Research Paper ■

The Employment of an Iterative Design Process to Develop a Pulmonary Graphical Display

S. Blake Wachter, BS, Jim Agutter, MArch, Noah Syroid, MS, Frank Drews, PhD, Matthew B. Weinger, MD, Dwayne Westenskow, PhD

Abstract **Objective:** Data representations on today's medical monitors need to be improved to advance clinical awareness and prevent data vigilance errors. Simply building graphical displays does not ensure an improvement in clinical performance because displays have to be consistent with the user's clinical processes and mental models. In this report, the development of an original pulmonary graphical display for anesthesia is used as an example to show an iterative design process with built-in usability testing.

Design: The process reported here is rapid, inexpensive, and requires a minimal number of subjects per development cycle. Three paper-based tests evaluated the anatomic, variable mapping, and graphical diagnostic meaning of the pulmonary display.

Measurements: A confusion matrix compared the designer's intended answer with the subject's chosen answer. Considering deviations off the diagonal of the confusion matrix as design weaknesses, the pulmonary display was modified and retested. The iterative cycle continued until the anatomic and variable mapping cumulative test scores for a chosen design scored above 90% and the graphical diagnostic meaning test scored above 75%.

Results: The iterative development test resulted in five design iterations. The final graphical pulmonary display improved the overall intuitiveness by 18%. The display was tested in three categories: anatomic features, variable mapping, and diagnostic accuracy. The anatomic intuitiveness increased by 25%, variable mapping intuitiveness increased by 34%, and diagnostic accuracy decreased slightly by 4%.

Conclusion: With this rapid iterative development process, an intuitive graphical display can be developed inexpensively prior to formal testing in an experimental setting.

■ **J Am Med Inform Assoc.** 2003;10:363–372. DOI 10.1197/jamia.M1207.

Despite today's technologic advances, human error is responsible for the majority of accidents and mishaps across all industries.[1] In anesthesia, Cooper et al.[2] concluded that an alarming 82% of preventable patient injuries are caused by human error. Patient injury due to human error falls into many categories,[3] including human error associated with vigilance and clinical monitoring.[3,4] Therefore, it is easy to understand human error when a critical intensive care clinician must assess more than 200 variables to diagnose a problem and treat a patient before injury occurs.

The presentation of medical data can be improved. Critical-incident studies have identified the improvement of monitoring devices as a factor when considering reducing human error and improving patient safety.[3,5–7] Human factor studies in aviation and power plant management have shown that better monitors can improve detection, control, and prediction of future states.[5,8–11]

Many medical graphical displays have been developed using human factors design principles. Two of the common graphical displays found in the literature are configural and ecological displays.[9,12] Configural displays are graphical

Affiliations of the authors: Medical Informatics, University of Utah, Salt Lake City, Utah (SBW); Architecture, University of Utah, Salt Lake City, Utah (JA); Anesthesiology, University of Utah, Salt Lake City, Utah (NS, DW); Cognitive Psychology, University of Utah, Salt Lake City, Utah (FD); Department of Anesthesiology, University of California, San Diego, and the San Diego Center for Patient Safety at the VA San Diego Medical Center, La Jolla, California (MBW).

This research is funded by a research grant from the National Institute of Biomedical Imaging and Bioengineering of the National Institute of Health (8 R01 EB 00294) and the NASA Space Grant Program. The authors thank Butch Loeb, MD, from the University of Arizona for his support and participation.

Correspondence and reprints: S. Blake Wachter, University of Utah, Department of Anesthesiology, 3C444 SOM, 30 North 1900 East, Salt Lake City, UT 84132-2304; e-mail: <blake@relia.net>.

Received for publication: 08/16/02; accepted for publication: 01/13/03.

displays that map parameters to geometric shapes designed to accentuate emergent features. Emergent features are portions of the display that become more salient when parameters change. Ecological displays are graphical displays that directly represent all system parameters and the relations among them (Fig. 1).[13] In critical care medicine, both configural and ecological displays have been shown to improve detection, control, and prediction of patient status.[14–20] Just as the art and science of medicine are multifaceted and ambiguous, configural and ecological displays have tended to be unnecessarily complex, requiring comprehensive training of the intended users.

Complex and unintuitive displays add additional cognitive workload to the user.[5–13] Human factor studies have shown that mistakes occur more often when the user experiences an elevated cognitive workload.[5,13,21] A display designed according to human factors and basic findings of cognitive psychology should be intuitive to the user and thereby reduce cognitive workload.[12,22] To make a display intuitive, a representative group of users should be involved during the development and testing of the display.[23,24] When designing medical displays, however, few clinicians have the time or desire to develop a new graphical display. Further, it is difficult to recruit a large enough number of clinicians to provide reliable input to the design process. Thus, most researchers do not address the intuitiveness of the displays and depend only on formal testing of a mature display.[25] A reason may be that the time and expense of developing a display discourage refinement and continued testing. A design process that involves the users during the development of the display while remaining cost-effective and time-efficient needs to be used.

Similar to software developers who optimize code development through the use of iterative design methodologies, an iterative development cycle can also be used to design information displays.[23] Such processes are utilized to minimize alterations to the requirements and the design late in the development life cycle when a change is more costly.[23] In addition, design changes often are made subjectively or with (unintentional) cognitive bias from the designers.[26] For instance, designers might have selective biases for their design, which may hinder the designer from evaluating it effectively. Therefore, it is critical to evaluate the information display by the intended users as it is being designed—well before implementation and formal evaluation.[23] This allows the design process to be optimized while gaining an understanding of the intuitiveness and usability of the display.

Our interdisciplinary team (architects, cognitive psychologists, bioengineers, computer scientists, human factors experts, clinicians, and medical informaticians) has adopted a process for developing information displays that promotes design as a function of human behavior and the interaction between subjects and the display. The design approach was based on the concept of a "hermeneutical circle," described by Snodgrass and Coyne[27] as an iterative process of implementing a design, learning and understanding from discussion and feedback, and subsequent design refinement. With the iterative rapid design development, the tests of the display focused on usability and intuitiveness of the display. Users were recruited as subjects during each testing phase per development cycle.

