Abstract
In this paper, we examine how users interact with a computer-based infusion device adapted for terbutaline infusion to treat preterm labor in women experiencing high-risk pregnancies. This study examines: (1) the HCI deficiencies in the device as related to this context of use, (2) how the device characteristics increase the potential for error, and (3) the tailoring strategies developed by users to insulate themselves from failure. Interviews with nurses and tests of the behavior of the infusion device in different conditions identified several classic HCI deficiencies: complex and arbitrary sequences of operation, mode errors due to poor differentiation of multiple operating modes intended for different contexts, ambiguous alarms, getting lost in multiple displays, and poor feedback on device state and behavior.
INTRODUCTION

Technological change and economic pressure are moving medical practice out of hospitals and into the home or other alternative health care settings. Patients with chronic conditions may be able to move out of the hospital through the use of infusion devices that support self-administration of drugs. For example, diabetics may use these infusion devices for insulin therapy, or women with high risk pregnancies may use these devices to self-administer drugs that control preterm labor.

These changes are made possible by changes in medical technology--automated infusion devices. But these new computer-based devices, if designed poorly from a user-centered point of view (Norman, 1988), can induce erroneous actions. Previous studies of computer-based medical devices in critical care medicine have found that computer-based medical devices often exhibit a variety of classic human-computer interaction (HCI) deficiencies such as poor feedback about device state and behavior, complex and ambiguous sequences of operation, multiple poorly distinguished modes, and ambiguous alarms (Cook, Potter, Woods and McDonald, 1991; Moll van Charante, Cook, Woods, Yue, and Howie, 1993; Cook and Woods, in press). These deficiencies are important because they have been shown to increase the potential for erroneous actions and to impair the physician’s ability to detect and recover from errors (e.g., Cook, Woods, and Howie, 1992).

In this paper we extend the results of those studies of physician interaction with computer-based medical devices to the home health care context. We examined how nurses and patient/operators interact with a computer-based infusion device used for terbutaline infusion to treat preterm labor in women experiencing high-risk pregnancies. This device was originally used in insulin administration for diabetics, but was adapted to assist in the control of pre-term labor.

The purpose of the study was to investigate how nurses and patient/operators used the device in the context of control of preterm labor and to identify characteristics of the device that make its operation difficult and prone to error. Our investigations also focused on how the perinatal nurses developed strategies to work around or guard against the human-computer interaction (HCI) deficiencies in the device (Cook and Woods, in press). These adaptations or tailoring strategies occur because patients and their nurse caregivers were responsible for achieving their own goals -- for the patient to remain at home during a difficult pregnancy, and to have a successful delivery as close to term as possible, regardless of the design of the computer-based device.

Three kinds of investigations were carried out: (a) interviews with nurses about how they used the device and about how patients/operators used the device, (b) “bench” tests that explored how the device behaved, how the displays represented those states and activities, and the control sequences needed to interact with the device across a range of tasks and contexts relevant to terbutaline therapy for preterm labor, and (c) observations of nurses programming the device to accomplish different tasks.
This paper describes three aspects of user-device interaction: (1) the HCI deficiencies in the device as related to this context of use, (2) how the device characteristics increase the potential for error, and (3) the tailoring strategies developed by users to insulate themselves from failure.

THE CONTEXT:
TERBUTALINE THERAPY FOR CONTROL OF PRETERM LABOR

The focus of this investigation was the use of a computer-based infusion pump for terbutaline therapy with pregnant women experiencing preterm labor. Terbutaline is a member of the drug class--beta-adrenergics, and affects adrenergic receptor sites. The drug interacts with the beta receptor sites leading to uterine relaxation.

Control of pre-term labor is a 24 hour a day therapy, and patient/operators may need to interact with the infusion pump at any time (e.g., changing an empty syringe in the middle of the night). The therapy consists of a background base rate and with periodic larger doses which are adjusted empirically for each patient to achieve control of pre-term labor. Therapy plans need to be adjusted over time for each patient to avoid recurrence of preterm labor because physical changes occur as pregnancy advances (i.e., changes in contractions) and because desensitization of the beta receptor sites occurs due to terbutaline use.

The medication can affect the patient’s physiological and mental state at the same time that the patient is an active user of the device. Adverse side effects of terbutaline are increased heart rate, increased cardiac contractility, tachycardia, palpitations, tremors, restlessness, anxiety, and nausea (Sala and Moise, 1989). In addition to having to cope with the side effects of the medication, the patient/operator also experiences stress due to the uncertainty of her situation, the question of the medical risks to herself and to her baby, and confinement to bed rest.

Under-administration of the medication may lead to under control of contractions. Over-administration can produce toxicity (acutely or cumulatively) which can be a very serious threat to the health of the baby and to the mother. Therapy plans are developed, monitored, and modified to control contractions without exceeding toxicity limits.

