Multivariate Relational Visualization of Complex Clinical Datasets in a Critical Care Setting: A Data Visualization Interactive Prototype

Anthony Faiola, Simon Hillier
School of Informatics and School of Medicine, Indiana University
afaiola@iupui.edu, shillie@iupui.edu

Abstract

One mission of medical informatics is to provide physicians, nurses, and other health care providers with the technology and tools for interpreting large and diverse data sets, so that appropriate critical care decisions can be facilitated. Ideally, medical data visualization provides the means to transform data into information and contextual knowledge suitable for interpretation and decision-making [31, 9]. The authors propose a model through which data is organized into multivariate multidimensional critical care patient data visualizations (CPDV). It does this as the primary means to represent and manage complex context-based patient data at various user-defined temporal resolutions. Furthermore, user-defined spatial organization of multiple (clinically related) datasets allows rapid visualization of significant trends that are related to several co-variables. Currently, anticipated findings from usability testing support the notion that the proposed model will facilitate medical decision making in a critical care environment.

Keywords—**Medical data, visualization, human-computer interaction, multivariate, multidimensional, health care**.

1. Introduction

The intensive care unit (ICU) is a data rich environment. By definition, ICU patients have significant organ system derangement and have very little physiologic reserve. Optimal medical care requires intensive monitoring of organ function, frequent multimodal diagnostic testing, and many consultations from subspecialty physicians and other health care providers. In this kind of a critical care environment, it is often difficult to make rapid evaluations of a patient's condition because of the overwhelming volume of data that is continuously generated. Sources of patient-generated data include continuous automatic physiologic monitoring and intermittently determined data that is gathered by bedside care providers and from various diagnostic testing sources. In addition to patient-generated data, there are vast arrays of clinical data generated that document the treatment that is received by the patient. This data includes drug therapy, respiratory therapy, physical therapy and all other clinical interventions [16].

The best clinical decisions pertaining to the care of these patients are made when physicians can easily organize and understand the vast flood of data from these various sources.[1] Unfortunately, physicians and other health care staff have to retrieve this critical data from multiple locations and organize it into a cohesive profile of the patient's current condition. Furthermore, the data that is retrieved from these diverse sources is usually presented in a form that does not allow trends and relationships between co-variables to be immediately recognized. To make the situation worse, critical care physicians (CCP) work in a highly stressful environment and are usually pressed for time, frequently looking after several critically ill patients.

As a result, the process of monitoring, evaluating and treating complex medical problems is labor intensive and time-consuming; it demands that physicians analyze data in text and numeric form. Although the monitoring devices of a modern ICU are intended to support the instantaneous recognition of problems, only a limited amount of past data can be reviewed on the monitor itself. For example, although current monitoring devices, such as HP's CareVue 9000, maintains a comprehensive log of a patient's entire critical care period, its output is in conventional spreadsheet form, with, at best, a few parameters displayed as X-Y plots.

Medical data visualization (MDV) can provide valuable assistance for data analysis and decision making for critical care physicians (CCPs). How CCPs perceive and interact with patient data can affect their ability to arrive at the best critical care solution. As seen with the earliest visualizations of computerized tomography data, the medical field has provided excellent opportunities for the application of data visualization technology with the potential to help improve healthcare that can improve health-related outcomes [33]. In a real-life critical care context , the clinical decision making process should always be supported by relevant and reliable clinical data [2]. In

1 The primary sources of data include: Bedside physiologic monitoring devices (at least 6 physiologic parameters),Ventilator (6 - 15 parameters plus flow loop display) Laboratory data (at least 25 parameters), Radiology (x-ray imaging and text dictation of interpretations), Nursing records (vital signs, nursing progress notes, patient care documentation), anesthetic records, surgical procedure notes, physical therapy, respiratory therapy, social work, dietician, clinical pharmacist, chaplains text entries, subspecialty consult notes, and previous medical history and old clinic data.

Proceedings of the Information Visualization (IV'06)
0-7695-2602-0/06 $20.00 © 2006 IEEE

this case, the use of MDV can be considered a data analysis method in which physicians are able to select the relevant variables to model and decide on representation or encoding of factors to better obtain the analysis results and draw conclusions [2].

