A Dynamic Visual Analytics Framework for Complex Temporal Environments

by

Rishikesan Kamaleswaran

A thesis Submitted in Partial Fulfillment
Of the Requirement for the Degree of

Doctor of Philosophy in Computer Science

in

The Faculty of Business and Information Technology

Computer Science

University of Ontario Institute of Technology

August, 2016

Abstract

Introduction: Data streams are produced by sensors that sample an external system at a periodic interval. As the cost of developing sensors continues to fall, an increasing number of data stream acquisition systems have been deployed to take advantage of the volume and velocity of data streams. An overabundance of information in complex environments have been attributed to information overload, a state of exposure to overwhelming and excessive information. The use of visual analytics provides leverage over potential information overload challenges. Apart from automated online analysis, interactive visual tools provide significant leverage for human-driven trend analysis and pattern recognition. To facilitate analysis and knowledge discovery in the space of multidimensional big data, research is warranted for an online visual analytic framework that supports human-driven exploration and consumption of complex data streams.

Method: A novel framework was developed called the temporal Tri-event parameter based Dynamic Visual Analytics (TDVA). The TDVA framework was instantiated in two case studies, namely, a case study involving a hypothesis generation scenario, and a second case study involving a cohort-based hypothesis testing scenario. Two evaluations were conducted for each case study involving expert participants. This framework is demonstrated in a neonatal intensive care unit case study. The hypothesis generation phase of the pipeline is conducted through a multidimensional and in-depth one subject study using PhysioEx, a novel visual analytic tool for physiologic data stream analysis. The cohort-based hypothesis testing component of the analytic pipeline is validated through CoRAD, a visual analytic tool for performing case-controlled studies.

Results: The results of both evaluations show improved task performance, and subjective satisfaction with the use of PhysioEx and CoRAD. Results from the evaluation of PhysioEx reveals insight about current limitations for supporting single subject studies in complex environments, and areas for future research in that space. Results from CoRAD also support the need for additional research to explore complex multi-dimensional patterns across multiple observations. From an information systems approach, the efficacy and feasibility of the TDVA framework is demonstrated by the instantiation and evaluation of PhysioEx and CoRAD.

Conclusion: This research, introduces the TDVA framework and provides results to validate the deployment of online dynamic visual analytics in complex environments. The TDVA framework was instantiated in two case studies derived from an environment where dynamic and complex data streams were available. The first instantiation enabled the end-user to rapidly extract information from complex data streams to conduct in-depth analysis. The second allowed the end-user to test emerging patterns across multiple observations. To both ends, this thesis provides knowledge that can be used to improve the visual analytic pipeline in dynamic and complex environments.

Acknowledgements

Since I began this journey, the learning has never stopped. I want to thank first God, who has served as my spiritual foundation throughout my life. I want to thank my family, who have been extremely supportive. I want to thank my dear girlfriend Pooja, who has stuck with me throughout the emotional roller coaster. I want to thank my best friend, Sri Ramanan, who never forgets to call and find out about my well-being. Thank you to all of you who have been a part of my life thus far. I pray that I continue to receive the warmth throughout the next part of my journey.

I want to thank my PhD supervisor, Dr. Carolyn McGregor who has personally seen me through my growth as an independent researcher from my undergrad, through my masters and PhD. I would not have reached this far without your encouragements and support.

I want to thank Dr. Chris Collins for taking a deep interest in my research growth and providing incredible guidance and encouragement. I was afraid of code, but now my hands crave for it! Thank you for giving that gentle push to achieve new success.

I want to thank Dr. Jennifer Percival for pushing me to become more meticulous in my scientific methodology. I want to thank Dr. Andrew James, who has taught me an incredible amount about medicine, and exposed me to the great world of neonatology. I would finally like to thank all my colleagues and lab mates, your support has been deeply impactful.

Contents

List of Figures

Publications Related To This Dissertation

R. Kamaleswaran, R. Wehbe, J. E. Pugh, L. Nacke, C. Mcgregor, and A. James, “Collaborative Multi-Touch Clinical Handover System for the Neonatal Intensive Care Unit,” electronic Journal of Health Informatics, vol. (In Print), 2015.

R. Kamaleswaran and C. McGregor, “A Review of Visual Representations of Physiologic Data,” JMIR Medical Informatics (in review).

R. Kamaleswaran, C. Collins, A. G. James, and C. Mcgregor, “PhysioEx: Visual Analysis of Physiological Event Streams,” Eurographics Conference on Visualization (EuroVis) 2016, vol. 35, no. 3, 2016.

R. Kamaleswaran, C. Collins, A. James, and C. McGregor, “CoRAD: Visual Analytics for Cohort Analysis,” in IEEE International Conference on Healthcare Informatics 2016 (ICHI 2016) (accepted)

R. Kamaleswaran, J. E. Pugh, A. Thommandram, A. James, and C. Mcgregor, “Visualizing Neonatal Spells: Temporal Visualization of High Frequency Cardiorespiratory Physiological Event Streams,” in Proc. of IEEE VIS 2014 Workshop on Visualization of Electronic Health Records, 2014.

R. Kamaleswaran, C McGregor. A Real-Time Multi-dimensional Visualization Framework for Critical and Complex Environments. Computer-Based Medical Systems (CBMS), 2014 IEEE 27th International Symposium on. 2014. p. 325–8.

1. Introduction

Unprecedented growth in networked sensors in recent years has generated large repositories of heterogeneous data streams in applications across many domains, from intelligence analysis to market research. This is also true in the neonatal intensive care unit (NICU), where clinical researchers need to identify patterns that extend across multiple high frequency physiologic data streams and clinical observations [1]. Identifying events of interest in such an environment is a challenging and common problem [2]. Patterns identified in the data stream, along with clinical information generated by humans are converted to knowledge. The insight generated from this knowledge can be utilised to support the care of critically ill infants. Insight is perceived knowledge that may be novel and surprising, thus allowing the user to enhance their cognitive understanding [3].

In the NICU, physiologic signals change over time as infants grow and mature, and new normal ranges are established week-by-week [4]. A single physiologic data stream can contain a variety of complex events that collectively represent real-world conditions in critically ill patients. One of the aims of a clinical researcher is to identify precursors in the data stream that may be applied in models predicting to the onset of adverse clinical conditions. In order to enhance this task, the researcher requires the ability to visually explore multi-faceted patterns in a single patient [5], and by using that insight, perform an observational research method, such as case-controlled analysis against patient cohorts [6]. Supporting that requirement necessitates novel visual solutions that can reduce the cognitive burden of clinical researchers, while also involving the human in the analysis pipeline [7].

This thesis presents extensions to the traditional data warehouse architecture by introducing the temporal Tri-event parameter Dynamic Visual Analytic (TDVA) framework. The TDVA framework, as output, generates instantiated dynamic visual analytic marts. These marts serve as the visual interface allowing researchers to interactively explore heterogeneous data streams. TDVA combines consumption and exploratory (research) requirements for dynamic and complex environments using heterogeneous data streams. This thesis contributes to the later requirement through two major instantiations, namely the *Physiologic Explorer* (PhysioEx) and the *Cohort Relative Alignment Display* (CoRAD). While presented as a proof-of-concept (Heart Rate Variability graph) in this thesis, the consumption requirement is left for future work.

1.1 Researcher and Consumer

There are two prevailing groups of users that perform frequent analysis of heterogeneous data streams. The first group seeks to actively explore or passively discover new knowledge that may be useful in achieving a business objective. Active exploration involves the user utilising one of several direct manipulation tools, while a user performing passive discovery may suddenly identify novel insight using means such as serendipity [8], [9]. The second group, in contrast, is interested in the consumption of that data, in order to gain insight given limited cognitive resources [10], [11]. These two groups of analysts will hereafter be addressed as researchers and consumers respectively. The researcher seeks to discover knowledge using data and information, while the consumer uses knowledge to accomplish a task.

According to Davenport and Prusak, data, within an organizational context is defined as: ‘a set of discrete, objective facts about events’ [12]. Information as defined by Oppenheim and

Stenson, may largely be based on data, and in some cases expressed as an identifiable and communicated entity, which encapsulates both the intention of the sender and the expectation of the receiver [13]. Knowledge is defined as: 'large structures of related information' [14]. Oppenheim and Stenson, propose a pathway for the transformation of information to knowledge through comparison, consequence, connections, and conversation [13]. Comparison involves identifying how a particular set of information compare to other known sets, consequence acknowledges the implication of this new set with respect to business decisions and actions. Connections identify relationships to existing knowledge, and conversation places an emphasis on what other people think about this particular set of information. Through careful completion of each of the four transformation processes, a user arrives at actionable knowledge, that is, knowledge which can be applied in the context of real-world situations. Visual systems act as a medium, which aids the human in producing actionable knowledge by means of interactive and efficient displays [15].

The consumer, with knowledge requirements, may be a critical care physician in an intensive care unit, or an operator in a nuclear power plant. In the case of the former, the critical care physician achieves situational awareness by exploiting visual perception abilities to consume knowledge from numerous information systems in a timely manner to support a critically ill patient. To support that form of instantaneous knowledge delivery, systems must be capable of consuming, analysing, and visually communicating insight. A component of the TDVA framework may be extended to support this workflow.

Meanwhile, the researcher, such as in the case of a network security analyst, uses highly interactive tools to explore large volumes of data to identify security breaches [16]. Moreover, in complex real-world settings, the researcher has an additional requirement to perform in-depth and multi-dimensional exploration of complex information belonging to an individual system. These activities are commonly referred to as the N-of-1 analysis, and commonly used in psychiatry [17], [18], education [19], and individualized medicine [20], [21]. N-of-1 analysis has been used to support generation of hypothesis in complex environments [5], [22], [23]. Once a hypothesis has been generated in a single subject, an attempt is made to compare these patterns across similar population cohorts. This study design is referred to as the case-controlled analysis, this design is commonly used in epidemiology [24]. In this thesis I focus on the researcher to support both N-of-1 and case-controlled methodologies, while allowing for extendibility to support the consumer.

Researcher: The researcher requires highly interactive functionalities to support N-of-1 tasks, such as multi-faceted and in-depth analysis of heterogeneous data streams for the purpose of hypothesis generation. To that end, the researcher requires appropriate interactions techniques for selecting, filtering and retrieving relevant details while conducting hypothesis generation. Subsequently, to test that hypothesis the researcher requires tools that support case-controlled analysis tasks across multiple observations.

Consumer: The consumer requires rapid access to knowledge. This involves the communication of key temporal parameters of dynamic events. Moreover, the level of interactivity must be reduced as to remove complexities in the interface. The emphasis is placed on communicating salient deviations as determined by online algorithms processing the data stream.

1.2 Hierarchy of Events

One of the key distinctions that will be carried throughout this thesis, is the notion of heterogeneous data streams. In many domains, such as in network surveillance, for instance, a single data stream of intrusions at a firewall can be replicated across multiple nodes, thereby producing a collection of heterogeneous intrusion streams. In contrast, in sensor networks, such as in smart homes that use smart water, electric meters, weather sensors, represent unique views of a single system. Each data stream generated from those sensors are then analysed in part or collectively to identify anomalies. In this thesis, the latter of the two heterogeneous data streams assumes the subject of interest. Similar to home sensors, patient sensors sample data that generate information about independent physiological systems. Collectively, these streams form a heterogeneous physiologic collections that can be analysed for identifying and preventing pathological clinical conditions.

The availability of high density data streams continues to motivate research in the areas of acquisition, event processing and visualization. A data stream is defined as a real-time, continuous, ordered sequences of events [25]. An event is defined as an occurrence of interest in time, and can be reduced to either primitive or complex, which will be defined below [26]. While that distinction generalize occurrences within a single stream, additional hierarchies are required when developing visual representations of salient occurrences between and outside of data streams.

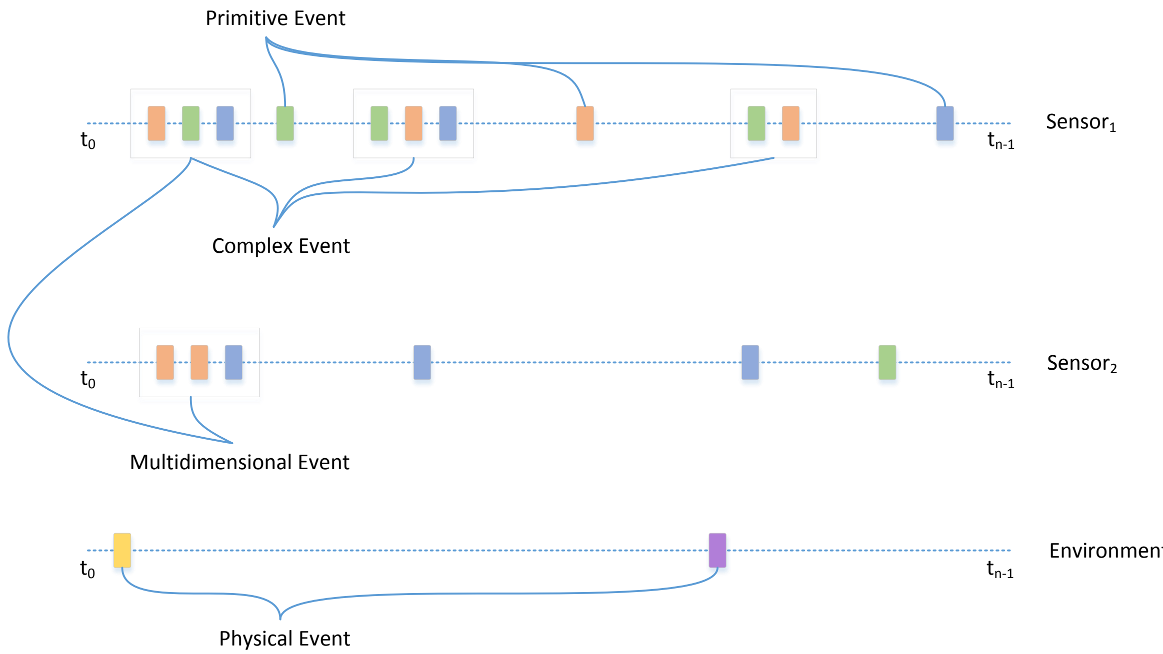


Figure 1: Hierarchy of Events: The primitive event is a single observance of interest in a data stream. A complex event contains primitive events over a period of time. A multidimensional event involves salient co-occurrences across data streams. A physical event manifest in the environment can generate the lower three events.

This thesis further introduces a hierarchy of events to expose the depth of analytic complexity, and also to motivate tools that support these distinctions. The hierarchy is illustrated in Figure 1. Events fall into one of four classes, including primitive, complex, multidimensional, and physical. A primitive event is an atomic occurrence of interest, observed in a data stream at a point in time. This may include abrupt changes in the signal baseline, such as asynchronous acceleration or deceleration to an abnormal state. A primitive event includes a start and stop time, and the duration of that event may itself contain salient information.

A complex event includes a series or a unique sequence of primitive events in a single data stream over a period of time. The genesis of an abrupt acceleration followed by a prolonged deceleration is an example of a complex event that includes two primitive events. A multidimensional event is the observance of two or more complex or primitive events across multiple data streams over a period of time. A multidimensional event may come from significant external disturbances on the system. Finally, a physical event is the manifestation of a real-world condition that acts as a cause for the disturbance. An example of a physical event may be a disturbance induced by an earth quake impacting a network of environmental sensors. A clear delineation of events allows for extensibility of the TDVA framework into other domains that share similar analysis and consumption requirements with data streams.

1.3 Temporal Tri-Event Parameters

This thesis introduces the concept of critical temporal variables, namely: trajectory, frequency, and duration. Researchers assessing events produced in dynamic and complex data streams rely on those temporal critical variables to identify meaningful insight [27]. These critical temporal variables have been extensively explored in the temporal data mining literature and often implicitly identified in the intensive care environment [28]. The I-Pass system, for instance, consists of a minimal list of medical event information that can support the transfer of patient care (handover) between clinical shifts [29]. Much of these handover processes contain common variables, such as those that capture the number of observed events, the time and duration of the condition, and the presence of a positive or negative trend [30].

In this thesis, trajectory, frequency, and duration are collectively identified as the tri-event parameters. The details of these metrics are discussed below.

To identify the significance of any event, the most commonly utilized variable is the event's trajectory pathway [31]. Trajectory of events can alert the domain expert of abnormal deviations from a normal baseline. The second variable is the frequency of an event within an epoch, such as the start and end of a monitoring period. Identifying the frequency of abnormal events can help communicate appropriate knowledge about the stability and health of the system. For instance, the frequency of pauses in breath over an hour can inform the clinician of potential pathology such as apnoea, or the need to modify some ongoing intervention [32]. The third variable is the total duration of any event. Duration of an event can be enumerated in seconds to several hours or days. Together these metrics are presented in this thesis, to convey important information of multidimensional events that can then be used within a dynamic visual analytic framework to support hypothesis generation and testing.

1.4 Clinical Challenge – Case Study Context

Critically ill premature infants admitted to the neonatal intensive care unit face a series of medical challenges that require ongoing monitoring and intervention. Most of the high frequency physiological and other medical device data are not used beyond the operational instantaneous vital organ monitoring [33]. More recently, researchers have begun to analyse and explore this dataset in order to generate new knowledge that can improve care at the bed-side [34]–[36]. There is great potential for real-time analytics techniques coupled with visual analytics techniques to service this need.

McGregor et al. developed the Artemis platform [37]; this platform provides advanced real-time health analytics for multiple patients, watching for the onset of multiple conditions using multiple physiological and other medical device data streams. The Artemis platform, is currently being used to progress research studies relating to the diagnosis of nosocomial infection [38], neonatal apnoea [39], and retinopathy of prematurity [1]. Nosocomial infection, also known as sepsis, is a common hospital-borne infection for babies receiving care in the unit [40]. Neonatal apnoea, on the other hand is a state in which the infant ceases to breathe for longer than 20 seconds, and the cause can be identified as belonging to one of several muscular or neurological conditions [39]. Traditionally these conditions have been identified by text in clinical notes, or aggregated by frequency with minimal information of the individual events. Through the instantiation of the TDVA framework, a series of visual mediums can be generated and used to provide support context-sensitive consumption and knowledge discovery activities.

In this thesis, two unique scenarios, one each for hypothesis generation and hypothesis testing, are used to instantiate the TDVA framework. PhysioEx, an N-of-1 hypothesis generation visual analytic tool is detailed in chapter seven, and CoRAD, a case-controlled hypothesis testing visual analytic tool is presented in chapter eight. An early support for contextual deployment is also contributed using the Exploration-Consumption continuum in §6.3.

1.5 Contributions

The contributions presented in this thesis cross multiple domains. There are four areas where knowledge is contributed. Three of these are disciplines within computer science and informatics, while the fourth is within a subspecialty of medicine called neonatology. The three disciplines within computer science are, health informatics, information systems, and visual analytics.

To the discipline of health informatics, this thesis makes five contributions, namely a systematic review of physiological visualizations, and development of four visual analytic tools: Heart Rate Variability Graph, Sequence of Events (SeqEvent), Physiologic Explorer (PhysioEx), and Cohort Relative Aligned Dashboard (CoRAD). These tools address health informatics challenges. For instance, the Heart Rate Variability Graph allow consumers to rapidly associate heart rate variability scores to pathologic states and verified using two domain experts. SeqEvent, allows researchers to identify novel sequence pathways in output generated by event stream processing algorithms and was verified using a single clinical researcher. PhysioEx allows researchers to identify salient features within a single patient to perform hypothesis generation tasks and was validated in a preliminary expert evaluation study. Finally, CoRAD allows researchers to test hypotheses across population cohorts to validate physiologic markers and was validated in an expert evaluation study.

Within information systems, four contributions are made. The contributions include: a theoretical Exploration-Consumption Continuum, a TDVA framework proposing novel methods of generating visual marts within an event stream processing environment, a TDVA methodology outlining an instantiation process, and a TDVA platform design for deploying the TDVA framework. These contributions are validated with applications deployed in the context of neonatal intensive care.

Three novel contributions are made to the discipline of visual analytics. All of the contributions to the domain are visual methods utilized in the PhysioEx and CoRAD. These visual techniques are: the temporal intensity map, SequenceGraph a bubble-like display, and the cohort-based relatively aligned dashboard which is a combination of a heatmap representation and a contextual bar that supplements data using an interactive technique. These contributions can be applied in complex domains where analysis of the temporal tri-event parameters may provide meaningful insight for the end-user.

To the discipline of neonatology, knowledge is contributed from evaluations that were conducted using PhysioEx and CoRAD. This knowledge advances the understanding of physiologic behaviours, and the relationships between neonatal spells & neonatal sepsis, and heart rate variability & increased bradycardias. While initial evidence is presented with the evaluations presented in this thesis, additional work is required to validate the findings with larger patient cohorts and domain experts.

1.6 Thesis Structure

In chapter 1, I provide an overview of three key concepts that underlie the TDVA framework, namely, distinctions between an explorer and a consumer, hierarchy of events, salient temporal properties in complex and critical settings, clinical challenges and the contributions of this thesis. Chapter 2 details the clinical environment and the nature of challenges introduced in the previous chapter. The clinical challenge serves as a motivation for this research, however the challenges outlined in the clinical context contain numerous properties that allow the tools to be generalizable to other domains. This generalizability will be made explicit in chapter 3 where I present related works. Chapter 3, contains results from a systematic review of physiologic visual interfaces, and the contents are interpolated from a publication that is currently under review [41].

Chapter 5 presents results from a qualitative explorative study that motivates the proposed framework. That chapter also contains materials interpolated from two publications [42], [43]. Chapter 6 introduces the Tri-event Parameter Dynamic Visual Analytic (TDVA) framework, TDVA methodology and TDVA platform design, which serves as the foundational contribution of this thesis. The prototypes of TDVA marts: Heart Rate Variability Graph and SeqEvent, are presented in §6.5.1, and §6.5.2 respectively. Chapters 7 and 8 presents two validated instantiations, PhysioEx [44] and CoRAD [45] respectively. PhysioEx and CoRAD are instantiations of a TDVA mart, and were created by following the TDVA framework and methodology. These TDVA marts are within the TDVA platform which instantiates the TDVA framework. PhysioEx has been published, while CoRAD under print. Chapter 9, summarizes this thesis and provide some future research directions.

2. Neonatal Intensive Care

About 10% of the world's babies are born premature [46]. In the developed world, premature babies are usually admitted to the neonatal intensive care unit (NICU). Babies within the NICU, have continuous monitoring of their heart rate, blood oxygen saturations, breathing and blood pressure levels. A critically ill infant admitted to the NICU may produce several gigabytes of complex physiological, neurological and pharmaceutical data every day.

The electrocardiograph (ECG) sensor samples data at 1000 hertz (Hz) to produce waveform signals of the infant's real-time heart rhythm, along with 1 Hz readings of derived heart rate (HR). Breathing is detected using respiratory impedance wave plethysmography (IRW) sampled at 65.8 Hz. Blood oxygen saturation (SpO_2) is gathered as a 1 Hz value along with a plethysmography wave at 65.8 Hz. This has the potential to produce upwards of 390 million data points per patient per week [1]. These multidimensional signals are actively consumed in the critical care environment and serve as an indication of patients' vital statues. Nurses record these vital signals in hourly or half-hourly intervals to derive longitudinal temporal trends. Concretely, the amount of data collected for each patient, in a typical 40 – 60 patient unit, can be overwhelming. To that end, graphical displays in medicine have long been used to support bed-side decision making and clinical research [47].

Patient monitors, including the ECG, SpO_2, and similar modules at the bed-side have long displayed a series of critical physiological variables in real time, using a combination of waveform and numerical formats. These patient monitors were initially designed to support monitoring needs of the consumer. Hence, the design was intentionally minimal, with a focus on instantaneous values. One of the earliest designs of the *modern* bed-side physiologic monitor is the Siemens Sirecust 404-1 [48] (Figure 2a) introduced in the early 1980s. Since then, very little physical changes have been made. The Phillips MP70 displayed in Figure 2b shows only minor differences, specifically the introduction of colour and varying size of text. The modern patient monitor however, can list up to 36 critical physiological variables in real time following the same SSSI paradigm [49]. These systems have not converted low level data to information and knowledge as required at the point of care.

(a) (b)

Figure 2: Medical Monitors retain the same interface design. (a) A Drager Serie SIRECUST monitor displaying demo data [48]. (b) A Phillips MP70 monitor displaying demo data [306].

While bed-side devices were introduced as a means of managing the overwhelming throughput, they do not sufficiently serve the need of a researcher. For instance, the data presentation and analytic faculties provided by the manual charts they replace were found to be superior when compared to the patient monitors [50]. Hence, the practise became norm to transcribe physiologic values from the monitor onto manual charts, or electronic patient charts.

