Practical Visualization of Multivariate Time Series Data in a Neonatal ICU

Christoph U Lehmann, MD[1]
Johns Hopkins University

Patricia Ordóñez Rozo[2]
University of Maryland, Baltimore County

Jim Facker, MD[3]
Johns Hopkins University

Kathryn Holmes, MD[4]
Johns Hopkins University

ABSTRACT

Visualization of electronic medical data in the Neonatal Intensive Care Unit (NICU) is mainly tabular or in the form of stacked univariate plots of variables over time. In the NICU, norm values differ significantly from adult values, which determine scales and alarm limits in current clinical displays. Thus, the value of information displayed in traditional interfaces is limited by standard visualizations. Providers have difficulties identifying pertinent changes in the patient's condition resulting in delayed diagnosis and harm. We developed a novel interface that allows clinicians to visualize variables critical in the detection of a patent ductus arteriosus (PDA) in a neonate. The interface was designed to allow users to quickly determine changes in variables and the direction of the change. By providing a personalized view that normalizes data points to the patient's state over the total time period reviewed, minor changes in the patient's condition are more easily detected and may allow for earlier diagnosis and treatment of a PDA. By allowing providers to experience the changes in multiple variables simultaneously, we hope to identify patterns that can be recognized by providers as changes in patient status (no PDA vs. PDA). We present the design of a multivariate time series visualization that is interactive and animated, and personalized to an individual patient, such that medical personnel can quickly and efficiently recognize significant changes in the patient's condition.

Keywords: multivariate, time series, visual analytics, computational physiology, human-computer interfaces

INDEX TERMS: J.3 [Life and Medical Sciences]: Medical Information Systems

1 INTRODUCTION

Medical data such as vital signs (i.e. temperatures, blood pressures, heart rates) or laboratory values are typically presented as a series of static numbers listed in tables. These are usually viewed over time in isolation or as a group of parameters on a simple two-dimensional axis. Often the x-axis presents only data points and not their exact location in time. (Figure 1)

Component	Low Range	High Range	Range Units	08/24/2010 02:34:00	08/23/2010 17:48:00	08/23/2010 01:24:00	08/22/2010 18:07:00	08/22/2010 15:08:00	08/22/2010 00:09:00	08/21/2010 17:19:00	08/21/2010 02:48:00	08/20/2010
COMPREHENSIVE PANLEM												
SODIUM	135	148	mEq/L	135 A	135 A	134 A	135 A	136 A	135	139 A	138	141
POTASSIUM,SERUM	4.0	5.9	mEq/L	4.3	5.5 A	6.5 * A	7.4 * A	6.8 * A	TEXT *	7.0 * A	4.8	4.9
CHLORIDE	99	111	mEq/L	99	99 A	97 A	98 A	100 A	98	99 A	98	99
CO2	21	31	mEq/L	23	20 A	24 A	24 A	24 A	21	26 A	26	27
GLUCOSE	60	99	mg/dL	132 *	53 * A	76 * A	91 * A	92 * A	96 *	109 * A	86 *	100 *
UREA NITROGEN	7	22	mg/dL	46	45 A	43 A	41 A	41 A	36	33 A	31	32
CREATININE, SERUM	0.5	1.2	mg/dL	0.7	0.7 A	0.7 A	0.7 A	0.8 A	0.6	0.6 A	0.6	0.7
CALCIUM	8.4	10.5	mg/dL	9.0	8.1 * A	9.4 A	9.8 A	9.8 * A	11.5	12.2 A	11.5	10.8 *
TOTAL PROTEIN	6.0	8.2	g/dL	4.4	4.3 A	4.2 A	4.8 A	3.8 A	4.3	4.0 A	3.5	3.9
ALBUMIN	3.5	5.3	g/dL	2.8	2.8 A	3.1 A	3.7 * A	2.4 A	2.7	2.8 A	2.3	2.5
TOTAL BILIRUBIN	0.1	1.2	mg/dL	3.8	4.1 A	4.4 A	4.6 A	4.7 A	4.2	3.8 A	4.1	4.4
ALKALINE PHOSPHATASE	150	420	U/L	264	260 A	233 A	239 A	240 A	250	222 A	254 *	184
ASPARTATE AMINO TRAN	0	31	U/L	TEXT *	TEXT * A	235 A	214 A	230 A	TEXT *	TEXT * A	91	73
ALANINE AMINO TRANS	0	31	U/L	94	101 A	102 A	91 A	101 * A	55	42 A	41	36
ANION GAP	7	16	mEq/L	13	16 A	13 A	13 A	12 A	16	14 A	14	15
SUN/CREAT RATIO				66	64 A	61 A	59 A	51 A	60	55 A	52	46
AST/ALT RATIO				TEXT *	TEXT * A	2.3 A	2.4 A	2.3 A	TEXT *	TEXT * A	2.2	2.0