In sum, because of large datasets of different forms and sources of data, CCP`s find it difficult and cumbersome to make rapid critical evaluations of a patient's condition. What CCP`s need is an integrated platform through which data can be organized into a uniform multivariate and multidimensional data visualization as the primary means to represent and manage complex patient critical care data. This includes an intuitive human-computer interaction (HCI) that allows access to various time resolutions and subsets of patient data, with control at each level of detail, presented in a manner that allows the examination of past events and trends and enables the physician to explore their relationship to the current clinical status.

The authors propose a critical care data visualization (CPDV) model that is currently being developed and tested at the Indiana University School of Informatics and School of Medicine (IUPUI). The unit will provide a way of integrating quantitative and qualitative data that is encoded by a display with the capability for output in various visualization formats (e.g., PDF) and on a range of devices including computer displays, printers, cell phones, and specialty devices.

2. Background

Most data visualization research has centered on personal medical histories. As Plaisant et al. [23] suggested, medical records can be highly complex with data intervals ranging from minutes to decades. They also noted that clinical data can be divided into a range of scalable values for simplicity of interpretation in which critical events can be viewed relative to a patient's past and current status. Plaisant et al. and other earlier attempts at data tracking systems for personal and biographical histories used an intuitive approach to visualizing histories with graphical time scales [12, 34, 4].

The development of medical computing has evolved over four decades with considerable progress in the areas of interactive computer-based clinical record systems for the practitioner and data visualization for medical and science researchers [37, 36]. Needs among practitioners have been in the area of both longitudinal and contemporaneous data to better support clinical epidemiology, risk management, and critical care assessment. Applications of computing power have provided those involved with advances in medical data analysis. Nevertheless, there has been limited success in longitudinal patient data tracking (comprehensive record systems) because of the multiple sources and complexity of data. Health care professionals have complained about the lack of a well-organized, comprehensive, and easy-to-use patient record system. Even more problematic has been the lack of success in capturing relational or context-based data that would allow physicians to expedite an immediate comparative analysis during a patient episode in a critical care setting.

Moreover, other problems have been identified that encumber physicians ability to acquire, record, and organise patient care data efficiently. Many of the issues have concerned the interaction design of graphical user interfaces of medical data visualization products. Yoder [36] suggests that the earliest computerized medical information systems were designed primarily to improve qualitative services to patients by providing some meaningful data to aid in the prognosis and treatment of disease.

Since the 1970s, a small number of scholars and researchers [33, 28, 27, 7, 8, 6] have been developing techniques of visual representation using primarily two-dimensional (statistic-based) forms, with only a few hundred data points (with mostly black to white values). The first stage of exploratory data analysis [33, 15] dealt with either one or two variate data, in two-dimensional plots with which scientists were able to encode data with multiple parameters or multidimensional data points. Tukey's data analysis technique signified the beginning of a new trend in scientific visualization, in which anyone could visually decode information more efficiently from data in real time. Of course, the availability of high resolution color displays and proprietary software gave way to the study of multidimensional multivariate (MDMV) data visualization. MDMV is a sub-field of scientific visualization that deals with the analysis of data with multiple parameters and key relationships among them [35].

The evolution of data visualization was seen at the 1987 NSF workshop on Visualization in Scientific Computing [20] and the IEEE Visualization 1991 conference [22], which became a watershed event defining the next two major stages in the discipline. Wong and Bergeron (1997) state that during this next stage, "two and three-dimensional spatial data were the most common data types being studied, although multivariate data started gaining more attention" (p. 4).

Our proposed visualization method is an off-shoot of these techniques, whereby correlations are shown between variants, using 1) points and lines to generate a scatter-plot matrix, and 2) a reference grid to better facilitate data point location [10]. However, CPDV is both multivariate and multidimensional. Multivariate refers to the dimensionality of the dependent variables, while multidimensional refers to the dimensionality of the independent variables and the study of relationships among multiple (dependent and independent) parameters (factors) [3, 18].

3. Human-Computer Interaction and Cognitive Theory

Human-computer interaction theory and practice can play a significant part in the creation of MDV tools and visual systems that can support physician decision making [26]. Theories on cognition and perception provide system designers with an understanding of information processing [13, 14]. This is important, because the science and craft of designing interaction behaviors for MDV systems continues to be a daunting task for software developers.