After successfully developing a cardiovascular[28] and drug display[29] using this iterative development cycle, we theorized that an initial pulmonary graphical display, shown in Figure 2, could be refined to become more intuitive by using our iterative development cycle. This report documents this process for developing information displays by using a pulmonary graphical display as an example.

The primary focus of developing and evaluating the pulmonary graphical design was to increase the success of future testing of the design while minimizing design evaluation costs. As a result, this report represents the

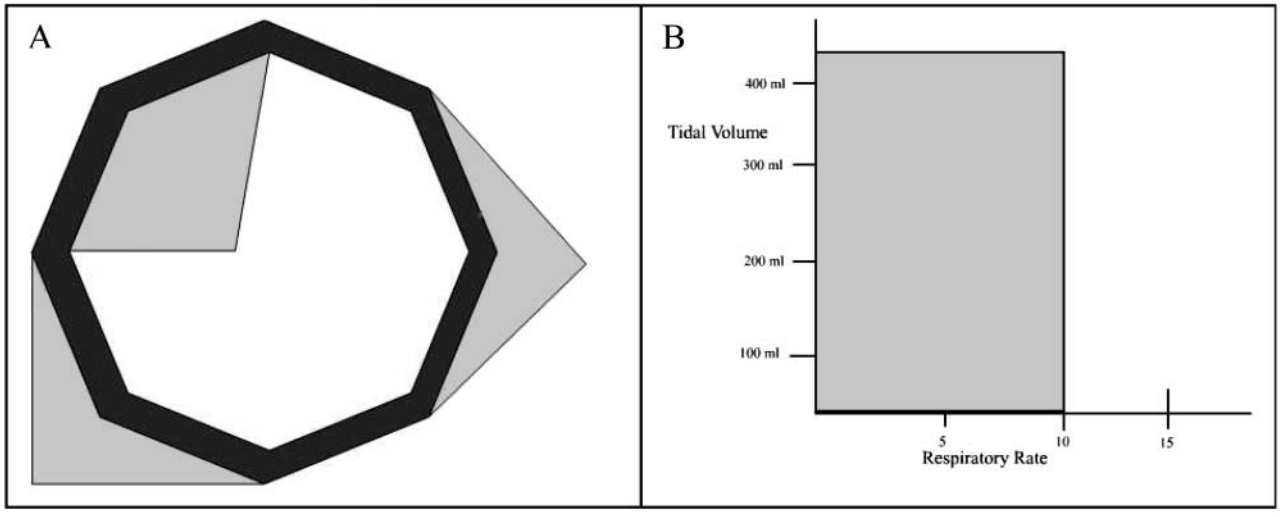


Figure 1. (A) Configural displays map parameters to geometric shapes. As parameters change, the shape distorts. (B) Ecological displays focus on the relationships between parameters. In this example, the rectangle shape will change height and width with respect to volume and respiratory rate.

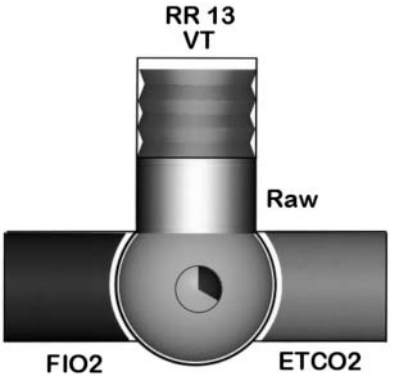


Figure 2. The initial pulmonary design that began our iterative development cycle. (Bellows : blue, inspired gas : green, lungs : green, expired gas : gray, airway : metallic.)

initial testing as the first part of a three-part study: the development process to design a pulmonary graphical display in anesthesia. The second stage will bring the pulmonary graphical display to life in an anesthesia simulator with evolving scenarios in a realistic environment. Finally, the third phase will test the display in the operating room.

Background

The pulmonary graphical display related variables of respiratory physiology and anatomy that are monitored during general anesthesia. By combining aspects of both configural and ecological graphical displays, the pulmonary display focused on data representation, emergent features, and reference frames. Through unique combinations of simple shapes and colors, our goal was to develop a pulmonary display that had an intuitive look and feel of the pulmonary system. A second goal was to add clinical relevance to the pulmonary display through presentation of pertinent information at just the right moment to support diagnosis. Distinguishing normal from abnormal, a reference frame that surrounds each emergent feature defined the current state and allowed the user to identify changes from normal.

Data Representation

As an example of an ecological display, Cole and Stewart[16,17] described their graphical display to be an intuitive representation of the data. They developed a graphical display to represent pulmonary ventilator information in the intensive care unit (ICU) by a series of rectangles representing a patient's tidal volume and respiratory rate at specific points in time. The shape of each rectangle integrated tidal volume and respiratory rate defining the rectangle's height and width (Fig. 1B). With their graphical display, Cole and Stewart[17] showed that subjects intuitively recognized the dark rectangles as representing mechanically ventilated patients and the light rectangles as representing spontaneously breathing patients. The volume rectangles of Cole and Stewart[17] are an example of a simple ecological display. The ventilator rectangles were intuitive and representative because the subjects were able to "see" the data in the context of ventilator weaning.

