Interfaces to Medical Information Systems: Supporting Evidenced Based Practice

Timothy R. McEwen
Underwriters Laboratories
Concord, MA USA
tmcewen@gmail.com

John M. Flach
Department of Psychology
Wright State University
Dayton, OH USA
john.flach@wright.edu

Nancy C. Elder
Department of Family Medicine
University of Cincinnati
Cincinnati, OH, USA
eldernc@ucmail.uc.edu

Abstract—This paper describes the development and initial evaluation of an interface to medical information related to cardiovascular health. The interface, which is designed to be used on a touch tablet, uses principles of Ecological Interface Design to create an interface that addresses cardiovascular health at multiple functional levels: specific test values, general state of health based on a published epidemiological model of risk, and treatment guides based on published standards of best practice. The evaluation focused on the value of providing multiple functional levels for making judgments about cardiovascular risk.

Keywords—Cognitive Systems Engineering; Ecological Interface Design; Electronic Medical Information Systems

I. INTRODUCTION

Electronic Medical Information Systems provide a unique opportunity to make patient data available to practicing physicians. However, as has been seen in other domains, access to data is generally not sufficient to achieve the goals of productive thinking. In fact, for complex domains like nuclear power, there is a danger that people can be overloaded by too much data, making it very difficult for them to diagnose faults and to prescribe effective interventions. However, there appears to be an antidote to the problem of too much data. Research has shown that performance can be greatly enhanced when the data are organized or configured in ways that reflect functional process constraints [1].

Thus, the problem of data overload is rarely an unavoidable consequence of limits in human information processing capacity. Rather, the problems typically reflect poor design choices with respect to the way data is represented. Even for complex domains, productive thinking is possible when data is organized to reflect both the deep structure of the work domain and the pattern recognition capabilities of humans [2,3,4].

The development of Electronic Medical Information Systems not only facilitates access and handling of medical information, but it also opens up the possibility to take advantage of the capabilities of digital display devices (e.g., computers, tablets, and smart phones). These capabilities include the use of color, graphics, and motion to build configural representations. Thus, we began exploring the potential to utilize these capabilities to build representations of medical information with the explicit goal of supporting evidence-based medical decision making. The goal was to organize patient health data to reflect the underlying scientific evidence-base. The hypothesis was that this organization would not only reduce the possibility that physicians would miss critical test results, but that additionally it would enhance the meaningfulness or significance of the data relative to evidence-based standards of best practice. Our goal was to create a representation that would let the physician ‘see’ the implications of the data relative to risks and treatment options associated with cardiovascular health.

II. ECOLOGICAL INTERFACE DESIGN APPROACH

A. Overview

The Ecological Interface Design approach is based on a triadic model of the semiotic system (Fig. 1) [1,6]. This model explicitly recognizes that there are three components to effective communication. These components are: the information source, the representation, and the problem solver. In the case of medical systems the information source is the results from medical examinations and tests assessing a patient’s health. The problem solver is the physician. Note that it is clear that effective problem solution depends on collaboration between patient and physician. And we believe that one of the goals for an effective representation of medical information is to facilitate communication between physicians and patients. However, the current design and evaluation focused specifically on physicians and their role in making decisions about the course of treatment.

Fig. 1. The triadic semiotic model illustrates that the representation should reflect structure of the problem domain (e.g., patient health) in a way that is coherent to a physician and that leads to judgments and actions that correspond with functionally significant aspects of the problem.
In designing an effective representation it is necessary to consider both the information source and the problem solver. The goal is to design the representation so that it will shape the mental model of the physician to become a valid model of the patient’s health. Where ‘valid’ means a model of health that is consistent with the latest medical evidence-base and that leads to effective treatment. Note that this is somewhat different than more conventional human factors and HCI approaches that focus on designing interfaces that match the mental models of the users. The key difference is that the EID approach does not assume that all users (e.g., all physicians) have valid models. The EID approach requires that designers must not only model the users, but they must also base their design on explicit normative models of the work domain (in this case, the medical evidence base). Thus, the EID approach is formative in that a goal is to shape mental models toward more productive ways of thinking [7]. In medical domain terms, the goal is to ensure that clinical decision making is guided by the scientific evidence base.