Some of the earliest research in cognitive psychology research showed that visualizations enhanced problem solving capabilities by processing information that is unidimensional or multidimensional [1, 25]. Because the mind is not adept at processing large sums of data, it prefers to simplify complex information into patterns or configurations that are much easier for it to assimilate [24]. For example, in Miller's [21] early work in three dimensions and color, he concluded that "the span of absolute judgment and the span of immediate memory impose severe limitations on the amount of information that we are able to receive, process, and remember" (p. 97). His work contributed to our current understanding of how to organize stimuli into chunks to decrease the "informational bottleneck" (p. 97). This addresses the specific problem of the limits on human information perception when attempting to understanding complex multivariate data; a technique for filtering large data sets to better serve cognitive strategizing and data comprehension. This process would provide for an explicit attempt to preserve dataset relationships and therefore allow for an easier deciphering of data [10].

Kosslyn [19] also concurs that images provide a better means than words for expressing subtle spatial, color, or textural relations. Dreyfus [11] stated that this is because of the brain's ability to store dense amounts of visual information, while escaping the limits of text or numbers. Most importantly, empirical findings have shown that properly designed visualizations have the potential to give a clear picture at a glance. This is because the laws of semantic configuration must be learned, but images provide automatic comparisons and obvious relationships [5]. Using cognitive fit theory, Vessey [32] attempted to illustrate the process of problem-solving as an outcome of the relationship between representing and solving a problem. Thus, there must be a match between the problem representation and the task in order to enhance decision-making [32]. Kellen [17] continues this notion by stating that visualizations are the best fit for spatial or relational tasks where value interpolation or perceiving relationships of data are used. Therefore, when developing visualizations of complex and multivariate data, the designer should give full attention to the tasks performed by the decision-maker to optimize the solution.

Tufte's [29] model, although not multivariate or multidimensional, is one example of an attempt at consolidating, integrating, and visualizing a broad array of data. Unfortunately, little is known of the impact of his models in the real world, because of a lack of empirical data.

4. The Critical Care Patient Data Visualization System

The goal of the proposed computer-based critical care patient data visualization system (CPDV) is to collect online and manually entered data within a specified time frame and then to push the data to a visualizer engine for online graphic representation. Using CPDV, data visualization can assist physicians to assess a specific situation of a patient quickly by recognizing essential changes to multiple physiological readings over a designated time frame. Using selection menus, physicians control the necessary data sources, time periods, and time resolutions to narrow down their diagnosis and final assessment of a patient's condition. Using the CPDV, physicians are able to focus on the most relevant parameters for characterizing a patient's critical state at that moment.

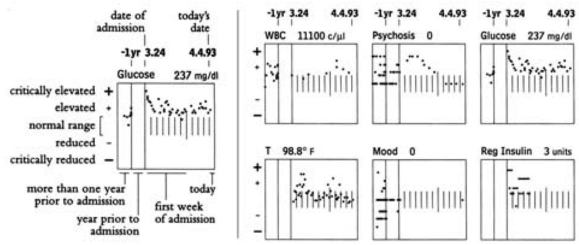

Figure 1. MDV designed by Edward R. Tufte. The larger illustration on the top explains the meaning of the components [29, p. 110-111].

4.1. Rational for Maximizing User Interface Control of Context-Related Data

The CPDV user interface has been designed to maximize the clinicians' ability to control what data is visualized during a specific context-related patient episode or general periodic review. Achieving a degree of control to scrutinize relational data in a multivariate visualization is imperative for physicians (See Figure 2).

For example, a child who has recently undergone cardiac surgery may experience instability of the cardiovascular system. When an anesthesiologist or cardiovascular surgeon is asked to evaluate that patient for cardiovascular instability, they will need to review particular datasets in order to understand the current clinical situation. However, a nephrologist (kidney specialist) who is asked to review the same patient, because of concerns about the function of the kidney, will review a somewhat different clinical dataset. In both cases it is important that the clinician have data presented in a way that clearly displays the relationship between important variables. In current practice, it is uncommon to present "treatment data" (e.g., drug doses, ventilator settings) adjacent to "response data" (e.g., heart rate and other physiologic data).

We propose that clinical data be presented in a "context-related" fashion. In this way, the response to interventions can be readily visualized. Data should be grouped according to its relevance to the particular clinical problem rather than according to its source (See Figure 2). Clearly, the process of presenting or visualizing data should be customizable. An anesthesiologist will have different data visualization priorities than a nephrologist. Similarly, a child who has undergone cardiac surgery will have (in most cases) different data presentation requirements than a 70-year-old man who has an acute exacerbation of chronic respiratory disease. Thus, data presentation can be tailored to the patient, the disease, the consulting specialty, or the acute medical problem under analysis at any point in critical care.