DEVICE OPERATION:
A COMPUTER-BASED INFUSION PUMP IN HOME HEALTH CARE

Nurse caregivers perform the initial set-up of the infusion pump for terbutaline therapy for each patient. They program the infusion device to deliver doses of medication with certain time intervals between doses as well as setting an underlying basal rate if it is needed. The dose and delivery intervals are based on each individual patient's medical requirements. Table 1 lists some of the operations the nurse must perform as part of her tasks.
Once the device has been programmed by the nurse for a specific patient, the patient, at home, must perform regular tasks to ensure that she receives the therapy as prescribed and when needed. Table 2 illustrates a sample of the normal daily procedures required of the patient and the operations that she must perform in order to implement those procedures.

Improper programming or use of the device can result in under- or over-administration of terbutaline. Failure to successfully operate the infusion device or failure to control preterm labor can have varying degrees of impact: a call to the nurse, a trip to the hospital, a prolonged hospital stay, or premature delivery. Another type of failure occurs when a patient is screened out as a potential candidate for terbutaline therapy at home because they are unable to operate the pump successfully.

The specific device in question is a portable, battery-operated, electronic infusion pump (Model 404-SP, MiniMed® Technologies, Sylmar, Calif.). It is one example of a class of infusion devices, i.e., syringe pumps, for use with therapies requiring delivery of small volumes of high concentration medication (Figure 1). The pump is used with 3-mL syringes (3 mg of terbutaline sulfate per 3-mL). The medication is pumped through a 42-inch long infusion set (tubing used to deliver medication from the pump into the tissue) with a flexible teflon cannula (a needleless tubing attached to the infusion set) that is inserted into the tissue just beneath the surface of the skin (subcutaneously). Infusion sites are selected in the upper or lower abdomen or the anterior thigh and changed every three to four days.

Users interact with the device through four multi-function buttons: select (SEL), activate (ACT), up arrow, and down arrow. In principle, the SESelect button allows the user to "page" through the different programming displays. The ACTivate button allows the user to "activate" the various screen displays to change the settings (e.g., the time, amount of a demand dose, profile settings). The up and down arrow buttons allow the user to increase or decrease the setting (for example, to change the time or to increase or decrease the medication dose).

Multiple Modes

The device operates in a hierarchy of multiple modes. Two regulate drug delivery: one is used for insulin therapy (rate mode); the other for terbutaline therapy (interval mode). If the pump is set in the rate mode, medication is delivered continuously at a programmed rate (from .000 ml/hr. to .720 ml/hr. in .002 ml/hr. increments). When the pump is set in the interval mode, medication is delivered intermittently at a programmed dose size or bolus (from .000 ml to .998 ml in .002 ml increments) with a time interval between doses.
mode, it is also possible to set an underlying basal rate (from .000 to .250 ml/hr.), which enables flexibility in the therapy prescribed. Figures 2 and 3 graphically illustrate examples of therapy with each delivery mode. These two kinds of therapy are significantly different. They are designed for different medications treating different kinds of medical conditions.

Insert Figures 2 and 3 approximately here.

Some of the other modes that can be used while the pump is in either rate or interval mode are profile mode and lock-out mode. Profile mode allows the user to program up to six different dose/time interval settings (profiles) during a 24-hour period. These profiles will repeat every 24 hours until the user changes the programming.

The lock-out modes block user access to device functionality. There are three lock-out levels. In Lock-Out level 0, all functions are accessible. Lock-Out level 1 allows for all functionality except for setting maximum dosage levels and for resetting the “Totals” display. In this mode, the user can set the Dose and Interval or Rate, set the “Demand Dose” to preset maximums, deliver a demand dose, put the pump in and out of “Suspend” mode (this pauses drug delivery, for example, while changing the syringe), and reset the time. The only functions available to the patient/operator with Lock-Out level 2 is to deliver a preset “Demand Dose”, and to place the pump in and out of “Suspend” mode.

Multi-Functions Keys

The four buttons actually perform multiple functions depending upon the sequence of key presses. For example, to deliver a demand dose, the patient/operator must press the SELect button twice and the ACTivate button twice. To put the device in Suspend mode (e.g., when changing the syringe), she must press the SELect button five times (if the device is in interval mode) and the ACTivate button twice.

Multiple Displays

Information is provided to users through an LCD panel (approximately 3/4” by 1”). We constructed a map of the possible displays that could be called up and viewed on the LCD panel as one part of our investigation of the device. We found that the LCD serves as the viewport to multiple screen displays nested at two levels. The basic operations of the device for terbutaline therapy are arranged under seven different screen displays. Under each of those displays are one to seven different displays. Figure 4 maps a portion of the display space (it illustrates the screen displays that are nested under the Maximum Settings Display as an example). Note that users can see only one of these display at a time.

Insert Figure 4 approximately here.

Alarms
Alarm messages appear on the LCD panel accompanied by auditory alarms that consist of a number of beeps. The number and rate of beeps are meant to indicate different device states and abnormalities.

INVESTIGATING DEVICE USE IN CONTEXT

Three kinds of investigations were carried out. Nurses were interviewed about how they used the device and about their experiences with how patients/operators used the device. The investigators conducted “bench” tests of the device that explored how it behaved, how the displays represented those states and activities, and the control sequences needed to interact with the device across a range of tasks and contexts relevant to terbutaline therapy for preterm labor. Nurses were observed while programming the device to accomplish different tasks.