Intuitive pulmonary graphics can also be found in medical textbooks. Figure 3A shows airway restriction by a thickening or a bulging of the tube edges. This picture is intuitive to the reader because it looks anatomically and physically like the diagnoses being discussed. The user would need little training to associate this image with airway restriction. Figure 3B shows a more complex, although still intuitive, image of anatomic features. This display may cause

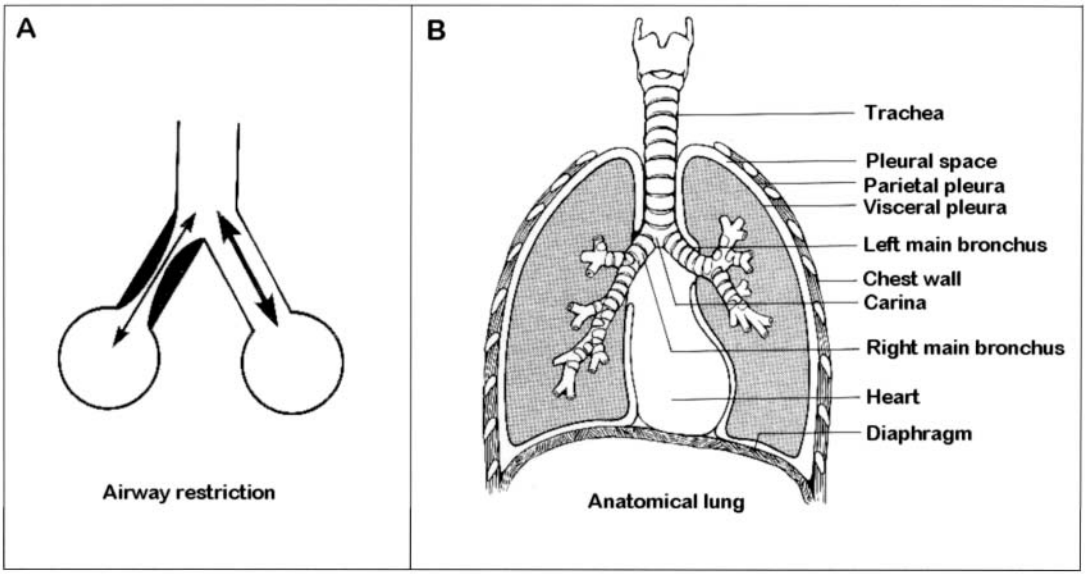


Figure 3. (A) A simple and intuitive image shows airway obstruction. (B) Complex and intuitive image shows a diagram of the anatomic lung. These images are borrowed with permission from the medical textbook, Grippi's *Pulmonary Pathophysiology.*[35]

confusion to the user when interpreting and associating measured data. Therefore, the difficulty of designing a successful display lies in discovering the balance between intuitiveness and simplicity while maintaining meaning.

Normal Reference Frame

A reference frame is a graphical indication of an initial or baseline state. When the state changes, a reference frame indicates how much of a change has occurred. A display with a reference frame emphasizes change by the distortion of the normal shape.[13]

A study showed that a display that incorporated reference frames enabled anesthesiologists to perform a task more accurately. Jungk et al.[19] compared subject performances using a traditional waveform trending display and a type of configural display. The configural display indicated a normal reference frame by a vertical horizontal line. Deviations from normal were shown by horizontal histograms. Subjects were able to recognize the normal state of the patient with less error when using the configural display as opposed to the traditional waveform trending display.[19] Also adding support for normal reference frames, Blike[30] compared variations of an ecological display and a traditional alpha-numeric display representing hemodynamic information. Subjects were asked to identify five basic patterns of hypotensive shock in a flash-card style computer-based test. Blike[30] concluded that subjects were able to perform better with the ecological displays compared with the alpha-numeric display due to the presence of boundary information indicating normal ranges of parameters.

Emergent Features

In the domain of graphical displays designed for patient care, emergent features provide additional information about the patient's status. In similar studies, noted previously, Blike et al.[18,31] also attributed the success of subjects' ability to improve their diagnostic accuracy to emergent features shown on the researchers' ecological displays. Blike concluded that the shape (emergent feature) of the overall object was accountable for the improved accuracy.[31] Yet, he also cautioned that the emergent features must have clinical meaning to provide similar results. Emergent features should emphasize a clinically relevant change in patient status, thereby enhancing the detection and diagnostic power of the display.

Methods

The pulmonary display development process began by listing the variables that were considered by clinicians to be important in diagnosing pulmonary events (Tables 1 and 2).[32] With expert consultation and literature reviews, the scope of the pulmonary display included the following set of variables: tidal volume (V_T), respiratory rate (RR), fractional inspired oxygen (F_{IO_2}), end tidal carbon dioxide ($_{ET}CO_2$), fractional alveoli oxygen (F_{AO_2}), upper airway resistance (Upper R_{AW}), lower airway resistance (Lower R_{AW}), and compliance (C_L).[32] These variables were chosen because they are used commonly to diagnose pulmonary events.[32] With these variables and the concepts of configural and ecological displays in mind, we began to design a pulmonary display. Many initial designs were considered (Fig. 4) before the development team decided on the initial shape for the pulmonary graphical display shown in Figure 2.

Table 1 ■ A Subset of Measured Pulmonary Variables Determined by Experts to Be Useful Information When Diagnosing Pulmonary Events in Anesthesia

Variables	Meaning
VT	Tidal volume
RR	Respiratory rate
I:E	Inspiratory to expiratory rate
F_{IO_2}	Fractional inspired O_2
PEEP	Peak end expiratory pressure
Sp_{O_2}	O_2 saturation of arterial blood measured by pulse oximeter (infers Sa_{O_2})
Sa_{O_2}	O_2 saturation in arterial blood
Sa_{CO_2}	CO_2 saturation in arterial blood
Pa_{O_2}	O_2 partial pressure in arterial blood
Pa_{CO_2}	CO_2 partial pressure in arterial blood
Sa_{O_2}	O_2 saturation in alveoli
Sa_{CO_2}	CO_2 saturation in alveoli
$_{ET}CO_2$	End tidal CO_2
FRC	Forced residual capacity
pH	Acidity of blood
PIP	Peak inspiratory pressure (indirectly measures R_{AW} and CL)
Vd	Dead space
CL	Compliance
Lower R_{AW}	Lower airway resistance
Upper R_{AW}	Upper airway resistance
F_{AO_2}	Fractional alveolar oxygen
V/Q	Volume/flow match
Shunt	Shunting of blood
A-a_{O_2}	Arterial/alveoli O_2 gradient

A three-step testing protocol was used. It consisted of three paper-based tests intended to evaluate the pulmonary display design's intuitiveness and the ability to support the diagnosis of pulmonary events. For each of the tests, the subjects were informed that the display represented data from a mechanically ventilated patient. Subjects, comprised of the intended users of the display, attempted to identify the anatomic, physiologic, and graphical meaning of the display from a predetermined set of choices. A set of choices for each test was comprised of both intended answers and related possible choices to assess the subject's interpretation of the display elements or intended diagnoses. Test results were compared with the designers' intentions for the display.