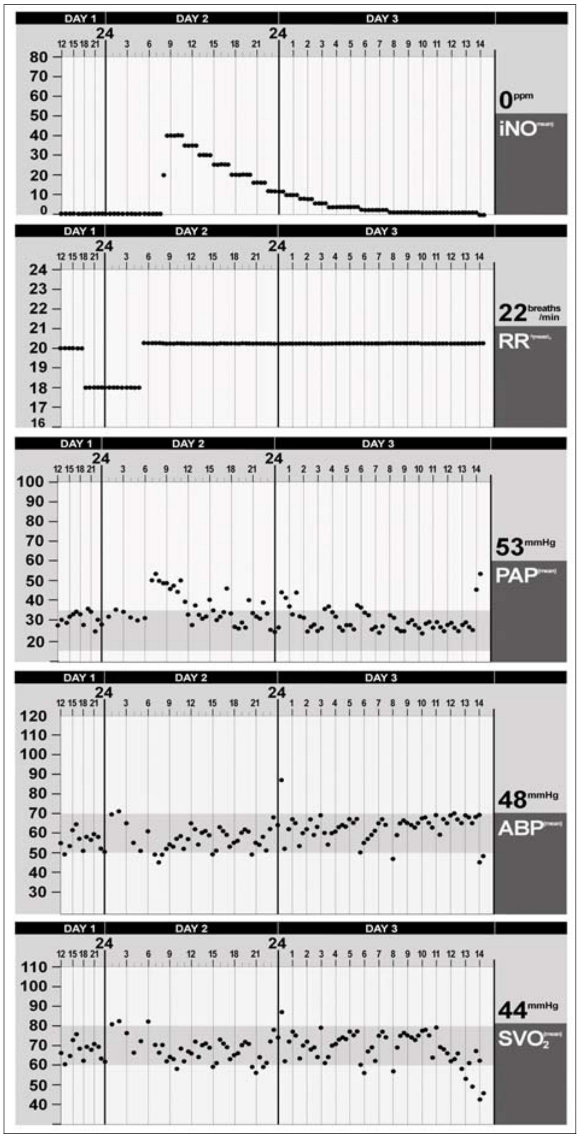

Figure 2. Five relational critical care datasets from a fictitious clinical scenario. Physicians view this data using CPDV when asked to resolve acute problems.

4.2. The CPDV Interface

CPDV uses a data visualization system that allows a multivariate and relational (multi-data-source) display to present the contextual change in the patient's status over pre-selected periods of time resolutions. A multivariate comparative analysis is vital in a relational diagnosis to observe the interrelated effects and interactions of physiological conditions. Each visualization displays the selected critical care data parameter from each bedside monitoring device with hyperlinks to manually-entered data from each healthcare practitioner.

Figure 3 shows the control interfaces. The CCP must first select the *data source* tab to select (check) which of the types of datasets are required for the relational analysis (e.g., Ca, iNO, or CO2). Once these are selected the physician uses the *data resolution* tab to indicate the time period to be visualized. Because the purpose of the CPDV is only to show a patient's recent history, the CCP can only view a maximum of 11 days. Each day is divided into 3 hour increments by default, but the CCP must still indicate which hours of those days are to be examined. Once the day and hours are selected, the CCP must decide whether to view the MDV in increments of 30 minutes, 15 minutes, 1 minute, or 30 seconds (see Figure 3).

Figure 4 shows three examples of the CPDV interface. The full interface consists of the patient's profile on the top with the data visualization on the bottom. The illustration shows the interface with the control panel and visualization underneath. The middle illustration shows the full size of one visualization after activating the control panel. The bottom illustration shows the reduced size of five data visualization so a CCP can view them relationally. Individual visualization can be dragged and dropped at arbitrary positions to rapidly compare patient data.

