none of these errors was detected and intercepted. Thus, substantial improvements in patient safety may be achieved by focusing on adverse events occurring during the administration of analgesics using infusion devices. The user programming of patient-controlled

<sup>1</sup>Present address: Logicon, Inc., Dayton, Ohio.

<sup>2</sup>To whom correspondence and reprint requests should be addressed at Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, Ontario, Canada M5S 3G8. E-mail: vicente@mie.utoronto.ca. URL: [www.mie.utoronto.ca/labs/cel/](http://www.mie.utoronto.ca/labs/cel/).

analgesia (PCA) infusion pumps satisfies this criterion, and thus is a suitable choice for research on patient safety.

### *Patient-Controlled Analgesia*

Traditionally, analgesia was only delivered in large, infrequent doses by a nurse who had to retrieve, prepare, and administer the drug manually. This process can be ineffective because patients may become oversedated by large doses, and inefficient because nurses must perform several time-consuming steps. PCA pumps have the potential to improve pain management for patients and reduce workload for nurses by using automation technology to help patients self-administer more frequent, but smaller doses of analgesia—usually morphine [4, 5]. For most patients, the PCA device is easy to use. Whenever they are in pain, or are about to do something likely to be painful (e.g., get out of bed), patients press a push-button on a pendant. If the patient is eligible to get a dose (as determined by the computer program in the pump), a dose of analgesic is pumped into an intravenous line by an automatic controller. If insufficient time has elapsed since the last dose (the “lockout period”), the program denies the patient’s request.

To make sure that patients do not receive too much analgesic, a nurse is required to program the PCA pump using a human–computer interface consisting of displays and a keypad. First, the nurse manually inserts a drug vial into the machine. Then, the nurse programs several parameters, including: the concentration of the drug in the vial, the dosage that should be given to the patient upon each request, the lockout interval specifying the minimum time between doses, and sometimes, the maximum allowable dose over any 4-h period. The parameters that are entered by the nurse govern the behavior of the PCA pump (the patient is not allowed to change the programmed parameters for safety reasons). There is no way for the pump to verify independently if the settings are correct because it cannot sense the concentration nor the type of analgesic in the vial. Thus, a programming error could lead to an underdelivery or overdelivery of analgesic. An overdelivery can, in turn, lead to patient injury or death. According to one review of Food and Drug Administration medical device reports 67% of all problems associated with PCA pumps are attributable to user error, including programming errors [6].

### *Abbott LifeCare<sup>3</sup> 4100 PCA Plus II Infusion Pump*

To conduct research on how to make PCA pumps less error-prone, the Abbott LifeCare 4100 PCA Plus II infusion

pump was chosen. This device is used daily at our local teaching hospitals. Anecdotal reports from nurses indicated that the device was difficult to program, and thus could perhaps benefit from being redesigned. The LifeCare 4100 PCA pump is the market leader, has been used on over 22 million patients, and “represents approximately 75% of all PCA use in the U.S.” [7]. The device is also used around the world, and in nearly 4000 hospitals in the United States alone [8]. These figures suggest that the LifeCare 4100 is frequently and widely used, making it a suitable choice for research on PCA infusion pumps.

### *Adverse Drug Event Reports with LifeCare 4100*

After the experiment described below was completed, we learned that several patients had allegedly died or been injured while connected to the LifeCare 4100 [9–14]. The incidence of such adverse drug event reports appears to be low [7]. However, it is well known that adverse events in general, and adverse drug events in particular, are both vastly underreported. Epidemiological studies have observed reporting rates ranging from a low of 1.2% to a high of 7.7% [15–18]. Using these estimates, for every adverse drug event that is reported, there are an additional 12 to 82 that go unreported. “This suggests that using [incident reports] as a primary data source to study drug-related complications will be misleading” [15].

From a human factors engineering point of view, an important question is whether the risk of patient injury or death can be further reduced systematically [19]. To answer this question, it is important to know if the aforementioned alleged patient injuries and deaths occurred under similar circumstances. Reports of these adverse drug events have repeatedly stated that errors were made in programming drug concentration [9–14]. The relationship between drug concentration errors and patient safety is counterintuitive because setting the drug concentration at a *lower* level than intended can result in repeated *overdeliveries* of analgesic [5]. For example, if the vial concentration is set to 1.0 mg/mL and a dosage of 1.0 mg of morphine is requested, then the pump will infuse 1 mL of liquid (1.0 mg [desired mass of morphine] ÷ 1.0 mg/mL [programmed mass of morphine per unit volume of liquid in vial] = 1 mL [volume of liquid delivered to patient]). But if the actual concentration of the vial is 5.0 mg/mL, then an infusion of 1 mL of liquid will deliver 5 mg of morphine rather than the intended 1 mg (1 mL [volume of liquid delivered to patient] × 5.0 mg/mL [actual mass of morphine per unit volume of liquid in vial] = 5 mg [actual mass of morphine delivered to patient]).

<sup>3</sup>LifeCare is a registered trademark of Abbott Laboratories.

In addition to being counterintuitive, the effects of drug concentration errors are also potentially safety-critical because a single error can have multiple and continued negative effects. The value that is entered for drug concentration is used by the device to calculate how much liquid to infuse for every subsequent bolus dose, how much liquid to infuse for every subsequent PCA dose, how much liquid to infuse as continuously delivered analgesic (if chosen, see below), and when the 4-h safety limit should be invoked. Thus, entering a lower than intended drug concentration once can result in up to four enduring effects: (a) an overdelivery of every subsequent bolus dose, (b) an overdelivery of every subsequent PCA dose, (c) an overdelivery of the amount of continuously delivered analgesic, and (d) an increase in the amount of analgesic that the device will allow to be infused during a 4-h period. Any one of these effects alone can pose a threat to patient safety, but the last effect is particularly important because it relaxes one of the design features that PCA pumps have to safeguard patient safety [11].

