

# Evaluating and Predicting Patient Safety for Medical Devices with Integral Information Technology

Jiajie Zhang, Vimla L. Patel, Todd R. Johnson,  
Philip Chung, James P. Turley

## Abstract

Human errors in medical device use account for a large portion of medical errors. Most of these errors are due to inappropriate designs for user interactions, rather than mechanical failures. Evaluating and predicting patient safety in medical device use is critical for developing interventions to reduce such errors either by redesigning the devices or, if redesign is not an option, by training the users on the identified trouble spots in the devices. We developed two methods for evaluating and predicting patient safety in medical devices with integral information technology, then applied and tested them on several infusion pumps. The first method is a modified discount-usability method called heuristic evaluation. The method was used to evaluate and compare the safety of two 1-channel volumetric infusion pumps. The results show that heuristic evaluation, when modified for medical devices, is a useful, efficient, and low-cost method for evaluating patient safety features of medical devices through the identification of usability problems and their severities. The second method is an extended hierarchical task analysis (EHTA), devised to predict medical errors in medical device use. EHTA divides the task space between the external world of the device interface and the internal cognitive world of the user, allowing for descriptive predictions of potential user errors at the human device level. Its use is demonstrated in the analysis of two infusion pumps. The estimates of the likelihood of user errors with the two pumps are consistent with the corresponding reported use errors in the Federal Drug Administration (FDA)'s Manufacturer and User Device Experience (MAUDE) database, thus demonstrating the usefulness of this tool for predicting medical device use errors.

## Introduction

Since the medical error report from the Institute of Medicine in 2000<sup>1</sup> and subsequent increased funding for research on medical errors from U.S. government and private institutions, many studies on medical errors have been conducted and reported, including many from the medical informatics and cognitive science communities.<sup>2,3</sup> Our research has focused on developing methods for evaluating and predicting medical errors in medical device use from the perspectives of cognitive science and user interface design.<sup>4-6</sup>

# Report Documentation Page

*Form Approved  
OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

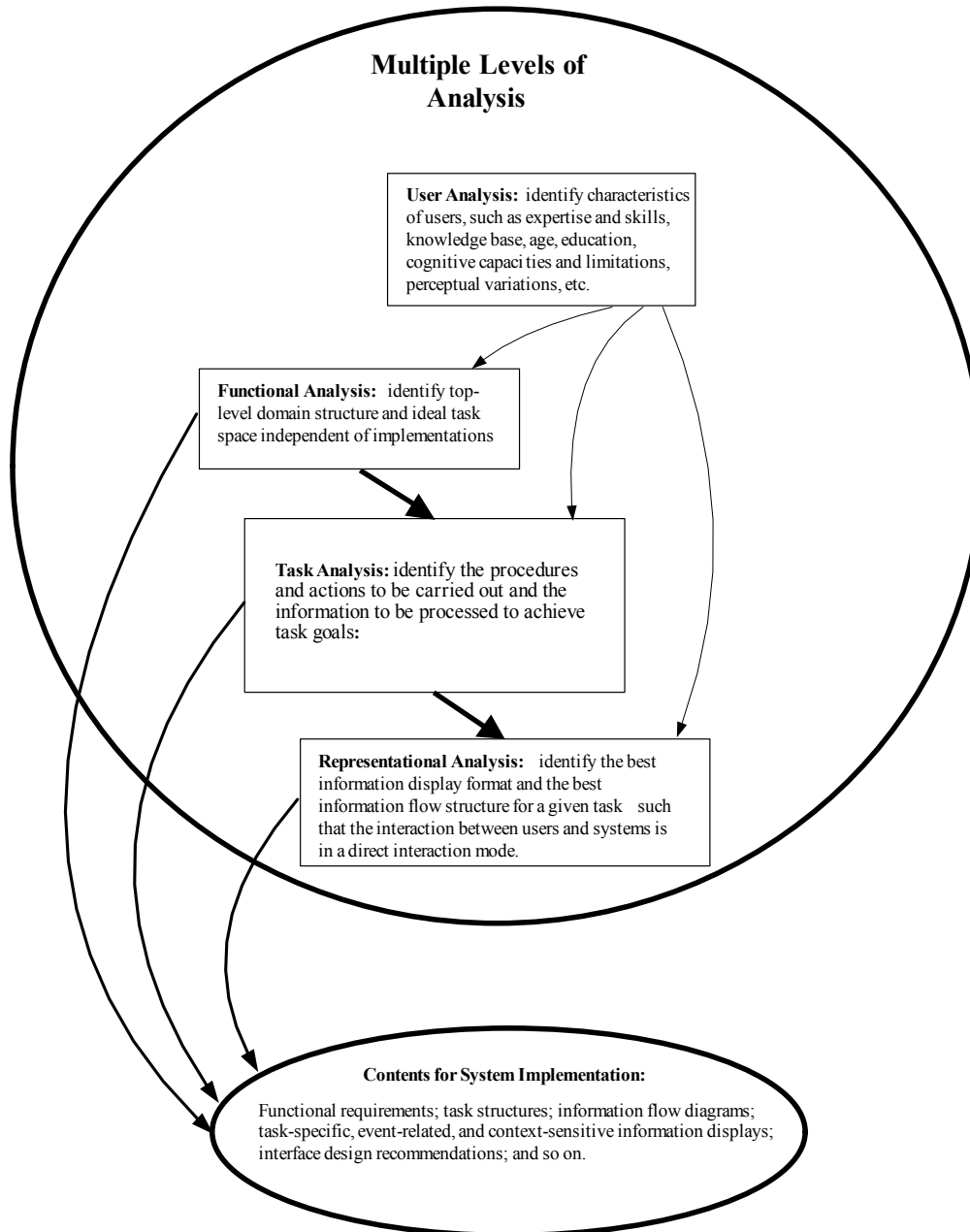
1. REPORT DATE <b>2005</b>	2. REPORT TYPE <b>N/A</b>	3. DATES COVERED <b>-</b>			
4. TITLE AND SUBTITLE <b>Evaluating and Predicting Patient Safety for Medical Devices With Integral Information Technology</b>		5a. CONTRACT NUMBER			
		5b. GRANT NUMBER			
		5c. PROGRAM ELEMENT NUMBER			
6. AUTHOR(S)		5d. PROJECT NUMBER			
		5e. TASK NUMBER			
		5f. WORK UNIT NUMBER			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Agency for Healthcare Research and Quality 540 Gaither Road, Suite 2000 Rockville, MD 20850</b>		8. PERFORMING ORGANIZATION REPORT NUMBER			
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)			
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)			
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release, distribution unlimited</b>					
13. SUPPLEMENTARY NOTES <b>Published in Advances in Patient Safety: From Research to Implementation. Volumes 1-4, AHRQ Publication Nos. 050021 (1-4). February 2005. Agency for Healthcare Research and Quality, Rockville, MD. <a href="http://www.ahrq.gov/qual/advances/">http://www.ahrq.gov/qual/advances/</a></b>					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>UU</b>	18. NUMBER OF PAGES <b>14</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