Shabot et al. [51] recognizing the need for better graphics and clinician driven visual representation of critical data, introduced a platform for non-programmers to enhance data presentation in the generated reports. The system displayed in Figure 3 shows an early form of graphical displays containing patients' cardiovascular history. These graphical systems were more useful for researchers to identify abnormal conditions and trajectories over longer durations.

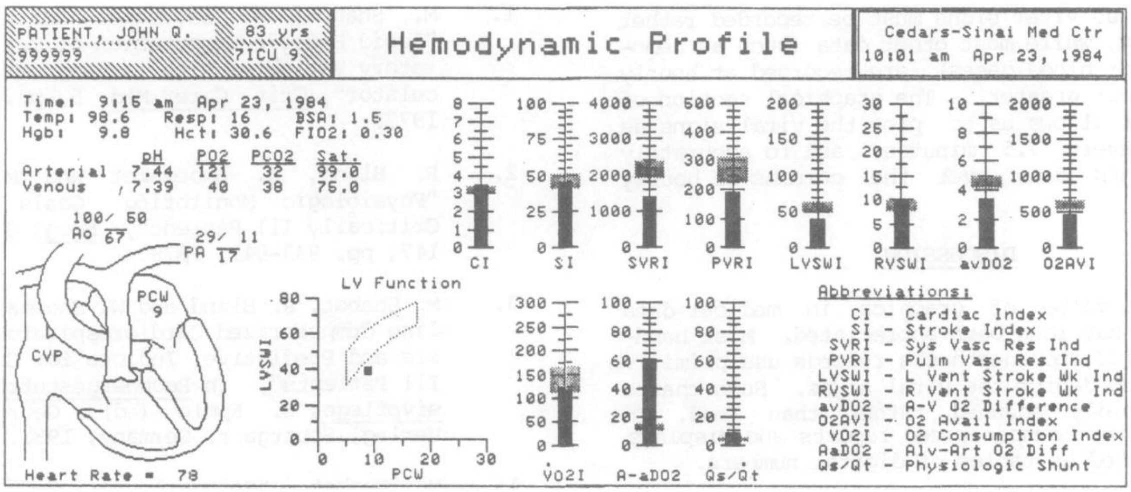


Figure 3: Hemodynamic Profile Display [51]

The goal of clinical graphical interface designers in the past decade has been to create novel methods for presenting integrated and multidimensional information on a single screen. Yet, current graphical display, have been designed to be strictly consumption driven, and thereby, support only minimal interaction and limit any manipulation of that data. These displays do not offer clinicians the ability to explore the physiological data space to extract novel information about salient physiological behaviours. Analysts in this domain seek simple and intuitive graphical displays so as to elicit rapid knowledge from multidimensional and multi-modal network sensor data to confirm hypotheses about clinical conditions and their associations [52].

2.1 Contextual Challenge

The monitoring period of an ill patient can greatly vary. An anaesthesiologist may attend to a patient for several continuous hours [53], while a critical care team may care for a patient over several days, weeks or even months as usual in the case of a premature infant. Moreover, while a patient is receiving care in the critical care unit, they may also develop new conditions that were not present at admission. Two of these conditions are apnoea of prematurity, and the onset of nosocomial sepsis.

2.1.1 Apnoea of Prematurity

Premature infants, who are born at less than 35 weeks' gestation have neonatal spells, which are episodes of cardiorespiratory events denoted by a breathing pause [54]. The breathing pause is a consequence of the infant's immature brain and its inability to maintain respiratory control [54]. An increase in the frequency of spells may be associated with other serious conditions such as a blood stream infection, lung collapse or a seizure. The term neonatal spells is commonly used in NICUs for cardiorespiratory events that include pauses in breathing, fall in heart rate, or fall in blood oxygen saturation [39]. The reason for this term is that no raw signal and no monitor can cope with the complex definition of apnoea. Infants cannot be observed by bed-side staffs continuously, it is not always possible to visually assess what happens when a monitoring alarm was initiated. Hence, the term spell has come into common usage to mitigate the uncertainty.

Apnoea is defined as a pause in breathing for 20 seconds or more or a pause in breathing that is associated with a change in heart rate, colour or muscular tone. A temporal pause in breathing is relatively easy to detect from a single stream of data. The second portion of the definition is more complex. This relies on the temporal integration of the three physiological streams of heart rate, breathing, and blood oxygen saturation, which serves as surrogate of colour. Clearly, this omits one of the portions of the definition, which is muscular tone. However, clinically, this is the least consequential as it rarely occurs in isolation from changes in heart rate or blood oxygen saturation [39].

Thommandram et al. developed three distinct sets of algorithms to accurately identify various types of neonatal spells [39]. By combining these three algorithms and looking for temporal patterns in the raw data, it is possible to accurately identify the various types of neonatal spells. Moreover, this algorithm produces significant secondary data, such as, the start and end times of an event, duration, and percentage from baseline of events across each raw data stream. Typically this information is hidden from the user to avoid information overload, however, in the proposed visual analytic framework, large volumes of secondary data will be shown to support the end-user in identifying contextual detail, and higher level insight.

The proposed framework will be used to demonstrate both the active and exploratory detection of large volumes of features extracted from this algorithm.

2.1.2 Late Onset Nosocomial Sepsis

Late onset neonatal sepsis (LONS) is a major health problem, in which the infant is faced with severe sepsis, a form of hospital acquired infection, requiring antibiotic therapy [35]. It is typically defined as sepsis acquired as early as four days before, or up to 28 days after birth.

Studies have shown that heart rate variability may be used as an early indication of the onset of sepsis [55]. Currently it is very difficult to detect using non-invasive methods, such as by bed-side monitoring. Clinicians rely on qualitative observational methods for identifying signs on this illness. When LONS is suspected, blood samples are drawn and required to confirm any diagnosis. However, neither methods have been found to be reliable [56].

McGregor et al. developed an algorithm that produces real-time heart rate variability scoring for neonatal infants [57]. This scoring can be used to identify temporal areas where there is reduced heart rate variability that indicates some sign of illness. Flower et al. [58], find that there may also be increased instances of periodic deceleration of heart rate when heart rate variability is reduced. Buchman [59] previously demonstrate that the presence of clusters of decelerations are correlated to neonatal sepsis.

In this doctoral research, the TDVA framework will be used to identify novel means of displaying heart rate variability and decelerations in heart rate information to support rapid detection of low heart rate variability areas, and within them, areas that may be indicative of sepsis across a cohort population. The visual analytic tool enables interactive exploration of cohort patient populations to perform hypothesis testing.

2.2 Clinical Contributions

The proposed research contributes key outcomes that advance information visualization and analysis in the neonatal intensive care environment. Firstly, the proposed framework supports the notion of coordinated exploration of temporal tri-event parameters to perform explanatory and exploratory research. By inheriting a strong temporal event-based approach, the proposed framework is able to highlight key metrics including frequency, duration, and trajectory of event features and event classifications. This approach allows the clinical researchers and domain experts to see past, present and anticipate future direction of events as they are relatively aligned to clinical events that manifest at the bed-side.

Secondly, the proposed framework has access to three types of data, namely, secondary and meta-data from the online analysis engine, the algorithmic event features and event classifications, and human-generated input. Graphical representation of these low-level and high-level datasets has not been previously demonstrated with clinical data streams.

The last contribution of the framework is tied to the visual design that displays information and knowledge to the end-user. The use of novel visual designs employing tri-event parameters, such as the Temporal Intensity Map, Sequence Graph, and other coordinated graphs, contribute unique approaches to addressing the clinical analyst's information overload and situational awareness in the intensive care environment. While the Cohort Relatively Aligned Dashboard contributes a visual analytic tool for performing cohort studies involving physiologic data.

2.3 Chapter Summary

This chapter presents an overview of the context in which the research of this thesis is applied. In modern intensive care units, physiological data is generated for instantaneous consumption, and after some time, it is flushed from the device. This limits the ability of the analyst to perform any exploratory analysis or hypothesis testing. A visual analytic framework for complex and critical environments can enhance the ability of the user in analysing large volumes of data streams from a single patient or consisting of data from multiple patients. The design considerations for such framework, must additionally incorporate important temporal properties of streaming data, such as the trajectory of the signal, prior trends, and adopt effective modalities to highlight regions of interest.

The next chapter presents the literature review of computer science research that forms the basis of this thesis. Chapter 4 presents prior work in the space of physiological visual representations, and from Chapter 5 – 8, this thesis contributes knowledge that can address some aspects of the contextual challenges identified in §2.1.

3. Literature Review

The design of analytic systems have a rich history dating to the early 1950s, one of the seminal works in management decision support, titled “The New Science of Management Decision” by Simon (1977) [60] introduced a structured approach to the development of usable decision support systems. Simon proposes a methodology involving four major phases: intelligence, design, choice and implementation. The intelligence phase allows the designer to identify constraints, the design phase involves formulating a model of the system and setting criteria for choice, the choice phase includes active selection of alternatives and finally the implementation of the system occurs. However that approach does not consider salience of high-throughput sensor data systems. Much of the processes are manual and human-driven, and the methodology inherently assumes static nature of the data. In order to extend that work into the sensor data analytic domain, a framework and methodology are proposed. The framework incorporates several aspects of Simon’s processes, but also extends areas where data-driven automation can occur.

This chapter presents details that support the development of the TDVA framework, methodology, and platform. The chapter is structured as follows: §3.1 presents unique requirements of sensor data, §3.2 introduces the concept of complex data, §3.3 identifies prior data warehouse approaches, §3.4 introduces information visualization, §3.5 presents several related works in visual analytics that support interactive analysis of a large variety of data formats, §3.6 presents work in dynamic visual analytics and identifies open research in that space. §3.7 concludes this chapter with a summary.

3.1 Sensor Network Data Streams

Sensor networks are a collection of sensors that produce streaming data [61]. Streaming data is a collection of synchronous data elements that provide a continuous and detailed snapshot of a system's current state. Multiple data streams can be generated from a system by sampling various components using networks of sensors, such as their health parameters and respective states. With every 'movement' of the data objects, greater duration and specificity of the systems states are captured by those network sensors [62]. Continued movement of this data stream allows trends to be captured, which may be used to discover the temporal stability or instability of that system. The aggregation of significant volumes of data streams that strains traditional search, filter, and analysis methods is collectively defined as big data [63].

Advances in information management have produced novel methods for collecting and analysing network streaming information [64], [65]. Location sensors are used to map episodic movement of multiple devices between points in an area map. The movement may be captured using information retrieved from static sensors, during a transaction, or using mobile sensors carried by the user [66]. In each scenario, streaming data must be associated temporally and identified to physical locations. In prior work, bio sensors have been used to explore motion and other physical parameters for monitoring health statuses of critically ill patients [67]. While advances have also other sensor data environments, such as in automated and visual analysis of location [66], [68] and social-media based streaming data [69]–[72]. Further contributions are required in areas where subtle behaviours in data streams require analysis to discover new patterns and generate hypotheses.

An example of complex big data may involve a single residual data stream integrated from variability features extracted from a stream of heart rate, oxygen saturation and respiration. Individually, these base streams may provide clues of their individual organ, but when integrated that residual allows for a high-level appreciation of the cardiorespiratory compensatory system. However, the results of that data are then used to infer the condition of organs, tissues, and prepare prognosis for future states and trajectories. Another example of complex big data is the efforts by Electronic Medical Record and Genomics (eMERGE) consortium to integrate with the Pharmacogenomics Research Network to support clinician practice [73]. Pharmacogenomics data provides significant volumes of casual and inferential data which provides knowledge in form of definitive diagnosis and potential treatment pathways [74], [75].

Continuous monitoring of multi-modal and multidimensional physiological data is an integral aspect of evidence-informed care and is largely composed of complex big data. After acquired through numerous invasive and non-invasive medical sensors attached directly to the patient, this complex data is reduced to modest periodic samples for the purpose of cognitive ease. However, the machine sampling frequency for these sensors can be in excess of several kilohertz [37]. Clinicians attempt to gain situational awareness at the cost of dimensional reduction, that is, kilohertz frequency is reduced to hourly or daily values [76]. Persistent generation and consumption of continuously monitored physiological data consequently, presents challenges to the clinician who must comprehend short and longer term trends to makes appropriate decisions in line with accurate interventions. The next sections present related work in the area of complex big data, supporting decision-making in complex environments, and concludes with visual analytic as an emerging field in knowledge discovery and information gathering.

3.2 Complex Big Data

The previous subsection introduced large volumes of data streams and their accumulated pool as big data. In this section, the concept of complexity will be introduced, and Complex Big Data will be discussed.

A large group of complex big data sets are scientific in nature; these sets are gained from simulations of physical phenomenon like climate change, biological interactions, chemical composition, and medical scanners [77]–[79]. They contain hidden underlying associations or compositions which are difficult to ascertain by simply looking at their raw data. However, when integrated across multiple modes or data streams, interesting features can be identified [77]. However, before these features are observed by the user, complex data undergoes several stages of pre-processing such as signal processing, artifact filtration, and model projections [28], [80].

Pedersen & Jensen propose nine requirements as discrete signatures of complex multidimensional data [81]. These requirements were then evaluated against existing data models to assert that none of the data models could support all nine requirements. Three of the primary requirements proposed by the authors were attributed to the absence of visualizing hierarchy of each multidimensional data, along with displaying uncertainty and granularity in the low level data. Keim, further separates transactional and analytics aspects of data. He proposed that the dimensionality and type of the data be used in identifying usage and potential [82]. He illustrates examples to show data can exist in one of many basic types, for instance: one-dimensional, two-dimensional or multidimensional.

One-dimensional data are those which form a sequence of analog or digital signals existing in a single dimension. A prominent group of one dimensional data is temporal data, which includes: audio, video, or location tracking information [83]. Two-dimensional data include mapping information such as geo-spatial information and diagnostic images [84]. Multidimensional and multi-modal complex data are those in which several one-dimensional complex or homogenous data objects are aggregated to produce multidimensional associations from multiple unique sources [85]. Physiological data, in particular those involving network of biomedical and imaging sensors can be seen as multidimensional.

Karasti, Baker & Halkola, 2006, present several additional characteristics of data. The authors differentiate between deep-simple, wide-complex, and deep-complex data volumes [86]. Data variations are unique to individual disciplines of basic and translational research. Genomics consist of deep volumes of data yet the data themselves are sequential and non-complex [73]. Field observations are examples of wide-complex data, in which many independent observations contain enriched information and contain pre-filtered data. Deep-complex data, such as those generated from multi-modal and multi-run 3D simulations of computed tomography imaging, contains significant intrinsic complexity and involve large data volumes that can extend to petabytes in volume [87], [88]. Physiologic data resides in between wide-complex and deep-complex. While the sheer volume may not extend to those generated from diagnostic imaging, it demonstrates deep-complex potential when considered over long period of time, including the lifetime of a human.

An extreme example to illustrate the differences between big data and complex big data is to consider the discovery of the Higgs boson, in which approximately 200 petabytes (2×10^{17} bytes) were collected and analysed [89]. Although significant amounts of data were captured, the data sampled just three parameters measuring the mass, direction and energy of the particles. Hence, as illustrated in that scenario, by volume alone, complexity can be introduced. With complex big data, volume is just an aspect of the environment, these data sources also include attributes such as hidden variables, and complex linked graphs [90].

3.3 Traditional Data Warehouse Model

Traditional data warehouses provide the analyst with fast and convenient access to data gathered from a variety of operational data sources to support decision making by analysts [91]. A traditional approach to data warehouse function is illustrated in Figure 4.

The traditional data warehousing model collects data from several operational source systems, such as sales, customer interactions, marketing, and finance for the purpose of enabling rapid querying and analysis. The frequency of extraction from these data sources are highly variable, while some data marts are updated once a week, others such as the customer relationship management applications retrieve data when a customer calls in for support [91]. Therefore, the traditional data warehouse does not support near real-time or instantaneous response as required by users in complex environments [92].

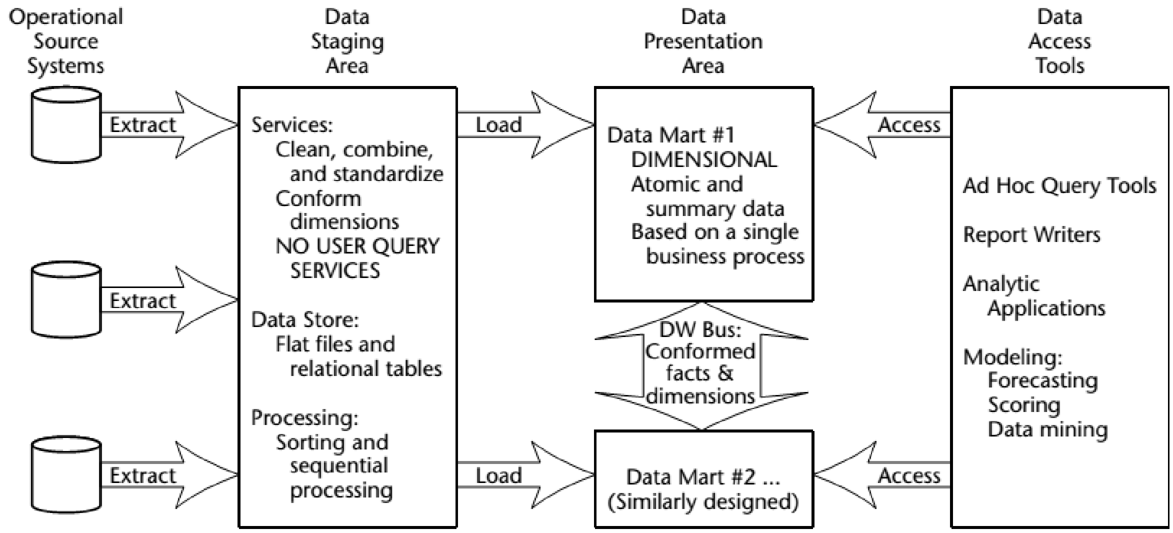


Figure 4: Traditional Data warehouse model by Kimball & Ross, 2011 [91]

The traditional data warehouse model, according to Kimball & Ross [91] begins by (1) retrieving source data from operational systems, (2) cleaning and standardizing the operational data at the data staging layer, (3) storing this data in a standardized data marts in the presentation layer, and for improved efficiency and speed when querying specialized sets of data by providing (4) dedicated portals for ad hoc data queries. Once the data marts are ready, the traditional models accept connections from a suite of applications grossly simplified at the presentation layer [93], [94]. This often involves reducing front-end function to browsing and shifting between pre-determined views.

That approach can be seen as following an explanatory analytic approach, wherein, knowledge pertaining to data modeling in included a priori. Concretely, this a priori knowledge is kept within online analysis process (OLAP) cubes by pre-specified dimensions augmented against a static dataset [95]. These OLAP cubes containing arrays of data are then used by business intelligence suites as dashboards or reports for analysts to view and analyse largely static data.

3.3.1 Pre-built OLAP Reports

OLAP generated reports and streamlined graphical dashboards that are used to extract knowledge from low level data are limited in their ability to support explorative analysis of relationships and behaviours in complex data [96]. For instance, much of the currently deployed custom-reports in medicine have limited temporal and human-factor considerations to support dynamic analytics [97]. A large number of systems supporting analysis of physiologic data use dashboards and methods that were transformed from displays developed for the enterprise [98]. Still, much of business dashboard data are abstract and oriented towards retrospective reports. Hence, this method presents challenges for gaining insight or performing deep analysis of multidimensional temporal behaviour observed throughout the course of intensive care [99].

Another complex data intensive domain is network management, where some notable visual analysis tools have recently been contributed [16], [100]. In this domain, the dynamic analyst captures real-time events as they occur across local or wide area computer networks. Numerous hosts are monitored in real-time to identify compromised hosts as well as to prevent precipitous attacks. The network data stream itself, however, is not as complex as physiological data streams. Data streams in the network management systems contain events are captured by performing edge detection on binary streams [101], and usually captured at longer intervals [16]. Similar to the clinical domain, network management systems have used static custom reports using pre-built queries [16]. As in the clinical domain, these custom and pre-built query reports limit the ability of the dynamic analyst in this domain to compare across epochs and perform basic functions such as aggregation.

Moreover, these pre-built and custom reporting interfaces suffer from poor representations of sequences, and trajectory of multidimensional events [16]. In the clinical domain for instance, changes in blood pressure and the residual influence on bodily fluids are not intuitively represented in the custom reports due to the interface's visual segmentation by unique organ systems [102]. In the case of network management, traditionally dynamic analysts needed to manually siphon for salient features amidst an average of thousands of events that are presented over a typical shift [16], [100]. These method of compartmentalisation, and tabular presentation are derived from the division of functional units of the enterprise, and hence, appear obtuse when applied to dynamic environments [103]. Ultimately, this limits the holistic view required by researchers and analysts to adequately appreciate dynamic behaviours occurring across complex heterogeneous data from network physiological sensors and human-derived data [104].

3.3.2 Reusable Online Analytic Systems

Online analytics system present modular components designed to process large volumes of data. Figure 5 presents the Solution Manager Service (SMS) [105] architecture for an event-based data warehouse that sources data from network of electronic sensors and traditional data source systems. In that architecture, data streams are collected in the web service interfaces layer where all necessary data staging occurs (Figure 5). Following this, data streams are sent to the event stream processor for near real-time analysis and trend detection.

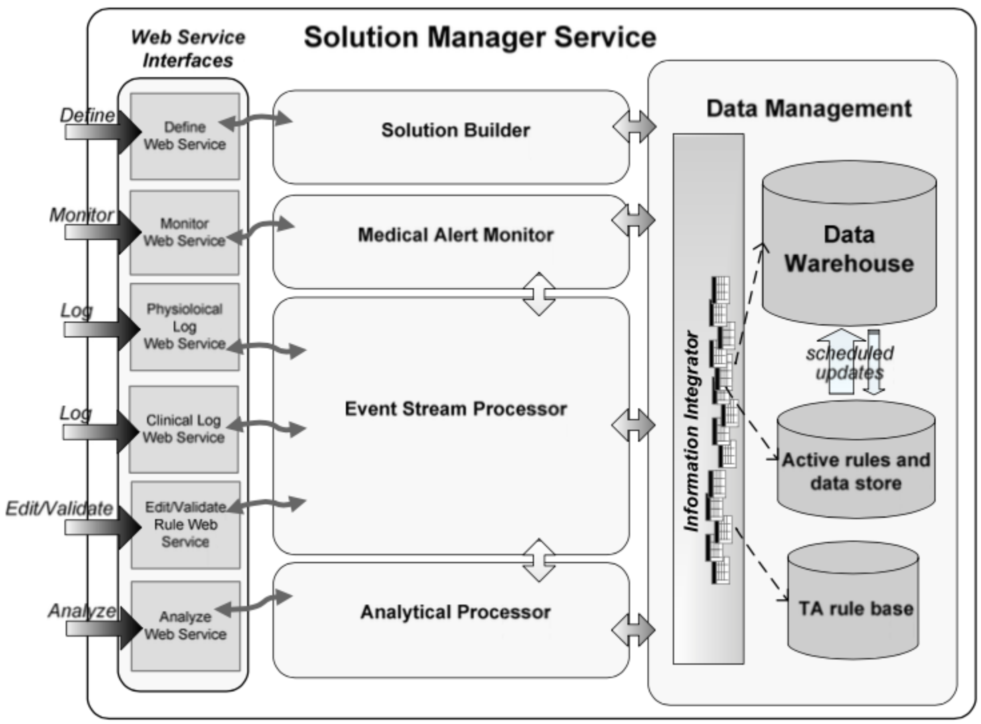


Figure 5: Solution Manager Service Architecture [105]

After the analytics has been performed, information is integrated and stored via ELT in a data warehouse, post analytics, the SMS model then allows reporting software and other front-end business intelligence tools to access data for presenting the data to the analyst. The presentation layer is seen as an external service that must access data using traditional query and ELT methods. While the SMS architecture, as opposed to the traditional data warehouse model, allow for near real-time response rates, the SMS model does not support session based data modeling.

Unstructured Information Management (UIM) applications are software systems that allow for the analysis of large volumes of unstructured data [106]. Similar to the SMS architecture, the UIM system supports interoperability based on a common data structure, and loosely coupled components known as annotators. These annotators are used to analyse unstructured data, largely text based to produce assertions about the data. The UIM system has gone on to support other specialized analytic systems, such as the IBM Watson system [107]. The key advantage of UIM is the decoupling of data storage, visualization and analytics. The modular and reusable approaches utilized by UIM applications afford great flexibility and scalability. The UIM framework places emphasis on a large section of unstructured datasets, such as text, audio and video, however it does not support data streams from dynamic sensors streams.