B. Graphical User Interface (GUI) Design

The design goal was to leverage the power of graphical computer displays to provide a representation that would improve physicians’ ability to ‘see’ a patient’s data in the context of current medical research on cardiovascular disease (CVD). This is consistent with previous arguments we have made - that the value of electronic healthcare systems will not rest with replacing fallible humans, but rather with enriching the perception-action coupling with smart humans, through enhanced visualizations of the complex domain of medicine [11]. The first step toward meeting that challenge with respect to CVD risk was to discover the deep structure that could potentially guide physician decision-making.

The means to discovering the deep structure is to do work analysis [7,8]. In contrast to conventional task analysis, the focus of work analysis is not on mental activities, but rather the goal is to identify the functional constraints that shape the activities or that bound the field of possible and/or safe activities. With conventional task analysis the emphasis is descriptive – that is describing what people think and do, with work analysis the emphasis is formative – that is describing what could be done (i.e., physical and regulatory constraints on behavior) and what should be done (functional goals and values).

The work analysis underlying the CVD display has been described in a previous publication [9]. This work analysis involved interviews with domain experts and review of the medical literature for models of cardiovascular health and for published treatment standards. In this paper, we will present key discoveries from the work analysis in explaining the logic of our interface design shown in Fig. 2.

In this section we will describe the logic of our design decisions with specific emphasis on how our choices were motivated by the deep structure uncovered in the work domain analysis. The GUI show in Fig. 2 is divided into four distinct regions. Region 1 provides a complete summary of the raw data driving the display. Region 2 illustrates how individual variables contribute to the overall risk of a cardiac event, using the Framingham model of general CVD risk [11]. Region 3 illustrates published guidelines for treating hyperlipidemia [12] and Region 4 illustrates published guidelines for treating hypertension [13]. Each region will be described in greater detail in the following sections.

Color is used in the GUI in two ways. Shades of Green, Yellow, and Red are used to indicate the status of a variable with respect to health and the potential need for treatment. Green indicates good health (no treatment required), yellow indicates a low threat to health suggesting a need for change in life style, and red indicates high threat to health indicating the need for pharmaceutical treatment options. Other colors are used to map individual variables into the Framingham model. Each variable in Region 1 is color-coded and the contribution of that variable to the Framingham score is indicated using that color in the contribution graphic in Region 2 that will be described more fully later.

In Region 1 of the GUI, the basic data relative to the risk of CVD is presented. The data formats include number lines for continuous variables and switches for discrete, dichotomous variables. Continuous variables include age, Cholesterol values (Total, LDL, HDL, and triglycerides), blood pressure (Systolic and Diastolic), and the total risk of CVD based on the Framingham model.

Each continuous variable is presented on a number line that is color coded to reflect clinical norms as described above: bright red indicates a dangerous level; bright green indicates a healthy level; and yellow reflects moderate levels. Shadings of these three colors are used to indicate intermediate levels where appropriate. As previously noted, the colors on the labels are used to link individual variables to other regions of the display. This is true for all labels accept the label on the bottom line assessing overall risk of CVD. The color of that label changes as a function of total risk using the red, yellow, green convention.

Specific values of a variable are indicated by the position of indicators (either triangles or squares) on the number lines. The position information is supplemented with the numerical value specified digitally above the indicators. The triangle indicators are slider controls that can be directly manipulated to change a value. The square indicators cannot be directly manipulated, but are instead used for representing variables that are
computed based on the values of lower order variables. For example, total cholesterol cannot be entered directly, but it is determined as the weighted sum of the lower order cholesterol variables [e.g., LDL + HDL + (5 x Triglycerides)].