Medical device use errors are a common source of patient injury and death. Numerous research reports, medical error reports, and other documents show a clear link between usability problems and user error.<sup>7, 8</sup> For example, a physician treating an infant with oxygen set the flow knob between 1 and 2 liters per minute, then later noticed that the infant was not receiving any oxygen. Even though the knob rotated smoothly, the device was designed to deliver oxygen only when the knob was set on a number, not between numbers. Adding detents to the knob, so that it would click onto a number, and providing visible feedback of the rate of flow could have greatly decreased the chance of this type of error. This is a clear example of medical errors caused by the poor design of the device's user interface. FDA data show that nearly half of all device recalls were because of poor user interface design of the products.<sup>9, 10</sup> Other research also suggests that injuries resulting from medical device use errors far exceed injuries arising from device failures.<sup>11</sup> A poorly designed user interface, even operated by a well-trained and competent user, can lead to errors and operating inefficiencies. In response, the FDA has revised its Good Manufacturing Practice (GMP) regulations to include specific requirements for product usability.<sup>9</sup> It has also published guidelines for interface design and usability testing<sup>10</sup> and produced a continuing education article that specifically covers usability issues.<sup>12</sup> Some nonprofit organizations routinely perform and publish safety evaluations of medical devices. For example, Emergency Care Research Institute (ECRI) published several reports on the evaluations of a series of infusion pumps.<sup>13</sup> These reports are typically comprehensive, covering reviews of features, functionalities, and testing results, and making recommendations based on assessment. Human factors are also considered in many of the evaluations, although they are typically brief and at a high level without much detailed information. The methods in our current article can potentially expand the human factors evaluations in these types of reports, providing detailed descriptions of usability problems, factors causing the problems, suggested solutions, and even predictions of the likelihood of errors.

Human factors engineering is a discipline that seeks to design devices, software, and systems to meet the needs, capabilities, and limitations of the users, rather than expecting the users to adapt to the design. A complete human factors engineering analysis for medical devices or software systems includes four major components (Figure 1): user; functional; task; and representational analyses.<sup>14</sup> User analysis is the process of identifying the characteristics of existing and potential users, such as their expertise and skills; knowledge; educational background; cognitive capacities and limitations; perceptual variations; age-related skills; cultural background; personality; time available for learning and training; frequency of system use; and so on. User analysis can help us design systems that have the knowledge and information structure that match that of the users. Functional analysis is the process of identifying top-level domain structures and goals that are critical for the system, but largely independent of implementations. Task analysis is the process of identifying the procedures and actions to be carried out, and the information to be processed to achieve task goals. Representational analysis is the process of identifying an appropriate

information-display format for a given task performed by a specific type of user such that the interaction between the users and the system is as direct and transparent as possible. With direct interaction interfaces, users can directly, completely, and efficiently engage in the primary tasks they intend to perform, not the housekeeping interface tasks that are barriers between users and systems. The file browser in Microsoft Windows uses a direct interaction interface to move, delete, and rename files, whereas command line systems (e.g., MS DOS) do not.

**Figure 1. Four components of a complete human factors engineering analysis**



These four types of analyses, when combined and applied to a single product, can reveal the full range of usability issues, which are essential for an understanding of patient safety implications of the product. In this chapter, we describe two methods we developed:<sup>4, 15</sup> heuristic evaluation and extended hierarchical task analysis (EHTA). They correspond to two of the four analyses of human factors engineering—heuristic evaluation is largely a type of representational analysis and extended hierarchical task analysis is a typical task analysis.