Note that the counterintuitive relationship between drug concentration errors and patient safety reflects the physics of analgesic delivery. Thus, it holds for any PCA pump, not just the LifeCare 4100 (although some pump interface designs may be more likely to induce this type of error than others). In sum, the key implication from the adverse drug event reports is that minimizing the incidence of drug concentration errors is particularly critical to ensuring patient safety with PCA pumps. Can this class of errors be reduced by device redesign?

The design of the interface has been consistently cited as a contributing factor in adverse drug event reports with the LifeCare 4100 [9–14]. For example, the Emergency Care Research Institute stated: “the likelihood of this sort of misprogramming is increased by the fact that the user interface and programming logic of the pump are particularly complex and tedious. We believe that the likelihood of user error is increased by the repetitive and time-consuming programming process required by this pump” [11]. Thus, the LifeCare 4100 interface can be used as an experimental testbed to see if the application of human factors engineering techniques can make PCA pumps even less prone to error, particularly to drug concentration errors.

### *Human Factors Engineering*

Other safety-critical industries, such as nuclear power and aviation, have reduced human error by applying techniques from *human factors engineering* [20]. This discipline focuses on the interaction between technology, people, and their

work context. Human factors has sometimes been narrowly associated only with human–computer interface design guidelines, such as “minimize the load on users’ memory,” but a comprehensive human factors design process is much broader and consists of several activities, including [21]:

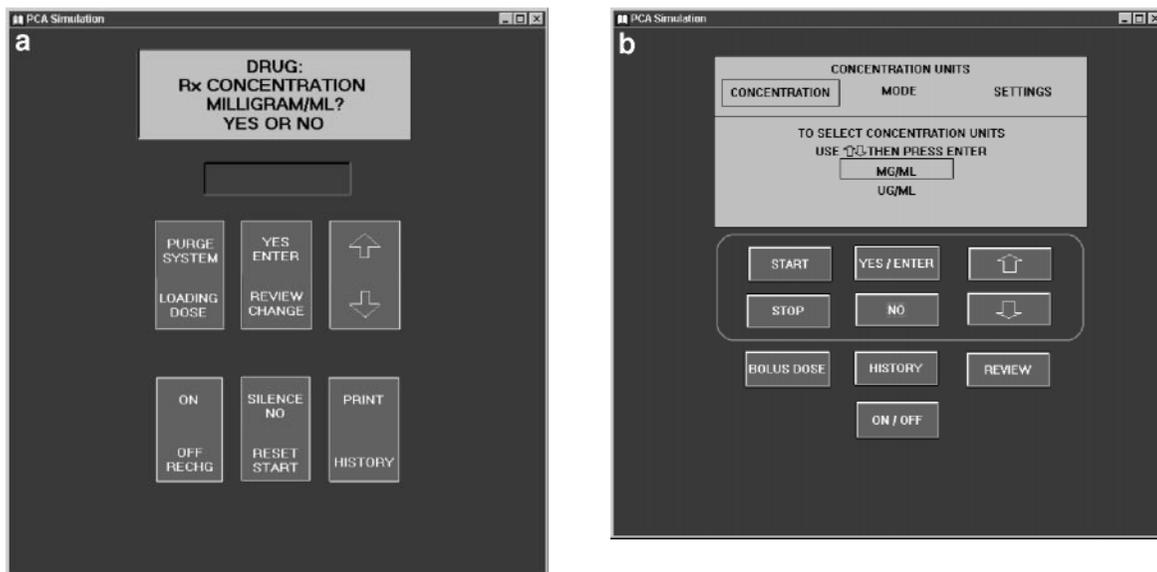
- Continuous involvement of representative users throughout the design process, starting right at the very beginning;
- Field observations to understand the actual conditions under which users work with technology;
- Task analysis to identify the job demands and performance bottlenecks faced by users;
- The use of human factors design guidelines (see below for examples) to ensure that technology is compatible with human capabilities and limitations;
- An iterative design process that leverages user feedback to improve initial design concepts;
- Experimental evaluations of detailed design prototypes with representative users based on objective measures of performance, not just subjective opinion;
- Post-market surveillance (e.g., review of adverse event reports) to identify unanticipated threats to safety that can be removed or reduced by redesign.

Given the success obtained in other safety-critical industries, medical researchers have often noted that a comprehensive human factors engineering process may be the key to improving patient safety [22–25]. However, direct empirical evidence of the potential benefits of applying human factors to medicine is still relatively scant. The primary aim of this article is to investigate whether PCA pump interfaces could be made less error-prone by redesigning them using a human factors engineering approach.

### *Previous Research*

Before learning of any adverse drug event reports, Lin and colleagues had already redesigned the commercially available interface for the LifeCare 4100, depicted in Fig. 1a, using human factors engineering techniques [26]. First, comments were obtained from recovery room nurses who had extensive experience with the device, field observations were conducted in the postsurgical recovery room at a teaching hospital, and a task analysis was conducted using bench tests. These activities identified information requirements that could serve as a basis for a device redesign.