Figure 1. Tabular data of laboratory values of an NICU patient. Note occasionally there is more than one entry per day

[1]clehmann@jhmi.edu, [2]patti.ordonez@umbc.edu, [3]fackler@jhmi.edu, [4]kwholmes@jhmi.edu

Although the methods to collect data in the Intensive Care Unit are more sophisticated and the volume of data produced by them is greater now than at any point in the history of medicine, the analytic and diagnostic processes are slow and inhibited by the information overload that providers face [3].

In order for physicians to get a "gestalt" or "snapshot" view of a patient, they are required to quickly and efficiently summarize 20-40 pieces of tabular data, then synthesize this data with radiological images, the physical examination, and patient history and then compare the result to known patterns of disease states. A psychological study in 2005 found that humans can effectively process up to four independent variables in bar graphs accurately and efficiently. When a fifth variable is introduced, accuracy decreases to no better than random [2]. It can be argued that master clinicians obtain their expertise by acquiring the skills of synthesizing and analyzing the medical, radiological, and physical examination to predict most advantageous therapeutic intervention. However, especially for trainees the synthesis of multiple tabular data elements over a period of time is a serious challenge and hinders timely assessment of a patient. More importantly, critical information may be hiding from analysis in plain view because all data is displayed similarly despite its varying value of information (importance) for the case.

Existing visualizations may aid providers in the analysis of physiological data. These visualizations consist of a table or plot of values for a particular parameter over time and primarily focus on analyzing the data along multiple single dimensions (Figure 2). Poorly designed interfaces space values in even intervals despite the actual time differences between the observations.

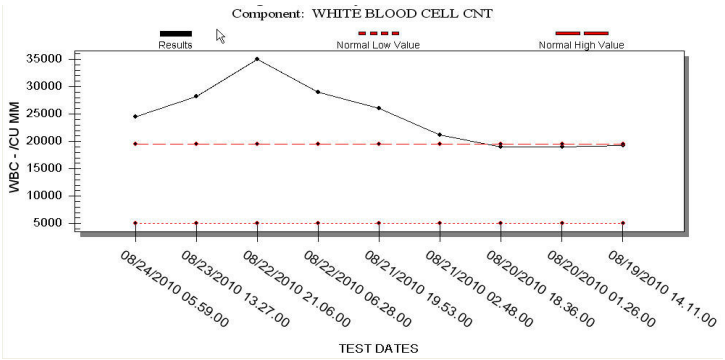

Figure 2. Example from a medical information system. Note that the time interval on the X axis is equally spaced and does not reflect actual time scale relative to the other data points.

In a complex organism, parameters are interdependent. For example in a human, a low blood pressure will result in an increase in heart rate in most circumstances. Thus, it is our hypothesis that examining the vital signs and laboratory results in a multivariate temporal representation shows promise for providing greater insight into how the body and its vital organs interact and function as a whole.

In this paper, we describe a very common medical problem in the neonatal intensive care unit, diagnosis of a patent ductus arteriosus (PDA). We then demonstrate the current methodology available for visualization. Finally, we propose a multivariate

visualization to assist providers in analyzing of the large amounts of data over time. In addition, this visualization captures the relationships between the variables and the rate of change of the variables over time in an animated map of selected variables.