The first paper-based test asked the user to identify the anatomic features of the design. The anatomic test was designed to assess the subject's ability to associate the element of the display with the display's underlying intended representation. For instance, in Figure 5, the designers' intent was for the accordion-shaped cylinder to represent the bellows of the anesthesia machine.

Table 2 ■ A Subset of Adverse Pulmonary Events that May Be Encountered during Anesthesia and Diagnosed Using Monitored Pulmonary Variables*

Event	Variables Used to Diagnose
Hypoventilation	Tidal volume, $_{ET}CO_2$, F_{AO_2}, RR, I:E, Sa_{O_2}
Hyperventilation	Tidal volume, $_{ET}CO_2$, F_{AO_2}, RR, I:E, Sa_{O_2}
COPD	Tidal volume, $_{ET}CO_2$, F_{AO_2}, compliance, Sa_{O_2}
Intrinsic PEEP	Tidal volume, $_{ET}CO_2$, PEEP, RR, I:E
Ventilator malfunctions	Tidal volume, $_{ET}CO_2$, F_{AO_2}, RR, I:E, R_{AW}, compliance, Sa_{O_2}
Stiff lung due to decreased compliance	Tidal volume, $_{ET}CO_2$, F_{AO_2}, compliance
Pneumothorax	Tidal volume, $_{ET}CO_2$, F_{AO_2}, compliance, Sa_{O_2}
Bronchospasm	Tidal volume, $_{ET}CO_2$, F_{AO_2}, lower R_{AW}, Sa_{O_2}
Obstructed ETT	Tidal volume, $_{ET}CO_2$, F_{AO_2}, upper R_{AW}, Sa_{O_2}
Esophageal intubation	Tidal volume, $_{ET}CO_2$, F_{AO_2}, compliance, Sa_{O_2}
Hypoxemia	Tidal volume, $_{ET}CO_2$, F_{AO_2}, Sa_{O_2}
Endobronchial intubation	Tidal volume, $_{ET}CO_2$, F_{AO_2}, compliance
Hypercarbia	Tidal volume, $_{ET}CO_2$, F_{AO_2}, Sa_{O_2}
Pulmonary fibrosis	Tidal volume, $_{ET}CO_2$, F_{AO_2}, compliance

*In many of the events listed, auscultation would be used to diagnose the event. Auscultation was not included in the list of variables because it is not measured.

The second paper-based test asked the user to link the pulmonary variables with the graphical features of the design. For instance, in Figure 5, the height of the bellows reflected the patient's tidal volume.

The third paper-based test asked the user to diagnose pulmonary events based on the emergent features of the pulmonary display. The display represented an adverse pulmonary event with portions of the display deviating