Second, the commercially available interface was analyzed using some of the most basic design principles of human factors engineering:



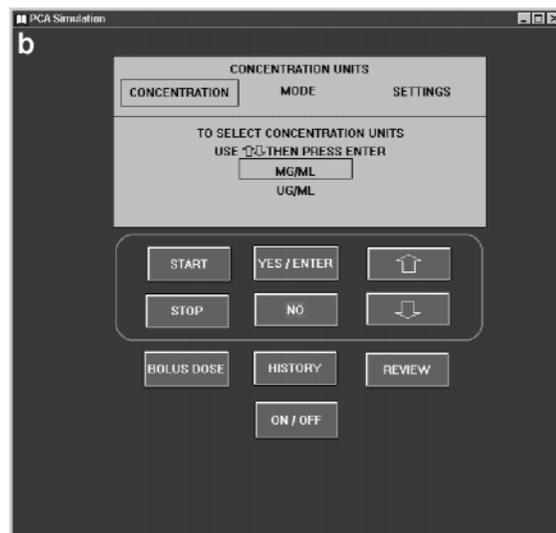
**FIG. 1.** (a) Old interface for LifeCare 4100, a simulation of the commercially available design. (b) New prototype interface for LifeCare 4100. Copyright 1997 by Laura Lin [27]. All rights reserved.

- Provide users with prompt, salient feedback after each action;
- Make the functions of the various controls clear and obvious;
- Make the displayed messages easy to understand;
- Minimize the load on the users' memory;
- Provide users with reliable shortcuts to increase efficiency;
- Provide clearly marked exits for the user to leave the system.

This analysis identified several aspects of the commercially available interface that could perhaps be improved (e.g., complex dialogue structure, limited feedback, inconsistent control functionality, confusing layout of controls, and confusing message displays).

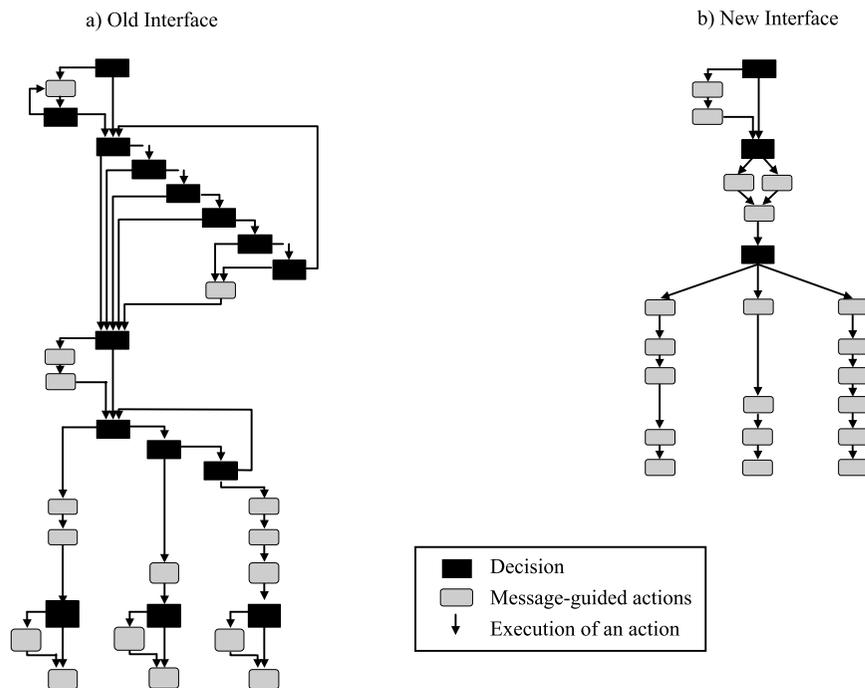
Third, Lin *et al.* redesigned the interface for the LifeCare 4100 using human factors design principles and the information requirements they had identified [26]. The result was a New prototype interface, shown in Fig. 1b, that has essentially the same functionality as the commercially available interface, but includes several design changes, such as fewer programming steps, an overview display showing the user's location in the programming sequence, more display feedback, a more logical layout of controls, and clearer wording for labels and messages.

Two concrete examples can be used to illustrate how a human factors approach helped to uncover aspects of the



commercially available design that could be improved. The LifeCare 4100 can be operated in three different modes: PCA, Continuous, and PCA + Continuous (see below for a description of each). In the commercially available interface, the programming sequence presents these options one at a time. At each step, nurses can either accept the currently displayed option and set the mode, or they can reject that option and go on to the next one. Lin *et al.* observed that the three modes are not independent; they are three alternative ways of operating the device. By presenting the three options serially, the commercially available design may be more complex than it needs to be, requiring users to make up to three decisions to select the device mode. Accordingly, Lin *et al.* used the principle of "Provide users with reliable shortcuts to increase efficiency" to redesign the interface. In the New interface, related options are presented in parallel wherever possible, so that only one decision is required to choose among the available related alternatives. Consequently, the programming sequence of the New interface appears simpler than that of the Old, as illustrated in Fig. 2. (Detailed descriptions of each step in these thumbnail sketches are provided by Lin [27].)

A second example of how human factors was used to identify opportunities for redesign involves the setting of drug concentration. In the factory-set configuration (see below), the LifeCare 4100 sequentially offers four default options for programming drug concentration (i.e., morphine



**FIG. 2.** Thumbnail sketch of the programming sequences: (a) For the Old interface, and (b) for the New interface. Copyright 1997 by Laura Lin [27]. All rights reserved.

1 mg/mL, morphine 5 mg/mL, morphine 0.5 mg/mL, and meperidine 10 mg/mL). Lin *et al.* noted that: “The default values provided on the current interface conflict with the standard operating values used at TGH [Toronto General Hospital]. For example, the most widely used concentration is 2.0 mg/mL. Users must currently toggle through four screens of default values before being able to enter this concentration. Making suitable default settings available would eliminate many of these unnecessary programming steps” [26]. Accordingly, Lin *et al.* again used the principle of “Provide users with reliable shortcuts to increase efficiency” to redesign the interface. The initially displayed concentration value was changed to match the value that is most frequently used at TGH. As a result, the first concentration value that users encounter in the programming sequence with the New interface would be more likely to be the desired value.