We focus on heuristic evaluation because it has been shown to be one of the most cost-effective methods of finding usability problems. This method is for the evaluation of usability problems in medical devices. Through the identification of usability problems, we can indirectly identify medical devices' potential trouble spots that are likely to cause medical errors. We will describe the usability heuristics, the scale for severity rating of usability problem, and the procedure of carrying out a heuristic evaluation. Then we will demonstrate how to use this method by applying it to the evaluation of two 1-channel volumetric infusion pumps.

Likewise, we focus on EHTA because it can reveal the deep structures of a task that typically show the root causes of a usability problem. This is a method developed to predict medical errors with medical devices. EHTA divides the task space into the external world of the device interface and the internal cognitive world of the users, allowing for descriptive predictions of potential user errors at the human-device level. We will demonstrate this method by applying it to the comparison of the likelihood of medical errors with two infusion pumps.

## **Heuristic evaluation**

Heuristic evaluation is an evaluation technique that identifies major usability problems of a product.<sup>16-19</sup> This method has become extremely popular in the realm of usability evaluation due to its low cost, low time commitment, and ease of application.<sup>18</sup> This technique typically requires three or more expert usability evaluators to independently apply a set of usability heuristics to a product, identify violations of the heuristics, and assess the severity of each violation. In general, evaluators can conduct the evaluation in a few hours with minimal training. For a complex medical device such as the infusion pump, the evaluators should also have some clinical knowledge about the devices. Evaluators should already possess this knowledge or can obtain it through training at a level that is sufficient for the understanding and use of the device. Ideally, double experts trained in both usability and the target clinical domain should be the evaluators. In reality, such double experts are in short supply and the priority is typically given to the expertise of usability over the expertise of the clinical domain.

Heuristic evaluation has been traditionally used to evaluate web sites, as well as desktop software applications, and it is typically used to point out software interface difficulties to be addressed in the design process. It can be applied to paper or electronic mock-ups or prototypes, as well as completely implemented

designs. We modified this method to address three issues in the evaluation of medical devices.<sup>4, 20</sup> First, we used it to discover usability problems that are likely to cause medical errors. Second, we used it for comparison of patient safety features of alternative medical devices, which is often helpful in the purchasing process of medical devices. Third, we tried to demonstrate that heuristic evaluation is a good tool for medical device manufacturers to improve the patient safety features of their products during the design and redesign processes. We will describe this method and its application in the evaluation of two infusion pumps.

## Usability heuristics

Nielsen<sup>18</sup> described 10 major heuristics that should be followed for good user interface design. Shneiderman<sup>21</sup> also described 8 golden rules that all good user interface designs should feature. Based on the work of Nielsen and Shneiderman, we created 14 heuristics—aptly named Nielsen-Shneiderman heuristics—with a focus on medical devices.<sup>4</sup> More details of these heuristics are provided by Zhang, et al;<sup>4</sup> here, we only give a top-level description:

1. [**Consistency**] **Consistency and standards.** Users should not have to wonder whether different words, situations, or actions mean the same thing. Standards and conventions in product design should be followed.
2. [**Visibility**] **Visibility of system state.** Users should always be informed what is going on with the system through appropriate feedback and display of information.
3. [**Match**] **Match between system and world.** The image of the system perceived by users should match the model the users have.
4. [**Minimalist**] **Minimalist.** Any extraneous information is a distraction and a slowdown.
5. [**Memory**] **Minimize memory load.** Users should not be required to memorize a lot of information to carry out tasks. Memory load reduces users' capacity to carry out the main tasks.
6. [**Feedback**] **Informative feedback.** Users should be given prompt and informative feedback about their actions.
7. [**Flexibility**] **Flexibility and efficiency.** Users always learn and users are always different. Give users the flexibility of creating customization and shortcuts to accelerate their performance.
8. [**Message**] **Good error messages.** The messages should be informative enough such that users can understand the nature of errors, learn from errors, and recover from errors.
9. [**Error**] **Prevent errors.** It is always better to design interfaces that prevent errors from happening in the first place.
10. [**Closure**] **Clear closure.** Every task has a beginning and an end. Users should be clearly notified about the completion of a task.

11. [**Undo**] **Reversible actions.** Users should be allowed to recover from errors. Reversible actions also encourage exploratory learning.
12. [**Language**] **Use users' language.** The language should be always presented in a form understandable by the intended users.
13. [**Control**] **Users in control.** Don't give users the impression that they are controlled by the systems.
14. [**Document**] **Help and documentation.** Always provide help when needed.

The heuristics are used to check the interface of the device design. If a heuristic is violated, it is given a severity rating based on the following scales:<sup>17</sup>

- 0** = Not a usability problem at all.
- 1** = Cosmetic problem only. Need not be fixed unless extra time is available.
- 2** = Minor usability problem. Fixing should be given low priority.
- 3** = Major usability problem. Important to fix and should be given high priority.
- 4** = Usability catastrophe. Imperative to fix before product can be released.