The discrete variables include whether the person has existing heart disease; whether there is a family history of heart disease; whether glucose levels are above 110 mg; whether the person is diabetic; whether the person smokes; whether the person is obese (waist > 40 in.); and whether the person is under treatment for blood pressure. These variables are indicated by switches: Y indicates that the discrete criterion is satisfied (e.g., yes the person smokes; or yes the person is diabetic): N indicates that the discrete criterion is not satisfied (e.g., there is no family history of heart disease).

Note that the choice to represent specific variables as continuous or discrete was made based on how they are treated in the Framingham model. We are reassessing this decision and are considering changing some of the discrete representations (e.g., glucose level and waist size) to continuous representations. We believe that the added information provided by continuous representations may be valuable, even though that information will not directly impact the Framingham score.

The graphic in Region 2 was designed to illustrate the Framingham Risk model [11]. The overall risk of CVD, based on the Framingham model is represented in the vertical axis of the graph. This axis is color coded to reflect three general categories of risk (red – high risk; yellow – moderate risk; green – low risk). These categories are based on decision criteria associated with the published guidelines for treating hyperlipidemia [12]. The horizontal axis of the graph is the sum of the weighted variables that are used in the Framingham model. Below the horizontal axis is a contribution graphic to illustrate the impact of specific variables on the Framingham score. Variables where the product of the weighting coefficient from the Framingham model and the value relative to the population mean is positive are represented in the bottom portion of the contribution graphic with the arrow pointing to the right. These variables contribute to increase the risk of CVD. Variables where the product of the weighting coefficient and the value relative to the population mean is negative are represented in the top portion of the contribution graphic with the arrow pointing to the left. These variables contribute to the reduction of CVD risk.

The contribution graphic is configured so that the sum of the positive and negative contributions is aligned with the left edge of the top portion of the contribution graphic. This line then projects through the horizontal axis to intersect with the modeled risk function. From the intersection point with the risk function, the line is projected parallel to the horizontal axis to intersect with the total risk value on the vertical axis. The total risk is also indicated with a redundant digital display that changes color to correspond with the category of total risk.

The actual risk function in the graph is a function of the population used to empirically derive the correlation function. In the case of the Framingham model, the correlations were computed separately for males and females. Thus, when the gender switch is changed in Region 1, the actual graphic function will change in Region 2. The function for females is much shallower than the function for males, indicating that females have lower overall risks of CVD.

As noted earlier, color is used to link the individual variables in Region 1 with their contribution to total risk in Region 2. For example, the cholesterol variables are colored in shades of blue and the blood pressure variables are colored in shades of purple. For the example in Fig. 2 total cholesterol is a bit high, increasing risk. However, HDL cholesterol levels are at a good level, reducing the overall risk.

Due to space limitations in the display it can be difficult to parse the contribution graphic to easily identify the contributions of each specific variable to the Framingham score. To compensate for this, a circular, cyan colored toggle button was added (labeled with an “i”) that opens (and closes) an exploded view of the contributions of each specific variable as shown in Fig. 3. With the exploded view, the variables are listed in order from the variable making the greatest contribution to increasing risk to the variable making the greatest contribution to reducing risk. In the example, systolic blood pressure is the biggest risk factor, with total cholesterol also contributing to increased risk. Strong factors in this persons favor include the fact that he doesn’t smoke and his age.

Region 3 is designed to reflect the decision criteria associated with the Third Report of the Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (ATP-III) [12]. There are three treatment categories associated with the ATP-III guidelines: no treatment; therapeutic lifestyle change; and drug treatment. These three treatment categories are illustrated as the three color regions in the graph to the right of the Framingham graph in Region 2: red – drug therapy; yellow – therapeutic lifestyle change; and green – no treatment.