Finally, Lin *et al.* conducted an experiment comparing computer simulations of the commercially available and New prototype interfaces with nursing students as participants. The results showed that the New prototype led to statistically significant reductions in errors, programming times, and mental workload ratings [26].

The individual findings reviewed thus far have each been previously reported in the literature. By collecting and integrating this body of evidence in the same article for the first time, the significance and implications of these findings for patient safety become more clear. In the remainder of the article, we describe the primary novel contribution of the present work—an experiment comparing simulations of the commercially available and new interfaces for the LifeCare 4100 with professional nurses who already had extensive experience programming the commercially available interface. This activity is an essential part of a comprehensive human factors engineering process [21]. The most important question was whether the New interface would reduce the incidence of drug concentration programming errors.

## METHOD

### Participants

Twelve recovery room nurses from TGH participated in the study. They averaged 5.15 years (SD = 1.19) of clinical

experience with the LifeCare 4100. These nurses were on a rotational schedule in which they programmed PCA pumps for 1 to 3 weeks each month. When they were on the assigned rotation, a nurse would program the LifeCare 4100 from 2 to 10 times per day. Institutional review board approval and informed consent were obtained.

### Materials

Simulations of the New prototype and commercially available interfaces running on an IBM-compatible personal computer were used [26]. The simulation of the commercially available interface was a high-fidelity visual replica, and will be referred to as the Old interface. Its functionality was essentially the same as that of the commercially available interface (see below). The primary difference was that the Old interface required participants to input data using a mouse, whereas the commercially available interface requires nurses to use touch keypads. Given these minor differences, and the fact that functional fidelity plays a stronger role than physical fidelity in transfer of training for procedural cognitive tasks [28], the participants' extensive experience with the commercially available interface was expected to generalize to the Old interface.

Both the Old and New interfaces have eight programming functions. The *purge* function would allow users to remove air from the intravenous line before the line is connected to the patient. The *concentration* function specifies the numerical value and the units of the drug concentration (e.g., 1 mg/mL). The *bolus* function would allow users to administer a one-time infusion of analgesic (e.g., 1 mg). The *mode* function would allow users to choose between the three different modes in which the device can be operated. In the PCA mode, the device would deliver only discrete, patient-requested doses of analgesic. In the Continuous mode, the device would function as a normal infusion pump (i.e., provide a continuous delivery of analgesic with no possibility for patient-requested doses). In the PCA + Continuous mode, the device would function in a hybrid mode with a continuous delivery of analgesic as well as the possibility for additional delivery of discrete, patient-requested doses of analgesic. The *lockout interval* function specifies the minimum time between patient-administered doses (e.g., 5 min). The *4 h limit* function specifies the maximum amount of analgesic that could be delivered during any 4-h period (e.g., 30 mg). The *PCA dose* function specifies the amount of analgesic that would be delivered for each patient-requested dose (e.g., 1 mg). The *continuous dose* function specifies the amount of analgesic that would be continuously delivered per unit time (e.g., 1 mg/h).

The commercially available interface for the LifeCare 4100 can be configured to operate in several different ways [10]. The configuration in regular use at TGH and simulated by the Old interface is the factory-set standard. In this configuration, the pump sequentially offers four options for programming the concentration of the drug in the vial: morphine 1 mg/mL, morphine 5 mg/mL, morphine 0.5 mg/mL, and meperidine 10 mg/mL. At each step, nurses can either accept the currently displayed option and set the drug concentration, or they can reject that option and go on to the next one. If all of these are rejected, then several additional options are presented sequentially. This latter set of options can be selected if a drug other than morphine or meperidine is desired.

### Experimental Design

A  $2 \times 2 \times 3 \times 2$  mixed design was adopted with Order of training (Old first and New first) as the between-participants factor, and Interface (Old and New), Programming Task Mode (PCA, Continuous, and PCA + Continuous), and Repetition (1st and 2nd) as the within-participant factors. The orders of the Interfaces and the Modes were both counterbalanced.

### Procedure

Participants read a set of instructions on how to use one of the PCA interfaces and performed three practice tasks. Every participant then completed a total of six different tasks (two Repetitions of each of the three Mode types). For each task, participants were given a TGH PCA Order Form and were asked to program the interface according to the operating values specified on the Form. The same six Order Forms were used for both interfaces. The quantitative values of several programming parameters (e.g., drug concentration, PCA dose, continuous dose, lockout interval, and 4-h limit) varied across the six Order Forms. To reduce confusion between interfaces, participants were brought in for a second day, and were instructed and tested on the alternate interface using a comparable procedure.

### Dependent Variables

There is only one correct action sequence for each programming task, so an error was defined as any deviation from this gold standard, regardless of whether it was subsequently

corrected. Errors were classified according to the eight programming functions described earlier. Examples of errors include answering a YES/NO question incorrectly, entering an incorrect numerical value for one of the programming parameters, and choosing the incorrect option among several displayed alternatives. Task completion time was measured as the interval between the time that participants turned the simulation on and the time of the last programming action. To evaluate the degree of mental work that participants expended during each programming task, subjective mental workload ratings were collected after each trial using the NASA-TLX rating scale [29]. This scale is a sensitive and reliable measure of mental workload that is well-accepted and frequently used in the human factors engineering literature. The NASA-TLX method requires participants to rate their experience on six different subscales that have been found to contribute to mental workload (i.e., Mental Demands, Physical Demands, Temporal Demands, Own Performance, Effort, and Frustration). The participants' ratings on these individual subscales are then combined to obtain an overall workload rating for each participant for each trial. At the end of the experiment, participants were debriefed and their subjective preferences and qualitative comments were both recorded.