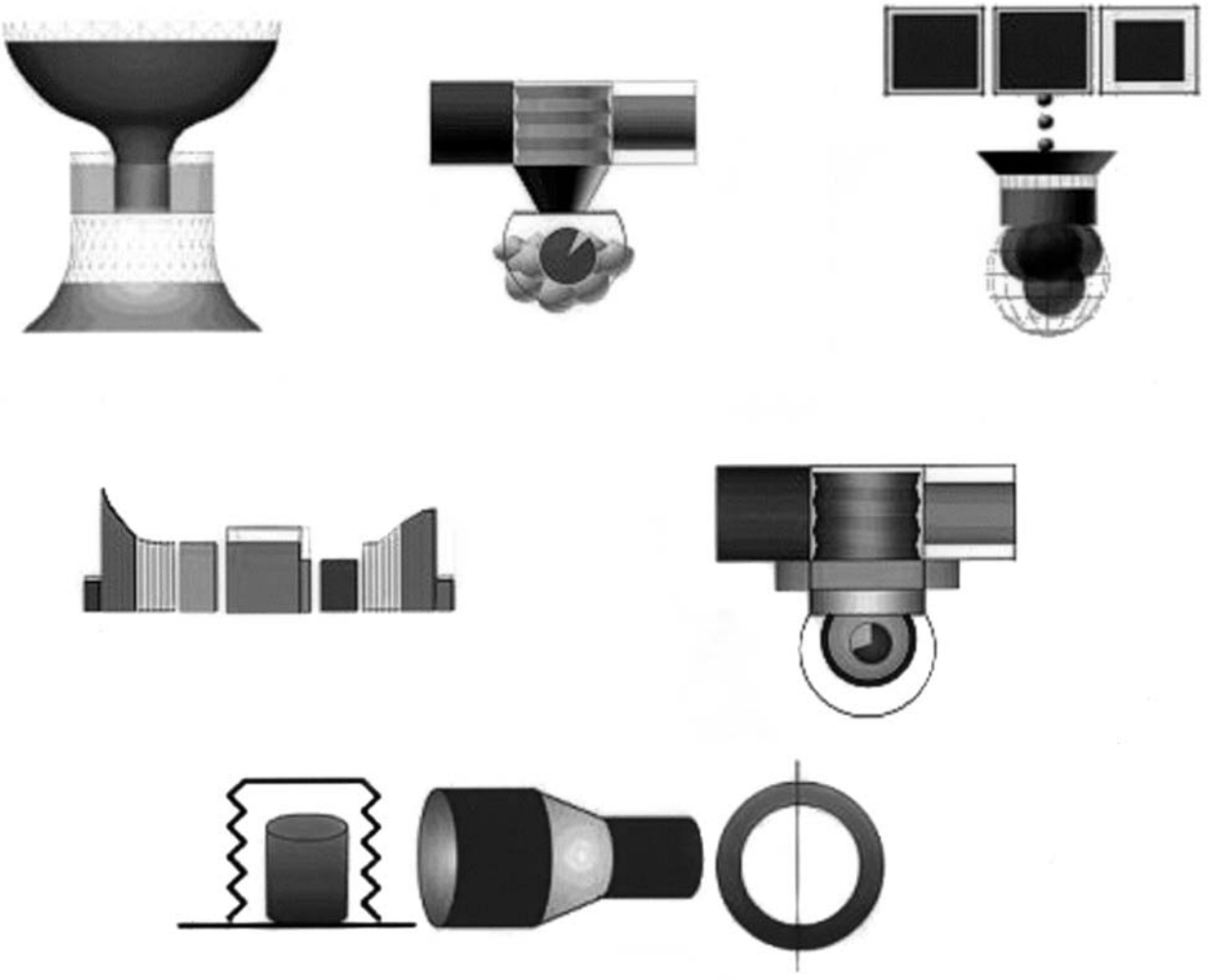


Figure 4. Some preliminary pulmonary designs.

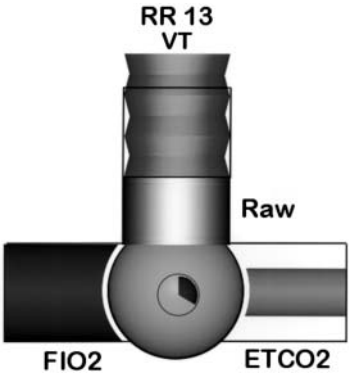


Figure 5. Hyperventilation pulmonary event.

from the normal reference frame. For example, the bellows object increased in height (increased tidal volume) with respect to the normal reference line, while the $ETCO_2$ box decreased in height (decreased $ETCO_2$) to represent a hyperventilation pulmonary event (Fig. 5).

After obtaining an institutional review board (IRB) approval for a multicentered study, subjects were recruited from the University of Utah Hospital, VA San Diego Medical Center, and University of Arizona Hospital. The 46 study subjects included anesthesiologists (n = 22), nurse anesthetists (n = 1), residents (n = 18), and medical students (n = 5).

Results

The iterative process led to five designs that were evaluated sequentially (Table 3). Each design anatomically represented the same objects (lungs, airway, bellows, inspired gas, and expired gas) and measured variables (fractional alveoli oxygen, airway resistance, tidal volume, fractional inspired oxygen, and end tidal carbon dioxide). Each of the designs was tested as described above. It quickly became apparent after only a few subjects that designs 1 and 2 were not yet optimized.

Design changes were based on analysis of the data from the three paper-based tests that were assembled into confusion matrices (Tables 4 and 5). The confusion matrices listed the designer's intended answers in the left column and the user's actual answers along the top row for each of the three paper-based test types per design cycle. An intended answer that matched the subject's answer contributed a data point on the matrix diagonal. Thus, answers deviating from the diagonal were tagged as potentially confusing features. With this analysis, potential design weaknesses were seen easily. As an example, the anatomic testing confusion matrixes of design 5 and design 3 are shown in Table 4 and Table 5.

On identifying a suboptimal feature of the design, the display was altered and retested with a new group of subjects. For example, the airway representation in design 3 (Table 4) was confusing. On changing the design to an anatomic picture of the airway in design 5, the pulmonary design appeared to become more intuitive. The iterative cycle of design and testing continued until the anatomic and variable mapping test scores reached above 90%, and the diagnostic test score reached above 75%. We used a lower diagnostic test score criterion because the diagnostic images were snapshots of pulmonary events in time, and a time reference was not included in the diagnostic test. At the point of meeting the tests' thresholds, the team determined that the design had completed the iterative development cycle and was ready for phase two of the study that will focus on more formal testing.

The cumulative percentages of each of the three tests for each of the five designs are shown in Figure 6. Using the overall trend that emerged with each design cycle as a guide, the design progressed, and the elements of the display that were addressed according to the design and testing process became easier to recognize.

With our iterative design process, the pulmonary display began with an initial design with an average of 70% recognition rate of all three tests and finished with a final design with an average of 88%. The design iterations improved the anatomic recognition of the display from 73% to 98% and of the mapping of variables from 57% to 91%. The final pulmonary display design resulted in diagnosis of eight pulmonary events with 79% accuracy.

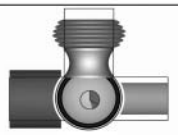
The final pulmonary display, Figure 7, anatomically represented the bellows, airway, lungs, inspired gas, and expired gas. The accordion-shaped bellows object was similar to the bellows in the anesthesia ventilator machine. The pulmonary graphical bellows moved vertically representing changes in tidal volume with respect to normal. The airway graphic resembled the trachea and the branched bronchi shown in Figure 3A. Also, the lung object was shaped as a bisemielliptical sphere similar to the anatomic picture of the lungs shown in Figure 3B. Colors were chosen according to the operating room standards used in the United States: green was used to indicate oxygen and gray to indicate carbon dioxide. The shade of green was mapped to the amount of oxygen in the alveoli (FAO_2). The green box (*upper left*) represented inhaled gas, and the shade of green is mapped to FIO_2. The gray box (*upper right*) represented exhaled CO_2 and the height of the box is mapped to $ETCO_2$ (Fig. 7A).

The pulmonary display contained emergent features to show lung compliance and airway resistance. As shown in figure 7B, a billowy enlargement of the lung depicted an increase in lung compliance, such as in emphysema. On the other hand, a decrease in lung compliance, as in pulmonary fibrosis, was depicted by a black mesh outlining the normal reference frame of the lung (Fig. 7C). An obstruction in the lower airways (such as bronchospasm) was depicted as a narrowing in both of the two lower bronchi (Fig. 7D). An obstruction in the upper airway (such as obstructed endotracheal tube) was depicted as a narrowing of the upper trachea (Fig. 7E).