We iterated across these types of investigations in order to identify (1) error prone tasks or situations (mode error), (2) characteristics of the device that create or enhance the potential for error (e.g., multiple modes with poor feedback about device state) contribute to error prone and difficult to observe, (3) characteristics of the context of terbutaline therapy that interact with the device characteristics to provide opportunities for error, and (4) the tailoring strategies developed by users over time to work around error prone tasks and device deficiencies.

In the bench tests one of the authors (JHO) operated the infusion device in situations that are likely to occur in the context of terbutaline therapy. Device indications and behavior were explored in all of the situations noted in Tables 1 and 2. As part of the bench tests, we mapped the organization of displays that could appear on the LCD panel (Figure 4). Control sequences for typical user tasks were also identified. The bench tests explored device behavior when errors occur in these control sequences.

The three types of investigations were iterative and intermixed. For example, results from interviews would define situations where confusions seemed to occur. We would then conduct a bench test to define exactly how the device behaved in that situation including the consequences of erroneous entries (e.g., the control sequences needed, the displayed indications and alarms that resulted, the device activity that resulted). Armed with this background information, we would then observe how a nurse uses the device by presenting her with a context where she needed to interact with the device, observing her behavior, and following up with a discussion of difficulties she experienced or had seen others experience. Another type of iteration occurred when we identified areas in the bench tests where one might expect user problems to occur. We would then use this information to query nurse users about their experiences and experiences of the patients they supervised. For these cases we also might observe several nurses programming the device. Data were combined across these different sources to specify places where users would be expected to have trouble using the device. In these activities we paid particular attention to strategies that nurses or patients had developed to protect themselves from HCI difficulties.

One of the authors had used this device as a patient/operator when the system first went into use in this region of the country. We were able to use this
experience as a baseline to see how the community of practice (nurses and patient/operators) had developed tailoring strategies to cope with problematic aspects of device design. In the interviews with perinatal nurses, we specifically examined how users learned to train, inform, and proceduralize the tasks so that nurses and patient/operators could use this device despite its difficulties from an HCI perspective.

We were able to do this study only because of the cooperation of individual nurses. One of the nurses we worked with was responsible for training the nurses who are the primary caregivers for patients using the device. This nurse has been involved in the use of the device since it was first introduced for terbutaline therapy in her company. The other nurses we interviewed have worked with this device an average of three years, are responsible for programming the therapy and for all other aspects of care for the patient experiencing preterm labor.

FINDINGS: HCI DEFICIENCIES AND USER TAILORING

HCI Deficiencies
The investigations of device behavior in different conditions identified several classic human-computer interaction deficiencies in the infusion device (Norman, 1988; Cook et al., 1991).

Complex and arbitrary sequences of operation. All of the user’s normal tasks are accomplished using just four buttons. But these buttons are used in many different sequences to accomplish these tasks. To set up or modify a therapy plan, requires a complex sequence keystrokes. For example, when programming the patient’s profiles in the interval mode (different dose-interval settings), the nurse must press the SESelect button seven times, the ACTivate button twice, the SESelect button once, the ACTivate button three times, the SESelect button once, then the arrow keys are used to adjust the interval to the desired setting. Once the desired dose and interval are on the display, the nurse must press the ACTivate button, set the time for the next profile, press the ACTivate button, use the arrows to set the dose and interval for that profile, and so on until the desired number of profiles are programmed (up to a maximum of six for a 24-hour period). As a result of the interface design, users must remember the sequence of keystrokes needed to accomplish a task and where they are in the sequence of keystrokes for this task. This creates frequent opportunities for misoperation.

Furthermore, the interface exhibits low error tolerance. For example, hitting one of the buttons the wrong number of times (six instead of seven) may produce a legal sequence of commands resulting a different result than expected or desired. It is also possible with one erroneous keystroke to destroy all the previous programming. If one presses the wrong button, for example, at a late stage in entering a series of profiles, she will have to begin re-programming from the first profile to correct the erroneous action.
**Different operating modes intended for different contexts.** The multiple modes create the potential for mode errors. A mode error is a basic type of erroneous action that a human user can commit by executing an intention in a way appropriate to one mode of the device when the device is actually in another mode (Norman, 1988). This is a critical issue since programming the pump while it is in the wrong mode can result in an incorrect delivery of medication. For example, if one intends to program in interval mode (dose levels and time intervals between doses) but the device is actually in rate mode, the device will accept user input but interpret it as specification of different infusion rates to be delivered over different intervals. The device will deliver a continuous rate of medication rather than boluses of medication being delivered at pre-specified times as intended. Figure 9 illustrates how this mode error can change the medication therapy significantly.

The potential for mode error exists in part because the displays provide only very weak indications about which mode the device is in at any given time. For example, one display, the Normal Operating Screen, provides some indication of what medication delivery mode the device is in, but no delivery mode indications are available on any of the other display pages. There are indications of other modes (lockout level, suspend mode) on displays but in many cases these indications may not be very salient or observable to users given the context (stress), their training, and other demands.

If a mode error occurs that effects the drug infusion pattern, there is no feedback about actual device behavior available to help patient/operators or nurses monitor whether actual delivery of medication matches the desired therapy plan.