Event stream processing engines, such as IBM InfoSphere Streams [37] and Apache Storm [108] ingest data streams in real-time to extract event information. While the definition of events are loosely defined, these systems enable the use of temporal windows to perform event alignments, and compare sequences. The event stream processor can identify events according to the temporal behaviour exhibited by the data stream. In contrast to OLAP cubes which operates on static data, event stream processors apply instantaneous transformations on dynamic data, and include several layers of fault tolerance in order to guarantee each tuple is processed [109]. The results of that analysis can produce any one of the primitive, complex, multidimensional events to make predictions about the onset of a physical event. An event stream processing engine is utilized in the TDVA framework to support dynamic and static processing of physiologic data in this thesis to generate a collection of primitive, complex and multidimensional events. This information is provided to a user who discerns whether a physical event, such as the presence of infection can be predicted using the event stream processors output.

The traditional data warehouse model contains four major limitations that prevent dynamic visual analysis. Firstly, OLAP cube based analytic model provides limited access to dynamic real-time analysis of streaming data [93]. The cube based design assumes buckets of data are available for subsequent cubes to be created. Secondly, in the dynamic real-time environment, the life of data is measured in minutes and hours, unlike the traditional business case wherein life is measured across days and weeks [94],[110]. Thirdly, to enable dynamic analysis, the presentation layer cannot be restricted to simply ingesting data. Interactive and direct manipulation of visual interfaces driven by the human should allow for analysis to be performed on demand, and, data dimensions to be remodelled according to the requirements of the individual analyst.

Finally, unlike in a traditional business use case, domains where real-world data is captured require an additional level of pre-processing and temporal emphasis [111]. This includes individualized supporting components that convey meaningful changes in event features and highlight abnormal distance measures. In dynamic and complex environments where temporal interactions in data streams assume significance, visual analytic systems employing multi-dimensional temporal representations, such as trajectory, frequency, and duration may be useful for supporting the analytic requirement of the user [42]. This final constraint requires systems that are flexible, support functions such as dimensional aggregation, and, provide a dynamic link between the event stream processor and visual analysis interface for interactive exploration of the analytic space [95].

3.3.3 Artemis Platform

Artemis, illustrated in Figure 6 is an online health analytic platform that was developed, to source, analyze, and perform real-time feature detection on multiple physiological data streams, for multiple conditions in multiple patients [37]. Artemis supports clinical researchers in identifying and demonstrating evidence for new earlier onset detection features in physiological data to help identify clinical conditions earlier and create evidence for new care practices in NICUs.

Artemis supports the deployment of real-time event stream processing algorithms. In this thesis, an algorithm for neonatal spells was executed to detect and classify neonatal spells into 10 broad event classifications. The details of the neonatal spells algorithm have been previously published [32]. Results from the online analysis are then sent to a database and also available for real-time streaming. The output are then processed and sent to a visualization framework that produces visual marts, such as PhysioEx and CoRAD. Chapters seven and eight demonstrate PhysioEx and CoRAD in the context of supporting neonatal research by presenting trends and pattern information in neonatal physiological data detected by Artemis.

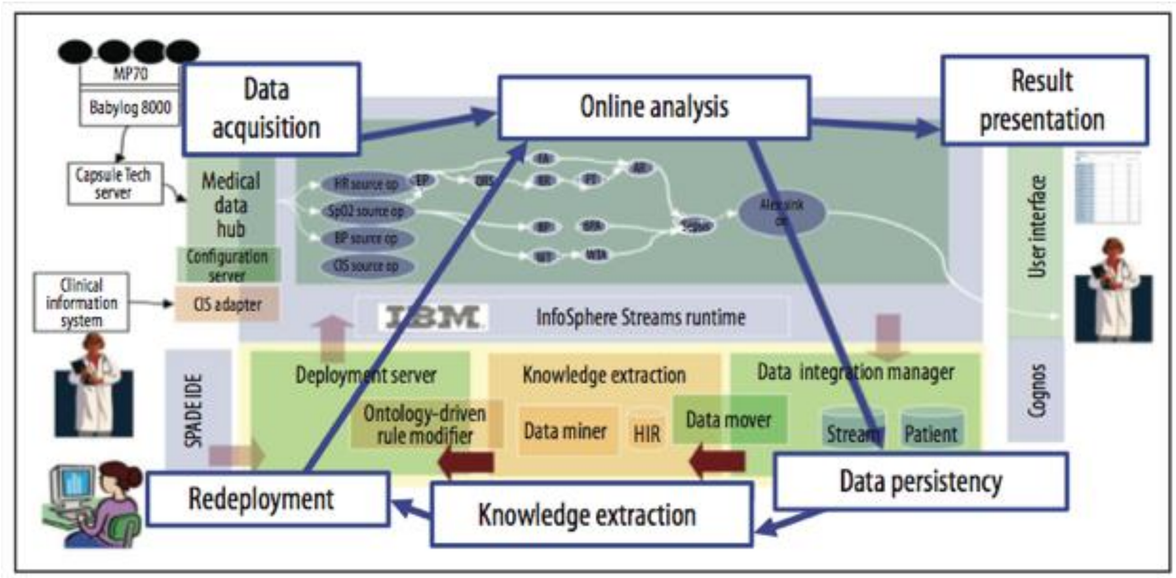


Figure 6: The Artemis Framework [37]

This thesis extends the Artemis platform through the TDVA platform that instantiations dynamic visual analytic visual marts. These instantiated components then interface the event stream processing engine to provide an analysis pipeline to support data stream driven complex research. The next two parts of this chapter from §3.4 – §3.5 introduce prior work in the area of information visualization and visual analytics. The final section in this chapter, §3.6 introduces the dynamic visual analytic paradigm which serves as a link between the content discussed earlier in §3.1 – §3.3 and §3.4 – §3.5. The TDVA framework, integrates the paradigm of dynamic visual analytics within the online analytic environment described in this section (§3.3.3).

3.4 Information Visualization

For millennia, humans have relied on graphical representation to make tractable very complex and critical information [112]. The human visual perception potentially supports higher bandwidth of information than any of the other senses [113].

The human visual system has remarkable capacity to capture fine and coarse changes, while rapidly distinguish between various form, colour and shape of objects [112]. Furthermore, the human can recall objects seen during an earlier exposure, and compare that pattern with current observations to glean knowledge [113]. These capacities of the human visual system have been exploited in several classic graphical representations of complex problems. Examples of these techniques include the Pythagorean Theorem, Cartography, and William Playfair's demonstration of graphical statistics [114]. Information visualization continues to provide a means to enable rapid communication of information and knowledge.

3.4.1 Visualizing Temporal Data

Time is a unique measure, apart from acting in itself as a variable; it also conveys the initiation, evolution, and termination of other parameters [83]. The primary objective of information visualization in the time domain is to recognize and compare data points located in two separate positions on the time axis. Extending from this, are other objectives to gain further insight through various trend analysis and subsequent transforms. Endsley [11], tightly incorporates the concept of time in situational awareness. Endsley's model is seen to be a continuum existing across three levels of knowledge, wherein the consumer (1) sources, (2) synthesizes, and finally (3) imposes contextual meaning to predict future trends or states.

Information visualization research has provided numerous tools to support users who exist in the temporal domain, and require one of the three aspects of Endsley's situational awareness. These tools are described below.

A method that provides an aesthetic view of time-series is the method called Flocking boids by Vande Moere [115]. Here, Vande `Moere demonstrates a method to visualize the progression of stock market prices using simulation and animation of flocking behaviour. This provides a simple and convenient method to track each ticker and its trajectory. The movement of the ticker is governed by behaviour rules in which each boid follows the trajectory of earlier time points, notably with animations presented in 3D. This method however is prominently univariant which presents considerable limitations.

TimeWheel is another method introduced by Tominski et al. [116]. In this method, the approach is to convey information simultaneously in several axis to perform multi-variant and multidimensional analysis without compromising the aesthetics of the view. It is however a static model, which feeds information from a predefined data set and would not be used to convey any dynamic real-time temporal information. However, by positioning the time axis in the centre and then defining additional interdependent axis around the circumference, the aesthetics of the graph is greatly enhanced.

While traditional time-slices have been represented using histograms, Havre et al. introduce an innovative temporal chart called ThemeRiver (Figure 7) [117]. This representation allows continuous representation of discrete values over a timeline without interruptions between each node. Coloured currents flow within this 'river' and express strength or weakness by the thickening of their width.

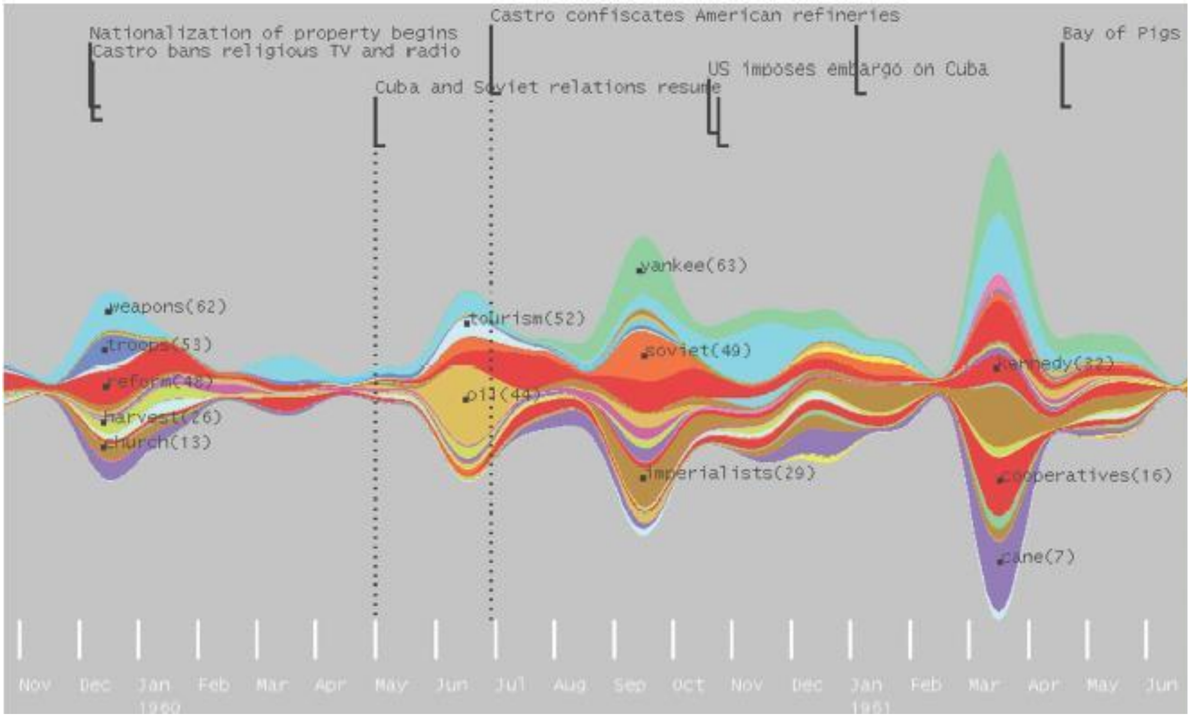


Figure 7: ThemeRiver [117], identifying patterns over time.

The goal of this method is to convey patterns rapidly. Hence, when a current flowing through the river suddenly widens, the user can appreciate its strength or significance has increased.

CircleView (Figure 8) [118], introduces another novel method for exposing temporal information. In this representation, each layer represents an additional segment of time, and colour is used to encode a severity value between low, high and above threshold. This representation displays a modifiable evolution of temporal trends the user can identify by following each layer, from the innermost to the outer or vice versa.

Techniques like ThemeRiver and CloudView, as well as other techniques which will be discussed in the following section provide a means for rapid knowledge dissemination, which was previously difficult to conduct and required time-consuming methods.

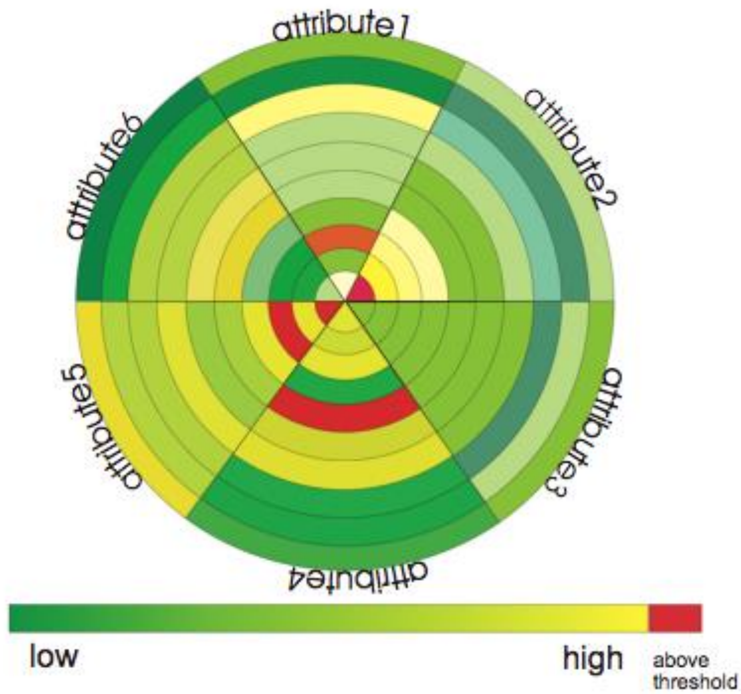


Figure 8: CircleView [118]: Showing evolution of multiple attributes over time.

Graphical tools, apart from enabling knowledge discovery, also foster new insight. These tools enable domain experts to formulate rich collaborative discussions around a display and generate new hypothesises reflecting observed patterns.

3.4.2 Information Visualization in the ICU

Health visualizations have seen a growing body of work with several novel representations proposed in the area of physiologic analysis [119]–[123],[124]. Health visualizations can be broadly classed into three groups, tabular, graph, and metaphoric. Many legacy visualizations have adopted a tabular display, such as flowsheets [125]. Integrated graph-based physiologic tools such as CareCruiser [119], and those presented by Anders et al. [4], and Koch et al. [6], illustrate methods to identify signal changes in waveform. These three visualizations are utilized in monitoring use-cases where the aim is to detect abrupt changes from baseline, a task commonly performed at the bed-side.

Albert et al. [121], Wachter et al. [123], and Zhang et al. [76] all present metaphoric visualizations to convey abstract principles derived from the real-world. Albert et al. [121] present the graphical cardiovascular display, a pipe-like visualization that presents up to 12 hemodynamic variables, including stroke volume, central venous pressure, pulmonary vascular resistance, heart rate, mean arterial blood pressure, ST segment depression of the ECG waveform among others [121]. Wachter et al. [123] isolate a sizable list of variables from the ventilator and develop an anatomical display, utilizing the abstract metaphoric principles. Finally, Zhang et al. [76] develop a balloon metaphor display for hemodynamic monitoring that shows the balloon expanding and constricting as pressure and volume change. These visual designs highlight the diversity of representations that exist in the medical domain.

While many methods have been developed, existing health visualization methods are limited in the analytic facilities, along with limited interaction with the visual interfaces. Much of the work has gone into designing real-time systems for surgery, in which the time duration is limited to immediate past and present [53], [126]–[128]. Novel methods to display clinical information using simplified and integrated displays have been proposed but have not be implemented in the clinical environment [28], [129]–[132].

Case-control studies (cohort) remain an important aspect of clinical research [133]. A case-control study involves retrospective analysis that separates patients based on the presence of a condition [134]. Differences are studied and hypotheses are generated based on the analysis that support future more rigorous research. However visualizations that support these efforts in physiologic data remain elusive. A major contribution of this thesis is a novel cohort-based relatively aligned dashboard for comparing physiologic variation within patient populations.

In the general space of health-based cohort analytics, some recent work has resulted in high fidelity visualizations. TimeSpan [135], provides an interactive dashboard for identifying door-to-needle time for stroke patients at a large tertiary hospital in Calgary. LifeLines presents graphical summaries of patients [136]. The Cohort Comparison (CoCo) tool, illustrated in Figure 9 provides a simple interface for exploring statistical correlations across multiple clinical datasets [137].

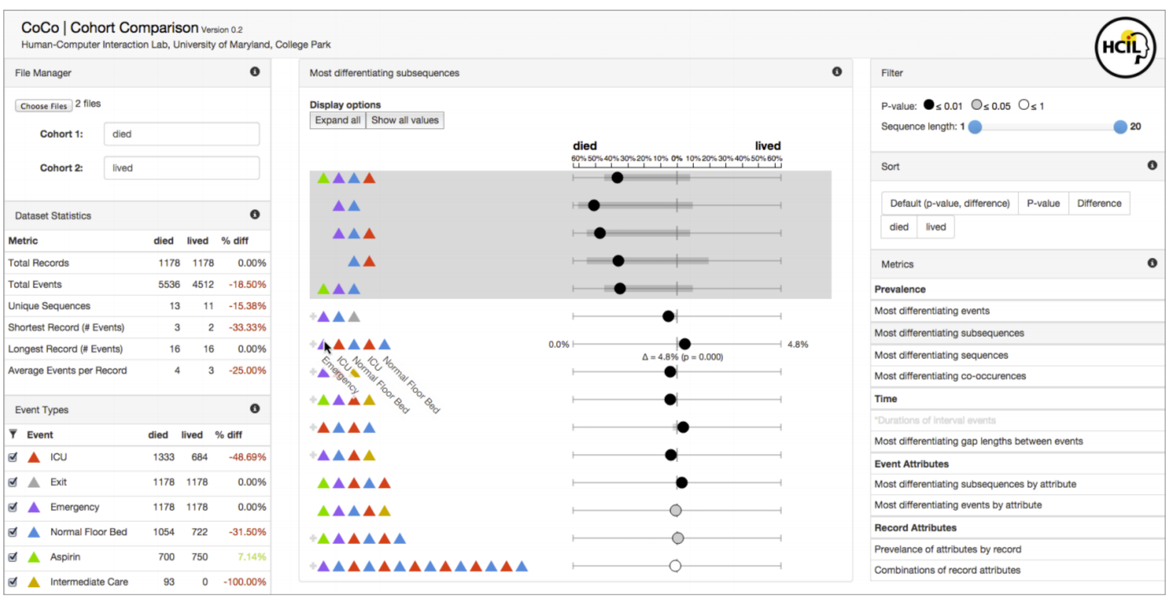


Figure 9: Cohort Comparison of Event Sequences [137], an exploratory interface for clinical data

DecisionFlow presents graphical summaries of patients who developed heart failure relative to a population [138]. VISITORS is a dashboard for analysing clinical temporal abstractions in oncology patients [139]. EventFlow presents a method to simplify event sequence information to rapidly identify abnormalities [140]. While all of these visualizations introduce cohort analysis of patients using clinical information, there is a need for research in representing temporal abstractions of physiologic data across cohorts due to the lack of contributions in that space. The visualization should allow for temporal alignment, enabling the user to gain contextual awareness using low and higher-level summarizations of data.

A systematic survey was conducted as part of the research program to investigate attributes of existing physiologic displays. The survey also sought to identify common elements used in the visual displays, along with their support for the tri-event temporal parameters and case-controlled studies. The methodology and results from the systematic review are presented in the subsequent chapter.

3.4.3 Methods to Evaluate Visualizations

The primary goal of visualization research is to study the ability of a user to discover insight, through directed task, or serendipitously [3]. To determine the level of insight that is generated by a visualization, four thematic areas of evaluations exist, namely, controlled experiments; formative usability testing; metrics, heuristics, and models; and longitudinal, case, and field studies in realistic settings [141], [142]. Of the four, longitudinal case, and field studies in realistic settings have been noted to come closest to examining the open-ended usage of visualization to derive insight [143]. The challenge of evaluating visualizations for their ability to generate insight has been long understood [144].

In order to provide knowledge in meaningful and insightful visualizations, a variety of visualization research methods have been introduced [145]. One such method proposed by Isenberg et al. is grounded evaluation [146]. In grounded evaluation, an increased emphasis is placed on qualitative inquiry early in the visualization development life cycle. The goal is to collect data that describe meaning and rich contextual information such as subjective experiences. Observations, interviews, documents (written artifacts), and audio-visual materials are four basic sources of qualitative data [147]. The grounded methodology begins with identifying the problems, tasks, and pre-existing visualization tools. It is standard to use a small cohort of domain experts for gathering requirements and then going back to those experts to present novel visualizations. This practise has been called human-centric visualization design in the literature [148], [149]. The understanding gained from the pre-design study are then used as evaluation criteria. Sample size are often lower than for quantitative studies, and continually sourced until no new data can be gained [150]. There isn't currently a method to identify when that saturation has occurred [151]. The qualitative data is then analysed using the thematic analysis approach, where themes are identified through the review of observational data and coded. Open-coding is often used to draw conclusions from the raw data [147]. Finally, as in qualitative research, the grounded evaluation methodology acknowledges the researcher's views, research context, and interpretations are a key part of the research process such that it is grounded in the collected data [146]. The grounded evaluation methodology is used in an exploratory study presented in this thesis (§5.1).

Problem-driven work, also known as a design study in the visualization literature, aims to work with real users to solve real-world problems [149], [152]. Sedlmair et al. identify three classes of contributions in that space [148]. The first-class contribution is identified as “problem characterization and abstraction”, in which a domain problem is characterized through an abstraction into tasks and data elements. Knowledge from a problem characterization and abstraction can be used by other researchers to develop fully automated approaches. The second-class contribution is a validated contribution, in which a tool is evaluated and evidence is presented. The last form of contribution is a reflection about lessons learned to improve future designs. In this thesis, the visual analytics contributions, namely the PhysioEx, and CoRAD fall within the first two classes of contributions.

Characterizing insight is a challenge problem. To address that challenge, a novel insight coding methodology was proposed [142]. That methodology introduced a series of general insight characteristics that can be used to better distinguish the nature and depth of insight. Insights can be characterized as, observation: any single finding about the data; time: the amount of time taken to reach the insight; domain value: perceived importance and significance of the insight; hypotheses: the ability of an insight to generate new hypotheses; correctness of the insight, breadth vs depth: the depth of the insight that was generated; a categorization of insight based on one of four types (overview, patterns, groups, and details). Those insight characteristics can be used as a code for assisting evaluators in determining the degree of insight generated by the user [149], [153]. In this thesis, the insight-based methodology is used to report the nature and depth of insight that was generated while using both PhysioEx and CoRAD.

3.5 Visual Analytics

The benefits of involving a human as part of an automated chain of analysis appear to be a rewarding and highly efficient practice [154]. While algorithms process quantitative data, the human applies qualitative expertise. This two-step analytics adds value to existing analytical practices by providing easy access to manipulating complex data and directly executing arithmetic functions on interesting subsets [83].

High densities of data, have been a challenge for traditional methods of information visualization, particularly to communicate changing temporal events [155]. Thomas and Cook, 2005 in a book titled Illuminating the path: The research and development agenda for visual analytics, introduce a method of human-driven visual exploration of data to discover pattern, trends, and knowledge. They call this method visual analytics and define it as the science of analytical reasoning facilitated by interactive visual interfaces [15].

Keim et al. refine this definition by proposing visual analytics as the combination of automated analysis techniques with interactive visualization for effective understanding, reasoning and decision making for big data [156]. Keim further states that the overarching vision for visual analytics is to turn the information overload problem into an opportunity, that is, to make the method by which data and information is processed more transparent for analytical discourse. This involves (1) synthesizing information and deriving insight from massive data; (2) detecting expected and discovering unexpected events; (3) provide timely, defensible, and understandable assessments; and finally (4) communicate assessment effectively for action [156].

3.5.1 Visual Analytics Process

The Visual Information Seeking Mantra "Overview first, Zoom/Filter, Details on demand" for information visualization was proposed by Shneiderman to assist in explorative activities [157]. Keim expanded that definition to include user interactivity and automation as key elements of the exploration paradigm, and thus, formed the Visual Analytic Mantra "Analyze First - Show the Important - Zoom, Filter, and Analyze Further - Details on Demand" [158].

A researcher performs two basic functions; first she actively searches through the information space to identify relevant data objects. After aggregation, she attempts to identify trends, or features within that data set. If she is not satisfied with the existent view, she modifies the analytical space to produce new views that may have further potential for gaining insight. This process is continuous and cyclic until she has satisfied her objective. While information visualization assists in the first aspect of information seeking, the two latter and mainly analytical components are challenging when using traditional visual methods. In order to derive some hidden knowledge the user, in addition to being able to see the data, also requires tools to manually search, compare, and query the data in an interactive user-computer interface.

The visual analytic pathway proposed by Keim et al. is illustrated in Figure 10 [159]. Data transformation occurs followed by either a simple visualization of the data, or a pathway leading to the generation of data mining driven models. These models are subsequently visualized to produce knowledge.