Discussion

The pulmonary display's intuitiveness resulted from extensive iterative usability testing. If the display is intuitive,

Table 3 ■ The Design Pictures and Changes of Each of the Five Design Iterations

Design #	Design Images	Design changes in response to usability testing	N (Att, CRNA, Res, Med Stu)
1		❖ initial design	3 (3,0,0,0)
2		❖ added intrinsic PEEP diagnostic event ❖ removed FAO2 center dial and replaced with shade of green of round ball	6 (3,0,2,1)
3		❖ moved compliance line to outside of circle rather than inside circle for stiff lung diagnostic event in diagnostic test and variable mapping test	10 (4,0,3,3)
4		❖ change airway to yellow color instead of gray/silver ❖ intrinsic peep shows breath stacking, ❖ stiff lung shows emergent cage around green ball	14 (4,1,8,1)
5		❖ Changed the lungs to be more anatomical ❖ Changed the airway to be more anatomical ❖ Separated the upper and lower resistance ❖ Final design	13 (8,0,5,0)

Abbreviations: Att, attending anesthesiologists; Res, residents (first-to third-year); CRNA, certified registered nurse anesthetists; Med Stu, medical students (third- and fourth-year).

the user will need less training to learn to use the display and should be able to quickly and accurately diagnose adverse pulmonary events through pattern recognition.[5,13]

Our iterative development cycle allowed the refinement of the display before expensive formal testing. As needed with any new display for medicine, the development process of the pulmonary graphical display will continue with detailed testing of the display in realistic clinical environments. For our formal testing, we will use high-fidelity anesthesia simulators and animated pulmonary events to test the display's utility for helping make accurate and rapid diagnoses. Operating room evaluations will be a final predictor of the display's success.

Referring to Figure 6, it is interesting to note that although the recognition of the anatomic and variable mapping features of the design improved with each iteration, the accuracy of the diagnoses decreased until the final design was developed. One possible reason could be that the display during the first four iterations lacked gestalt characteristics. The individual parts of the display were anatomically intuitive, but the integrated graphical display of the patient's condition did not appear emergent to the users. Also, noting the subtle changes of designs 1 through 4, our team had experienced bias toward our pulmonary display design 1. Admittedly, we wanted to believe that design 1 would be intuitive to the user, and therefore we changed the design only subtly to address the user's confusion points evident in the confusion matrices. After

Table 4 ■ The Anatomic Test's Confusion Matrix for Design 3*

	Subjects' Choices								
Designers' Intentions	Inspired Gas	Expired Gas	Bellows	Airway	Lungs/SOGE†	CO_2 Absorber	Diaphragm	Mouth	Blank
Inspired gas	8	1			1				
Expired gas	1	9							
Bellows			10						
Airway				5		2	1	1	1
Lungs/SOGE					10				

*All users identified the designer's intentions for the lungs and bellows. Most users identified the inspired and expired gas objects. The results indicated, however, that the airway might require additional design refinement.
†SOGE = site of gas exchange.

exhausting many design options, we were forced to concede that designs 1 through 4 were not sufficiently intuitive and began to consider a complete redesign before the group decided to test design 5. This design yielded a diagnosis test score close to the original with an improved anatomic and variable mapping test score.

Design 5 resulted in the desirable gestalt characteristics. The airway was modified to look anatomically like an airway and, thus, probably brought coherence to the entire design. Although the final diagnostic test score of design 5 calculated to be 79% compared with 83% in the first iteration, we felt that because all three scores had an upward trend, the anatomic score of 98%, and a variable mapping score of 91%, the design had reached a stage at which it would be appropriate to begin formal testing.

The change from a gray cylinder for the airway to a yellow cylinder in design 4 may have caused the dip in scores for all of the three tests during the fourth design iteration. The change of color was suggested due to the consistent association of the airway representation with the CO_2 absorber in previous design iterations. It is interesting to note that the yellow color may have even further enforced association with the CO_2 absorber. In one case, the subject commented that the object looked strongly like a CO_2 absorber full of CO_2 because, in his mind, CO_2 is an undesirable gas, and yellow is an undesirable color. Further, the design team noted that the use of color as an indicator may not be optimal because some users may be colorblind. For instance, colorblind users may have problems detecting the shade of green of the lung object mapped to the amount of measured oxygen in the lung (FAO_2). The colors were chosen because they add intuitiveness to the display by conforming to the accepted color choices to represent these gases in U.S. operating rooms. Colors that might be undetected by colorblind users are labeled redundantly.

Although peak inspiratory pressure (PIP) was included as one of the important variables used to diagnose pulmonary events, we chose not to visualize PIP in our pulmonary display. First, our group determined through usability testing that we were not portraying PIP adequately to the anesthesiologist with our graphical displays. Second, realizing that PIP was used as an indirect indicator of resistance or compliance, we decided to directly portray the clinical information through our resistance and compliance elements in the pulmonary display.

The design process described in this report was used to develop the pulmonary graphical display for anesthesia. Recognizing the need for formal testing of medical displays, this process should be used prior to formal testing to increase the likelihood of a successful outcome. Through the rapid and inexpensive development process described here, the pulmonary display was enhanced and refined through several stages (from design 1 shown in Figure 2 to design 5 shown in Figure 7) to a design deemed suitable for clinical testing.

Study Limitations and Conclusions

A multiple-choice format was used for the three paper-based tests to ensure answers were unchallenged as correct or incorrect. The disadvantage of the testing format is that the set of possible answers may have influenced the subjects to choose an answer from the list, rather than what came to mind immediately on viewing the display image. A hybrid test format consisting of multiple-choice matching and

Table 5 ■ The Anatomical Test's Confusion Matrix for Design 5*

	Subjects' Choices					
Designers' Intentions	Inspired Gas	Expired Gas	Bellows	Airway	Lungs/SOGE†	CO_2 Absorption
Inspired gas	13					
Expired gas		12				1
Bellows			13			
Airway				13		
Lungs/SOGE					13	

*With one exception, all graphical features were correctly identified by the subjects.
†SOGE = site of gas exchange.