In general 3–5 usability experts independently evaluate the user interface of a product and each of them generates a separate list of heuristic violations according to the 14 heuristics described above. A single usability problem identified by an evaluator can be violations of multiple heuristics, which means that the number of heuristic violations is typically more than the number of usability problems identified. For example, the oxygen flow control-knob problem described earlier violates consistency and standards (most devices with smoothly rotating knobs work at any position); visibility of system status (there was no indication of oxygen flow); match between the system and the world (users expect smoothly rotating knobs to work at all positions); and prevent errors (the design of the knob and lack of feedback increase the chance of error). In other words, the oxygen flow control is one usability problem corresponding to four heuristic violations. Once each evaluator has identified potential usability problems, the separate lists are compiled into a single master list. The master list is then given back to the evaluators who independently assess the severity of each violation. The ratings from the individual evaluators are then averaged. Once the master list with severity ratings is generated, it can be sorted in different ways, typically by places of occurrences of violations and by severity ratings.

Heuristic evaluation is relatively easy to do, even for those not trained in usability. We have found that two to three hours of training, combined with clear examples, and a practice evaluation with feedback, is often sufficient to begin using the technique. Typically one evaluator can only catch 35 percent of the usability problems, but 3–5 evaluators can detect 60-75 percent.<sup>18</sup> Heuristic evaluation also has a number of limitations. It does not indicate the elements of the interface that correctly follow usability guidelines. Nor does it reveal major

missing functionality. The second method we describe in this chapter, Extended Hierarchical Task Analysis, partially remedies this limitation.

## Heuristic evaluation of two 1-channel infusion pumps

We applied the 14 heuristics and the procedures to evaluate the usability problems of two 1-channel volumetric infusion pumps from two different vendors.<sup>4</sup> We focused on identifying usability problems that might be potential triggers for medical errors. We also compared the two infusion pumps in their usability and patient safety features.

Figure 2 shows the numbers of heuristic violations for the two pumps across the 14 heuristics. For Pump 1, heuristics were violated a total of 192 times. Consistency and visibility were the two most frequently violated heuristics (53 and 28, respectively). Feedback and match were the next most common violations (22 and 21). These four heuristics account for 64 percent of the violations. For Pump 2, heuristics were violated a total of 121 times. Visibility was the most frequently violated heuristic (29 violations). Memory and consistency were the next most common violations (19 and 17, respectively). These three heuristics comprised 54 percent of the violations. An example of a violation of the visibility heuristic would be: “When the ‘enter’ button is not pressed, after entering part or all of the value for ‘Rate’ and ‘VTBI’ (Volume to be Infused), a message appears that reads ‘complete entry.’” It is not clear what it means. A better phrasing would be “Press ‘enter’ to confirm value.” In this case, users are apt to become confused by no clear delineation as to what action would come next. In terms of the quantity of heuristic violations, the results in Figure 2 indicate that Pump 1 has more usability problems and thus may accordingly have a higher chance of generating medical errors.

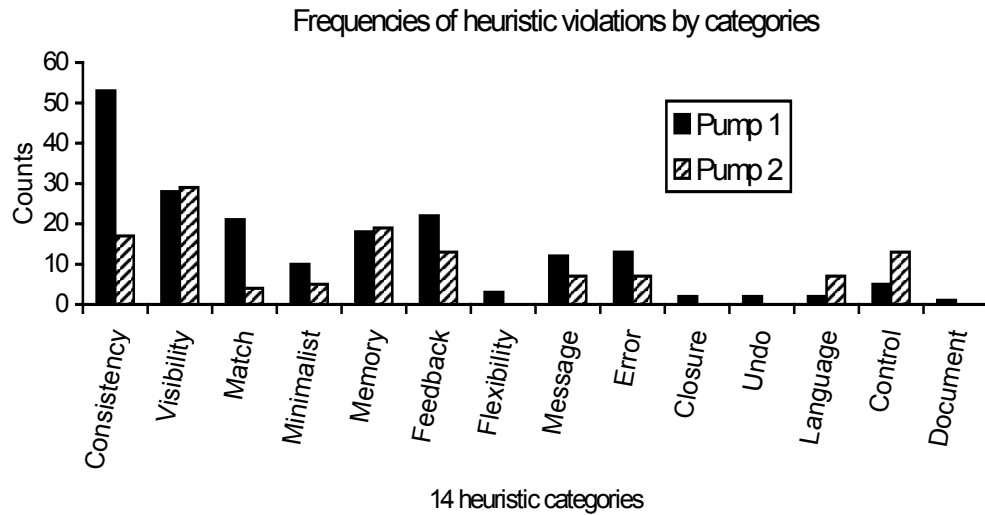
Figure 3 summarizes the severity of the problems found in Pump 1 and Pump 2. The severity ratings were divided into four regions: a severity rating equal or above 3.5 is catastrophic; a severity rating equal or above 2.5 but below 3.5 is major; a severity rating equal or above 1.5 but below 2.5 is minor; and a severity rating below 1.5 is cosmetic. For Pump 1, there were 2 catastrophic, 38 major, 49 minor, and zero cosmetic usability problems. For Pump 2, there was 1 catastrophic, 26 major, 26 minor, and zero cosmetic usability problems. In terms of the severity of usability problems, the results in Figure 2 indicate that Pump 1 has a larger number of more severe usability problems and thus it is likely to cause more medical errors than Pump 2. This conclusion is consistent with the measure in terms of the quantity of heuristic violations: Pump 1 has more heuristic violations than Pump 2. In short, Pump 1 has not only a larger number of heuristic violations in total, but also a larger number of more severe heuristic violations.