The combination of multiple modes with weak feedback about device state makes mode errors a predictable consequence of the design. We found in testing the device in realistic scenarios that it is easy for mode errors to occur. Nurses were aware of the potential for mode errors, at least in some of the cases (confusing rate and interval modes).

**Ambiguous alarms.** The infusion device has 13 alarm states, seven of which are signaled by the same auditory alarm. The visual display for the alarms is cryptic and ambiguous (e.g., E01 is displayed when the device has encountered and recovered from a system error, and E2 is displayed when the device has encountered a system error that may result in an over-infusion). One result is that users treat the alarms as a generic master caution signal.

**Getting lost.** This device is activated by a sequence of button presses. But depending on the screen that is displayed or the level of lock-out the device is in, the same actions can produce different results, putting the user into different displays. Given the arbitrary command sequences and the lack of feedback, it is easy for users to enter a command and then find themselves looking at a different display than the one they were expecting. They get lost in the complex command sequences. When this occurs, they have to re-orient by escaping from the task they were attempting to carry out and starting the interaction over from the beginning.

Pressing the SELect button once and the ACTivate button once is one example of how the same sequence of actions will put the user into different displays depending on device state. When the lock-out level is 0 and the Normal
Operating Screen is displayed, pressing the SELEct button once and the ACTivate button once will call up the Dose-Interval Page with the dose value blinking (which indicates that the dose can be changed). If the starting point is the “Suspend” screen display (i.e., the device is in Suspend mode and is not delivering medication), pressing the SELEct button once and the ACTivate button once will result in the display of the Normal Operating Screen and a resumption of medication delivery.

If the lock-out level is 1 or 2 and the user again starts from the Normal Operating Screen display page, the same action sequence produces a third result. In this situation when the user presses SELEct once and ACTivate once, nothing happens for six to seven seconds and then the Normal Operating Screen is displayed. What has occurred is that the user has taken a sequence of operations that are not allowed for the lock-out level, but the user is never informed that her action is illegal. When a user makes an illegal entry, the system does nothing for 6 to 7 seconds and then reverts to the Normal Operating Screen. The user receives no other feedback that the machine has judged her inputs to be illegal. The appearance of the Normal Operating Screen can come as a surprise leaving the user confused and wondering how she got there; did she inadvertently enter a command to go to this surprising display or did the system do something automatically.

**Poor feedback on device state and behavior.** A user’s perception of a device depends on an interaction between its capabilities and the feedback mechanisms that influence what is observable about system behavior in relation to events in the environment. What feedback is available depends upon the “image” the device presents to users (Norman, 1988). Effective feedback or observability is more than mere data availability; observability depends on the cognitive work needed to extract meaning from the data available (Sarter, Woods, and Billings, in press). Poor feedback or low observability occurs when it is difficult to notice, attend to or process the available indications. Factors that affect how difficult it is to interpret available indications include how much background knowledge users must bring to bear, how much users must integrate multiple pieces of data from different places, how easy it is to recognize “interesting” changes or events, and how attentional demands affect data search -- how users know when to look where.

This device provides little or no feedback about its state or behavior. If a user is in the wrong mode (rate versus interval), only the initial screen display differs. The other displays used in programming the device are identical for the two modes and provide no indication of which mode the device is in. For example, lock-out mode is indicated by a small “L” in a relatively crowded display which indicates that the device is in either Lock-out level 1 or 2. The absence of the L indicates Lock-out level 0. Furthermore, patients/operators are not informed or trained to check for the correct lock out mode.

When an erroneous action occurs, the device often provides little or no feedback to help the user realize that an error has occurred or to aid the user in understanding what has led to surprising changes in device behavior. There is little feedback about the amount of medication being delivered and whether it matches the therapy plan. The complexities of the device make it possible that
erroneous actions can inadvertently reprogram the therapy without the user being aware of the consequences of their actions, for example, if the device is not in Lock-out level 2. In general the device has low observability and this lack of feedback exacerbates or contributes to many of the above problems.

For example, consider the situation where the pump is in Lock-out level 2 and the user intends to stop the delivery of medication by putting the pump in Suspend mode. This is a typical device state for patient/operators and this action is a part of the sequence of activities for changing the syringe or the infusion set (a frequent task for patient/operators). The user presses the SELect button five times and then, not attending to or not understanding the meaning of what is displayed, presses the ACTivate button. If only four of those button presses successfully changed the displays, the patient would be at the Totals screen display when she pressed the ACTivate button. In Lock-out level 2, this action results in the Normal Operating Screen being displayed because in this mode, the Totals screen display can be viewed but cannot be activated. So, the patient will find herself in a display that she had not expected, the action intended was not taken (suspension of medication delivery), and the patient has no feedback as to why.

The limited feedback provided to the user about device state and behavior is particularly troubling (a) because it limits error recovery (Woods et al., 1994) and (b) because some of the error traps inherent in device design can lead to the actual delivery of the medication being different from what the user thinks she has triggered. Depending on the specific error, the device may function but over- or under-medicate. Since there is no clear feedback available about the devices activities, it is difficult to detect over-medication. Under-medication is indicated by the failure to control preterm contractions. With other errors the patient/operator may be unable to get the device to function precipitating calls for assistance and running the risk of failure to control preterm contractions.