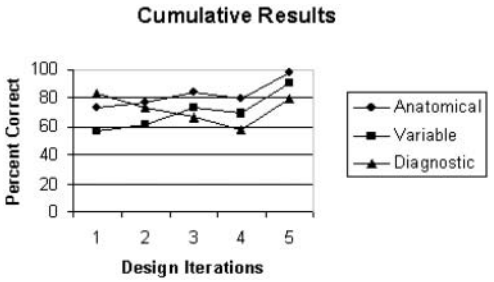


Figure 6. The cumulative test scores for each of the five design iterations.

a section to freely contribute ideas may be an improvement to the current test design.

Also, the small number of subjects in each of the study groups may be considered a limitation. We felt it was essential to test the intuitiveness of the design with the actual users of the system. Therefore, the number of users at our participating facilities was a limiting factor. It was also necessary to test subjects who had not seen any form of the display prior to testing. Studies have found that four to five subjects are adequate for identifying 80% of the usability issues with a proposed design,[33,34] and our later designs were tested with ten to 14 subjects.

Another important aspect is that the intuitiveness of this design may change when the pulmonary image is animated with real-time data. The testing described in this report used static images on paper without an indication of time elapsed. Future studies are needed to ensure the display remains intuitive and usable when animated with patient data.

In addition to graphical information, we recognize that the numeric data and waveforms are important information. These values were omitted in this preliminary testing to evaluate the value of the information provided by the pulmonary display alone. Future implementations of the pulmonary graphical display will be supplemented with standard clinical numeric values.

We describe a straightforward iterative design process involving frequent simple usability testing that resulted in a graphical pulmonary display with improved intuitiveness and usability. A similar approach has been used successfully to design clinically useful cardiovascular and drug displays prior to their testing and implementation.[28,29] The next phase of testing of the putative pulmonary display described here will be in an anesthesia simulator.

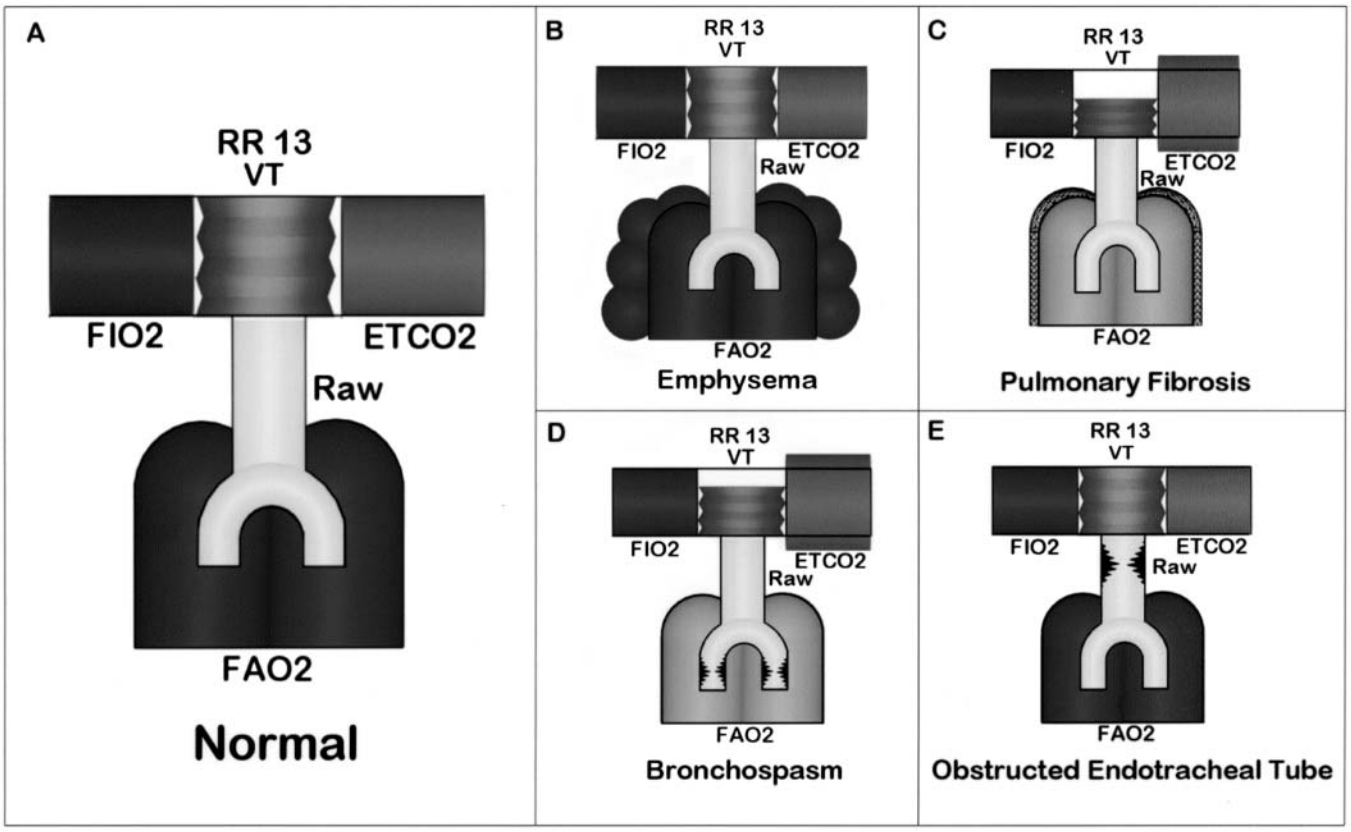


Figure 7. (A) The final design (design 5) of the pulmonary display in a normal state. (Inspired gas : green, bellows : blue, expired gas : gray, airway : pink, lungs : green). (B) Emergent features of the design show a lung with emphysema. The billowing lung depicting an increase in compliance was an emergent feature of the pulmonary display. (C) Different emergent features of the design show a lung with pulmonary fibrosis. The cage surrounding the lung depicted a decrease in compliance (stiff lung). Black triangular objects show lower (D) and upper (E) airway restriction.