## Summary of heuristic evaluation

As a discount usability technique, heuristic evaluation is easy to use and master; efficient; effective; and useful. It can be used to identify a great portion of major usability problems in a product in a timely manner with reasonable cost.



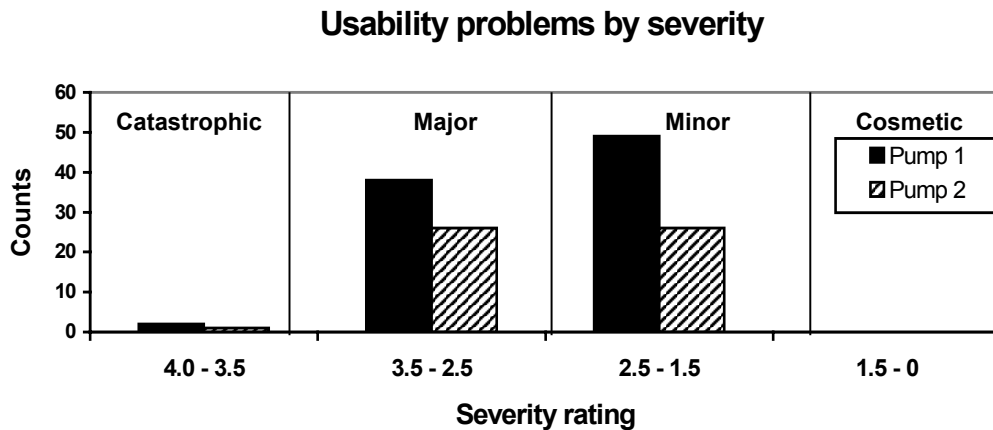
Figure 2. This figure shows the heuristic violations in two 1-channel infusion pumps



See text for details.

Reprinted with permission from Elsevier from Zhang J, Johnson TR, Patel VL, et al., Using usability heuristics to evaluate patient safety of medical devices. J. Biomed Inform. 2003; 36(1-2):23-30.

Figure 3. Severity ratings of usability problems for the two 1-channel infusion pumps



See text for details.

Reprinted with permission from Elsevier from Zhang J, Johnson TR, Patel VL, et al., Using usability heuristics to evaluate patient safety of medical devices. J Biomed Inform. 2003; 36(1-2):23-30.

Human errors in medical device use are largely due to interface design problems that can be potentially addressed through user-centered design. Since the quantity and severity of usability problems are usually correlated with the frequency of medical errors, heuristic evaluation is a method for indirectly assessing patient safety features in medical devices. Although it is limited in its scope of coverage of the full range of patient safety related features in medical devices, it is a

practical tool that should be adopted by medical device manufacturers for the design and modification of medical devices, and by health care institutions for the evaluation of medical devices. Usability engineering has become a niche industry with many professionals trained in human factors and related disciplines. Medical device manufacturers should routinely use the services offered by the usability industry or perform usability evaluations in-house by hiring usability specialists. The second method we describe below will complement the heuristic evaluations to address a different set of issues in medical errors.

## Extended hierarchical task analysis

The heuristic evaluation method we described in the last section is mainly for the evaluation of usability problems that may potentially cause medical errors. In this section, we describe an extended hierarchical task analysis method with a focus on predicting medical errors in medical device use.<sup>5</sup> While the prospect of a bulletproof method for predicting human error given an interface and task still seems distant, the current state of theory and research in the field of human-computer interaction (HCI) and cognitive science seems sufficient to provide much improvement over current techniques. These techniques will be priceless if successfully applied in the medical arena where human lives are consistently at stake. This work represents one such effort.

### Hierarchical task analysis and its extension

Hierarchical task analysis (HTA), one of the most widely used forms of task analysis, involves describing a task as a hierarchy of tasks and subtasks.<sup>22</sup> The first three columns of Figure 4 show a traditional hierarchical task analysis. First, major tasks are identified. Then the subtasks of each major task are identified, and subtasks of the subtasks are identified until meaningful subtasks are exhausted. This process establishes a hierarchy. Each task and subtask in the hierarchy is assigned a code that indicates the level of the task and the sequence in the task. In the example in Figure 4, the first level tasks are coded as 1.xx, 2.xx, etc.; the second level tasks are coded as 1.1x, 1.2x, etc.; and so on. By doing a traditional HTA, the structures of the tasks and the interrelations of tasks and subtasks can be explicitly described. This is a very useful process for the understanding and design of any user interfaces.

However, the traditional HTA is limited. It does not provide enough details about the cognitive processes and information processing that are typically crucial for user interface design. In our extended hierarchical task analysis (EHTA), we added a few more steps that are especially useful for the prediction of medical errors. In Figure 4, the columns after the first three are added steps in EHTA. These added steps are described below.