User Tailoring

Our investigations included analyzing how the system of people and artifacts evolved over time to produce generally successful performance (Cook and Woods, in press). This adaptation or tailoring process occurs because users are responsible not just for the device operation, but also for the larger performance goals of the overall system (i.e., for the patient to remain at home during a difficult pregnancy, and to have a successful delivery as close to term as possible). These stakeholders tailor their activities to insulate the larger system from device deficiencies (Woods et al., 1994).

Examples of the tailoring that have occurred since the introduction of the infusion device within the one perinatal services organization examined include the following:

*Developing a Patient Guide.* The perinatal nurses recognized that patients/operators were experiencing various difficulties interacting with the device and that the manuals provided by the manufacturers of the pump were inadequate to help patients operate the device. Based on their model of the sources of these difficulties, the nurses developed procedures, checklists and information in the form of handwritten and verbal instructions. Over several years the nursing staff
refined, broadened, and eventually formalized this information as a user help manual or what they call a patient guide.

*Modifying Procedures.* In addition to formalizing a patient guide, the nursing staff changed, modified, and eliminated existing procedures, as well as, introduced new procedures:

1. **Having the patient/operator change the syringe at the same time everyday.** Initially, the patient/operator was instructed to change the syringe only when the high-pressure alarm beeped, indicating that the syringe was empty. This often resulted in the patient/operator needing to replace the syringe after being awakened in the night. This is a sufficiently complex task to do even when wide awake. It is possible to incorrectly insert the syringe so that no medication is being delivered and yet have no feedback that this is occurring.

2. **Having the patient/operator give a daily medication total readout to their nurse caregiver.** The procedure of reporting daily medication totals began with the nurses having a vague realization that they were getting dosage errors from the device. In other words, they began to suspect that actual amount of medication delivered might be significantly different than that called for in the therapy plan. They recognized that they needed to figure out some way to monitor for these dosage errors so that they could then develop error recovery plans. They innovatively made use of some features of the device in developing a procedure to help them detect and recover from dosage errors.

   This procedure made use of a device feature that totaled the device’s assessment of what medication had been delivered over some period of time. To use this feature the patient has to go to a specific screen display (the “Totals” display) and inform the nurse (via telephone) of the value displayed under the label “medication totals”. The patient/operator then has to clear the value so that the machine’s running summary will represent a new total for the next interval (the time interval for this check was once each day at the same time). Nothing is provided to help the patient carry out this task (e.g., no written procedure). The nurse remotely provides the programming instructions to do this over the phone.

   The nurse then uses the data, in comparison to the therapy plan, to infer whether the device is delivering the medication as intended or whether misadministrations are occurring such as excessive demand doses (extra boluses of medication). This safety check sometimes reveals that the device was not delivering the amount intended and that patients were initiating more demand doses than they were authorized. But there are no formal aids to help the nursing staff perform this inference which requires several data transformations and comparisons.

3. **Eliminating “demand dose to air” procedure when changing the syringe.** This procedure was initially used every time the device ran out of medication and a new syringe was installed. The procedure involved delivering a demand dose whose purpose was to pump the medication to the tip of the syringe to take up any slack between the syringe driver and plunger. When .05 ml had been delivered, the patient had to place the pump in Suspend mode to stop the delivery. After the demand dose had been stopped, the patient then had to restart the delivery of terbutaline. This procedure was stressful for the patient to perform because while in the middle of the difficult procedure of changing the syringe, the patient had to
perform other programming tasks. This procedure was eliminated given the workload of the procedure to eliminate any slack and the potential for error, coupled with the fact that the possibility of reduced delivery of terbutaline is not critical at this point.

4. Using a paper clip to close the pages of documentation that discuss the delivery mode of medication not being used. In terbutaline therapy for women experiencing preterm labor, the interval mode of delivery is used. The display pages for all but the initial display are identical for both rate and interval modes, and, as a result, the programming instructions in the manual look the same although the detailed sequence of keystrokes varies slightly. However, following the procedure for the rate mode can significantly alter the therapy the nurse thinks she is initiating. Nurses discovered that they could be following the instructions in the manual for setting up rate mode therapies rather than the intended interval mode therapies. In other words, they could make a mode error in selecting programming instructions. To avoid programming terbutaline therapies in the wrong mode, they began to paper clip pages of the manual together.

Although the practitioners have tailored their strategies and behavior to cope with the complexities of this infusion device, their adaptations can be brittle, weak, or expose the system to other risks through side effects (Woods et al., 1994). For example, there are several potential side effects associated with the new procedure for reporting and evaluating daily medication totals (the intended or main effect is to detect dosage errors). For example, if the patient/operator does not correctly zero the machine’s running summary, the procedure will not provide accurate data for the nursing staff to detect misadministrations.