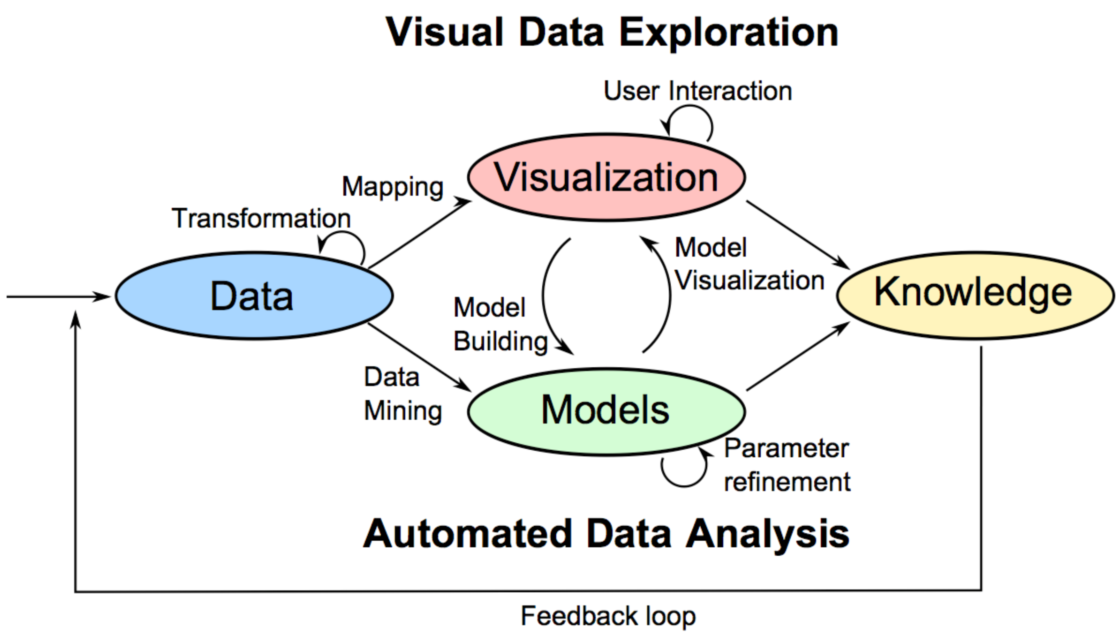


Figure 10: Visual model driven data mining for interactive decision-making [159]

Visual analytics is built around the principle of improving both the presentation and allowing direct manipulation of big data. Several methods have been developed by the community to execute queries entered by some combination of graphical sketches or gestures that act directly and manipulate data objects to perform functions defined by the user [160], [161]. Traditionally, physical and biological sciences have championed the collection and visual analysis of data [68], [162], [163]. Astronomy and genomics are examples of rapid data gathering and meticulous post-collection analytics. Limited applications of visual analytics has been adopted in those fields, these tools allow researchers to perform exploratory analysis on large volumes of data [89].

3.5.2 Visual Analytic Techniques

A number of prior works has focused on techniques for visual analysis of temporal big data, these works have relied on techniques such as hierarchical clustering [164], temporal nesting [165], alpha blending, sampling [71], and interactive brushing using parallel coordinates [166]. These techniques allow for significant number of shapes to be drawn on top of each other without hiding underlying shapes.

3.5.2.1 Techniques for Representing Density

Heatmaps have been widely used in visual analytics community to represent density or clustering of parameters from large volume data sources [167]–[170]. The use of a two-dimension matrix and intuitive temperature colour scales, have made heatmaps usable with minimal learning [171]. Not only are these displays effective at summarizing large volumes of information, but they also effectively allow the user to see the “big picture” [170]. Frequently, this technique has been implemented using two dimensional matrix of colour-shaded cubes, however, heatmaps can also been created along a single dimensions [172], and three dimensions as illustrated in Figure 11 [173]. Heatmaps are commonly utilized for gene expression and across scientific domains where large volumes of data must be reduced to manageable visual dimensions [167].

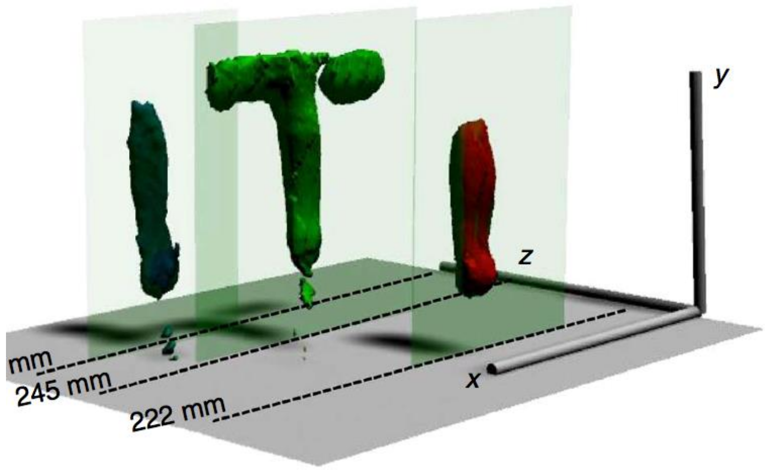


Figure 11: 3D heatmap of voxels [173] with depth controlled using colour scales

GScope [174] uses heatmaps to display biological microarray data, a domain where heatmaps are prevalently used. To generate the heatmap, they use hierarchical clustering based on up-regulated and down-regulated genes. The weakness with this representation is that the visualization is devoid of multisource and multidimensional attributes, and largely relies on Boolean aggregation of higher levels. In the example of reconstructing voxels [173], Velten et al. use novel 3D heatmap representations with shadowing to confer additional meaning. That representation however requires a degree of discretization, hence poses limitations when applied in the continuous temporal space. Krstajić et al. [175], introduce a visualization for tracking news streams using heatmaps. They introduce a timeline visualization which effectively allows the user to track changing trends over time. The timeline approach presents unique opportunities to highlight events across data streams. This approach can be further investigated within the space of heatmaps to contribute novel designs tailored for streaming events.

3.5.2.2 Temporal Event Flows

A large number of visual analytic software seeks to expose temporal changes in data. One example of a social media visual analysis tool is EvoRiver [176] illustrated in Figure 12. This technique visualizes analysis performed on a flow of numerous text streams to determine positive or negative collaboration within a social media case study. The technique itself, while retaining some features of ThemeRiver [117], provides the ability to interactively isolate single trends of discussion and follow them through their duration. As with ThemeRiver, this representation requires some level of cognitive attention to follow events; this can be a luxury in some critical domains.

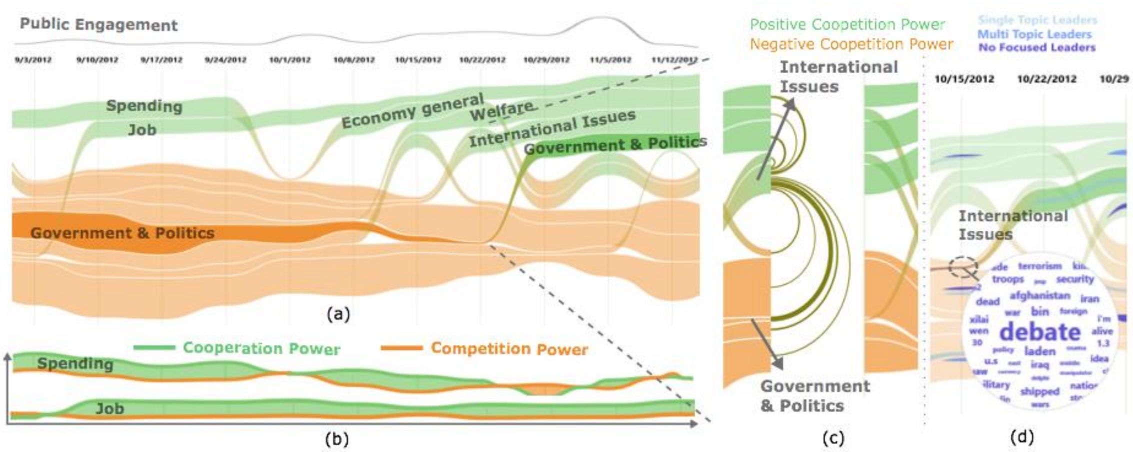


Figure 12: EvoRiver [176] a time-based social media visual analytic tool.

CloudVista (Figure 13) uses an interactive-video based method of displaying large quantities of data, however, since this only offers a single snapshot at one time, this technique does not allow for the identification of periodic and atypical patterns across time [177].

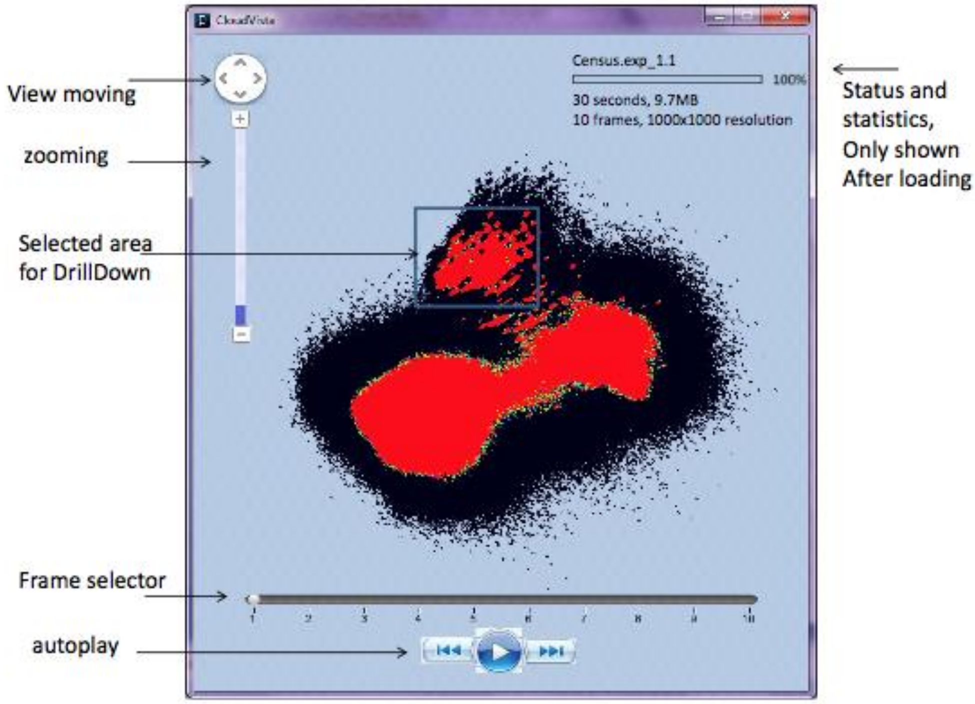


Figure 13: CloudVista [177] client-side interface

Over-plotting effects of time series data are a common problem, particularly in representing high frequency data streams. This challenge has been, to some extent, reduced by novel methods of sampling and clustering. For instance, hierarchical clustering have been used to allow the user to appreciate low level trends across multiple clusters [164]. Incremental sampling has been used at the data atomic level to render the visualization even before the full dataset is retrieved [71]. These techniques have their limitations, specifically in that they assume strong linearity in the data and a parametric distribution. When non-linear and non-parametric data is processed these techniques can hide potentially consequential details.

MotionExplorer (Figure 14) [67] is an example of an exploratory and analytical tool which was developed for assisting in domain experts in mapping and associating human movement. This platform presents the domain expert with interactive dendrogram visualizations to exploit hierarchical associations in the data, as well as visually search for related sequences. A significant aspect of this application is its ability to aggregate data by their temporal relevance. Event Visualizer [178], provides a user interfaces and tools for visual analysis against real-time or retrospective data streams. Multiple timelines are provided for integration and analysis of different streams of data. The authors have also provided the ability to semantically zoom in on regions of interest, as well as the ability to rank and filter through associated events. The interface can potentially overwhelm users with multiple runs of timelines, presenting usability concerns.

Much of visual analytic techniques for real-time deployment are developed by the time-series community, which the goal is to represent wavelets and univariant data. Techniques that have been used for exploratory time-series data are TimeSearcher, TimeSeer, RankExplorer, and ChronoLenses. TimeSearcher [179] also allows for exploration and pattern recognition in the static time-series data using an early visual analytic object called 'timebox'. The disadvantage of both these systems is that they offer poor graphical representations for communicating information. While ideal for wavelet monitoring, it lacks aggregated views to present summaries or predictive trends.

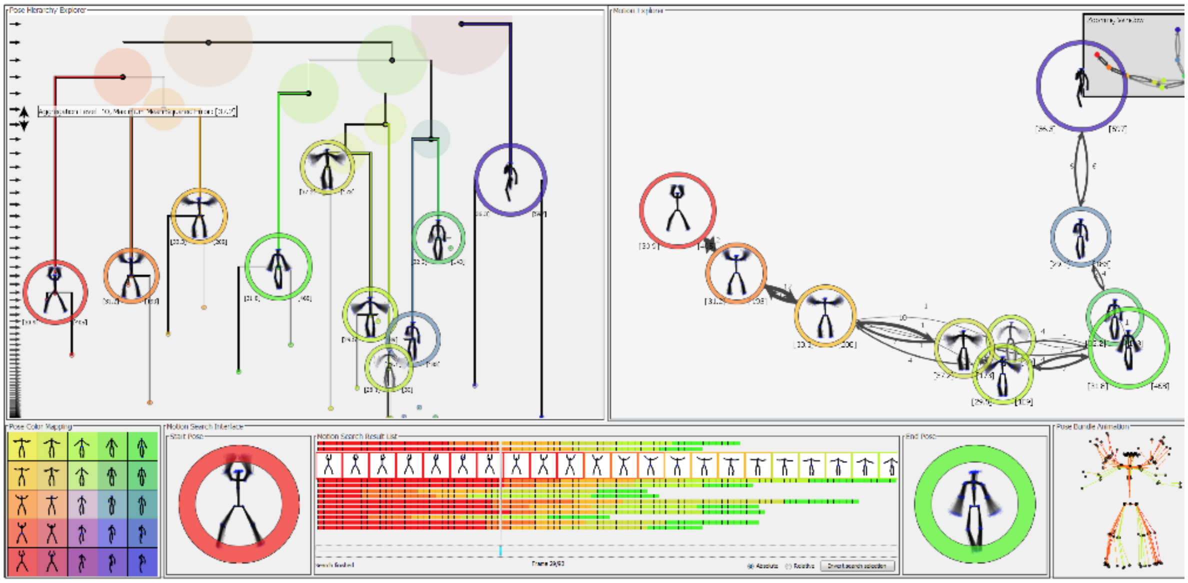


Figure 14: MotionExplorer [67], visual analytic tool for human motion

ChronoLenses (Figure 15) [180], allows the user to discover interesting regions in the univariate time-series data using the concept of lenses. The lenses allow the user to perform analytical tasks in sequence. TimeSeer [181] allows for automated modeling of high-dimensional temporal data, using a series of metrics called scagnostics to model the data. The display then represents estimated scagnostic values using a combination of line charts and scatterplot matrix.

RankExplorer [182] illustrated in Figure 16 is a visual analysis technique that combines the layered view of ThemeRiver, color bar, and glyph. RankExplorer introduces a novel method to identify changing ranks among multiple time-series data streams. Using traditional rank switching by controlling vertical order leads to visual clutter if the data stream contains dynamic and periodic systems which compete in rank frequently over a given time.

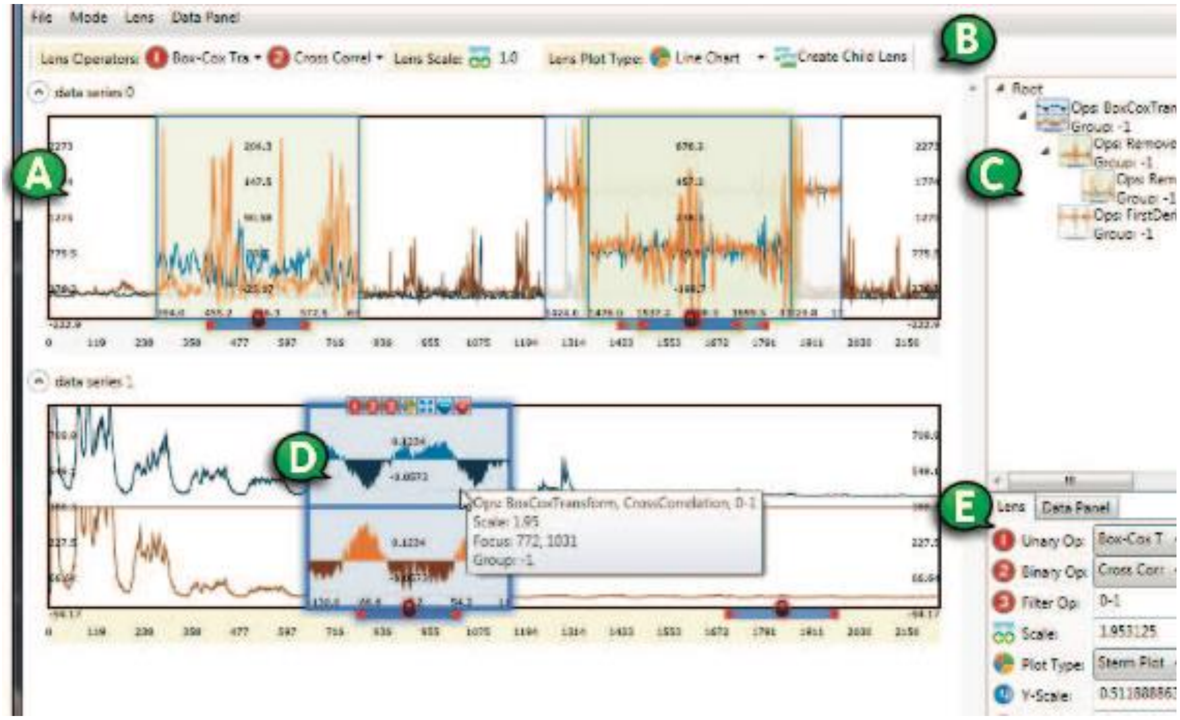


Figure 15: ChronoLenses [180], a visual analytic tool for signals.

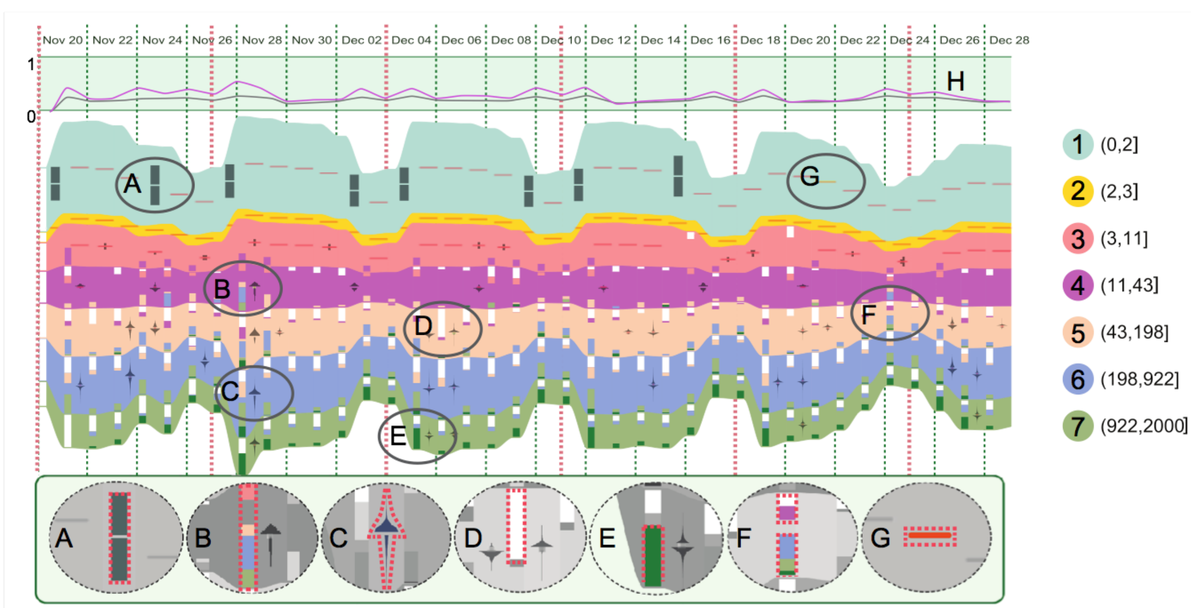


Figure 16: RankExplorer visualizes emerging ranks in data streams [182].

RankExplorer allows the user to appreciate rank changes within a data stream, and also across multiple streams. However this system has a limit of ten categories per view, and contributes to visual clutter when presented several categories of highly turbulent data streams.

Domain specific temporal visualizations have been shown to perform more advance functions in communicating anomalies to the end user. The VisAlert system [183], for example, provides situational awareness for network security analysts. It presents real-time information about the level at which an alert occurred, how long its duration was, as well as the alerts type. Since this system was developed as a means to communicate network attacks, there are limited analysis functionalities. Another system in the same domain is LiveRAC [16], this system supports additional exploratory features such as semantic zoom to search through the data set, and allows for side-by-side comparisons between different clusters. However, this system works on static data sets, and presents a complicated user interface with a greater amount of visual clutter.

Visual analysis techniques for explorations have also been extended to terabyte data. For instance, in visualizing particle-based simulation techniques such as smoothed particle hydrodynamics, an image covering continuous fields of particles needs to be reconstructed from data captured from all discrete particles. To accomplish this Reichl et al. [184], use octree grid and volume ray-casting to render each particle. The result is a considerably faster rendering time at high quality with a modest increase in memory compared to the raw set. Since the rendering is performed a priori, this technique does not support real-time interactions. Novel techniques for scientific temporal visual analytics recently introduced attempt to address issues of scalability and over-plotting, i.e. the issue of unreadable visualization.

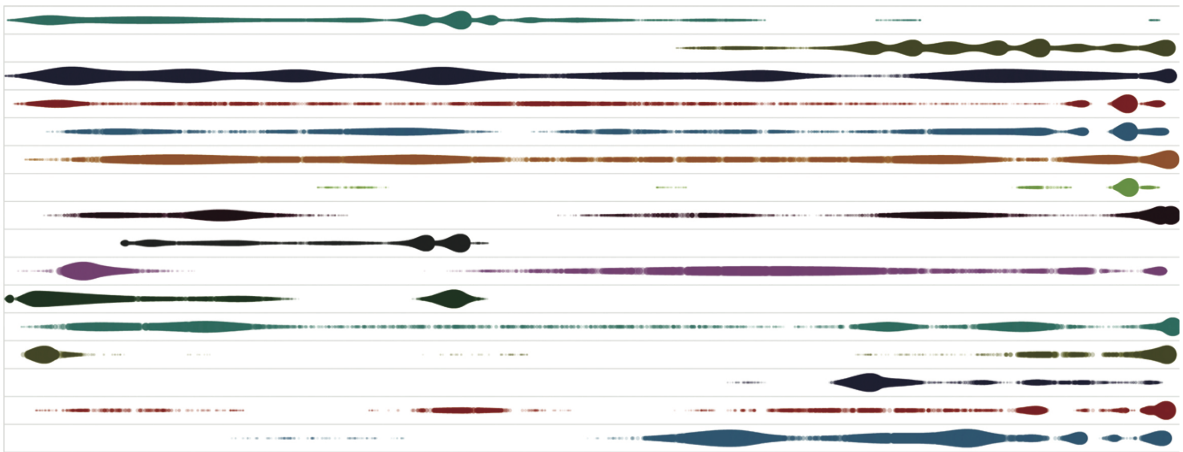


Figure 17: CloudLines [185], a tool for visualizing temporal flows.

Some prior work has been conducted to minimize *over plotting* errors using kernel density estimation (KDE). CloudLines [185] illustrated in Figure 17, introduces an incremental event visual analytic tool using KDE to amplify signals from highly dense areas and minimize low density areas. Lampe and Hauser, 2011 [186] extend the normal KDE method by introducing a weight factor to incorporate absolute densities, thereby allowing for single but significant events to gain prominence in the distribution.

Very limited prior work exists with considerations for workflows and interactions that are discontinuous and irregular [187]. The focus on continuity involves not only novel interactions to aid continuity, but also supporting basic actions such as initiating, pausing and resuming processes through visual cues and metaphors [158]. Some aspect of this requirement can be seen in collaborative interaction techniques. Isenberg and Fisher [188] introduce features that highlight actions performed by collaborators in the area of document analysis. Features such as, identifying who has read the document, time and frequency of its access allow a user to access rich information to support workflow continuity.

3.5.3 Visual Analytic and Sensemaking

Stasko et al. introduce Jigsaw [189] illustrated in Figure 18, as one of the earliest visual analytic tools to assist in exploration and sense-making. The theory of sense-making is a continuous self-motivated process that occurs when a user attempts to understand connections in order to anticipate their trajectory and act effectively [190].