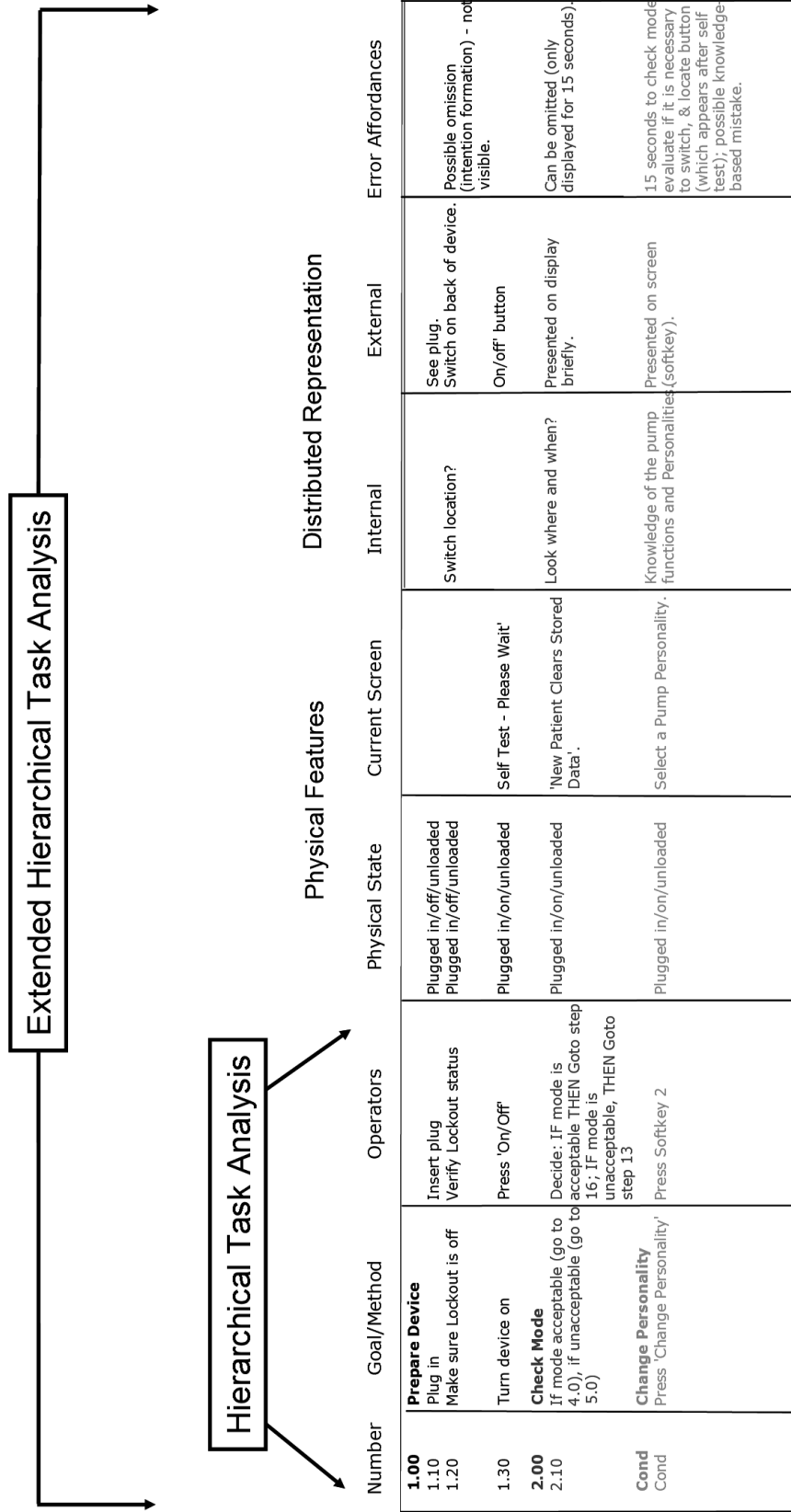
The physical features of a device are added for each step identified in the traditional HTA. These help determine the representations that are analyzed next. For example, a given state might indicate that the device is powered on. More detailed information can be provided, such as auditory cues, display messages,

soft key information, etc. Previous state and display information are also recorded, since these dictate the context or priming the user is given as they proceed to the next step. This information is used to decide what must be done on the subsequent screen, and thus has a great effect on the user's behavior.

The next step is to analyze whether the information needed to carry out each step in a task is represented as internal or external, using the distributed representation theory developed by Zhang and Norman.<sup>23</sup> Internal representation is information stored in users' memory, whereas external representation is available on the displays or other remote medium or devices. In the infusion pump example, internal representation is declarative knowledge or mental operators needed by a user to complete that step of the task, and external representation is the information available from the columns for soft keys, auditory warnings, display messages, etc., needed by the user to execute that step. According to Zhang and Norman, external representation—where the information is present on the device itself rather than in the head of the operator—is thought to drastically reduce or eliminate human error. Internal representation depends on cognitive processes for retrieval, which is error prone, but external representation relies on more efficient perceptual processes, hence mitigating cognitive load. In order to accurately evaluate this representation, it is necessary to consider all information presented in this manner. This includes auditory warnings, text messages, LEDs, physical controls, etc. Research shows that the specific wording and presentation of labels and feedback messages have a great impact on the actions users decide to take.<sup>24</sup> Listing all state and transitional information makes it possible for the evaluator to identify deficiencies in the external task space.

The next step is to analyze the error affordance for each step of a task. Error affordance is determined by a set of task characteristics found by psychological research to cause error.<sup>25, 26</sup> For example, an isolated step in a task affords error, such as the step of entering a hard key to find the “secondary” mode when programming an intravenous piggyback (IVPB) on an infusion pump. In general, the more action opportunities at each step, the higher the error affordance, i.e., the more likely a wrong action is selected. Error affordance is also associated with internal and external representations: the more information internal, the higher the error affordance because internal representations are more error prone. In the next subsection, we demonstrate EHTA by using it to analyze and compare two infusion pumps.