### Data Analysis

Errors and the subjective preference ratings were analyzed using  $\chi^2$  tests, except when there was a small number of observations, in which case a binomial test was used instead. Task completion time and workload ratings were analyzed using a mixed analysis of variance. All statistical tests that were significant at a level of  $P < 0.05$  are reported. No other tests reached statistical significance.

## RESULTS

Table 1 summarizes the error data. In total, more errors were made with the Old interface than with the New ( $\chi^2(1) = 6.1, P < 0.02$ ). Also, more errors were made in programming the mode function with the Old interface than with the New ( $\chi^2(1) = 4.6, P < 0.05$ ). Most importantly of all, no drug concentration errors were made with the New interface, whereas eight such errors were made with the Old (binomial test,  $P < 0.008$ ). Three of these errors were not detected and left uncorrected.

The statistically significant effects for task completion

TABLE 1  
Summary of Errors by Interface and by Programming Function

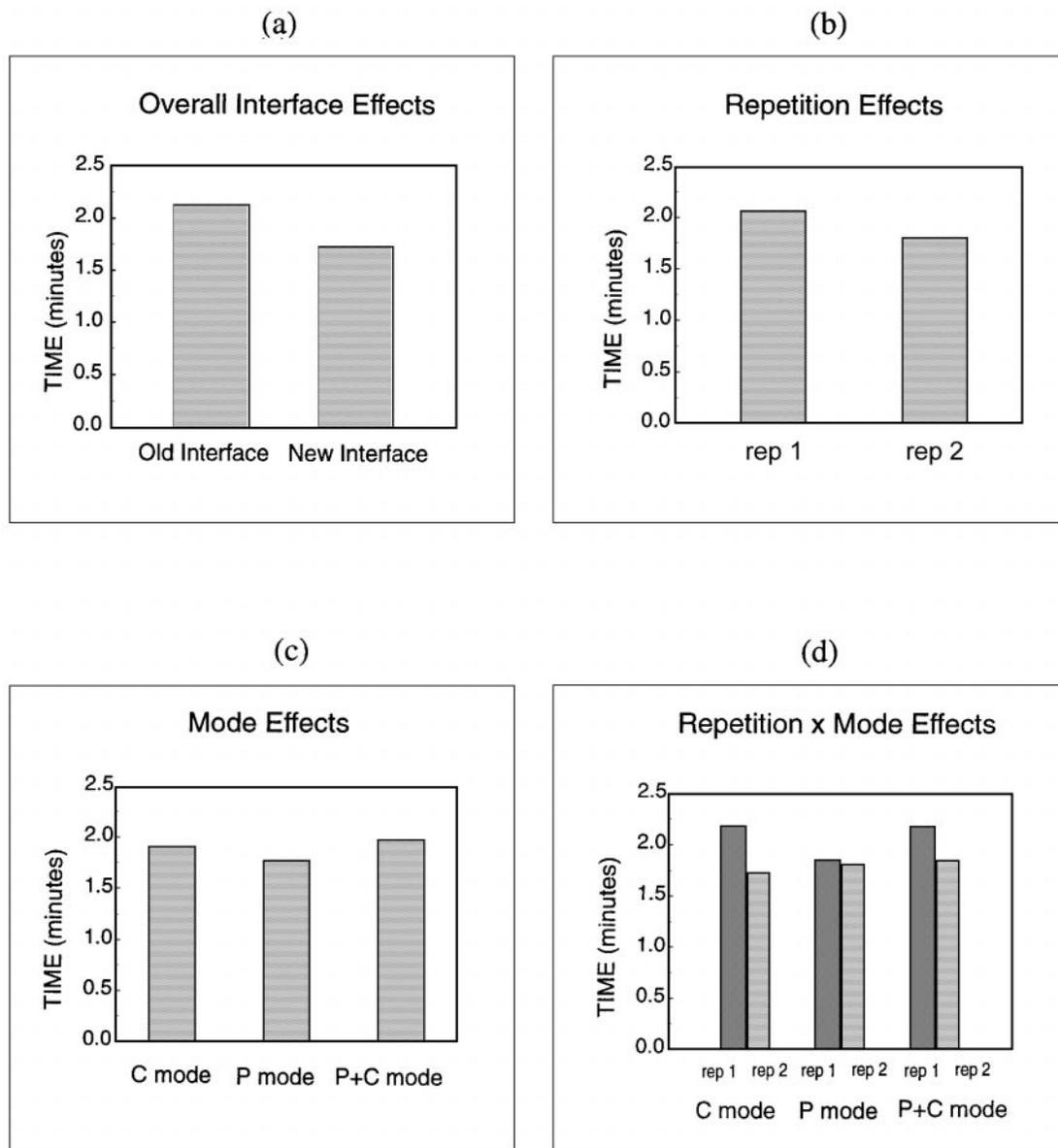
Function	Old	New
Purge	1	0
Concentration	8*	0
Bolus	4	8
Mode	11*	3
Lockout interval	0	0
4-h limit	1	0
PCA dose	4	1
Continuous dose	0	1
Total:	29*	13

\* Indicates a statistically significant effect,  $P < 0.05$ .

time are shown in Fig. 3. The New interface was faster than the Old ( $F(1, 10) = 12.17, P = 0.006$ ). Also, the second repetition was faster than the first ( $F(1, 10) = 21.81, P = 0.001$ ). In addition, the Modes differed, the PCA + Continuous Mode being the slowest and the PCA Mode being the fastest ( $F(2, 20) = 5.4, P = 0.013$ ). Finally, there was an interaction between Repetition and Mode ( $F(2, 20) = 5.64, P = 0.011$ ). The reduction in time with Repetition was smaller in the PCA Mode than in the other two Modes, probably because the nurses had the most experience with programming that Mode in clinical practice.