One of the primary factors to consider with respect to the ATP-III guidelines is the CVD risk score computed using the Framingham model. Thus, the vertical axis of the ATP-III graphic corresponds with the vertical axis in the Framingham graphic. These are linked by the horizontal line that projects to total risk on the vertical axes of both graphics. A second major factor to consider with respect to the ATP-III guidelines is the level of LDL cholesterol. This variable is represented on the
horizontal axis. The value for the patient is projected vertically to intersect with the Framingham score line. The color of the region in which these two variables intersect reflects which of the three treatment options is recommended by the evidence-base.

In addition to the Framingham Score and the level of LDL cholesterol, five risk factors are also considered in the ATP-III guidelines: smoking, whether blood pressure exceeds 140/90 mm/Hg or on a hypertensive medication; low HDL cholesterol (<40 mg/dL); family history of CVD; and age (men ≥ 45; women ≥ 55). The pentagon labeled Risk Factors indicates the state of each of these variables (green – factor present). The presence of any three of these factors can change the treatment criterion. This would be reflected in changes to the boundaries associated with the treatment options in the risk space (Framingham Score × LDL).

Finally, an additional five factors are considered in relation to a condition called Metabolic Syndrome. Factors associated with Metabolic Syndrome include HDL level, triglyceride levels, blood pressure, obesity, and blood sugar levels. The presence of any three of these factors would require additional attention to therapeutic lifestyle interventions and could impact decisions about drug treatment. The logic behind this graphic is similar to that for the Risk Factors. If a factor is present, the corresponding slice in the pentagon will change to green. If three or more factors exceed the criterion the Metabolic Syndrome label will be highlighted in red to alert the physician to the presence of this condition.

Region 4 is designed to reflect the decision criteria associated with the Seventh Report of the Joint National Committee on Prevention, Detection, Evaluations, and Treatment of High Blood Pressure (JNC-7) [13]. This is a fairly simple graphic that illustrates four categories of treatment that are contingent on blood pressure measures: No Treatment, Pre-Hypertension (monitoring); Stage 1 Hypertension (drug treatment); Stage 2 (more aggressive drug treatment). The key to this graphic is that the boundaries for each of these stages are contingent on either Systolic (SBP) or Diastolic (DBP) blood pressures. Note that by definition, SBP will always be higher than DBP. Thus, for example, Stage 1 Hypertension would be indicated by either SBP greater than 140 mm/Hg or DBP greater than 90 mm/Hg. In the graphic, different scales are used for SBP and DBP such that they align with the clinical categories of hypertension.

III. EVALUATION

The choice of a semiotic framework also has important implications for evaluation of a design. The focus on a specific problem domain requires care in choosing representative problem scenarios and representative participants. Also, scoring performance in a complex domain like medicine can be difficult, because there may be some uncertainty or disagreement about what constitutes a correct response.

A. Representative Experimental Design

In evaluating the CVD interface it was important to frame questions in terms of the medical domain and to presented cases that were representative of the kinds of patients that a clinical physician might encounter in their practice. For the initial evaluation of the CVD display we created 5 match patient pairs. The pairs were constructed in consultation with domain experts to be representative of prototypical cases. The patients in each pair were match in terms of the treatment recommendations based on ATP III and JNC-7 Guidelines. Table 2 provides summary descriptions of the particular pairs used in our initial evaluations. Note that in complex domains, there can be a wide range of realistic problem scenarios and the value of a particular interface design may vary widely from scenario to scenario. For example, the conventional paper display might be an adequate representation for making judgments about typical patients, but may make it difficult for the physician to see higher order patterns such as those associated with metabolic syndrome.

Table 2: Paired stimulus scenarios for the evaluation.