2 METHODOLOGY

A persistent patent ductus arteriosus (PDA) is the result of the reopening or the failure to close the ductus arteriosus after birth. In utero, a PDA is required to shunt blood away from the lungs. After delivery, it is critical that the ductus closes and remains closed to allow adequate perfusion of the lung and to prevent unoxygenated blood to perfuse the body. In premature infants, a patent ductus arteriosus can develop after initial duct closure and may contribute to considerable morbidity and mortality including high output heart failure, intracranial hemorrhage, respiratory failure, necrotizing enterocolitis, and renal insufficiency.

Diagnosis of a PDA is made clinically with evaluation of vital signs, blood analysis, urine output, radiographs, and physical examination with echocardiographic imaging as the gold standard. Early recognition of infants with a significant PDA is critical to avoid delay of care and early institution of the therapy, which may include prostaglandin inhibitors and surgical ligation. Given the ability to track patient hemodynamic and laboratory parameters via computerized data collection, we hypothesize that multivariate visualization may produce a pattern that is predictive of presence or absence of PDA. We also hypothesize that multivariate visualization may aid to identify patients unresponsive to pharmacological closure attempts.

Current methods to visualize medical data from a multivariate perspective focus on visualizing multiple individual parameters over time by either stacking plots of several univariate time series or overlaying them. For our patient population with a PDA, there are several physiological changes known to be associated with a symptomatic PDA. These parameters include elevated heart rate, low blood pressure, widened pulse pressure, decreased urine output, increased serum creatinine, and increased level of respiratory support. In our current system, Eclipsys Critical Care by Eclipsys Corporation, Atlanta, GA, we can display these individual parameters over time in a stacking plot as seen in Figure 3. There are two significant problems with this standard visualization:

1) The y-axis scale is designed for adult values such that small changes in blood pressure that are relevant in the neonatal population are not visualized properly by adjusting by the scale;

2) Plots for vital signs (heart rate and blood pressure) cannot be seen in the same screen as urine output or respiratory support.

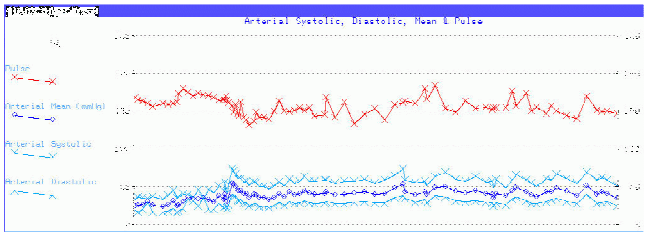

Figure 3. Eclipsys SCC demonstration of a typical neonatal intensive care unit patient. Note the y axis scale is designated for adult values (normal neonate systolic BP 80-45 mmHg).

2.1 Related Work

Pieczkiewicz et al. have shown that physicians prefer visualization of physiologic data over table representation [8]. Our visualization is not the first attempt to present NICU data in a novel graphic interface. Horn et al. used a structured metaphor graphic object to represent physiologic parameters in the NICU. However, their approach made the recognition of subtle changes of a parameter's value - important in PDA - hard to appreciate [4]. Others have designed exceptional visualizations of a patient's medical history from a multivariate perspective. Powsner and Tufte's visualization created a series of scatter plots for each parameter [10]. They also visualized the average and the standard deviation for each parameter. In contrast to our short time period visualization intended for visualization of rapidly changing physiologic data, the Lifelines project [9] displays medical data for several years by aligning all parameters on a single time line. The time lines for the different parameters are stacked on top of one another relative to their occurrence and grouped into expandable categories along the y-axis. The categories include Problems, Allergies, Diagnosis, Complaints, Lab-path, Imaging, Medications, Immunology, and Communication. This visualization is not intended to display rapidly changing physiologic processes. Also designed for data visualization over a longer time interval, the KNAVE-II prototype [11] displays a patient's history over several months to years. It retrieves clinically significant domain-specific temporal abstractions using multiple univariate time series. The visualization contains stacked views of the different states. A special feature of KNAVE-II is the customizable display of vital organs (such as the liver, pancreas, and kidney) states as well as the state of other physiologic parameters.