The statistically significant effects for mental workload are shown in Fig. 4. Workload was lower for the second repetition than for the first ( $F(1, 10) = 6.27, P = 0.03$ ). Also, there was a Repetition  $\times$  Mode  $\times$  Interface interaction ( $F(2, 20) = 8.62, P = 0.002$ ). The percentage reduction in workload with Repetition was larger with the New interface than with the Old for the Continuous and PCA + Continuous Modes. Participants seemed to find it progressively easier to program the two less familiar Modes with the New interface than with the Old.

Nine nurses favored the New interface and one preferred the Old ( $\chi^2(1) = 6.4, P < 0.02$ ). Two nurses expressed no preference. The qualitative comments provided by participants for the Old interface describe some of the problems that participants experienced initially in learning how to use the commercially available device in a clinical setting, the positive impact of frequency of use on ease of programming, and the importance of involving representative users in the design process. The comments for the New interface are generally more positive and consistent with the objective performance results, although there was one suggestion for how to improve the design of the bolus function.

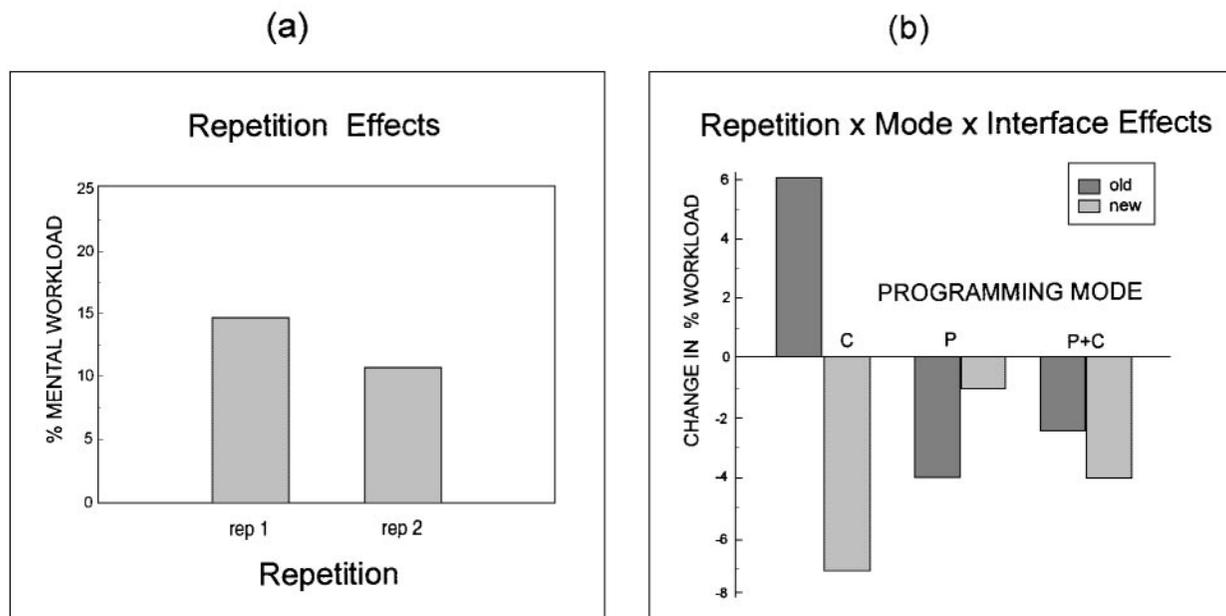


**FIG. 3.** Statistically significant effects for task completion time. Rep 1, first repetition; Rep 2, second repetition; P, PCA Mode; C, Continuous Mode; P + C, PCA + Continuous Mode. Copyright 1997 by Laura Lin [27]. All rights reserved.

## DISCUSSION

This experiment appears to be the first to evaluate the impact of a human factors engineering process on the redesign of a commercially available medical device with experienced users. The New prototype interface for the LifeCare

4100 eliminated drug concentration errors, whereas eight such errors were observed with the Old interface, three of which were uncorrected. This finding should be treated with caution due to the low number of observations per participant. Nevertheless, it is worth noting that the concentration errors were eliminated with the New prototype despite the



**FIG. 4.** Statistically significant effects for mental workload. Rep 1, first repetition; Rep 2, second repetition; P, PCA Mode; C, Continuous Mode; P + C, PCA + Continuous Mode. Copyright 1997 by Laura Lin [27]. All rights reserved.

fact that the nurses had only minimal exposure with this interface, but averaged over 5 years of experience with the commercially available interface. This result shows that presenting a new design to users who have extensive experience with an old design for the same device need not “cause a huge increase in errors that result from users exercising an over-learned, previously appropriate, response” [30].

In a clinical setting, the three uncorrected concentration errors could have resulted in repeated overdoses of analgesic to patients. For example, one participant mistakenly accepted the initially displayed concentration value of 1 mg/mL when the desired concentration value was 2 mg/mL. In a clinical setting, this error could have resulted in a twofold overdose of morphine every time a dose was delivered to the patient. Recall that the programmed concentration value is used to calculate when the 4-h limit should be invoked. Thus, this same concentration error could also have allowed 60 mg, rather than the intended limit of 30 mg, to be infused over any 4-h period, thereby relaxing one of the safety features of the device. Adverse event reports have repeatedly stated that concentration programming errors are an important potential threat to patient safety in PCA pump usage [9–14].