<table>
<thead>
<tr>
<th>Pair #</th>
<th>Cholesterol Tx</th>
<th>SBP Tx</th>
<th>Context/Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DT</td>
<td>DT</td>
<td>Patients could quit smoking to move from DT for cholesterol to NT</td>
</tr>
<tr>
<td>2</td>
<td>NT</td>
<td>TLC</td>
<td>Patients had enough risk factors to have metabolic syndrome</td>
</tr>
<tr>
<td>3</td>
<td>TLC</td>
<td>NT</td>
<td>Although patients have low LDL cholesterol that would not normally require treatment, their diabetes places them into the high risk area and requires TLC</td>
</tr>
<tr>
<td>4</td>
<td>TLC</td>
<td>NT</td>
<td>Patients have many treatment modifying risk factors that suggest TLC for cholesterol instead of NT if fewer factors were present</td>
</tr>
<tr>
<td>5</td>
<td>TLC</td>
<td>TLC</td>
<td>Patients with high 10-year risk due primarily to age</td>
</tr>
</tbody>
</table>

DT = Drug therapy; TLC = Therapeutic Lifestyle Change; NT = No Treatment Required

In addition to giving care to selecting representative problem scenarios, it can also be important to choose a representative sample of participants for the evaluation. In complex problem domains, no matter how well the interface is designed, some degree of expertise with the problem domains will often be necessary. Thus, one cannot simply ask generic subjects selected from generic pools of college students to make judgments about CVD. This can be a serious constraint on the evaluation process since access to experts can be very difficult.

For our initial evaluation of the CVD interface we were able to get four people who had sufficient expertise to understand the task. Two were faculty physicians with nearly 30 years of experience; one was associated with a department of family medicine and the other was associated with a department of internal medicine. Two additional participants were completing their residency in a family
Another important decision in designing an evaluation of a new display format is to decide on an appropriate baseline against which to evaluate performance. A natural choice is to compare the new design against the interfaces that are conventionally used within the particular domain. However, it is important to note that experts will typically have extensive experience with the conventional displays and will have organized their skills around that representation. On the other hand, the new display will be unfamiliar. In these cases, initial performance with a novel display, even one that is clearly superior, may be worse due to its unfamiliarity. It will often require extensive practice with a new representation before performance benefits are realized. And even in cases where there are clear performance benefits, experts will often prefer the conventional familiar representations. People often do not like change.

For the initial evaluation of the CVD interface, we chose not to use the conventional paper format as the baseline control. Rather, we chose to compare performance using only Region 1 of the CVD interface (NL = number line only) against performance using the full CVD interface including all four regions (NL+ = number line + semantics). The reason for this choice was that we wanted to assess the benefits of the semantic portions of the design (e.g., representing the Framingham model, the ATP III guidelines, and the JNC-7 Guidelines) above and beyond the benefits of basic improvements to the data format (i.e., the use of color coded number lines). Thus, for each pair of patients, participants in the study had one of the patients presented using only the number line portion of the display (only Region 1 was visible) and the other patient presented using the full display (all four regions were visible).

B. Results

We ended up with a total of 16 pairs of judgments. Fig. 3 shows the judgments in relation to conformity with the ATP III Guidelines as a function of whether the judgments with the alternative formats was the same or different. For 8 of the 16 pairs, the clinical judgments of the participants with regards to cholesterol were identical with the different display formats. Of these 8, three times the judgments conformed with the ATP III Guidelines. There were 8 cases where the judgments with the full interface (NL+) were different than the judgments based on the number lines alone (NL). In seven of these cases, the judgment with the full display (NL+) was consistent with the ATP III Guidelines. This suggests that the additional information in the full display helped to make these guidelines more salient.

Figure 4 shows the judgments in relation to conformity with the JNC-7 Guidelines as a function of whether the judgments with the alternative formats was the same or different. For 9 of the 16 pairs, the clinical judgments of the participants with regards to hypertension were identical with the different display formats. Of these 9, five times the judgments conformed with the JNC-7 Guidelines. There were 7 cases where the judgments with the full interface (NL+) were different than the judgments based on the number lines alone (NL). In four of these cases, the judgment with the full display (NL+) was consistent with the JNC-7 Guidelines. In one case, the judgment with the number lines alone agreed with the JNC-7 Guidelines.