Jigsaw allows for several interconnected data elements to be viewed interactively and automatically, providing the user with multiple perspectives on the connections between each node [189]. Jigsaw has been demonstrated as a useful tool for text and document analysis, allowing entities which appear in multiple documents to be tracked. There are some areas which are not supported by Jigsaw, including the ability to retrieve information in real-time, support for unpredictable data, support for multiple clusters of heterogeneous data, and finally complex inter-related semantics must be user-imposed.

When presented with sense-making in the scientific domain, there are several other dimensions which must be considered. One of the prominent challenges is to show evolution of scientific data or domain [191]. Chen et al. [191] present an approach to addressing the challenge through integration and streamlines of techniques such as spectral clustering and feature selection algorithms. The benefit of streamlining allows semantic strengths held

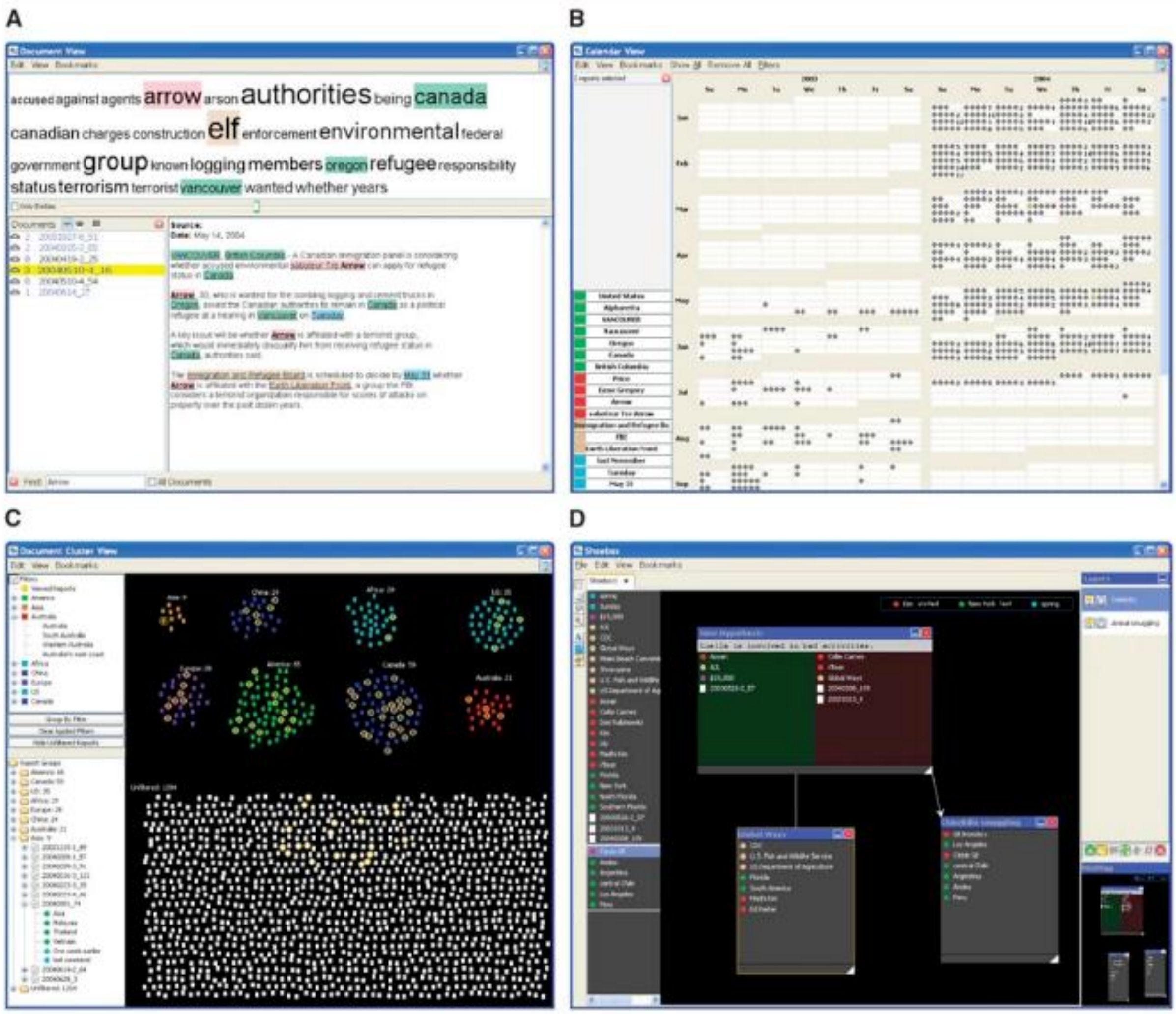


Figure 18: JigSaw [189], a document analysis tool.

between nodes to be reflected in their spatial arrangement, thereby, allowing the analyst to rapidly make sense of current and potential states. Underlying principles which support this theory is the Gestalt principle [192], wherein the author notes the importance of proximity between two objects as how humans perceive relationships and groups. However this system is also limited to static data and support for interactions is limited to the graph.

The role of sense-making [190] and Gestalt's principles [192] have provided a roadmap for past and present interaction techniques to address. However, there are several aspects of both principles which have yet to appear in one concerted visual analytics tool. The aforementioned papers provide a glimpse of major tools which integrate some requirements, but omit some critical aspects. For instance, in sense-making temporal continuity is important; yet most of the techniques only provide a snapshot of past events.

Gestalt's principle reinforces spatial arrangements as relative size as important visual cues for the human, yet one of the challenges of large graph visualization is the dilution of comparatively or complicated clutter introduced by over-plotting of nodes. To expect that a single visual analytic tool would address all challenges may be naïve, both principles metamorphoses in elaborate ways across distinct problem domains. In acknowledgement, many tools have been developed for specialized niches, for instance, LiveRAC [16], Jigsaw [189], and MotionExplorer [67] are three tools designed for computer network security, intelligence and text analysts, and human motion analysts respectively. A domain which has yet to be sufficiently exploited by the visual analytics community is critical care medicine.

3.6 Dynamic Visual Analytics

Real-time visual analytics is an ongoing research challenge [193]–[196]. Keim et al. [197] state one of the prominent challenges in visual analytics is data streams. Thomas et al. [15] also identify data streams as a difficult problem, especially in the area of data integration, data modeling, and supporting situational awareness. In typical data stream environments, the user faces an additional task of consuming large volumes of information to achieve situational awareness [11], by supporting basic functions such as monitoring and distributed knowledge sharing in complex real-time environments [198]. Rohrdantz et al. [195] identifies several open research challenges in this space. Much of these challenges can be separated into data stream management and building dynamic visualizations to support them. To that end Mannsman et al. [7] introduces a dynamic visual analytic pipeline (Figure 19) with dynamic data models and interactive visualizations.

The dynamic visual analytic pipeline [7], illustrated in Figure 19, identifies real-time users as requiring unique consumption and exploration tools that augment traditional exploratory visual analytic methods. That paper serves as an impetus for further research in the area of dynamic visual analytics to address unique requirements as exposed by real-time applications.

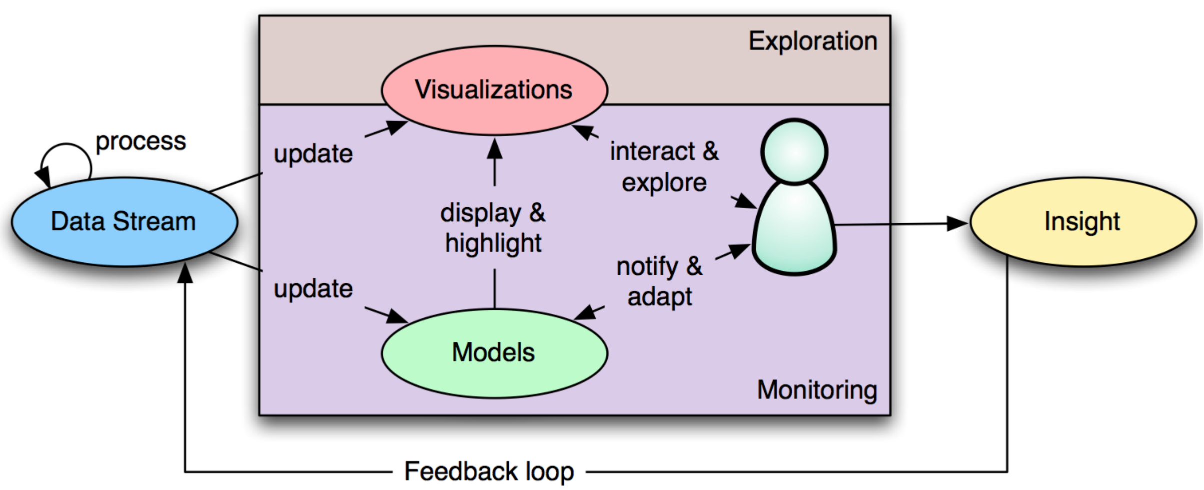


Figure 19: The Dynamic Visual Analytic Pipeline [7]

Further research in the area of dynamic visual analysis is necessary on account of real-time environments being frequently exposed to rapidly changing situations that directly impact the workflow and dynamics of the analyst. To approach the dynamic visual analytic challenge, several key works have produced novel methods to mitigating one or both of the underlying challenges [199]. The first of these involve approaches to supporting temporal data modeling and preparation for visualization.

imMens (Figure 20) [200] introduces a parallelized data management method to retrieve large volumes of data for the purpose of real-time visualization. The visual tool itself uses a combination of numerical, ordinal, temporal and geographic binning to achieve rapid data transmission.

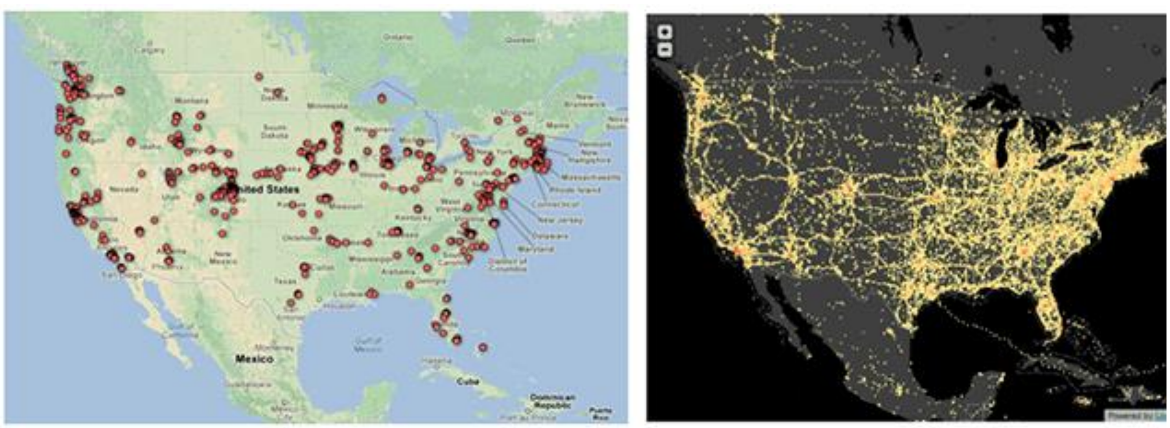


Figure 20: imMens [200] a real-time visual query tool

However, the disadvantage of this technique is that it does not support multi-dimensional alignment between distributed data sources, due to the pre-cubing that is performed at the underlying level. This presents challenges when utilized in unpredictable real-time environments. To demonstrate the underlying model, the authors show rapid data visualization of multiple graphs, such as heatmap and histogram to represent both cluster density and temporal change.

The second set of approaches resulted in the development of novel visual analytic user-facing systems that provide rich interactive mediums for users to engage with real-time data. The most prevalent examples of real-time dynamic visual analytics are found in the text event stream community. One example of this include News Stream Monitoring (Figure 21) [201] presented by Krstajić et al. In this system, news is automatically captured and visualized in real-time using a similarity algorithm.

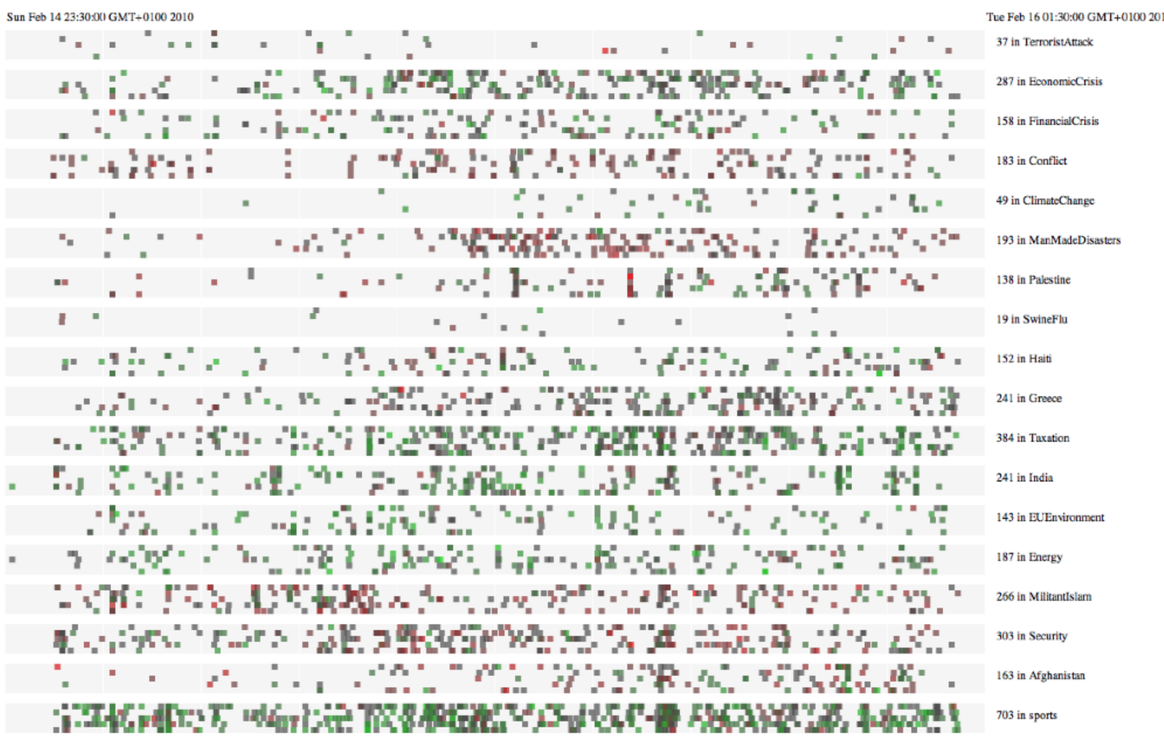


Figure 21: Real-Time News Stream Analysis [201]

However in its first incarnation, this system does not support exploratory interaction. The authors build on this principle to develop Story Tracker (Figure 22) [175] which supports rich exploration of topics as discovered in the event stream. The authors employ text mining principles to extract weakly or strongly related topics and either merge or split them over time. This creates a flow of topics that is visualized and can be interactively filtered by their duration and connectivity to explore the topics evolution.

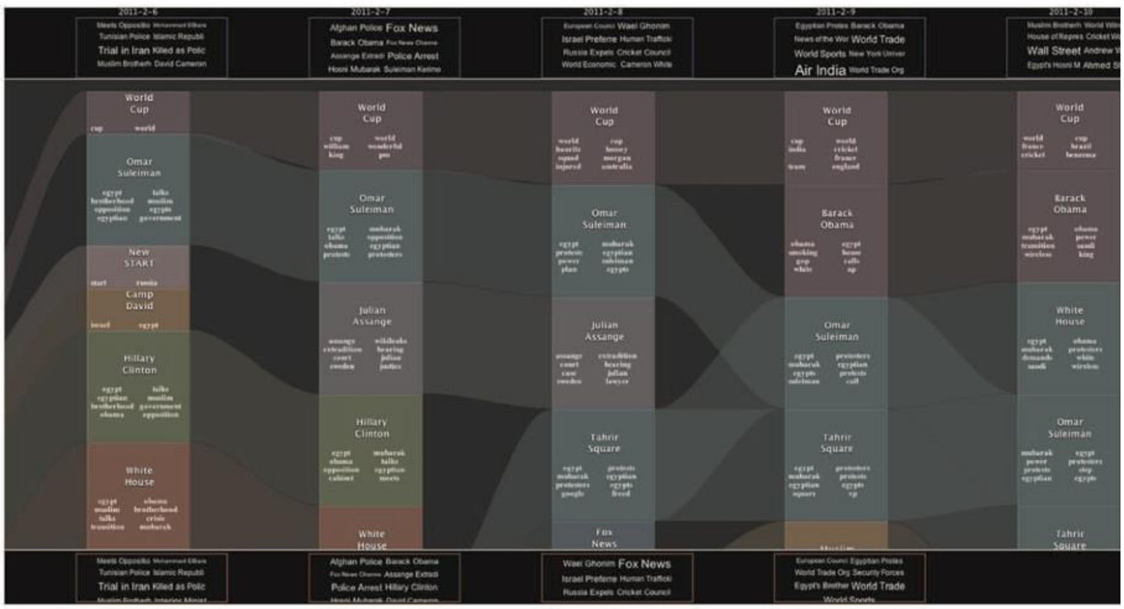


Figure 22: Story Tracker [175] a visual analytic tool for exploring topics in news.

Apart from text event streams, the majority of applications in real-time environments are reliant on temporal sensor data streams. Temporal characteristics are often expressed and represented using a combination of symbolic, waveform, or symbolic-waveform visualization methods [83]. There are a wide number of examples of non-interactive and interactive waveform visualizations. Non-interactive representations such as Density Maps [202], display data streams through controlling the colour of pixels on a display and sliding the window as new data is populated.

VizTree, an example of an interactive waveform representation (Figure 23) [203] allows the user to scan through large chunks of time-series data to identify patterns over the global or local set. It supports functions such as motif discovery, anomaly detection, and query by content. The time-series data is also visualized using a pattern frequency tree, enabling rapid monitoring of anomalies. VizTree uses a combination of symbolic and waveform displays. Symbolic methods attempt to reduce the volume of data by converting real number values to symbols that can be easily categorized by a visual tool and discerned by the analyst. Examples symbolic approaches include time-series bitmaps [204], and Symbolic Aggregate Approximation (SAX) [205].

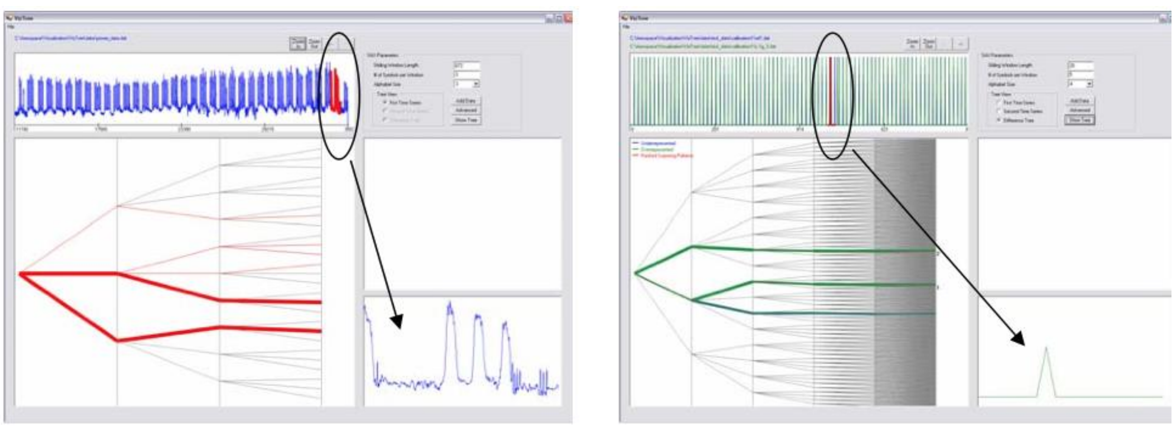


Figure 23: VizTree [307], anomaly detection in time-series.

Gschwandtner et al. introduce a visual analysis tool utilizing heatmap functionalities called CareCruiser [119], illustrated in Figure 24. CareCruiser is a visual analytic tool for exploration of clinical temporal data. CareCruiser allows clinical users to visualize effects of treatment plans on patients' health status. The tool utilizes heatmap like shading to convey metrics such as progress from initial value, distance to intended value, and slope of a parameter. Users can interactively select sections of clinical data to compare the trajectory over a time window. However a limitation identified by the authors was the lack of relative alignment of events to the grid, as is commonly performed in the clinical environment [119].

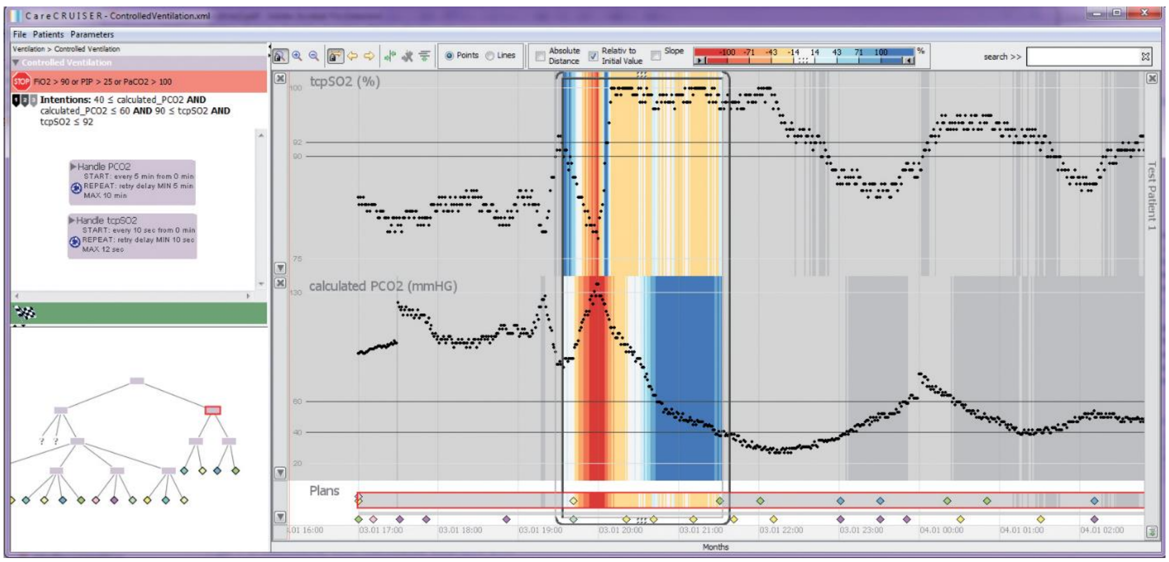


Figure 24: CareCruiser [119], a tool for visual analysis of clinical temporal data.

LiveRAC (Figure 25) demonstrates real-time visual analytic system with a matrix of re-orderable charts that provides necessary functionality to explore, and perform correlation between high density multidimensional time-series data. Yet a disadvantage is the depreciated aesthetic due to visual clutter, which may add to information overload. Moreover, this system does not perform complex operations across multi-dimensional data to produce higher level classifications. In addition, some views of this screen (Figure 25) has the potential for visual clutter when a large number of nodes are monitored.