References ■

1. Kohn LT, Corrigan JM, Donaldson MS. To Err Is Human: Building a Safer Health System (eds). Institute of Medicine. Committee on Quality of Health Care in America. Washington, DC: National Academy Press, 2000, p 312.
2. Cooper JB, Newbower RS, Long CD, McPeak B. Preventable anesthesia mishaps: a study of human factors. Anesthesiology. 1978;49:399–409.
3. Cooper JB, Newbower RS, Kitz RJ. An anlysis of major errors and equipment failures in anesthesia management: considerations for prevention and detction. Anesthesiology. 1984;60:34–43.
4. Loeb RG. A measure of intraoperative attention to monitor displays. Anesth Analg. 1993;76:337–41.
5. Allnutt MF. Human factors in accidents. Br J Anaesth. 1987;59:856–64.
6. Arnstein F. Catalogue of human error. Br J Anaesth. 1997;79:645–56.
7. Whitcher C, Ream AK, Parsons D, et al. Anesthetic mishaps and the cost of monitoring: a proposed standard for monitoring equipment. J Clin Monit. 1988;4(1):5–15.
8. Endsley M. Measurement of situation awareness in dynamic systems. Hum Factors. 1995;37:65–85.
9. Rasmussen J, Vincente KJ. Evaluation of a Rankine cycle display for nuclear power plant monitoring and diagnosis. Hum Factors. 1996;38:506–21.
10. Adams M, Tenney VJ, Pew RW. Situation awareness and the cognitive management of complex systems. Hum Factors. 1995;37:85–104.
11. Sotkes A, Wickens CD. Aviation displays. In: Human Factors in Aviation Displays. San Diego: Academic Press, 1988; pp 387–431.
12. Bennet KB, Walters B. Configural display design techniques considered at multiple levels of evaluation. Hum Factors. 2001;43:415–34.
13. Moray N. Conceptual frameworks. In: Durso FT (ed). Handbook of Applied Cognition. New York: John Wiley & Sons, 1999;215–43.
14. Gurushanthaiah K, Weinger MB, Englund CE. Visual display format affects the ability of anesthesiologists to detect acute physiologic changes. A laboratory study employing a clinical display simulator. Anesthesiology. 1995;83:1184–93.
15. Michels P, Gravenstein D, Westenskow DR. An integrated graphic data display improves detection and identification of critical events during anesthesia. J Clin Monit. 1997;13: 249–59.
16. Cole W, Stewart JG. Metaphor graphics to support integrated decision making with respiratory data. Int J Clin Comput. 1993;10(2):91–100.
17. Cole WG, Stewart JG. Human performance evaluation of a metaphor graphic display for respiratory data. Methods Inf Med. 1994;33:390–6.
18. Blike G, Surgenor SD, Whalen K. A graphical object display improves anesthesiologists' performane on a simulated diagnostic task. J Clin Monit. 1999;15:37–44.
19. Jungk A, Thull B, Hoeft A, Rau G. Ergonomic evaluation of an ecological interface and a profilogram display for hemodynamic monitoring. J Clin Monit. 1999;15:469–79.
20. Jungk A, Thull B, Hoeft A, Rau G. Evaluation of two new ecological interface approaches for the anesthesia workplace. J Clin Monit Comput. 2000;16:243–58.
21. Weinger MB, Herndon OW, Zornow MH, Paulus MP, Gaba DM, Dallen LT. An objective methodology for task analysis and workload assessment in anesthesia providers. Anesthesiology. 1994;80:77–92.
22. Cook RI, Potter SS, Woods DD, McDonald JS. Evaluating the human engineering of microprocessor-controlled operating room devices. J Clin Monit. 1991;7:217–26.
23. Patel VL, Kushniruk AW. Interface design for health care environments: the role of cognitive science. Proc AMIA Symp. 1998:29–37.
24. Gardner RM, Prakash O. Challenges and opportunities for computerizing the anesthesia record. J Clin Anesth. 1994;6: 333–41.
25. Nazir T, Beatty PC. Anaesthetists' attitudes to monitoring instrument design options. Br J Anaesth. 2000;85:781–4.
26. McPhee K. Design Theory and Software Design. Edmonton, Alberta: The University of Alberta, 1996, p 76.
27. Snodgrass A, Coyne R. Models, metpahors and the hermeneuticls of designing. Design Issues. 1992;9(1):56–74.
28. Agutter J, Syroid ND, Drews FA, Westenskow DR, Bermudez JC, Strayer DL. Graphic Data Display for Cardiovascular System. In: IEEE Symposium on Information Visualization, San Diego, 2001.
29. Syroid ND, Agutter J, Drews FA, et al. Development and evaluation of a graphical anesthesia drug display. Anesthesiology. 2002;96:565–75.
30. Blike G. The boundary information in a gray-scale object display and a color enhanced variant, improve problem recognition compared to an alpha-numeric display. Anesthesiology [abstract]. 1997;87 Sep(3A):A458.
31. Blike G. The emergent shape features in a gray-scale graphical display improve pattern recognition compared to an alpha-numeric display [abstract]. Anesthesiology. 1997;87 Sep(3A): A394.
32. Gaba D, Fish K, Howard S. Crisis management in Anesthesia. New York: Churchill Livingstone, 1994.
33. Virizi R. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors. 1992;34:457–68.
34. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. Proceedings of ACM INTERCHI'93 Conference. Amsterdam, The Netherlands: Apr 1993, pp 206–13.
35. Grippi MA. Pulmonary Pathophysiology. Philadelphia: Lippincott-Raven, 1995, 337.