In addition to measuring treatment judgments, the participants were asked for their subjective opinions about whether they would like to have the full version of the CVD display available for use in their daily practices. All participants indicated that they would like to use the display in their practices, however, not all wanted the full version of the CVD display to be the format in which test results were initially presented. Two of the physicians indicated that they wanted the test results returned to them from the laboratory in the new format initially, one faculty physician wanted results presented in the traditional alphanumeric format, and one resident wanted results presented in the number line format.

In spite of the mixed preferences, all participants felt that the full CVD display would be helpful to at least have easily accessible for “the complex patients.” A few physicians mentioned that although they would like to use (or have available) the new display, it would be necessary for the
patients’ data to be automatically populated (i.e., from an electronic medical record) into the display for them to use it for decision-making purposes. This suggests that the activity of entering data into the display would not fit into the physicians normal work flow.

One commonly reported reason for using the full CVD display was the ability to see the Framingham model, ATP-III, and JNC-7 guidelines presented in a visual format. Everyone made at least one positive comment on how the new display reduced the need to “try and remember the guidelines.” As one resident reported, the new display would “speed things up” for him because he would not “have to think of the guidelines in his head.” The other resident agreed on this point, and added that this format would “force people to use evidence-based practice.”

All participants also mentioned that the new display would be a helpful tool for patient education, particularly for the ability to make adjustments with the data and see real-time changes. As one enthusiastic faculty physician put it:

“I like the ability to show the patient the impact of the changes that we’re trying to make. Lovely for motivating patients. [...] This is powerful for a patient—they can actually see what they are doing. You’re giving the patient something. [...]This makes it easy and easy is really important. [...] It’s nice to have something simple and graphic in front of you to say [to the patient], ‘You’re doing the right thing,’ or ‘this is what we’re recommending’. Right now EMRs are designed to deal with billing as opposed to helping us take care of patients. This [the new display] helps you take care of patients.”

C. Conclusions

It is impossible to draw any definitive conclusions about the potential value of the new CVD interface based on this preliminary assessment. However, the performance differences suggest that the new format can lead to increased conformity of judgments with published medical guidelines, consistent with the goals of evidence-based medical practice. However, it is important to keep in mind that these are guidelines, NOT absolute truth, and, in fact, there are disagreements within the medical community, especially with regards to decisions about hyperlipidemia. Overall, however, we are encouraged by the enthusiasm and interest of the domain experts in the display. Thus, we are continuing to revise and evaluate the CVD interface as part of an iterative design process.

IV. GENERAL DISCUSSION

The CVD interface is very much a work in progress. There is a number of areas where improvements can be made. Both in terms of the data included in the display and the overall organization of the information. We are also exploring alternative designs targeted to the general population (i.e., assuming minimal medical training). Here preserving the underlying logic of the Framingham model will be less important. Rather the emphasis will be on representing the relative significance of each of multiple factors to overall health. The point of that interface will be to facilitate communications between patient and physicians, with the goal of increasing patient compliance with the physician’s treatment decisions.

Although the jury is out with respect to the specific CVD interface design described here, the EID approach to interface design has been well vetted and there are numerous examples in the literature where this approach has led to interface innovations that have resulted in performance enhancements. The evolution of Electronic Medical Information Systems opens the door to apply the EID approach in ways that can improve clinical decision making by presenting medical data in relation to the scientific evidence base and the standards for best practice. The key is for designers to appreciate that it is not sufficient to represent the medical data. It is imperative that the data be organized in the context of validated medical models and treatment standards.

ACKNOWLEDGMENT

Thanks to Randal Green who developed the Java code for the interface used in these experiments and who has also developed the Objective-C code to create a working prototype for implementation on an iPad.

REFERENCES