Figure 4. Traditional hierarchical task analysis and extended hierarchical task analysis. The example shows the first few steps in operating an infusion pump.



## An example of extended hierarchical task analysis

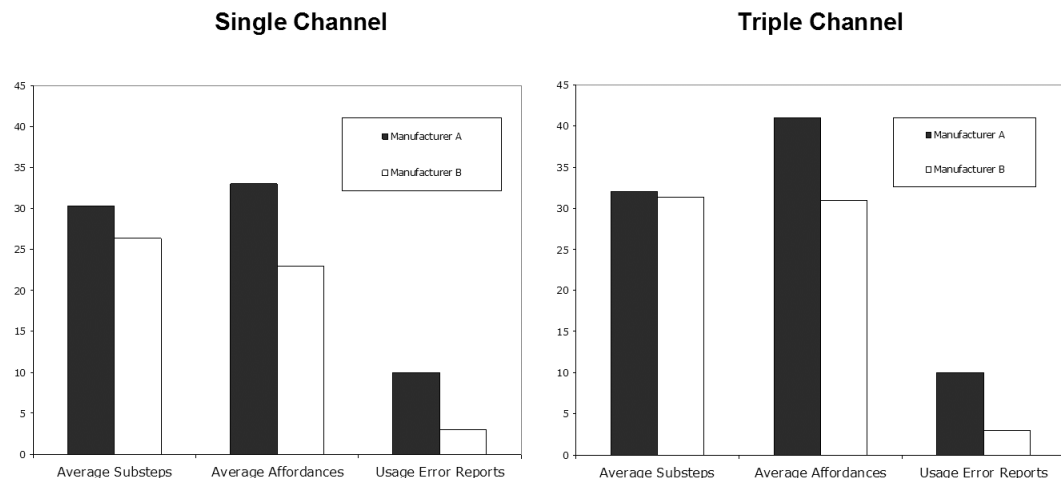
To make the methodology of EHTA more concrete, we did a sample analysis using 1-channel and 3-channel pumps from Manufacturer A and Manufacturer B.<sup>5</sup> Pumps from Manufacturer A are widely used, well recognized and touted for their human-centered design; the pumps are claimed to have several human factors improvements over the previous technology. Nonetheless, a quick search through FDA reports makes it clear that they fall short of being problem-free. Pumps from Manufacturer B are a competing product in the market.

Two 1-channel and two 3-channel pumps from the two manufacturers were analyzed on two different tasks to show their intrinsic differences. Task 1 was to infuse 1,000 ml normal saline at 125 ml per hour. Task 2 was to infuse 1,000 ml of normal saline at 125 ml per hour as a continuous infusion; give 1gm Ceftaxidime (Fortaz) IVPB every 8 hours; and administer over 30 minutes.

In the left panel of Figure 5, the first two columns show the average numbers of steps needed to perform Task 1 and Task 2; the second two columns the numbers of error affordances for 1-channel Pump A and Pump B. The third two columns show the actual usage error data from FDA’s MAUDE database (2001 and 2002). It shows that Pump A requires more steps than Pump B, and Pump A has more error affordances than Pump B. In addition, this pattern is consistent with the FDA data on the actual error cases for Pump A and Pump B. The right panel of Figure 5 is for the 3-channel pumps by Manufacturers A and B. The results are similar to those for the 1-channel pumps. Caution should be exercised here: there is the possibility of poor correlation between the data from EHTA and the data in the MAUDE database because the exact market share information of Pump A and Pump B is unknown to us. At the local Houston market where this study was conducted, it appears that Pump A was more popular than Pump B.

This analysis example shows that EHTA is a reasonably good method for predicting the likelihood of medical errors for a specific medical device. It

**Figure 5. Comparison of infusion pumps by extended hierarchical task analysis and data from FDA database. See text for details.**



established a strong correlation between error affordance and error frequency. However, our method is not foolproof for predicting human errors given an interface. Furthermore, it remains highly dependent upon the skill of the evaluator. Even so, it at least draws all major cognitive factors into consideration, extending present techniques and traditional treatments of human error to a deeper cognitive level. And our methodology is practical; as demonstrated, simply tabulating the external features of an interface side-by-side with the steps required can allow for quick visual comparisons of different pump models.

## Conclusion

In this paper, we described two human factors engineering methods for the evaluation and prediction of medical errors in medical device use. Human errors in medical device use account for a large proportion of medical errors. Most of these errors are due to inappropriate designs for user interactions, rather than mechanical failures. Evaluating and predicting patient safety in medical device use is critical for developing interventions to reduce such errors either by redesigning the devices or, if redesign is not an option, then by training the users on the identified trouble spots in the devices. We have shown initial success of these two methods for infusion pumps. We believe that these two methods, with little or no modification, can be applied to health information systems as well as other medical devices for a quick and inexpensive evaluation. With further refinement and validation of the methods, they have the potential to become solid tools for manufacturers, purchasers, and consumers to evaluate patient safety issues in various health related products.