Our visualization builds on the concept of the Zoom Star that was developed for multivariate time series data [6]. Zoom Star overlays star plots in a 2D space with different colors and levels of transparency, but does not attempt to animate the display to reflect the changes in state over time.

2.2 Prior Work

We have done work in the past on data from Pediatric Intensive Care Unit patients [7]. The primary goal of the visualization was to alert providers of a crisis in a patient with renal or respiratory or cardiac failure without the use of preset thresholds. The aim was to allow providers to efficiently examine changes in a patient's state over time in a multivariate fashion. We used a star plot [1] because of the capability of representing multiple variables over time in a small area. The challenges included the handling of fragmented data, determination of relevant time periods, and generation of a multivariate time-series representation that providers would find medically significant. In this previous work, the visualization was not animated.

3 VISUALIZATION

For our current dataset of neonatal patients, our goal is to alert the caregiver to the occurrence of a significant PDA with the long-term goal to achieve earlier intervention and reduce adverse outcomes. To achieve this goal, we had to determine the following factors: 1) Identification of parameters that should be contained within a "PDA map", 2) Determination of the time window(s) required to identify a PDA pattern, 3) Training of physicians to recognize the PDA pattern.

3.1 Preprocessing the data

The quality of medical data is far from ideal: Data collection for different variables occur at varying intervals and time points, human decision-making influences frequency and type of data collection, and workflow, and communication issues may result in missed or delayed data. A significant effort is required to preprocess medical data.

For this paper, we decided to use validated clinical values within a time window or interval that we thought realistic and

meaningful in a clinical setting based on expert opinion (CUL, KH). The time window represents the period of time considered long enough to allow a significant change in the parameter. Table 1 lists parameters included and the time interval length for each.

A critical obstacle in our case was that some of the parameters had missing data. Missing data means that for a given interval no data points were found. Other intervals had multiple data points. During preprocessing we linearly interpolated the missing data points in order for all variables to have a value within the intervals of 15, 30, or 60 minutes to assure a smooth display of our visualization.

Table 1. Physiological parameters and their categories

Parameter	Interval	Category
Heart Rate(HR)	60 min	Heart
Systolic Blood Pressure (SBP)	60 min	
Diastolic Blood Pressure (DBP)	60 min	
Delta Blood Pressure (SBP-DBP)	60 min	
Urine output (UO)	240 min	Kidney
Creatinine (Crea)	720 min	
Respiratory Rate (RR)	60 min	Lung

3.2 Personalization

Prior to rendering the visualization, a uniform scale for displaying the data had to be calculated on every axis for all parameters. We modified the concept of Piecewise Aggregate Approximation (PAA) for Symbolic Aggregate Approximation (SAX) [5]. PAA takes a univariate time series, divides it into equal sized intervals, and then approximates the value for each interval by averaging the values in the interval.

Using a previously described technique, we created a Multivariate Time Series Amalgam (MTSA) [7] by interleaving the values of all the parameters for a given interval in a consistent order. The parameters are grouped in categories according to their effect on one of four vital organs; the heart, liver, lung, and kidney, and within the category displayed alphabetically. (Table 1)

An MTSA represents a patient's average state during an interval in a star plot. The MTSAs for all time intervals in the visualization are plotted as a series of overlaid star plots to reflect the patient's changing state over time. The star plots have varying colors to represent the temporal changes in the MTSAs with the lightest color representing the earliest interval, and the darkest color the most recent interval.

A paper containing a more detailed description of the animated version of the visualization including two other views apart from the "Personalized View" has been accepted in the proceedings of the ACM's First International Healthcare Informatics Symposium (IHI 2010).