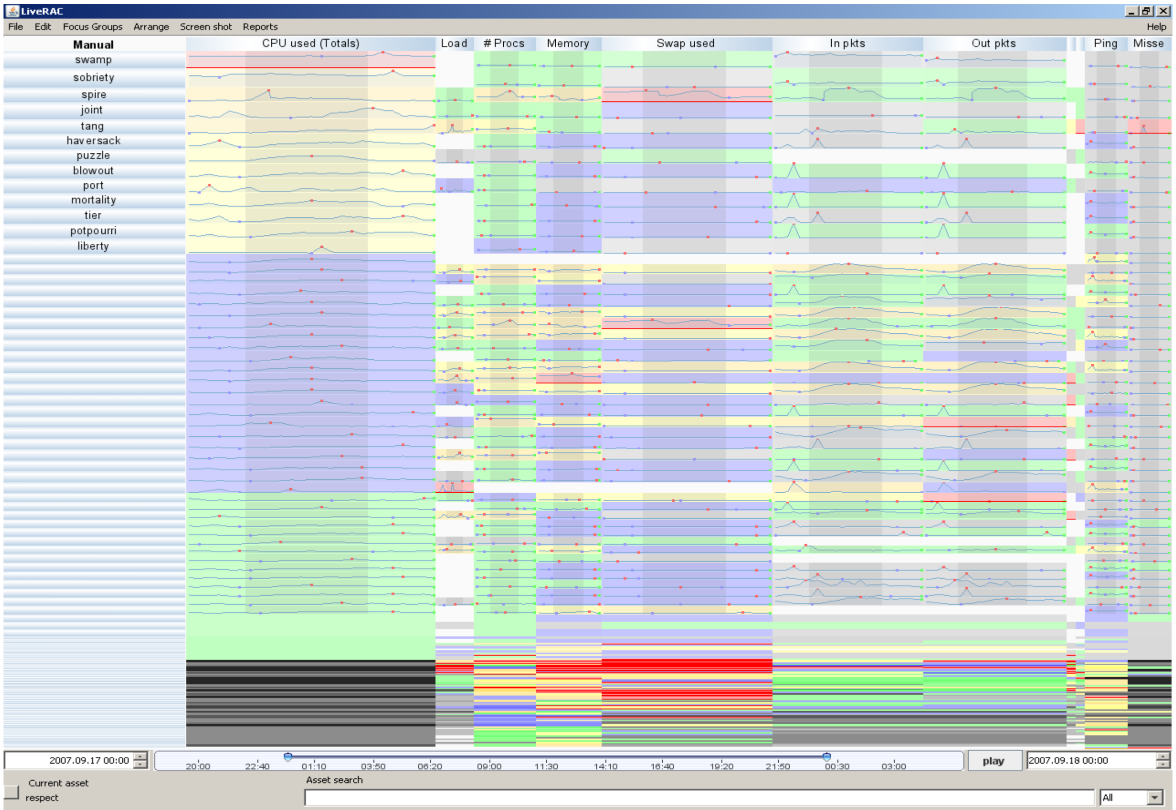


Figure 25: LiveRac Network Visualization [16]

BANKSAFE, illustrated in Figure 26 is a visual analytic tool built to address the VAST 2012 challenge case study of large scale computer-network monitoring [100]. The tool demonstrates several visualization modules, which employs treemaps, coloured matrix map, circular clock glyph and temporal timelines. The authors demonstrate the tool to detect network stress from intrusions, infections, and network stress. However, due to the nature of the hierarchical method of visualizing network data, BANKSAFE has limitations when the user attempts to identify meaningful data from individual hosts. Temporal trends are demonstrated at high level using the coloured matrix map, yet, this higher level aggregation of data may inadequate for domains where lower level event parameters may have prominence and consequence.

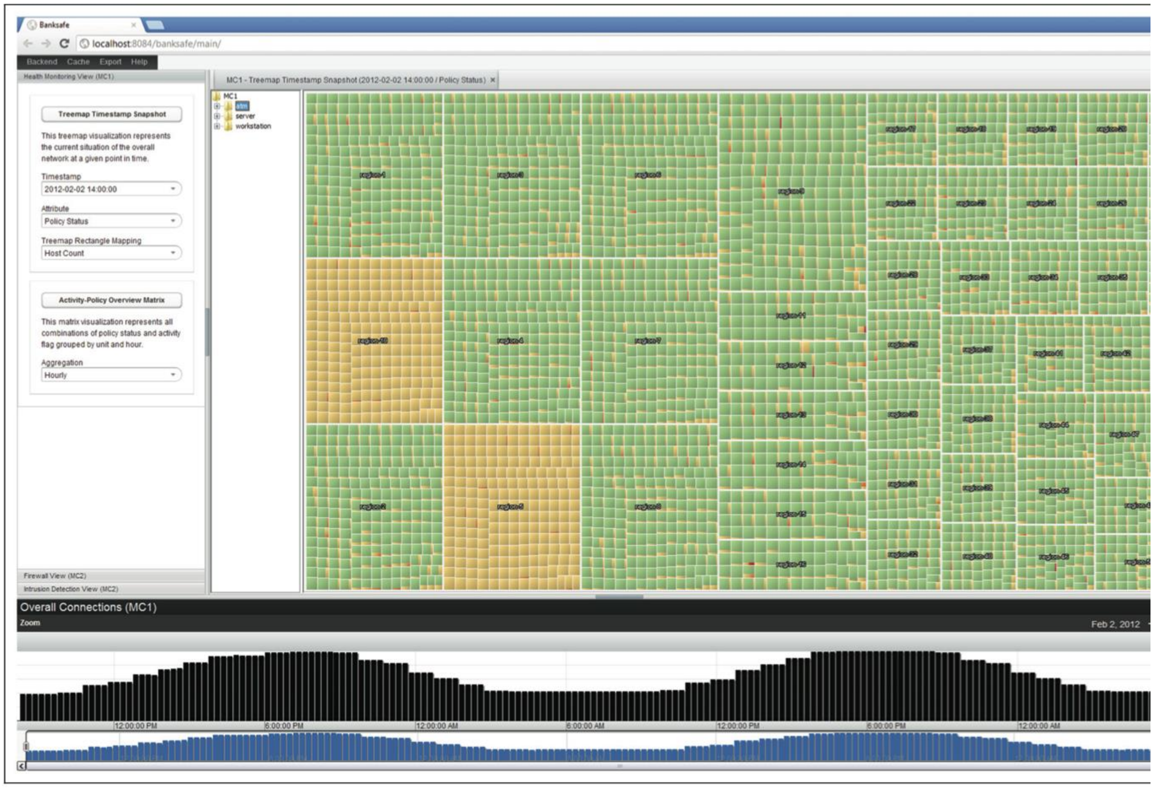


Figure 26: BANKSAFE: visual analytic tool for network monitoring [100].

While these methods have produced novel representations of temporal data streams, these representations do not represent temporal relationships and influence between and across multi-dimensional data streams. These methods also do not present an ability to analyse distinct data across multi-dimensional data streams. The ability to conduct relative temporal alignment to perform advanced multi-dimensional associations is also not supported.

Several aspects of dynamic visual analytics remain open research areas [156]. Among the open areas that require further research include the extension of the pipeline to include network sensors that integrate trends and patterns across different data streams. The requirements are further discussed in chapters 5, where a collection of temporal parameters called the temporal tri-event parameters are introduced.

3.7 Chapter Summary

This chapter started with an overview of sensor network generated data streams in §3.1. The following sections introduced key literature in the space of data warehouse architecture (§3.2), specifically highlighting modular information system designs that support analytics, it was noted that support for wide-complex and deep-complex data are still within the traditional data warehouse architecture. The Artemis platform, an online event analytics system utilized in this thesis is introduced in §3.3.3.

The subsequent sections introduced seminal work in information visualization (§3.4), health visualizations (§3.4.2), and visual analytics (§3.5). The final section introduced dynamic visual analytics (§3.6), and highlighted open areas for research in that space. Subsequent sections will introduce concepts meant to address some of these open areas, including the temporal tri-event parameters, and Exploration-Consumption continuum both of which are detailed in §5.2. The next section presents results from a systematic review of physiologic data visualizations. The systematic review provides important motivation for the works contributed through the instantiation of the TDVA framework.

4. A Review of Visual Representations of Physiologic Data

This chapter presents material from a publication currently under review [41]. The publication was co-authored with Carolyn McGregor.

4.1 Introduction

Two less formal reviews and one systematic review were published in the last decade observing positive impact of visual representations in the critical care setting. Sanderson et al. provides a forward looking analysis relating to representation of physiological data in anesthesiology [206]. Drews and Westenskow., reviews several graphical displays that facilitates rapid translation of physiological event knowledge for anesthesiologists [207]. Finally, a systematic review was published in 2007 by Görges and Staggers which surveys general physiologic data displays, with greater emphasis on surgical and anesthesiology specialities [103].

While those reviews provide important knowledge about the state of the art in physiologic data, they present only partial aggregation of results, and limited knowledge that could be used to enhance health visualizations. Furthermore, key elements such as the nature of visual encodings, support for interactivity, and particular visual elements of these representations are also not discussed. Moreover, all reviews focused on short-term patient monitoring. Visualizations supporting longitudinal monitoring and interactive visual analytics of physiological data were not sufficiently addressed.

There is also a need to better understand how temporal tri-event parameters are expressed by existing health visualizations. The support for these tri-event parameters gauges the ability of the visual representation to convey historical information, summarize duration of events, or present novel visualization techniques to illustrate trends. Temporal events while frequently synthesized from physiologic data, are rarely identified in existing reviews of physiologic displays.

To that end, this chapter presents an analysis of a broad spectrum of physiological graphical displays utilized at the bed-side, in the surgical ward, and for clinical research. We investigate the target problems, design approach, and implications of these representations and evaluate their support of tri-event temporal parameters, namely the display of trajectory, frequency and duration of events. Thereafter, a discussion of results gathered from the review is presented, including aspects which were promising and those that are still inconclusive.

The chapter is divided as follows: the second section presents methodology and reproducible search techniques, section three present results and provides comprehensive matrices summarizing survey results, section four presents discussion, and section five ends with conclusions.

4.2 Methods

4.2.1 Paper identification

We have adhered to the Cochrane Methodology for Systematic Reviews [208]. The identification of studies involved a pilot utilizing a single and subsequent search that used six prominent sources, including, one formal review [103], two informal reviews of physiological monitoring techniques [206], [207] and online search of several databases including IEEE Explore, ACM Digital Library, MEDLINE, EMBASE, ISI Web of Science, and Google Scholar. The first phase was a pilot of generic keywords applied to Google Scholar, which exposed great variability in self-identified keywords. From the pilot, 25 papers were found to be relevant and important citations from those papers were further evaluated to add to our list of important keywords. The pilot search was limited to the last 15 years and used a combination of keywords that were known to the author, such as, “(physiologic* OR clinical OR hemodynamic) AND (visual* OR graphic*) AND (interface OR display)”.

Following the pilot, a systematic review was initiated to ensure catchment of as many relevant literature as possible across clinical, engineering, and computing domains. Novel visual representations of physiological data can be highly diverse, yet still difficult to isolate as was evident from the pilot search. This was due to many novel displays being packaged as part of a clinical decision support system, or larger hospital system. Therefore to broaden the search to include as many displays as were possible, index terms were used to filter articles and included, “Data Display*”, “Diagnosis, Computer-Assisted”, “Monitoring, Physiologic/methods*”, “*Computer Graphics”, “user-computer interface”, “data display”, “interview* or discussion* or questionnaire* or “focus group*” or qualitative or ethnograph* or fieldwork or “field work” or “key informant”, “task performance and analysis”, “graphic* adj2 display*”.

Thereafter, a rigorous inclusion criteria was used that classified visualizations across three groups. The groups were (1) tabular displays, (2) waveform displays, (3) object displays and ecological displays. Inclusion criteria relating to outcome measures are divided into three sets of measures. They include, temporal and duration, human and qualitative factors, and quantitative measures.

We placed a restriction in years from January, 1, 1983 to August 1, 2014 limited to human studies in critical care, surgery, and anesthesia. We included snowballing of references and manual searching on Google Scholar and PubMed. This resulted in a total of 4,330 titles generated for review. Titles were classified as relevant or not using a rigorous inclusion criteria (Table 1). 620 titles were then designated for abstract review. Following that, 112 abstracts were selected for full review, and 38 papers were selected for inclusion in the analysis.

Table 1: Inclusion Criteria

Types of studies:

Randomized controlled trials, cohort, case-control, and design studies.

The review placed increasing preference for randomized control trials, followed by cohort, case-control, and finally design studies. Design studies are popular in the visualization community and were included to study results pertaining to user-evaluations.

Types of participants:

Critical care nurses and physicians.

Several studies have only tested interventions on physicians and excluded nurses, while other studies have used naive participants usually by recruiting undergraduates.

Types of interventions:

Novel knowledge representations, numeric, waveform or metaphor-based displays.

We focus on the intervention in which physiological display is not represented exclusively in waveform and/or static numerical forms.

Physiological parameters tested:

Central venous pressure (mm Hg)	Mean arterial blood pressure (mm Hg)
Mean left arterial pressure (mm Hg)	Pulmonary vascular resistance
Systemic vascular resistance	Cardiac output (mL/min)
ST segment depression of the ECG (mm)	Stroke volume (mL)
Arterial oxygen saturation (%)	Peripheral oxygen saturation (%)
Heart rate (bpm)	Respiratory Rate (rpm)

Respiratory Wave (impedance)	Pulse rate
End-tidal CO2	Mean pulmonary artery pressure (mm Hg)

Types of outcome measures*

Temporal metrics	Human-factors	Clinical relevance
Time to detection of adverse event (s)	NASA-TLX task load index score	Accuracy of diagnoses
Time to diagnose event (s)	Satisfaction of intervention (Likert scales)	Accuracy of treatment
Time to initiate treatment (s)	Number of participants	
	Clinical expertise of participants	
	Setting in which the trials were conducted	
	Noise level of the environment	
	Age of the participants	
	Caffeine intake	

Following the creation of the inclusion criteria, an online questionnaire was built using Google forms and used to evaluate all studies. The questionnaire consisted of six sections that were all identified as potential areas of interest for analysis. 74 questions were screened for each full paper reviewed by a single reviewer. Where value is significant, the data is charted. Questions to be included in the questionnaires were selected from themes identified in the pilot study. In particular, questions were generated to elicit detail about the study, design, and results from any human experiment or evaluations.

The details of the questionnaire are as follows:

1. **Study Background**: Questions 1 – 7 included information about author, year of publication, setting in which the study took place, the number of samples, and the provisioning of the data.
2. **Scenarios**: Questions 8 – 14 prompted for details of the scenarios developed for testing.
3. **Study Design**: Questions 15 – 19 prompted for details of study design. Question 20 – 22 was related to the expertise of the participants.
4. **Data Properties**: Question 23 – 28 were related to the data type and prototype.
5. **Design Properties**: Question 29 – 33 pertained to the prototype design.
    - i. **Functional level of Prototype**: Question 34 – 36 prompted for details of the prototypes function.
    - ii. **High level Findings**: Questions 37 – 42, specific findings were discussed relating to performance.
    - iii. **Situational Awareness**: Questions 43 – 51 elicit prototype's accuracy and detection abilities.
    - iv. **Time**: Questions 52 – 61, if time was collected as a measure, details are extracted here
    - v. **Significance values**: Question 62 – 68, Specific significance values are presented if they are found.
    - vi. **Cognitive Errors**: Question 68 – 73 collects details of workload and their perception.
    - vii. **Task goal and summary**: Question 74 asks for output items and goal of the prototype.

Where appropriate the questions were marked as either not reported if data was missing, or not applicable if the question was a follow-up of a prior conditional question.

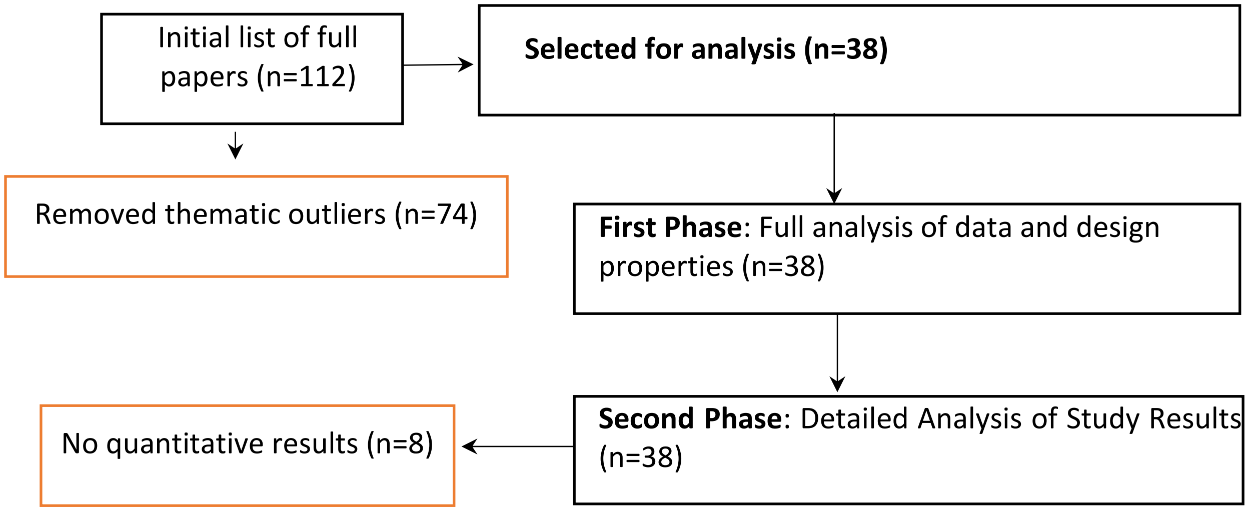


Figure 27: Analysis Methodology

4.2.2 Analysis Methodology

All papers included in the analysis were passed through the questionnaire, and resulted in an initial comprehensive matrix of results. Of 74 questions that were initially probed questions that yielded over 75% not reported, or not applicable across all papers analysed were removed. 39 variables were then selected for inclusion in the initial matrix. The analysis was conducted in two phases as illustrated in Figure 27.

Phase one results are summarized in the Comprehensive Matrix of Design Properties (Appendix 1) and phase two results summarized in the Comprehensive Matrix of Study Results (Appendix 2). The Matrix of Design properties contain 10 variables, in order they are, “Specialty”, “Year”, “Clinical Context”, “Number of Variables”, “Display Type”, “Colour Used”, “Pre-attentive processing”, “Gestalts”, “Interactive Controls”, and “Iterative Design”,. “Specialty” relates to the clinical specialty, such as critical care or anesthesiology, and “Year” refers to the approximate date the prototype was developed and tested, due to the difference between publication and evaluation dates this value was approximated based on the submission date of the article. “Clinical context” refers to the facility provided by the display to integrate clinical information, and “number of variables” refers to the total number of physiological or clinical variables that were visible in a single screen.

“Display type” refers to the four distinct groups of displays such as tabular, object, or metaphoric, and “colour used” identifies the particular use of hue where available. “Pre-attentive processing” identifies particular visual variables that were utilized such as, shape, size, and dimension. “Gestalts” refers to the grouping laws identified by Gestalt’s laws of perception, particularly, the use of proximity, similarity, closure, symmetry, and continuity as a means of discerning visual objects presented in the display [192]. Finally “interactive controls” refers to the ability for the display to support direct manipulation of one or more properties and “iterative design” identifies displays that were build using user-centred design approaches that include users into key decision making processes prior to the development of the display.

4.3 Results

A total of 38 papers were included in a dual phased analysis methodology (Figure 27). A questionnaire of survey questions were assessed for each paper and key thematic questions were chosen for further analysis. This resulted in two initial matrixes, the Comprehensive Matrix of Design Properties (Appendix 1) and Comprehensive Matrix of Study Results (Appendix 2). This section details results from both matrices and highlights key findings from the analysis of each matrix.

4.3.1 Comprehensive Matrix of Design Properties

The goal of the Comprehensive Matrix of Design Properties is to present design decisions that were followed to develop prototypes across all 38 papers analysed. Visual representations were found across mainly anesthesiology (n=17), critical care (n=19), and in some multi-discipline (n=2) environments. Only one display was developed as a tool for intensive nurses [209]. Multi-discipline environments consist of two or more speciality, such as integrated in-patient and out-patient systems. Visual displays started to become actively contributed from the early 1990s, then increasing every 10 years, 2014 – 2005 (n=15), 2004 – 1995 (n=14), 1994 – 1985 (n=8), and earlier than 1984 (n=1). Integrate clinical data was also found across some displays (n=15), while even greater number of displays were devoted to the display of physiological waveforms (n=23). Number of variables presented in a single screen was wide-ranging, most displays contained greater than 20 variables per screen (n=15), followed by 11 – 20 variables (n=11), while the remaining displays contained 0 – 4 (n=7), and 11-20 (n=5) variables per screen.

Surveyed visual representations included a mix of multiple display formats, such as tabular displays (TB) with waveform displays (WF) (n=5), waveform with metaphoric display (MT) (n=4), or a waveform with an object display (OB) (n=16). Metaphoric displays were most popular (n=19), followed by waveform displays (n=18), and object displays (n=9). One paper did not present any illustrations of the display or discuss the format of the graphical interface in their system [210]. Visual representations utilized at least two of the primary colours, red, blue or green (n=21), while colours yellow (n=11) and, teal (n=4) were also popular options. A number of papers did not specify the type of colour that was used (n=10). Pre-attentive processing of items were popularly exploited through manipulating visual variables such as, colour (n=23) and size (n=11), followed by dimension (n=6), and shape (n=5).

Visual representations also exploited some aspect of Gestalt's law of groupings, such as continuity (n=17) with waveform displays, closure (n=16) when identifying boundaries, symmetry (n=13) with visual metaphors and object-based displays, and proximity (n=6) to aid in higher level detection of abnormal events. The most popular interaction method that was supported was selection (n=13). Selection allows the user to select visual objects directly to reveal greater details. This was followed by interactive filtering (n=7) to select partial ranges such as short durations of time. Finally, in many cases designs were proposed without following user-centred design approaches (n=28). Seven papers reported using user-centred design processes, while three papers described a structured approach used in developing the proposed visual design [123], [211], [212]. The next subsection presents results from studies conducted using the proposed visual representations.

4.3.2 Comprehensive Matrix of Study Design

The Comprehensive Matrix of Study Design (Appendix 1) presents results that were reported by authors for any evaluation or experiment. While the search strategy yielded 38 full papers that were identified for analysis, only 28 of these papers contained primary study results from a case study, evaluation, or human experiment. Furthermore, those studies employed one of naïve, novice, or expert participants in the evaluation method. Naïve participants were generally undergraduate students with little or no prior clinical knowledge. Novice participants ranged from undergraduate nursing students to newly graduated clinical staff. Expert participants had at least 10 years of experience.

The number of participants exposed to test conditions were highly varied, however a majority of studies employed at least 15 participants. Six studies used a sample size greater than 20 to test for detection, diagnostic and treatment accuracy, with the minimum being 4 and maximum being 32 participants. Most displays integrated these systems in a single display using live or static representations (n=14), while displays that were presented as case studies (in situ) were connected to central monitoring systems. Some displays supported views of clinical information that integrated data from other clinical and laboratory systems (n=15) [213]. Most prototypes that were evaluated used more than one data streams, with exception of the studies that contained low-frequency updates (n=9). Most evaluation or experiment studies utilized more than one condition to test each display, yet some experiment, evaluation, and design studies, did not have any scenarios or patient conditions (n=9). A large number of studies also did not utilize data from more than one patient-source (n=26).

Most of the studies were conducted in laboratory environments (n=22), followed by evaluations or experiments in the intensive care unit (n=11). Some studies were evaluated over multiple specialities (n=2). Majority of studies used some form of experimentation to validate their designs (n=20), followed by case studies with clinicians (n=10), or evaluation using a subset of clinical staffs (n=6). Two studies were design papers without any validation methodology. Of the papers that reported results (n=28), most reported positive findings (n=24). One paper that employed a between-group design, yielded results that were site-dependent and skewed towards the site that produced the visual representation. For evaluations or experiments the source of data to support realism was spread across live simulations (n=18), live patient-origin data (n=11), or static patient-generated data (n=8). Much of the studies did not test for cognitive overload using ad hoc methods or traditional workload score metrics such as the NASA Task Load Index (NASA-TLX) (n=31). Where cognitive workload was reported (n=7), most were reported to have reduced cognitive overload (n=5) while others reported no change or mixed results (n=2).

Long-term historical values, specifically ranges exceeding 5 minutes of monitoring were not included in majority of the displays (n=28). Tri-event parameters, namely, trajectory, frequency, and duration, were seldom supported by visual representations, where these parameters were identified, trajectory was most frequently found (n=26). Temporal trajectory was encoded using either curves (n=23) or glyphs (n=3). In terms of duration, the second tri-event parameter was seen across nine displays, of which, glyphs (n=5) or text (n=4) representations were utilized. Duration, the last tri-event parameter was also seen in some visual representations encoded by glyph (n=5) or text (n=4) where supported. Where displays were validated through experimentation, both the display and scenarios were more often counterbalanced (n=10), while some experiments counterbalanced only the scenario (n=6) and others only the display (n=4). Finally, clinical scenarios were utilized across many studies utilizing experimentation or evaluation methodologies (n=17).

4.4 Discussion

A total of 18 novel visual representations were identified from the analysis of the literature. Novel displays were seen across four main groups, including tabular, waveform (graph-based), object, and metaphors. The latter two are aggregated together as ecologic displays.

4.4.1 Tabular Displays

The early-1990's saw growing interest in converting large-volumes of paper patient charts to 'virtual' records [214], [210], [215]–[217]. Initial representations adopted by these virtual patient records were largely tabular and text-dominant, and sometimes seen to contribute negatively to information overload [217]. Figure 28 presents an example of a traditional virtual patient chart that mimics a traditional paper flow chart. This review identified 14 such representations in papers that were published between 1952 till 1997.

The adoption of visual representations was identified with improvements in the general accessibility and usability of hospital [218]. Intensive care unit systems were seen as the primary speciality to promulgate integrate graphical electrocardiogram tracing [219] and diagnostic images [220] into hospital systems, now incarnated as a larger and more integrated electronic health record (EHR) system. These systems remained dominated by text and incrementally improved user engagement activities using the popular desktop-oriented, Windows-Icon-Mouse-Pointer (WIMP) interaction paradigm. Additional levels of interactions, such as multiple mouse clicks, were required to access disparate health data. Paper copies of data stored in the EHR remain prevalent, in part due to the simplicity and ease of reading paper charts, for instance, many intensive care units continue to utilize paper copies of EHR-based patient summaries [221].