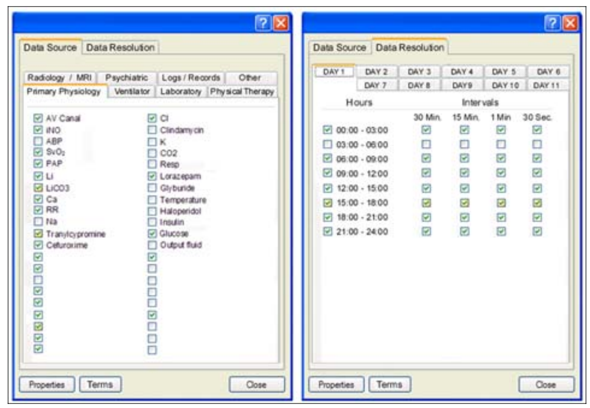

Figure 3. Illustration of MDV control panel with two primary tabs to indicate seven data sources and multiple variables of data resolution for the visualization.

5. Methodology

5.1. Subjects

Twenty senior anesthesia residents from the Indiana University School of Medicine participated in the study. The 20 were split into 2 test groups.

5.2. Treatment

The study included the design of preliminary MDV prototypes as solutions to fictitious clinical data provided by CCP at Indiana University School of Medicine (See Figure 2). The 20 residents were divided into two groups. Each group was given the same clinical scenario (see below) and then given nine questions to answer. However, group one was asked to answer the nine questions using the traditional form of medical data (in the form of numeric-based charts) to resolve the problems outlined in the scenario. Individuals from group two were asked to answer the nine questions using the MDV prototypes as their only source of critical care information. Individuals from both groups were timed to see how long was required to respond to each critical care question. Part two of the study consisted of a post-test questionnaire, which was completed while comparing the traditional form of data visualization with the newly developed MDV prototypes. Further inquiry in part two was necessary to determine which form of visualization enabled the participants to resolve the clinical scenario most rapidly.

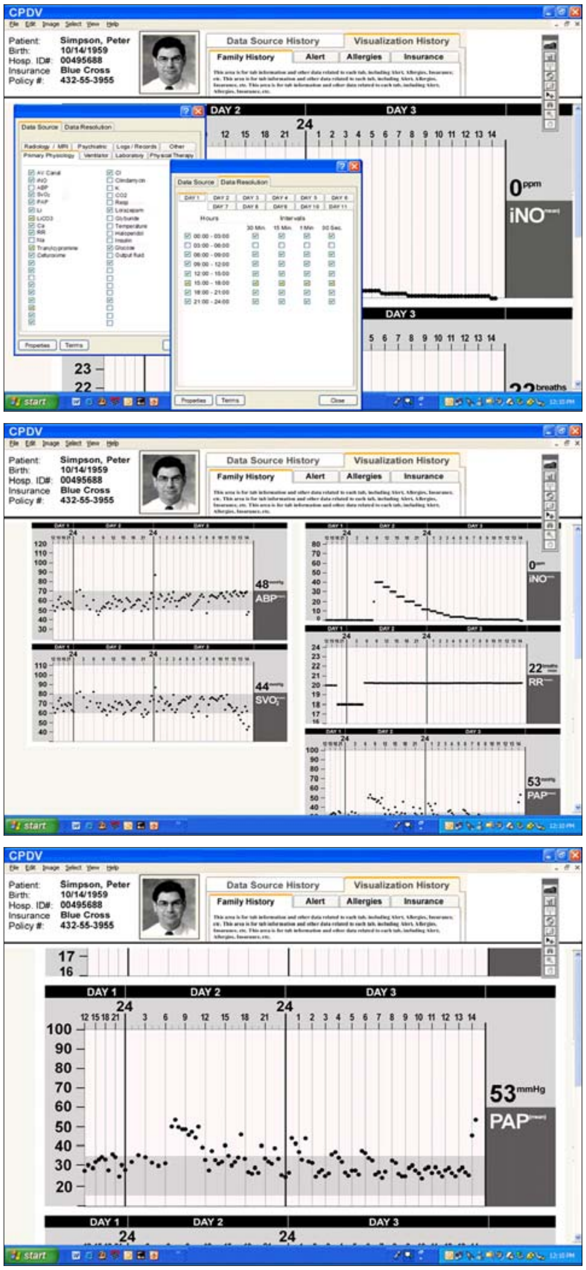

Figure 4. Three examples of the CPDV interface with profile data, control panel and data visualizations.