In addition, the procedure has side effects that may not be benign given the other HCI problems. One such side effect may occur in the following way. In order to reset the running summary to zero, the patient/operator must change lock-out level in order to receive the authority to change the values on the relevant screen display. Then after recording or reporting the daily medication total, the patient must reinitiate the stronger lock-out level (the nurse guides them through this remotely as part of the procedure). However, if the patient/operator errs in attempting to reinitiate Lock-out level 2, she has access to the full functionality and complexity of the device. Subsequent misprogramming can change significantly the therapy delivered. For example, the patient could find herself in displays with which she is unfamiliar, and in the attempt to escape from those displays to familiar ones, she could reprogram her drug regimen without realizing that she had done so. Furthermore, the indication of lock-out mode is weak (a small “L” appears in a relatively crowded display which indicates that the device is in either Lock-out level 1 or 2; the absence of the L indicates Lock-out level 0) and patients/operators are not informed or trained to check for the correct mode. Ironically, a procedure developed to make up for the lack of feedback is limited by the very same device problem.

Another example of the brittleness of user tailoring can be found in the nurse relying on a paper clip to ensure that she does not inadvertently program the pump in the interval mode while intending to follow the instructions for the rate mode. While this adaptation is typical of the resourcefulness of users to cope with
clumsy devices, this activity is obviously brittle as the paper clip may be lost or may be inadvertently placed on the wrong pages.

**DISCUSSION**

**HCI Deficiencies are Latent Failures**

Our investigations identified several classic human-computer interaction deficiencies in this infusion device given the context of terbutaline therapy to treat preterm labor at home: lack of feedback on device state and behavior, complex and arbitrary sequences of operation, ambiguous alarms, and multiple operating modes intended for different contexts. These problems occur repeatedly in computer-based devices in general (Norman, 1988) and in computer-based devices for medical applications (Cook et al., 1991; Moll van Charante et al., 1993).

These characteristics are problematic because they predictably create the potential for erroneous assessments and actions. For example, in this case as in others, multiple modes with weak indications of mode status will lead to mode errors. Thus, erroneous assessments and actions are not simply “human error;” rather, these errors are symptoms of underlying design deficiencies (Woods et al., 1994).

These design deficiencies matter because they produce the potential for errors that contribute to critical incidents and outcome failures -- in this case misadministration of therapy. Note that these design-induced errors do not always produce a misadministration. Other circumstances or factors must be present for an HCI problem to contribute to the chain of events leading to failure. In other words, HCI deficiencies are a kind of latent failure (Woods et al., 1994). It is a factor present in the system which can contribute to an outcome failure if other triggering and potentiating factors are present (Reason, 1990).

Because practitioners are responsible agents in their domain of practice, they actively work to insulate the larger system from device deficiencies as they perceive them (Cook and Woods, in press; Woods et al., 1994). We identified a variety of user tailoring strategies in this particular case. But notice how user tailoring can act to hide underlying design deficiencies from other parties. We also found evidence that user tailoring may be only partly successful. The adaptations, while useful in narrow contexts, can be brittle or produce unanticipated side effects which produce new paths towards failure.

The latent failure chain, where HCI design deficiencies can be one contributor to an accident sequence if other potentiating factors occur, has many implications for the development of computer-based medical devices. The most basic is implication is that avoiding design-induced human-computer breakdowns is important. Furthermore, it is part of the responsibility of development organizations in the medical technology field to avoid design-induced human-computer breakdowns as part of their design process. Similarly, development organizations and regulatory authorities need to begin to test and evaluate computer-based medical devices with respect to human-computer interaction issues. HCI is much more than a luxury factor or marketing edge; it is fundamental to patient safety and device efficacy.
Cook et al. (1991), and Moll van Charante et al. (1993) illustrate and discuss how to carry out evaluations of computer-based medical devices in order to identify HCI deficiencies and the potential consequences of those deficiencies. These studies show how it is critical to perform error analyses in the design of new computer-based systems (Norman, 1988). In other words, instead of showing how the device can work in textbook cases, we try to show how effective interactions can break down if complicating factors arise. By finding potential breakdown points, the device can be modified to make human-computer cooperation in context more robust.

Since the medical technology industry’s level of awareness about effective human-computer cooperation is an issue, it is appropriate to note that there was no organizational support for this research as an opportunity to understand device use or to redesign the device, training or procedures. In part this occurs because finding device deficiencies and exposing error traps is a politically, legally and financially charged enterprise. As a result, our ability to collect and report all of the kinds of data that we would want to fully analyze the potential for error was limited. For example, we would have liked to set up a mechanism to identify and analyze a corpus of actual incidents (e.g., as in Cook, Woods and McDonald, 1991), and we would have liked to observe patients or prospective patients during training and actual device use. Despite these larger organizational issues, individual nurses were willing to share information about their experiences and demonstrate device use to us.

Directions for Improved Feedback

Characterizing the problems and deficiencies in device use in context point to new design concepts. Some of these can be ways to make the current displays and control sequences more usable. However, understanding HCI in context can point to deeper implications for re-design. One of these is how to provide more effective feedback about device state and activities. The infusion pump is an automated system: given a set of instructions, the device will act to implement the programmed therapy and it will continue to act unless explicitly instructed to stop or change, even if that therapy regimen is not appropriate or is not what was intended. Miscommunications can occur where users can think that they have communicated one intention, but the device has interpreted the users inputs in a different way. Such miscommunications between users and automated systems have been one contributor to accidents in aviation (Sarter et al., in press; Billings, in press). Thus, it is very important to have an effective feedback mechanism that allows the practitioner to understand what the automation is actually doing to support the process of error detection and recovery.