3.3 The Personalized View

No patient is like the next patient. Human beings vary significantly in their baseline values for physiological or laboratory parameters. Factors that influence baseline values include age, gender, weight, metabolism, disease status, medical interventions etc. Thus to allow comparison of data to the patient's overall clinical state and interpretation of changes within that state, we had to assure that for the time period examined a "Personalized View" was achieved for each patient.

For the Personalized View, the values in the univariate time series for each parameter were normalized prior to constructing the MTSA so that the average was 0 and the standard deviation was 1 as seen in Figure 4. This Personalized View allows data to be compared to its whole set (the patient's general state) during the normalization process, addressing one of the visualization's primary objectives. (Figure 4) The dotted red circle represents the patient's average value for all variables over the entire period of time. Each point represents the deviation of a value from the average for its parameter.

3.4 Variable Relationships

Another objective of the visualization was to capture the relationships between the variables as well as the rate of change over time. By plotting the standard deviation of the values from the patient's average, the relationships between the variables and the rate of change of the variables over time are exaggerated. For example, creatinine mean for the whole period may be 0.7. A rise to 1.1 would be difficult to notice in a standard display; however since it represents a 2 standard deviation difference from the norm, now becomes easily perceived by a clinician. To allow a provider to correlate the relative deviations to the actual values, the minimum and maximum values for each parameter are displayed in square brackets next to the label for the axis.

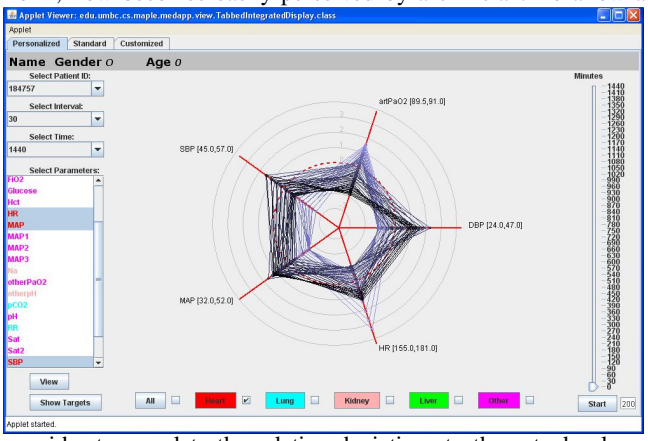

Figure 4. Visualization: Darker lines represent newer data. Dotted red circle represents mean value - grey circles represent standard deviations.

3.5 Animation

One of the primary objectives of the visualization is to display the change in a patient's state through animation. Patient's conditions are never static. Thus, any display of patient physiologic or laboratory values should reflect the dynamic state and allow a user to appreciate the changes over time.

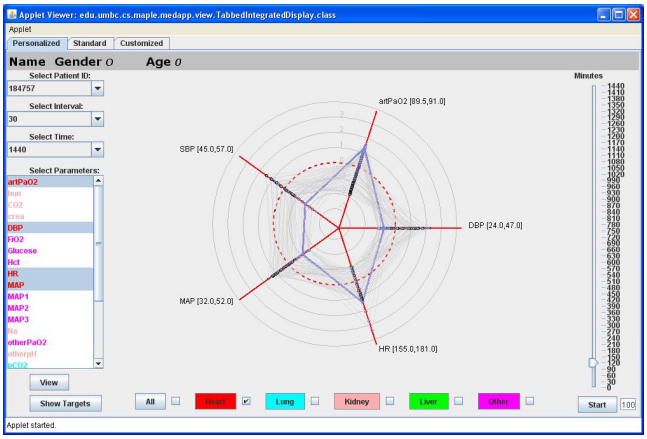

Figure 5. Animation halted to review data at a 120 min

Our visualization allows a user to "drive" through the period of time captured, stop at any point (Figure 5) to further investigate a specific moment in time and then proceed with the animation. This type of animation utilizes the human eye's ability for motion detection. In addition to perceiving the location of a data point, the user can perceive the change in directions and the speed associated with this change. It is our belief that with training, clinicians will be able to recognize patterns in the animation that reflect the opening or closing of a ductus, and thus allow for early diagnosis and treatment. A video of the visualization can be found at http://cs.umbc.edu/ordonez/GHC2010/MTSA.avi.