5.3. Clinical Scenario

A 6-month-old infant undergoes repair of an AV canal. The post-operative course is complicated by pulmonary hypertension, requiring nitric oxide (iNO) to be started on the second postoperative day. On the third postoperative day you are called to the bedside at 1415 because of an acute deterioration in the patient's condition. The bedside nurse states that the patient's mean arterial pressure (ABP) and mixed venous oxygen saturation (SvO_2) have declined and the mean pulmonary artery pressure (PAP) has increased over the past 15 minutes. Review the attached trend graphics and answer the following questions.

1. What is the current ventilator rate (RR)?
2. What was the mixed venous oxygen saturation at noon yesterday?
3. When did the pulmonary artery pressure begin to increase?
4. When was nitric oxide inhalation commenced?
5. What was the starting dose of inhaled nitric oxide?
6. What happened to the mixed venous oxygen saturation after starting nitric oxide?
7. What event precipitated the most recent acute deterioration?
8. How does the current mean PA pressure compare to the mean systemic pressure?
9. Identify a cause for the periodic increases in PA pressure.

Findings from these twenty subjects are being analyzed and used to create a final dynamic CPDV prototype.

Conclusions

We have shown that MDV provides a means to transform data into contextual knowledge, which can provide valuable assistance for data analysis and decision making during critical evaluations of a patient's condition. Data visualization is important to providing an immediate analysis of patient data, i.e., MDV allows a CCP the ability to more quickly assess, appraise, and execute decisive action in the critical care context.

The goal of our approach is to design a completely automated MDV platform with a control interface that allows CCPs access to each level of data and past events that are relevant to a patient's current clinical status. The CPDV model is a solution to organize and display multivariate data through visualization devices. Using CPDV, data visualization can provide physicians assistance to quickly assess a specific patient by recognizing associated changes in multiple physiological, diagnostic, and interventional variables. Using selection menus, CPDV allows physicians to control the necessary data sources, their time periods and temporal resolution. This approach will greatly accelerate the clinical data gathering process that is vital to critical care decision making. The goal of the CPDV project is to design high usability visualizations and interface behaviors that could allow CCPs a rapid and controlled analysis of various sources of data from a relational (context-based) perspective. A recent empirical study will provide findings that allow further development of the system beyond the prototype stage. We are convinced CPDV has the potential to enhance patient care by improving the quality of clinical decision making. This is because CCPs will waste less time foraging through multiple datasets when data is presented in a manner that is customizable to their needs, presented in an organized context-related format, and is in a single location.

References

[1] Baddely, A. (1990). Human memory: Theory and practice. Needham Heights, MA: Allyn and Bacon

[2] Bellazzi, R. & Zupan, B. (2001). Intelligent data analysis: Special issue. *Methods in Informatics Medicine*, *5*, 362-364.

[3] Bergeron, R. D., Cody, W., Hibbard, W., Kao, D. T., Miceli, K., Treinish, L., & Walther, S. (1994). Database issues for data visualization: Developing a data model. *Proceedings of the IEEE Visualization '93 Workshop on Database Issues for Data Visualization, 871*, 3-15.

[4] Boens, J., Borst, F., & Scherrer, J., (1992). Organizing the clinical data in the medical record. *MD Computing,* 9, 149-155.

[5] Bower, B. (1972). Mental imagery and associative learning. In *Cognition in learning and memory*. New York: John Wiley, 27-47.

[6] Chambers, J. M., Cleveland, W. S., Kleiner, B. & Tukey, P.A. (1976). *Graphical methods for data analysis*. New York: Chapman and Hall.

[7] Cleveland, W. S. (1985). *The elements of graphing data*. Monterey, CA: Wadsworth.

[8] Cleveland, W. S. (1993). *Visualizing Data*. Summit, NJ: Hobart Press.

[9] Coiera, E. (1997). *Guide to medical informatics, the Internet and telemedicine*. London: Chapman & Hall Medical.

[10] dos Santos, S. R. (2004). *A framework for the visualization of multidimensional and multivariate data.* Ph.D. Dissertation, University of Leeds, United Kingdom.

[11] Dreyfus, H. (1984). *Symbol sourcebook: An authoritative guide to international graphic symbols*. New York: Van Nostrand Reinhold.

[12] Factor, M., Gelernter, D. H., Kolb, C. E., Miller, P. L., & Sittig, D. F. (1991). Real-time data fusion in the intensive care unit. *IEEE Computer*, *24*, 45-54.