One implication of this research and other research on practitioner interaction with automated systems is that displays should help users perceive what the automation is doing and what it is going to do relative to expectations or plans (Woods et al., 1994, chapter 5). In this case, the graphic representations we used to illustrate the different therapy plans possible with this device provide with a starting point for developing more effective feedback (see Figures 2, 3 and 7). The key is the realization that human-computer interaction and device behavior are all about different bolus sizes, rates of infusion and time intervals. What is
informative are dose-time relationships. This provides the basic frame of reference for developing a more effective representation of device activity (what will it do? what is it doing?) and a more effective human-computer interaction (how does a user instruct the automation?).

Currently, users can see only one profile or dose-interval setting at a time. Each dose size, rate, or interval is entered one value at a time through a series of commands. Users cannot see the larger therapy plan as in Figures 2 or 3. Thus, they have to build up and maintain their own mental model of what has been programmed relative to the desired therapy plan one piece at a time. Because of the inability to see the therapy plan as it is being programmed and the inability to see the relationship between actual drug delivery and the programmed therapy, the potential for error increases: one can enter a therapy plan incorrectly or inadvertently modify the therapy without being aware of it.

The graphs of therapy plans in Figures 2 and 3 point us towards the kind of display that is needed to provide better feedback as a nurse sets up or modifies therapy plans -- an enhanced dose-interval graphic representation (see Yue et al., 1992 for another example of a new graphic concept for providing improved feedback on the activities of a different type of infusion pump). As the practitioner builds the therapy plan, we want to show a graph of that plan, mapping informative relationships within a larger frame of reference (Woods, in preparation). The relevant frame of reference is the relationship of dose level and intervals between doses. Within this frame of reference we can show a variety of important relationships analogically: basal rate, demand dose, intervals, and constraints on doses such as frequency or cumulative dose. Representing therapies in this way makes it clear that interval mode and rate mode are very different (compare Figures 2 and 3).

An enhanced dose-interval representation also enables us to simplify the instructions needed to set up or modify therapy plans (remember that in the current case programming errors can force users to start over again because of device intolerance to misentries or because of user uncertainty about what they have actually entered). The dose-interval graph provides the basis for designing a direct manipulation interaction (Hutchins, Hollan, and Norman, 1986) where users could directly indicate on this graph the doses, intervals, basal rates that they want to enter or modify (for example, via point and drag operations using a pen-based input device or some other pointing device).

Finally, an enhanced dose-interval graph provides a means to monitor the activities of the infusion device. This is done by following one of the graphic design principles from Woods (1995): “Highlight contrasts. Representations should highlight and support observer recognition of contrasts ... -- some departure from a reference or expected course. Representing contrast means that ... one shows how the actual course of behavior follows or departs from reference or expected sequence of behavior given the relevant context. Representing contrast signals both the contrasting states or behavior and their relationship (how behavior departs from or conforms to the contrasting case).”

We can represent contrast by plotting actual administrations against the therapy plan and constraints on the dose-interval graph. Figure 5 shows the target we would like to achieve for one case where a mode error results in the device
interpreting the nurse’s inputs being as specifying a rate type of therapy when she
intends to modify an interval therapy plan. The contrast between actual drug
delivery and therapy plan stands out. In this way, one makes it easy for observers
to see departures from the therapy plan, in effect, highlighting anomalies.

Insert Figure 5 approximately here.

Of course, developing enhanced and dynamic dose-interval graphs would
require a great deal of design work, wrestling with many interacting constraints,
and examining many different kinds of contexts and situations. However, the
concept illustrates how studying device use in context can point the way to new
design directions.

Critical Care In The Home

There often is a distinction made between home health care and critical care
medicine as being very different domains within the overall health care field (e.g.,
the general public as the user versus highly trained medical specialists; occasional
or temporary users versus experienced chronic users). The example of in-home
therapy for control of preterm labor illustrates that advances in technology are
making it possible to move aspects of medicine that involve critical care into the
home. The risks in this case do not disappear as care moves from the hospital into
the home setting. Therapy is still designed and adjusted empirically for each
patient based on feedback over time. Tasks associated with the collection and
management of information do not disappear.

What changes is not the criticality of the care, but the distributed system for
providing care. Health care is based on a system of multiple cooperating agents --
cognitive activities are distributed over a set of people and machine agents such as
this infusion device. Note that this system is larger than the device and the patient
or nurse.

The introduction of the infusion device and the shift from in-hospital to in-
home control of pre-term labor changes the roles and responsibilities of the
different participants in the therapy system (Figures 6 and 7 illustrate the
distributed system for each setting). The patient has a different role and becomes an
active participant in her therapy -- a patient/operator. She is required to (a) deliver
demand doses when uterine activity is greater than a pre-determined threshold, (b)
change the infusion site, (c) change the syringe when empty, and (d) monitor her
uterine activity, blood pressure, and heart rate. How the perinatal service nurse
gathers information about the impact of therapy and how the nurse adjusts
delivery of medication changes as well. A new component of supervisory control
is introduced into the nursing function as traditional nursing functions are
delegated in part to the patient/operators.