4 Conclusions and Future Work

We developed a novel interface that allows clinicians to visualize variables critical in the detection of PDA in neonates. The map interface allows users to quickly determine not only changes in variables and the direction of the change, but also allows for the interpretation of the rate of change of variables relative to one another over time.

For practicing clinicians, a single "snap shot" of a patient's data can be dangerous and lead to misdiagnoses. A patient that changes from one extreme state to another undergoes a period of seeming normality, where all vital signs and laboratory values may be briefly within normal range. Reducing the data provided to a clinician to this brief "normality" period - as standard display can do - may be deceptive and lead to error. Potential limitation of our novel approach includes the interpolation of data, which may mislead a physician to the patient's real status. However, the data points reflect clinical reality (a provider measured data in a certain interval). Thus, any misinterpretation is actually a function of the provider to collect data in an inappropriate interval and not a function of our methodology. Another limitation that must be evaluated is a potential increase in time/effort for clinicians to review our visualization compared to standard displays. Our visualization allows physicians you perceive not only the current state but also previous states and the direction (positive or negative) of change.

However, like any new visualization, training and a period that allows for the development of pattern recognition will be critical. Currently, the implementation of visualization is not in real time. We are using static data to analyze users' responses to our visualization. The next step will be to display the data in a traditional format and our new visualization and evaluate the time and effort it takes for clinicians to arrive at an accurate diagnosis.

Ultimately, beside monitors, Electronic Health Records will feed our visualization in real time and allow for bedside analysis. Adding additional clinical decision support to the visualization could alert clinicians to rapidly deteriorating patients and reduce response times, critical events, and errors.

References

[1] Chambers, J., Cleveland, W., Kliener, B., & Tukey, P. (1983). Graphical methods for data analysis. Boston: Wadsworth International, Duxbury Press.

[2] Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? Psychological Science, pp. 70-76.

[3] Heldt, T., Long, B., Verghese, G. C., Svolovits, P., & Mark, R. G. (2006). Integrating data, models, and reasoning in critical care. Engineering in Medicine and Biology Society, (pp. 350-353).

[4] Horn, W., Popow, C., Unterasinger, L. (2001). Support for fast comprehension of ICU Data: Visualization using Metaphor Graphics. Methods Med Inf 40, pp. 421-424.

[5] Lin, J., Keogh, E., & Lonardi, S. (2005). Visualization and discovering non-trivial patterns in large time series databases. Information Visualization, pp. 61-82.

[6] Noirhomme-Fraiture, M. (2002). Visualization of large data sets: the zoom star solution. Electronic Journal of Symbolic Data Analysis.

[7] Ordóñez, P., desJardins, M., Feltes, C., Lehmann, C. U., & Fackler, J. (2008). Visualizing multivariate time series data to detect specific medical conditions. AMIA Annual Symposium, pp. 530-534.

[8] Pieczkiewicz, D., Finkelstein, S. M., Hertz, M. I. (2007) Design and evaluation of a web-based interactive visualization system for lung transplant home monitoring data. AMIA Annu Symp Proc, pp. 598-602.

[9] Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., & Shneiderman, B. (1998). Lifelines: Using visualization to enhance navigation and analysis of patient records. AMIA Annual Symposium, pp. 76-80.

[10] Powsner, S., & Tufte, E. R. (1994). Graphical summary of patient status. Lancet, pp. 386-389.

[11] Shahar, Y., Goren-Bar, D., Boaz, D., & Tahan, G. (2006). KNAVE-II: A distributed architecture for interactive visualization and intelligent exploration of time-oriented clinical data. Artificial intelligence in medicine, pp. 115-135.