The New interface also led to faster completion times,

and was preferred by a strong majority of nurses. There was no significant difference in mental workload ratings between interfaces. The subjective comments obtained from the participants generally favored the New interface, and thus were in agreement with the objective performance data. For example, one participant stated that the New interface was simple, user friendly, and easier to learn to use than the Old interface. These subjective comments, and the observed reductions in programming errors and times, confirm the informal impression created by Fig. 2 that the New interface is simpler to program than the Old.

When combined with our earlier work, these results show the value of following a comprehensive human factors engineering process as opposed to only relying on human factors design principles alone. Lin *et al.* were unaware of any patient injuries and deaths with the LifeCare 4100, so they could not have based the New design on adverse drug event reports [26]. Furthermore, they had not conducted an experiment with professional nurses. Nevertheless, as mentioned earlier, Lin *et al.* redesigned the LifeCare 4100 interface so that the initially displayed concentration is a high, rather than a low, value because the higher concentration value was more frequently used in clinical practice.

Reviews of adverse events and experiments with representative users are essential parts of a human factors engineering system design process [21]. In this research, these two activities uncovered new safety-relevant evidence that had not been identified using design guidelines alone. This new evidence did not point to a need to change the prototype created by Lin *et al.* because the initially displayed concentration value had already been changed for reasons of efficiency. However, in other cases, new information identified by reviews of adverse events or testing with representative users may identify a need to make further design changes to the interface [31].

This research has limitations that suggest future research topics. First, the participants were recovery room nurses who program PCA pumps up to 10 times a day. Floor nurses also sometimes program PCA pumps, but typically far less frequently. Thus, it would be useful to replicate this experiment with floor nurses as participants. Given the participants' qualitative comments regarding the importance of frequency of use on ease of programming with the commercially available design, the impact of interface redesign may be even greater than that observed here. Second, nurses are frequently interrupted when they program PCA pumps in clinical practice, but they were not interrupted in this experiment. The impact of interruptions should be investigated. The benefits of the New interface may be even greater under such conditions because it provides more feedback than the Old. If nurses do not have to rely as much on memory, then they may be able to see, rather than remember, in what state they left the device after an interruption. Third, both simulated interfaces used a mouse for data input, whereas the commercially available device uses a keypad. More research is required to determine how the absolute level of performance would change if a keypad were used. Note, however, that the primary focus of this research was not the absolute level of performance of either interface, but rather the relative level of performance between interfaces. Therefore, the difference in input modality is not of great concern for our purposes because there is no reason to believe that using a mouse would affect one interface more than the other, particularly since Interface was a within-participants factor. Because both interfaces required all participants to use a mouse instead of a keypad, the safety and efficiency benefits of the New interface cannot be attributed to the input modality. Fourth, the experiment was conducted in a laboratory rather than in a field setting. Clinical trials must be conducted to see if the results obtained here generalize to actual use. Finally, this work only investigated one device, although the same process can be used to design, or redesign, virtually any biomedical device, product, or system.

## CONCLUSIONS

Adverse drug events are the leading threat to patient safety. This research showed that, by redesigning a commercially available PCA pump interface using a human factors engineering process, drug concentration errors and task completion times were both reduced for experienced users under laboratory conditions. The generalizability of these findings to other PCA pump interfaces, to other drug delivery devices, and to other medical products or systems can only be definitively established with more empirical research. However, the success of human factors engineering in improving safety in other domains, such as aviation and nuclear power, suggests that the benefits of this approach are likely to generalize widely [20].

To take just two biomedical informatics-related examples, previous human factors engineering research has shown that both an interface for a computer-based patient-monitoring system for use in cardiac anesthesia and an interface for a computer-based infusion device for terbutaline therapy to treat preterm labor suffer from unnecessarily complex dialogue structures that cause difficulties for device users [32, 33]. The research presented here goes two steps further by showing how these limitations might be systematically addressed by device redesign, and by suggesting that such changes may result in safety and efficiency improvements.

If these predictions are confirmed, then human factors engineering may have important implications for the design, regulation, and procurement of biomedical devices, products, or systems. Medical manufacturers might enhance patient safety by adopting a human factors engineering process in the design of their products. Government medical regulators might enhance patient safety by putting a greater emphasis on human factors engineering design criteria when making product approval and regulation decisions. Finally, hospital risk managers and procurement staff might enhance patient safety by adopting human factors engineering criteria—particularly user testing and evaluation—before making product purchasing decisions.

## ACKNOWLEDGMENTS

This research was funded by a grant from the Natural Sciences and Engineering Research Council of Canada to the second author. Thanks are due to the recovery room nurses at Toronto General Hospital for all of their time and invaluable input, and to Marilyn Sue Bogner, Klaus Christoffersen, Renée Chow, Michael Cohen, Pierre Duez, David

Gaba, Oscar Guerra, Greg Jamieson, Gill Hillel, John Lee, Matt Weinger, and the reviewers for their generous help. Laura Lin and Kim J. Vicente do not have any conflicts of interest to declare. After this research was completed, D. John Doyle provided assistance to plaintiff's lawyers whose clients were allegedly injured by Abbott's products.