[13] Faiola, A. (2002). A visualization pilot study for hypermedia: Developing cross-cultural user profiles for new media interfaces. *The Journal of Educational Multimedia and Hypermedia*. *11*(1), 51-71.

[14] Faiola, A., Groth, D., & Altom, T. (2005). Integrating the visualization of personal histories to enhance file search in 3D landscapes: Applying file search to 3D view. In G. Salvendy and J. Jacko (Ed.), *Human-Computer Interaction - Ergonomics and User Interfaces, Theory and Practice, Volume 4 - Theories Models and Processes in HCI. Proceedings of the 11th International Conference on Human-Computer Interaction.* Las Vegas: Nevada. (Mahwah, NJ: Lawrence Erlbaum, CD-ROM).

[15] Fisherkeller, M. A., Friedman, J. H. & Tukey, J. W. (1975). An interactive multidimensional data display and analysis system. In *Dynamic Graphics for Statistics*, Voilume?

[16] Horn, W., Popow, C. & Unterasinger, L. (2001). Support for fast comprehension of ICU data: Visualization using metaphor graphics. *Methods of Information in Medicine*, *40*, 421-424.

[17] Kellen, V. (2006). Decision making and information visualization: Research directions. Retrieved on March 23, 2006, from http://www.kellen.net/Visualization_Decision_Making.htm

[18] Keller, P. R. & Keller, M. M. (1993). *Visual cues, practical data visualization*. Los Alamitos, CA: IEEE Computer Society Press.

[19] Kosslyn, S. M. (1983). *Ghosts in the mind's machine: Creating and using images in the brain.* New York: W. W. Norton & Company.

[20] McCormick, B. H., DeFanti, T. A. & Brown, M. D. (1987). Visualization in scientific computing. *Computer Graphics*, *21*(6):1–14.

[21] Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information, *Psychological Review, 63*, 81-97

[22] Nielson, G. M. & Rosenblum, L. (Eds.). (1991). *Proceedings IEEE Visualization '91*, San Diego, CA: IEEE Computer Society Press.

[23] Plaisant, C., Milash, B., Rose, A., Widoff, S., & Shneiderman, B. (1996). LifeLines: Visualizing personal histories. *Proceeding of CHI`96*, ACM, New York, 221-227.

[24] Spence, R. (2001). *Information visualization*. New York: ACM Press.

[25] Tegarden, D. P. (1999). Business information visualization. *Communications of the Association for Information Systems*. 1, Article 4. Retrieved on March 22, 2006, from http://cais.isworld.org/articles/1-4/default.asp?View=html&x=22&y=6

[26] Tory, M. & Möller, T., (2004). Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics*, *10*(1), 72-84.

[27] Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphic Press.

[28] Tufte, E. R. (1990). *Envisioning information*. Cheshire, CT: Graphic Press.

[29] Tufte, E. R. (1997). *Visual explanations: Images and quantities, evidence and narrative.* Cheshire, CT: Graphic Press.

[30] Tukey, J. W. (177). *Exploratory data analysis*. New York: Addison-Wesley Press.

[31] Van Bemmel, J. H. & Musen, M. A., (Eds.). (1997). *Handbook of medical informatics*. Heidelberg, New York: Springer Verlag.

[32] Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences. 22*, 219-240.

[33] Vidal, F., Bello, K. W., Brodlie, N. W., John, D., Gould, R., & Phillips, N. J. A. (2006). Principles and applications of computer graphics in medicine, *Computer Graphics Forum, 25*(1), 113–137.

[34] Whiting-O'Keefe, Q. E., Simbork, D. W., Epstein, W. V., & Warger, A. A. (1985). Computerized summary medical record system can provide more information than the standard medical record. *JAMA*, *254*(9), 1185-1192.

[35] Wong, P. C. & Bergeron, R. D. (1997). Thirty years of multidimensional multivariate visualization. In G. Nielson, H. Hagan, & H. Muller (Eds.), *Scientific visualization: Overviews, methodologies and techniques*. Los Alamitos, CA: IEEE Computer Society Press, 3-33.

[36] Yoder, J. W., Schultz, D. F. & Williams, B. T. (1998). *Journal of Medical Systems, 22*(5), 325 - 337.

[37] Yoder, R. D., Evans, M. R. & Sweeney, J. W. (1967). Processing pictures with computers. An introduction, *Journal of the American Medical Association, 200*, 1171-1175.