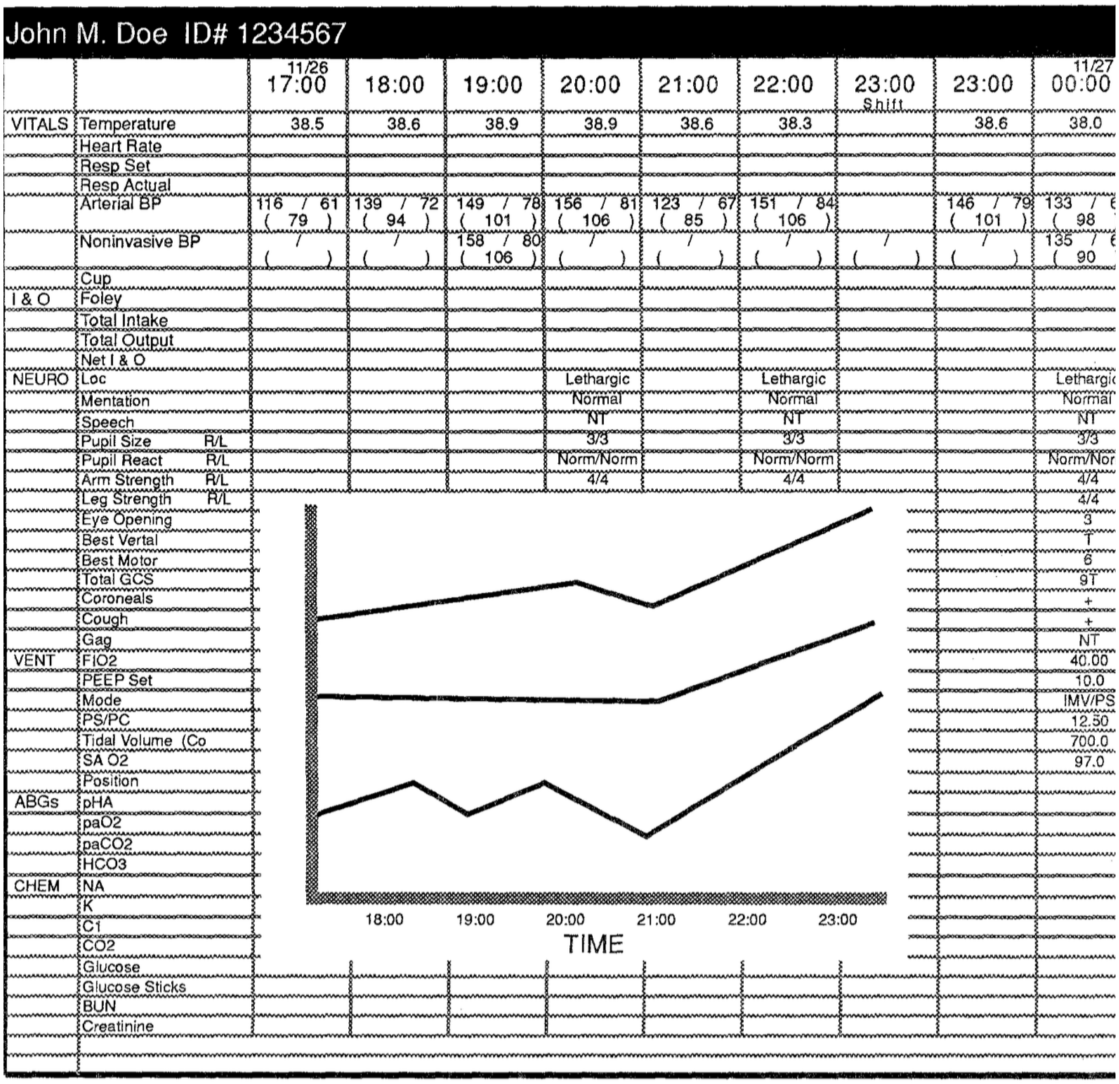


Figure 28: A tabular display that mimics traditional clinical flow-sheets. [308]

4.4.2 Waveform Displays

The intensive care environment utilizes significant waveform data, due to the presence of a plethora of real-time single-sensor-single-indicator (SSSI) monitoring apparatus. To support a limited scope of situational awareness, the SSSI paradigm assign each sensor output to its independent display space. Evidently, allowing for numerous isolated displays to proliferate. The SSSI method of displaying data utilizes wave-form graphs or aggregated numeric, with the former frequently exposed to artefacts and artificial smoothing [222]. The SIMON dashboard illustrated in Figure 29 shows a WIMP interface developed for accessing patient information in the intensive care unit [223]. SIMON, is developed for complex information environment, yet demonstrates significant visual influence from enterprise web systems.

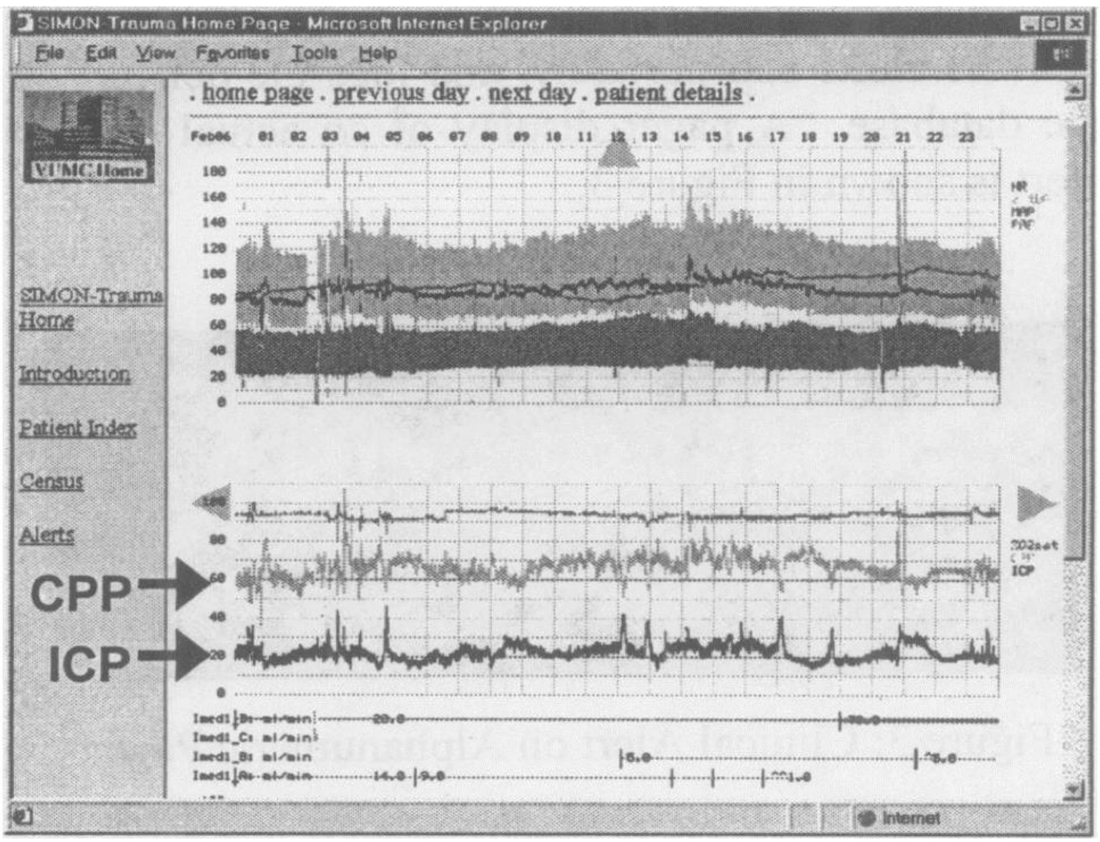


Figure 29: Snapshot of the Simon Clinical Graphical Interface [223].

The review analysis revealed more than 25% of studies utilized some form of live physiologic streams to display largely identical waveform representations. It was also noted that much of these waveform displays were integrated with other tabular and text representations. Three papers that presented waveform displays, also supported interactive capabilities including the ability to select regions of interest, filter based on patients, and generate screen captures [124], [224], [225]. Stylianides et al. present an engine for producing waveform graphics [226], their system however, serves the purpose of animating historic physiologic data streams.

Notwithstanding their ability to communicate acute time-sensitive events [227], waveform representations have numerous limitations [103], [212], [228]. One prime disadvantage of waveform displays is the potential to negatively impact cognitive load, that is, they require humans to monitor and consume large number of data point as they are produced to derive trends and higher level knowledge [229], [124], [129]. These waveforms display can convey several features in one frame, therefore easily disturb limited resources of the working memory capabilities [230]. The challenge of managing large volumes of data have been extensive studied in several domains, such as, information-overload [231], visual data mining [80], and addressing cognitive challenges related to interruptions, task performance, and decision making [11], [231]–[233].

Integrated methods of representing critical physiological information have been actively studied to reduce the internal mental processing requirement [129], [227], [234]–[237]. These integrated displays use a combination of text [238], [239], graphic [28], [103], [207], and waveform [219], [240] representations to summarize low-level information. Two such integrated displays were identified in the review, including Engelman et al. [225] illustrated in Figure 30, as well as Anders and colleague (2011) [124] who present a modern interactive waveform interface that allows clinicians to interactively select regions of interest while monitoring other forms of slow-changing clinical data. Other studies, seeking alternatives to the waveform visual encoding, propose novel and ecological methods to improve knowledge discovery and minimize cognitive overload.

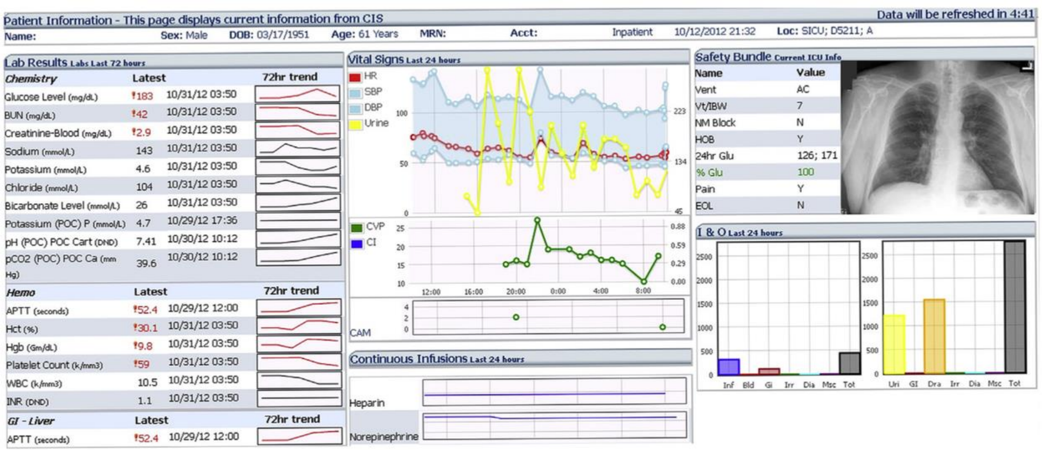


Figure 30: A modern dashboard utilizing waveform displays [225].

4.4.3 Ecological Displays

Ecologic displays attempt to integrate relationships existing across both workflows and semantics [131]. Among the primary goals of ecologic displays is to convey both the means-end relation, answering the particular means of arriving at that state and its ultimate consequence. From our review, two large classes of visual representations were identified that approach these objectives. Object-oriented displays, and metaphoric displays were seen to extend typical limitations found in text, tabular, and waveform displays by introducing novel information, such as spatial and temporal arrangements of closely related information.

Displays that utilize and manipulate 2-dimensional graphical objects, limited to basic shapes, and symmetries to produce emergent properties have been classed as object-oriented displays [206], [128], [241]. These displays follow demonstrated efficacy of graphical displays over traditional numeric displays observed in nuclear power station control stations [242]. Studies have shown a positive relationship with integrated displays and an overall improvement in diagnosis ability as well as a reduction in time to initiate treatment [132].

Blike et al. [243] show that subjects exposed to emergent features using novel graphics recognized a problem more rapidly, but their accuracy had not improved in comparison to the numeric display. Moreover, they show that the shape of the graphic, illustrated in Figure 31, improved detection of etiology compared to the numeric and control displays. While Blike et al. state an improved reaction and fewer errors when using the object-oriented display, the display was found to have been confusing and not ecological to naïve participants. Blike et al. also omit to study cognitive workload and the post-exposure effects on participants.

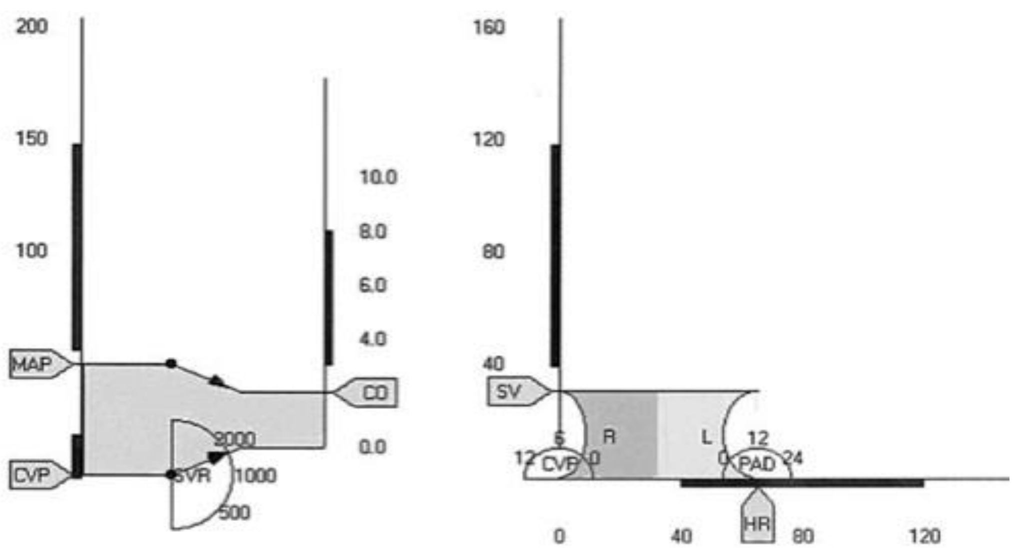


Figure 31: Advanced graphical display for hemodynamic monitoring [128].

Zhang et al. [76] reproduce the designs introduced by Blike et al. [243], and found that anesthesiologists were able to detect simple deviations faster, however no change was seen with detection times of more complex cardiovascular events. Jungk et al. [244] report on two ecological object displays that may have properties influencing the gaze fixation of individual systems, often at the expense of other key compartments.

A redesigned interface, illustrated in Figure 32, however, showed further improvements in the detection of complex events when the display was integrated with supporting alarms and contained improved graphics. Researchers have arrived at similar conclusions [127], [128],

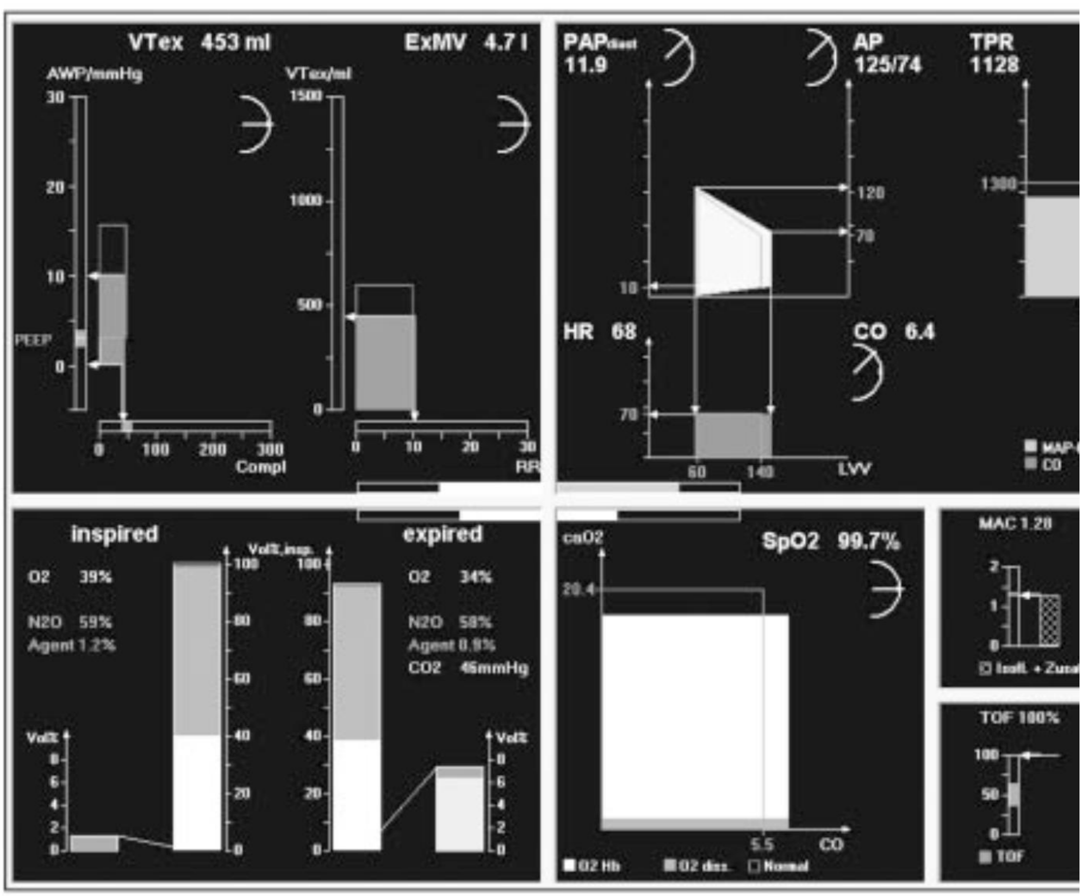


Figure 32: A refined anaesthesia workspace involving ecological design [244]

[132], [211], [245], showing a clear link between detection and reactionary time to the format

and features of the graphical display, exact mechanisms for these links have not been thoroughly studied and remain open investigation areas.

Not all studies have shown convincing evidence, for instance, some have shown negative links when presented object-oriented displays [246], [247]. The etiological potential display illustrated in Figure 33 by Effken et al. attempt to extract specific features of object displays that improve detection and diagnosis [248]. In that study, Effken et al. find no significance in the detection or diagnostic times, even when three abstract displays were tested. Two of these displays, not illustrated in Figure 33, required that features of the full prototype either be reorganized or removed.

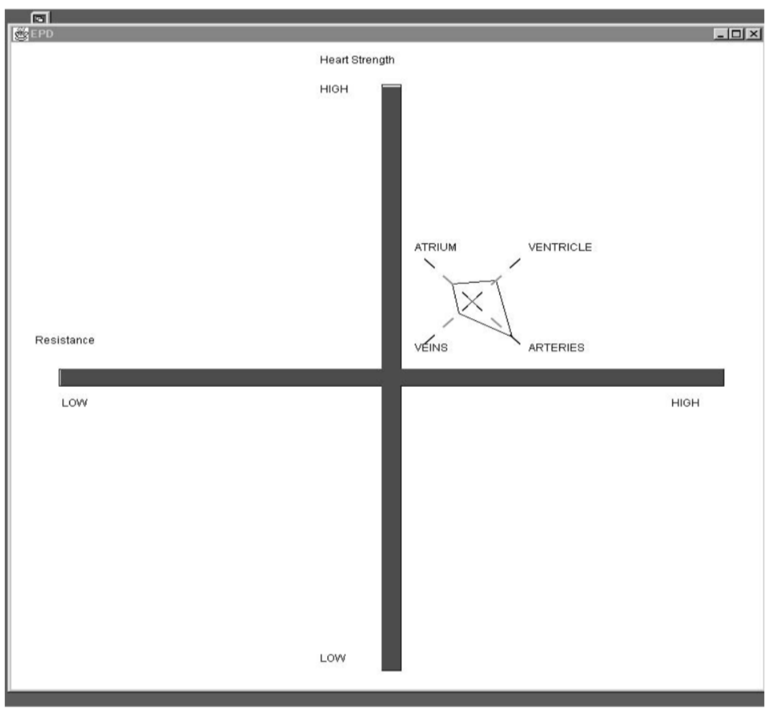


Figure 33: The etiological potential display moves an object across four quadrants of heart strength and resistance. The object in the top right quadrant is distorted to show relative depressions. [246]

Metaphoric displays, conversely, isolate semantic relationships existing in biological systems to invoke a sense of familiarity and associations to changing physiologic states. Visual metaphors have been presented extensively across multiple domains [67], [117], [249]–[251]. These visual metaphors often exploit salient aspects of physical systems to communicate underlying system states. For instance, extensive representation of a particular theme may result in greater areas for particular layers in a ThemeRiver graph [117]. The use of visual metaphors to address clinical problems has resulted in a handful of novel and practical representations. A total of 20 representations, representing over half of all visual representations analysed belonged to the metaphoric display group.

Most clinical metaphoric displays illustrate physiologic data in terms of organ-systems [76], [121], [123], [227]. Five papers presented metaphors that involved dynamic objects that exhibited behaviours similar to organ systems [120], [227], [246], [252], [211]. Cole and Stewart (1994) [252], introduce a visual representation (Figure 34) that consists of two volume rectangles that compress or expand similar to the respiratory system. This design was further improved in the work of Horn et al. [120]. Effken, Kim and Shaw present an integrated balloon display also resembling the respiratory system [246]. Agutter et al. introduced the Graphical Cardiovascular Display (Figure 35) that appends a pipe-like metaphor of the cardiovascular system, including representing high and low volumes across various arteries [211].

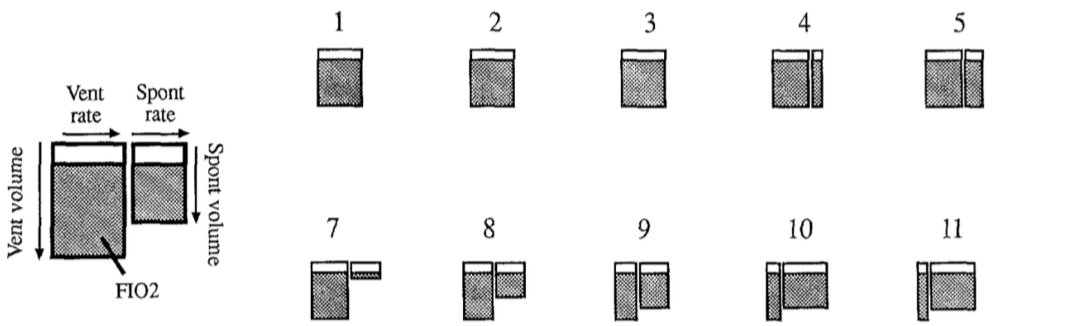


Figure 34: Volume triangles represent multivariate clinical data using a lung-expansion metaphor [308].

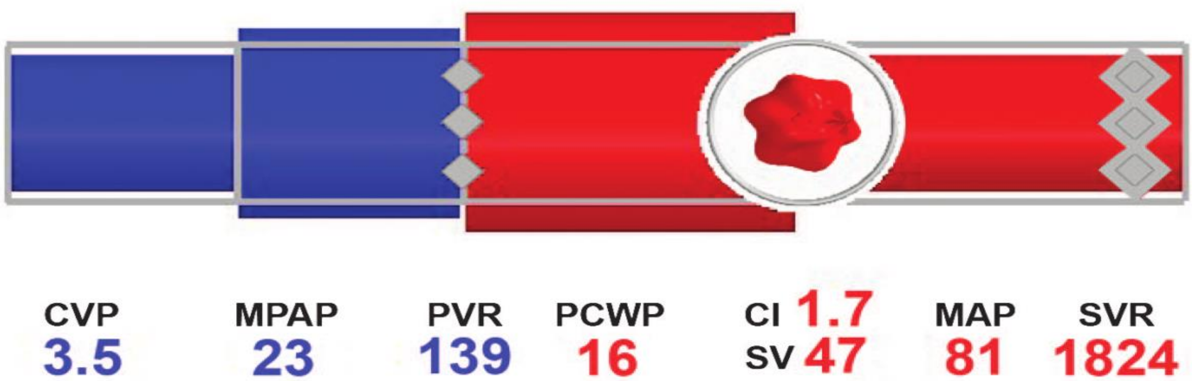


Figure 35: Graphical Cardiovascular Display [211], a metaphor of a pipes with volume and pressure.

Wachter et al. utilize an iterative design approach to develop a pulmonary graphical respiratory display (Figure 36) by exploiting knowledge about the anatomy and physiology of the human respiratory system. [123]. In that display, clinically relevant anatomic representations are used to highlight clinical conditions that are typically expressed in verbal handoff or transcribed notes [29].