Insert Figures 6 and 7 approximately here.

Information about the effects of therapy is critical to modifying the therapy
and to early recognition of problems. The use of automated infusion devices in the
home changes how this information is gathered and distributed to the people

16
responsible for problem recognition (e.g., approach to toxicity limits), and therapy
decisions (e.g., the need for closer monitoring in the hospital). The opaque system
image presented by the infusion device and the opportunities for misoperation
created by poor interface design impair this distributed therapy system’s ability to
detect potential problems.

Is the Human-Machine Ensemble Effective or Flawed?
When one investigates and discovers HCI deficiencies, stakeholders may ask,
"Do these problems mean that the device is ‘bad?’” or “Should its use be avoided?”
"Because there are HCI deficiencies in this infusion device, should we abandon
terbutaline therapy in the home?” This is not a useful way to see the implications
of such studies. Technology creates the opportunity for moving desirable and
medically useful functions such as the control of preterm labor out of the hospital
and into the home. However, this change in technology also transforms the
distributed cognitive system. People have new roles. The need for effective
coordination across the multiple agents goes up. The distributed system can break
down in new ways (Woods et al., 1994).

Technology-centered design misses these implications of changing
technology on the role and information needs of the people involved in medical
care. The infusion device studied here, like other technology-centered devices in
medicine and other fields, exhibits classic flaws in human-computer coordination.
These flaws predictably create the potential for certain kinds of erroneous actions
(e.g., mode errors) and misassessments. As responsible practitioners, the people in
the system attempt to tailor the device and their strategies to insulate themselves
from the potential difficulties. Despite their efforts, when other potentiating factors
are present, predictable forms of misassessments and erroneous actions can
contribute to the evolution of critical incidents.

This study adds to the growing body of work that shows how technology-
centered development of automation can lead to new types of difficulties in
operation (Billings, in press; Sarter et al., in press). This is not a problem of "over-
avtomatication," but rather a problem in the coordination between the automated
system and the various stakeholders who use the technology. Making the
automation function as a team player requires designers to attend to the context in
which the device is to be used and the new kinds of tasks users perform.

Making automation function as a team player requires designing the
distributed system of human and machine agents that manages the activity in
question. The challenge for the human factors community is to steer the
development of medical technology towards user-centered approaches in all facets
of medical care.
REFERENCES


Captions

Figure 1. External view of a portable, battery operated micro-processor-based infusion pump. This is one example of a class of automated infusion devices (syringe pumps) that deliver small volumes of high concentration medication.

Figure 2. Example therapy in Interval Mode. In this mode, the pump is programmed to deliver medication by setting dose size in milliliters (mL) and a time interval between doses. An underlying basal rate can also be set. Figure 2 illustrates a therapy plan with continuous basal rate of medication of .05 mL/hour with boluses (doses) of .25 mL delivered at different time points (programmed as “profiles”): at 12:00 a.m. (profile 1), at 4:00 a.m., 6:00 a.m. (profile 2 - 2 hour intervals), at 8:00 a.m. (profile 3), at 12:00 p.m., 2:00 p.m., 4:00 p.m., 6:00 p.m. (profile 4 - 2 hour intervals), and at 8:00 p.m. (profile 5).

Figure 3. Example therapy in Rate Mode. In this mode, the device is programmed to deliver medication by setting a rate in milliliters per hour, with the ability to have up to six different rates in a 24 hour period. Figure 3 illustrates a therapy plan with an underlying basal rate of .05 mL/hour and four profiles: increasing the rate to .10 mL/hour for a six hour interval (profile 1), decreasing the rate to .05 mL/hour for a four hour interval (profile 2), increasing the rate to .10 mL/hour for a four hour interval (profile 3), and decreasing the rate to .05 mL/hour for a six hour interval (profile 4).

Figure 4. An illustration of nested screen displays. The basic operations of the device for terbutaline therapy are arranged under seven different displays. Under each of those screen displays are one to seven different displays.

Figure 5. Mode error can affect the delivery of medication. In the case illustrated, the intended therapy was to be programmed in interval mode, but due to a mode error, the device was programmed while in rate mode. The contrast between actual drug delivery and the therapy plan stands out in this representation.

Figure 6. Distributed system for health care when the patient is in the hospital for control of pre-term labor. Compare with Figure 7. Note the patient is cared for with little participation in the health care process.

Figure 7. Distributed system for health care when the patient remains at home with an automated infusion device to control of pre-term labor. Compare with Figure 6. The introduction of the infusion device changes the distributed health care system. For example, the patient now has an active role in managing her own care by interacting with the device and informing the health practitioner of medication delivery, interventions, and their impact on her status.

Table 1. Procedures performed by the nurse during the initial set up of the device.
Table 2. Typical procedures performed by the patient/operator during the daily use/operation of the device.