## Acknowledgments

This project was sponsored by U.S. Department of Army under Cooperative Agreement #DAMD17-97-2-7016. The content of the information does not necessarily reflect the position/policy of the Government or NMTB, and no official endorsement should be inferred. It was also supported in part by grant P01 HS11544 from the Agency for Healthcare Research and Quality, and by grant No. NCC 2-1234 from National Aeronautics and Space Administration. We would like to thank Asma Ali, Julie Brixey, Daniel Page, and Xiaozhou Tang for their contributions.

## Author affiliations

School of Health Information Sciences, University of Texas Health Science Center at Houston (JZ, TRJ, JPT); Laboratory of Decision Making and Cognition, Department of Biomedical Informatics, Columbia University, New York (VLP); Department of Psychology, Rice University, Houston (PC).

*Address correspondence to:* Jiajie Zhang; The University of Texas School of Health Information Sciences, 7000 Fannin Suite 600, Houston, TX 77030. Phone: 713-500-3922; fax: 713-500-3929; e-mail: jiajie.zhang@uth.tmc.edu.

## References

1. Kohn LT, Corrigan JM, Donaldson MS, editors. To err is human: building a safer health system. A report of the Committee on Quality of Health Care in the America, Institute of Medicine, Washington, DC: National Academy Press; 2000.
2. Lenert LA, Bakken S. Enabling patient safety through informatics: works from the 2001 AMIA Annual Symposium and Educational Curricula for Enhancing Safety Awareness. *Journal of American Medical Informatics Association* 2002;9(6):S1–S132.
3. Patel VL, Bates DW. Cognition and measurement in patient safety research (guest editorial). *Journal of Biomedical Informatics* 2003;36(1–2):1–3.
4. Zhang J, Johnson TR, Patel VL, et al. Using usability heuristics to evaluate patient safety of medical devices. *Journal of Biomedical Informatics* 2003;36(1–2):23–30.
5. Chung PH, Zhang J, Johnson TR, et al. An Extended hierarchical task analysis for error prediction in medical devices. In: *AMIA Annu Symp Proc.* 2003;165–9.
6. Zhang J, Patel VL, Johnson TR, et al. A cognitive taxonomy of human errors in medicine. *Journal of American Medical Informatics Association* (in review).
7. Obradovich JH, Woods DD. Users as designers: how people cope with poor HCI design in computer-based medical devices. *Human Factors* 1996;38(4):574–92.
8. Lin L, Isla R, Doniz K, et al. Applying human factors to the design of medical equipment: patient-controlled analgesia. *J Clin Monit Comput* 1998;14(4):253–63.
9. Food and Drug Administration. Human factors implications of the new GMP rule. Overall Requirements of the New Quality System Regulations. [www.fda.gov/cdrh/humfac/hufacimp.html](http://www.fda.gov/cdrh/humfac/hufacimp.html) (Accessed 2004 Dec.)
10. Sawyer D, Aziz KJ, Backinger CL, et al. Do it by Design: an introduction to human factors in medical devices. In: U.S. Department of Health and Human Services, Public Health Service, Food and Drug Administration, Center for Devices and Radiological Health: Washington, DC; 1996.
11. Cooper JB, Newbower RS, Long CD, et al. Preventable anesthesia mishaps: a study of human factors. *Anaesthesia* 1978;46(1):399–406.
12. Food and Drug Administration. Improving patient care by reporting problems with medical devices: a MedWatch continuing education article 1997; <http://www.fda.gov/medwatch/articles/mdr/mdr.pdf>. (Accessed 2004 Dec.)
13. Emergency Care Research Institute. Evaluation: General-purpose pumps. *Health Devices* 2002;31(10).
14. Zhang J, Patel VL, Johnson KA, et al. Designing human-centered distributed information systems. *IEEE Intelligent Systems* 2002;17(5):42–47.
15. Chuah J, Zhang J, Johnson TR. The representational effect in complex systems: a distributed representation approach. In: *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*; Philadelphia 2000.
16. Nielsen J, Molich R. Heuristic evaluation of user interfaces. In: *Proceedings of ACM CHI'90*; pp.249–56.
17. Nielsen J, Mack R, editors. *Usability inspection methods*. New York: John Wiley & Sons, Inc.; 1994.
18. Nielsen J. *Usability engineering*. Boston: Academic Press; 1994.
19. Nielsen J. Finding usability problems through heuristic evaluation. In: *Proceedings of ACM Conference on Human Factors in Computing Systems: CHI'92, Monterey, CA*. New York: ACM Press; 1992. pp. 372–80.
20. Graham MJ, Kubose TK, Jordan D, et al. Heuristic evaluation of infusion pumps: implications for patient safety in intensive care units. *International Journal of Medical Informatics* 2004; 73(11–12):771–9.
21. Shneiderman B. *Designing the user interface*. 3rd ed. Reading, MA: Addison-Wesley; 1998.
22. Kirwan B, Ainsworth LK. *A guide to task analysis*. London: Taylor & Francis; 1992.
23. Zhang J, Norman DA. Representations in distributed cognitive tasks. *Cognitive Science* 1994;18(1):87–122.
24. Polson PG, Lewis CH. Theory-based design for easily learned interfaces. *Human-Computer Interaction* 1990;5:191–220.
25. Reason J. *Managing the risks of organizational accidents*. Aldershot. UK: Ashgate Publishing Company; 1997.
26. Gibson JJ. *The ecological approach to visual perception*. Boston: Houghton Mifflin; 1979.