## REFERENCES

1. Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, *et al.* The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II. *N Engl J Med* 1991; 324:377–84.
2. Bates DW, Spell N, Cullen DJ, Burdick E, Laird N, Petersen L, *et al.* The cost of adverse drug events in hospitalized patients. *J Am Med Assoc* 1997; 277:307–11.
3. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, *et al.* Incidence of adverse drug events and potential adverse drug events. *J Am Med Assoc* 1995; 274:29–34.
4. Heath ML, Thomas VJ. Patient-controlled analgesia: confidence in postoperative pain control. Oxford, England: Oxford Univ. Press, 1993.
5. Emergency Care Research Institute. Patient-controlled analgesic infusion pumps. *Health Devices* 2001; 30:157–85.
6. Callan M. An analysis of complaints and complications with patient-controlled analgesia. In Ferrante FM, Ostheimer GW, Covino BG, editors. Patient-controlled analgesia. Boston: Blackwell Scientific, 1990; 139–50.
7. McLeskey CH. Abbott addresses medication errors through advanced PCA technology. *APSF Newslett* 2000; 15:36–7. ([www.gasnet.org/societies/apsf/newsletter/2000/fall/07OpinionResponse.htm](http://www.gasnet.org/societies/apsf/newsletter/2000/fall/07OpinionResponse.htm), November 23, 2001)
8. Abbott Laboratories. LifeCare PCA Plus II Infuser, 2001. ([www.abbotthosp.com/PROD/pain/pcaplus.html](http://www.abbotthosp.com/PROD/pain/pcaplus.html), November 23, 2001).
9. Institute for Safe Medication Practices. Safety briefs. *ISMP Med Saf Alert* 1996; 1(12):2.
10. Institute for Safe Medication Practices. Safety briefs. *ISMP Med Saf Alert* 2000; 5(14):2.
11. Emergency Care Research Institute. Abbott PCA Plus II patient-controlled analgesia pumps prone to misprogramming resulting in narcotic overinfusions. *Health Devices* 1997; 26:389–391.
12. Institute for Safe Medication Practices. Evidence builds: lack of focus on human factors allows error-prone devices. *ISMP Med Saf Alert* 1999; 4(15):1–2.
13. Weinger MB, Pantiskas C, Wiklund ME, Carstensen P. Incorporating human factors into the design of medical devices. *J Am Med Assoc* 1998; 280:1484.
14. Doyle DJ, Vicente KJ. Patient-controlled analgesia. *Can Med Assoc J* 2001; 164:620.
15. Cullen DJ, Bates DW, Small SD, Cooper JB, Nemeskal AR, Leape LL. The incident reporting system does not detect adverse drug events: a problem for quality improvement. *Jt Comm J Qual Improv* 1995; 21(10):541–52.
16. Gardner S, Flack, M. Designing a medical device surveillance network. FDA report to Congress, 1999. ([www.fda.gov/cdrh/postsurv/medsun.html](http://www.fda.gov/cdrh/postsurv/medsun.html), November 23, 2001)
17. Classen DC, Pesotnik SL, Evans RS, Burke JP. Computerized surveillance of adverse drug events in hospital patients. *J Am Med Assoc* 1991; 266: 2847–51.
18. Jha AK, Kuperman GJ, Teich JM, Leape L, Shea B, Rittenberg E, Burdick E, Seger DL, Vander Vliet M, Bates DW. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. *J Am Med Inform Assoc* 1998; 5:305–14.
19. Vicente KJ. Patients perceived still at risk with PCA pump. *Anesthes Patient Saf Found Newslett* 2000; 15:60. (Available at [www.gasnet.org/societies/apsf/newsletter/2001/winter/20Opinion.htm](http://www.gasnet.org/societies/apsf/newsletter/2001/winter/20Opinion.htm), November 23, 2001)
20. Wickens CD, Gordon SE, Liu Y. An introduction to human factors engineering. Reading, MA: Addison-Wesley, 1997.
21. Association for the Advancement of Medical Instrumentation. Human factors design process for medical devices (ANSI/AAMI HE48:2001). Arlington, VA: AAMI, 2001.
22. Kohn LT, Corrigan JM, Donaldson MS. To err is human: building a safer health system. Washington, DC: National Academy Press, 1999.
23. Bogner MS. Human error in medicine. Hillsdale, NJ: Erlbaum, 1994.
24. Cooper JB, Newbower RS, Long CD, McPeck B. Preventable anesthesia mishaps. 1978; 49:399–406.
25. Leape LL. Error in medicine. *J Am Med Assoc* 1994; 272:1851–7.
26. Lin L, Isla R, Doniz D, Harkness H, Vicente KJ, Doyle DJ. Applying human factors to the design of medical equipment: patient-controlled analgesia. *J Clin Monitor Computing* 1998; 14:253–63.
27. Lin L. Applying human factors engineering to medical device design: an empirical evaluation of patient-controlled analgesia machine interfaces (unpublished Masters' thesis). Toronto: University of Toronto, 1997.
28. Patrick J. Training: research and practice. London: Academic Press, 1992.
29. Hart SG, Staveland LE. Development of the NASA-TLX: results of empirical and theoretical research. In Hancock PA, Meshkati N, editors. Human mental workload. New York: North Holland, 1988, 139–83.
30. Andre AD. Human factors consultant rebuts PCA pump danger charges. *Anesthes Patient Saf Found Newslett* 2000; 15:37–8. (Available at [www.gasnet.org/societies/apsf/newsletter/2000/fall/08Rebuttal.htm](http://www.gasnet.org/societies/apsf/newsletter/2000/fall/08Rebuttal.htm), November 23, 2001)
31. Garmer K, Liljegren E, Osvalder A, Dahlman S. Usability evaluation of a new user interface for an infusion pump developed with a human factors approach. Proceedings of the IEA/HFES 2000 Congress. Santa Monica, CA: HFES, 2000; 1-128–1-131.
32. Cook RI, Woods DD. Adapting to new technology in the operating room. *Hum Factors* 1996; 38:593–613.
33. Obradovich JH, Woods DD. Users as designers: how people cope with poor HCI design in computer-based medical devices. *Hum Factors* 1996; 38:574–92.