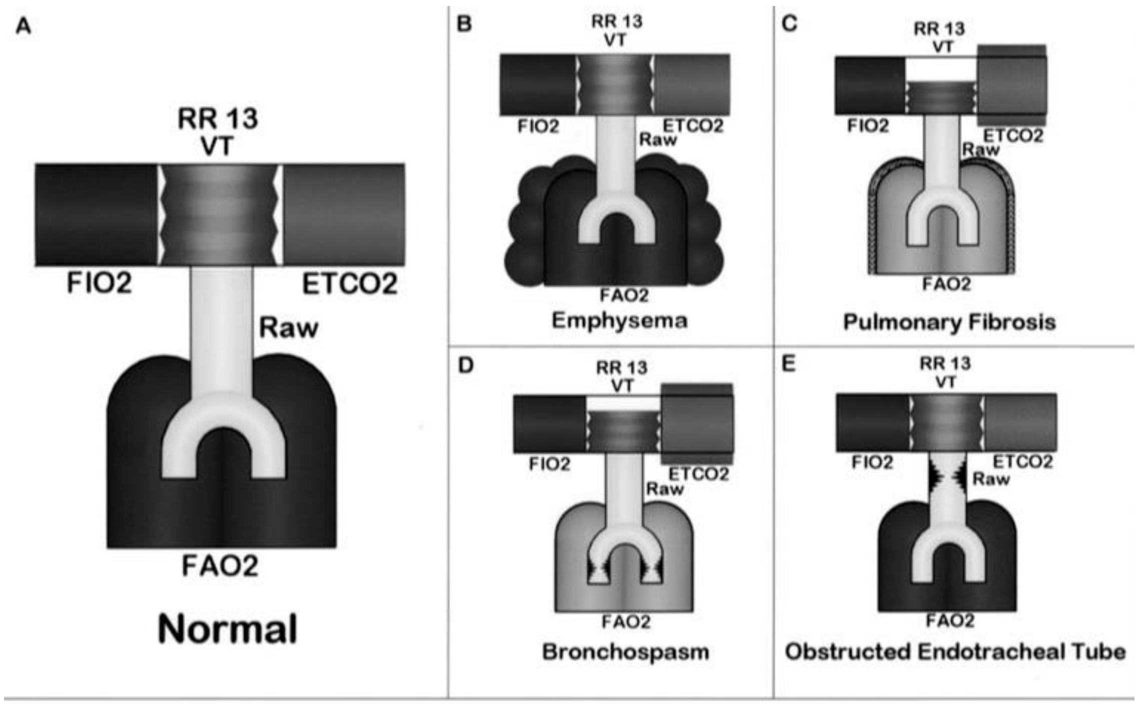


Figure 36: Pulmonary Graphical Display, an integrative visual metaphor for respiratory monitoring [123].

Other forms of visual metaphors however, rely on abstract principles to convey more abstract and general metaphors. Gorges et al., also from the department of anesthesiology at the University of Utah, introduce a series of visual metaphors to communicate visual signs to bed-side clinicians [253]. These displays adopt a clock metaphor illustrated in Figure 37 to convey salient features, such as temporal trends over the past 12-hours. There also exists a syringe metaphor to show information about current medication. Gorges et al. compare this visual metaphor with a traditional monitoring, along with another univariate bar display that plots histograms and line-charts to convey the same 12-h linear trends. They find however that the bar display outperformed both the visual metaphor and the control.

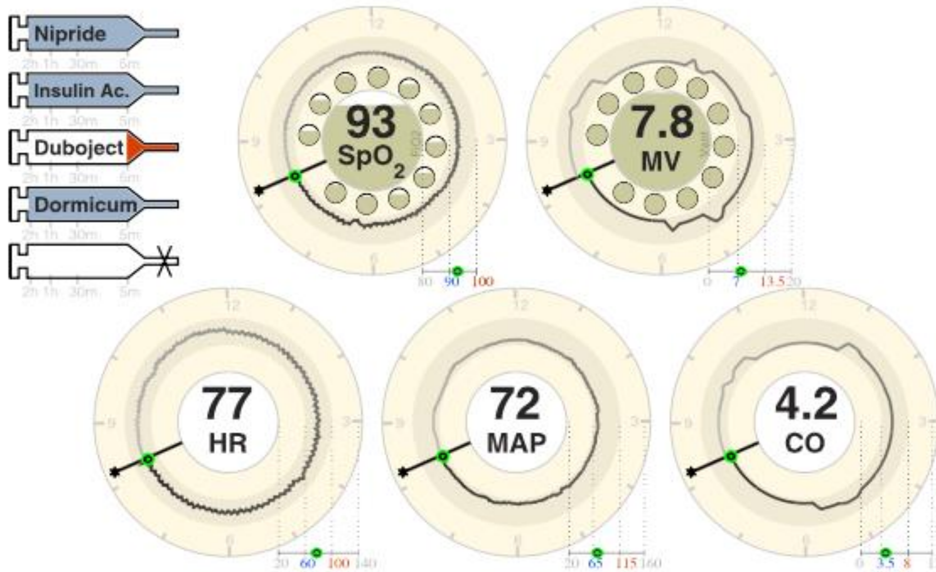
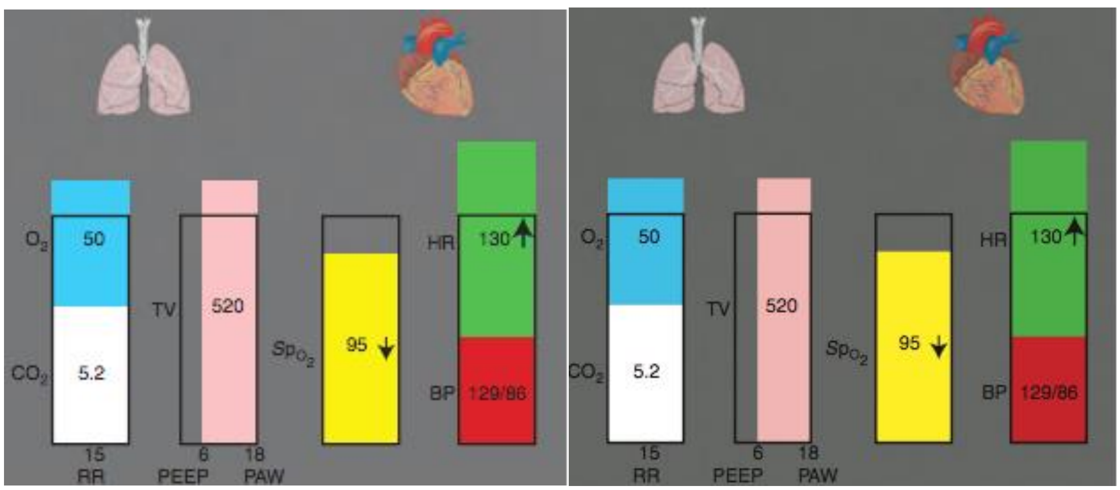


Figure 37: Far-view visual metaphors for triaging vital signs [253].

Charabati et al. from the Montreal General Hospital's depart of anesthesiology introduce a gauge metaphor to highlight normal and abnormal ranges, and conduct an evaluation across two-sites [229]. They find a combination of numeric and visual metaphors achieved the strongest advantage in detection, accuracy and workload. Tappan et al. evaluate visual metaphors by appending visual objects to traditional medical monitors [254]. They report significant improvements in detection of adverse events, with the visual metaphor having a 14.4 second advantage over traditional physiologic monitors. The visual metaphor was also found to reduce the number of missed events. However, these results, like previous studies were conducted in controlled environments.

Not all visual metaphors, however have seen similar success. Zhang et al. [76] introduce an integrated 3-dimensional balloon metaphor, building on the work of Blike et al. [243] with object displays. Zhang et al. find mix results after evaluations, with only 63% of scenarios have shorter detection than scenarios, and situational awareness being improved in one of four scenarios. Moreover, van Amsterdam et al. utilize customization features offered by vendor-based medical monitors to construct and evaluate a metaphoric display presented in Figure 38. The find however, that visual metaphors did not improve detection or accuracy of anesthesiologists [255]. Albeit, the visual metaphors developed by van Amsterdam et al. were modest and elementary, with limited design expressiveness and ecologic considerations.



(a) (b)

Figure 38: (a) Metaphorical anaesthesia interface and (b) Metaphorical interface with trend information (tMAI) [255].

Visual metaphors in medicine have a long history dating back over two decades. However, much of the explored modalities have been restricted to laboratory experiments, and even when utilized in real-world scenarios lasted only for short experimental phases. Studies have shown an overall positive association of visual metaphors for monitoring purposes, additional research is needed to evaluate the same to support analytic functions away from the bed-side. Finally, evaluation methodologies need to consider iterative design paradigms, and support several phased evaluations, or longer term case studies as potential means of understanding the usability and usefulness of visual metaphors in medicine.

Measurements incorporating situational awareness remains limited, and results from a lack of consideration during the development of the visual representation. However, some papers have produced insight that may prove useful in improving situational awareness in the clinical, including integrating multi-dimensional physiologic data to create novel visual representations [123], [211], [225], [227], [256], enabling greater visual interactivity to expose details [236], [212], [214], and eliminating silo views to prevent the user from having to perform the manual task of accessing critical information [124], [225].

Finally while ecologic representations were evaluated for diagnostic accuracy and speed, the challenges surrounding cognitive errors remain only a secondary concern in research involving visual representations. Less than 21% of papers analysed were identified to have measured for cognitive workload [76], [124], [129], [211], [229], [239], [254], [257]. Of the eight papers that measured for cognitive workload, four papers used a quantitative measure such as the NASA-TLX score [124], [129], [211], [239]. There are also limitations with the use of NASA-TLX, largely because it is a self-reported method of identifying perceived workload. Three of the eight papers were evaluated with critical care clinicians, however. Consequently, incorporating cognitive workload as a passive measure of potential cognitive error remains limited across visual representation research for clinical environments.

4.4.4 Hierarchies of Tasks

The analysis of papers included in the review also introduced a hierarchy of tasks. The identified tasks were (1) data gathering; (2) data cognition; (3) information retrieval; (4) information dissemination; (5) knowledge gathering; and (6) knowledge dissemination. Tabular displays, including those that were integrated into legacy hospital systems support the transformation and organization of data [210], [215]–[217]. Those displays format largely numeric data values into spreadsheets, and occasionally encode them with text decorations [223], [257], [258]. Data cognition was seemingly the implied goal of representations that adopted graphical trends and novel object displays. Effken's etiologic potential display is an example of a display that uses dimensional manipulation to convey abnormal physiology to a clinician [246]. The goal of that display is to alert the clinician and convey salience about the present physiologic state for implicit decision support. Information retrieval and information dissemination were tasks that emerged from integrated applications to support explicit decision support activities [124], [209], [238]. Those integrated systems utilize tabular representations that organize raw data, as well as present waveform, object or metaphor representations convey visual summaries of data. Finally, a group of visual representations were identified that required the user to perform actions based on explicit knowledge that was synthesized by online algorithms. For instance, Syroid et al. [259] present a waveform trend display for drug concentrations, an algorithm is used to generate predictive curves that inform future trends. The intent of these representations can be seen as requiring some action on the part of the clinician-- in cases where the abnormal trajectory needs to be averted—and that feedback loop encourages active interactions with the physical world in response to visuals appearing on the screen.

4.4.5 Limitations of the review

Ambiguity exists over the use of the term “visual representation”, traditional indexes have used a variety of index terms to classify studies involving visual representation. Due to that ambiguity, the definition of the term used in the review was purposefully inclusive. The original review design anticipated the use of two reviewers to perform the initial paper screening. Yet the withdrawal of the second reviewer meant that the initial screen could not be cross validated.

The review was limited by the quality of the included studies and a lack of randomized control trials. This represents the disadvantage of using technology as an intervention in medical research. Within the studies analysed, there was heterogeneity in the methods and metrics used to evaluate visual representations. For instance, some studies used the NASA-TLX after an observational study for reporting cognitive workload, while others used subjective feedback in a semi-qualitative study. Moreover, where objective results were produced, there was heterogeneity in the participants, including exclusive experts or naïve participants. For that reason, it was not possible to conduct robust analysis of the accuracy, and usefulness of physiologic visual representations. Finally, the lack of good evidence from real-world evaluations presents limitations to generalizability.

4.5 Conclusion

Visual representations of physiological data have been attempted several times as witnessed by the sheer size of prior work discussed in this chapter. Many have shown their potential to improve clinical care, while largely positive results have been released, there are still concerns as to efficacy both in reproducibility as well as translatability to the unit [206]. In particular, methods to identify the accuracy of actions post-treatment to the display remains a concern and an open area for further exploration [76].

Few clinical visualization papers studied associations of the treatment condition to the accuracy or accrued insight by the user [76], [206], [255]. It was also seen that most studies included detailed study of the time to diagnosis and its accuracy, however many of these studies included highly controlled scenarios with highly visible graphical distortions [211], [246], [254]. Additionally, few studies used real patient data to evaluate their prototypes. Hence, the frequency of events with clear and distinctive graphical patterns existing across real patient data remains untested. Detection was also another area where studies frequently report positive findings, however, in many cases these differences were marginal and found in narrow statistical ranges. It has yet to be proven whether these statistical significances are relevant in the clinical domain. Exact mechanisms inducing positive effect have yet to be explored within the prototypes studied [28], [260].

Visual representations show promise, however this review exposes several challenges relating to user-preference and interaction challenges. Results spanning two decades continue to show positive influence of graphical representations when they are used in simulated studies [103]. However, many of these studies have not used standardized metrics to test distinct controlled variables, or provide evidence of precisely which features of the graphical displays afford greater comprehension to the consumer. Questions still remain as to its efficacy in clinical practice, where, the availability of all data required by the representations may be limited. There are also the limitation of graphical representation failing to maintain interpretable coherence, when provided incorrect data [206].

Some studies have also demonstrated user involvement as an important factor which may have influenced results, in the design and development of the clinical system [213]. Future studies should focus on clinical validation as a means to identify real-life relevance. Clinical experiments are difficult in lieu of several considerations and their limitations. However, one study by Wachter, et al [245], demonstrates that observational studies although somewhat intrusive may produce some significant qualitative results. These studies need to be expanded, and clinical trials must ultimately demonstrate their efficacy. Cognitive errors also require additional research effort, specifically by including evaluation methodologies such as the NASA-TLX score [261] to allow end-users to self-report perceived workloads.

Finally, research in visual representations should include tri-event parameters as important design considerations for design development that communicate episodic information. These visual representations can then be used to better assess the influence of tri-event parameters on higher level workflows as well as in the progression of conditions. The inclusion of such parameters may address open challenges relating to consumption (monitoring) and exploration (analysis) activities in visualizations of physiologic data.

4.6 Chapter Summary

This chapter presented results from a systematic review that was undertaken to expose novel visual representations of physiologic data. While text-based tabular displays have been popular for several decades, some novel tabular representations have likewise demonstrated an ability to improve situational awareness [262], these displays were discussed in section §4.4.1. Similar to text-based displays, physiologic data has long been represented using waveform visualizations. These waveform displays were presented in §4.4.2.

Graphical interfaces utilizing visual objects or metaphors are frequently identified as ecological displays, they were discussed in §4.4.3. Ecological displays seek to make identified constraints in the real world, also visible in the interface [246]. Visual representations were further categorized by a hierarchy of tasks in §4.4.4, this hierarchy motivates codes that were used in the analysis of qualitative data presented in the next chapter. In this thesis, concepts extracted from ecologic displays are used to support the development of PhysioEx and CoRAD, both of which are presented in chapters seven and eight respectively.

5. Qualitative Study of the Neonatal Intensive Care Unit

This chapter presents a grounded qualitative study that was performed to identify general processes involved in information analysis in the neonatal intensive care unit [43]. The study consists of three phases. The first and second phase inform results that contribute a theoretic model called the Exploration-Consumption Continuum that is presented in §5.2.1. The third phase contributes knowledge that informs novel visual designs for clinical researchers with exploratory requirements. The study received approval of the UOIT Research Ethics Board (file 10-017). This chapter concludes with a discussion of the continuum's influence on the TDVA framework.

5.1 Methodology for General Information Analysis Processes

The association of human factors, analysis skill, and the visualization character can make studying information visualization scenarios difficult [144]. Traditional methods to gather empirical evidence have largely focused on controlled studies to evaluate visualization tools [263]. Due to the nature of those controlled studies, which often occur far from the actual workplace and the intended user, it is difficult to evaluate the generalizability of their results [144]. An alternative methodology was presented in §3.4.3 called grounded evaluation. That methodology seeks to use semi-structured qualitative studies to expose early design concepts that can be used to develop visualizations. The grounded methodology presents an opportunity to uncover insight early in the design phase, using a mix of observations and semi-structured interviews to develop targeted and potentially useful visualizations for real-world use.

The grounded evaluation methodology was used in exploring domain challenges and opportunities for dynamic visual analytic research in the neonatal intensive care unit. The aim of the qualitative study presented in this chapter, is to understand the general processes that occur during interactions between clinical staffs and technology in the neonatal intensive care unit. In addition to computers, patient monitors, and clinical information systems, physical worksheets that were printed from computer generated reports were also included as a form of technology. The initial phase contributed codes that would be used to code confirmatory qualitative data in the second phase. The literature was used as an additional data source to compare emerging categories in the third phase [264].

The remaining sections being with a discussion of the observation setting and participants for all phases (§5.1.1 & §5.1.2). The details of the initial phase is found in §5.1.3.1, the second phase study design is described in §5.1.3.2, and the third phase is described in §5.1.3.3.

5.1.1 Observational Setting

There were four main observational settings included in the qualitative study, and include: a large teaching room (morning handover); fellows' room (evening handover and radiology rounds); physician's office (think-aloud sessions); and the neonatal intensive care unit (bed-side rounds). Among many opportunities available to observe the critical care environment, the morning handover round serves as an ideal setting because it allows for observing rich interactions between clinicians and their artifacts (computers, paper worksheets, and personal notes). During this time, clinicians gather to discuss unanticipated events that occurred during the evening shift and plan care management strategies for the day. From a knowledge transfer perspective, this setting provided unique ethnographic insight into the heuristics of medical decision making.

At the study hospital, a large room, capable of seating up to 50 persons serves as the primary handover location. Two computers are always available in the room, and each computer serves as a designated terminal for retrieving diagnostic images or clinical information to support their analysis. A large projector screen was affixed to the wall, and six large rectangular tables were arranged in an open-ended “U” shape. Seats were arranged along table and the edges of the wall. Senior clinicians usually sat near the projector screen, while the leading clinical fellow, a senior trainee, took control of the presentation. A second location included the fellows’ room, where six computers were stationed on tables around the edge of the room. A dedicated radiology computer was available for viewing x-rays and other diagnostic imaging. The room was adjacent to the neonatal intensive care unit, so clinicians can quickly move to the unit should a critically ill infant require them.

The third location was the neonatal intensive care unit. There are 11 large rooms in the unit, of which 8 are used for actual patient care. The remaining rooms are used for storing medical devices. On average, a total of 38 patients may be in the unit, each room can support up to six infants (three along opposite walls). At least two nurses are present in each room, and medical residents, fellows, and staff physicians enter and exit each room as required. The final location is the physician’s personal office space, this location was secluded from the unit and was quiet. Semi-structured interviews, as well as the think-aloud sessions were held in the personal office space.

5.1.2 Participants

This section presents the distribution and characteristics of the participants that were involved in the initial, second and third phase of the qualitative study. Two different sets of participants were recruited to the initial and secondary phases of the qualitative study. A span of six months was used between the initial and secondary phases to capture participants across two academic calendars. The location remained the same, and clinical artefacts (computers, worksheet format, and medical devices) remained consistent.

5.1.2.1 Initial Phase

Due to the use of naturalistic observations, it was only possible to observe 15 – 20 participants at a time during morning handover and bed-side rounds, since participants were free to enter and exit throughout the duration of the handover to attend to urgent patient care. The range of experience was highly diverse, with some participants having less than one year of experience to 36 years for the lead physician. Participants were not excluded based on age, gender, or length of experience. The participant sample was restricted to residents, fellows, and staff physicians. Participants were sampled by convenience, and consisted of different staff physicians (selected using the clinical service roster) who were on duty for that day.

5.1.2.2 Second Phase

The second phase used the same setting and types of participants to confirm initial findings. A sample of convenience was also selected based on the availability of staff in the unit and their willingness to participate. Data from non-consenting participants were not logged. The directed observation involved between 30 -- 40 clinicians over the duration of the entire study. These participants were observed several times over a period of five weeks. The second phase included additional clinicians, such as nurse practitioners, dietitians, respiratory therapists, and nurses. Due to the dynamic nature of the unit, participants were free to enter and exit the meeting room to attend to the urgent care of their patients. A total of 185 hours were devoted to observing interactions with technology, and summarized verbal statements (Table 2). Data was collected until at least 1000 sample observations had been recorded to represent a statistically significant dataset. The data collection was not terminated until the end of the shift, and hence the total codes amounted to 1127, of which 1055 were used in the analysis. The remaining 72 codes were identified to be too context specific and hence was not included in the categorization of themes.

Table 2: Observation of the Domain

Observation Type	Amount	Hours	Total Hours	Codes
Morning Handover	15	2	30	204
Morning Bedside Rounds	28	4	112	247
Evening Handover	15	2	30	588
Weekly Radiology and Neurology Rounds	5	2.5	13	88
Grand Total			**185**	**1127**

5.1.2.3 Third Phase

Eight participants were recruited from the neonatal intensive care unit for the semi-structured interviews that produced knowledge about effective visual displays for exploratory tasks. Participants consisted of experienced staff physicians, clinical researchers, and trainees that were selected from the participant pool identified in the second phase. Semi-structured interviews were also conducted using convenience sample, of participants with time availability during the day of the study, to elicit more knowledge about the participant's methods to accessing and analysing information (n=8). Participants self-reported as being a clinician or a clinical researcher based on the nature of their roles and their involvement in clinical research. The sample size was informed by emerging results, new samples were recruited until redundant data began to appear [146]. All consenting participants were included in the analysis.

5.1.3 Procedure

This section presents details about the procedure used in each of the three phases. The initial phase produced knowledge that was confirmed in the second phase. While rich knowledge was gained from those two studies, there was still a need to understand how novel visual designs can address some of the challenges identified in the observations. To that end, a third phase was conducted, where think-aloud sessions were held with paper prototypes. The procedure of the initial phase is presented next, followed by the second phase (§5.1.3.2) and finally the third phase (§5.1.3.3).

5.1.3.1 Initial Phase

A total of two morning handovers were observed in the initial participatory observations, followed by two bed-side rounds, two evening handover and one weekly radiology consultation rounds (different participants in each) totalling 16 hours. The observation was transcribed by two researchers (author of this dissertation and another research team member). Observations were performed across these four different types of clinical rounds, as opposed to a single type of clinical round to capture frequent high-level and holistic processes that involved interactions with technology. Observations were informed by emerging themes following the coding of the qualitative data after each session and terminated after consistent patterns emerged. Each observation was coded using qualitative coding methods to identify emerging categories in the information requirement of the participants. In this initial phase, short transcriptions were hand recorded, sensitive patient information were not included in the transcription.

5.1.3.2 Second Phase

The initial phase provided knowledge about how participants interacted with technology and their peers to gather, synthesize and share information. The second phase of the study focused on interactions involving technology, in order to gain more data on specifically the use of technology in their daily workflow. Using the initial phase as a base, the second phase expanded the number of observations and utilised the coding structure (presented in the next section) to confirm themes identified in the previous phase, as well as to identify new themes [265]. The second phase was carried out by a single researcher (the author of the dissertation), who transcribed comprehensive statement summaries of interactions that were observed in real time. These comprehensive statements reduce the level of detail captured, but focused the data collection on general processes, which serves as the main objective of the qualitative study. Comprehensive statements were created describing situations, explicit quotes, as well as the length of time to interact with clinical information systems and whether the result was successful. Successful tasks were quantified when the user is able to identify elements of interest in the clinical information system, and use those results to continue their conversation.

5.1.3.3 Third Phase

Following the observation of the environment, a semi-structured interview was conducted (by the author of the dissertation) to better understand effects of novel visual designs on the participant's ability to gain insight. Several visual designs representing clinical information were presented to participants (Figure 39). A list of visual designs used in the think-aloud study is included in Appendix 3. They were then asked to partake in think-aloud sessions of paper prototypes that were shown to the clinician using the Microsoft PowerPoint software. Participants were encouraged to use the visual designs to explore patterns and identify new insight. Participants were asked to verbalize when they believed that an insight had been identified. Statements where the participants did not explicitly verbalize an insight, but resembled an insight, such as using statements of “I find this interesting”, and “Oh look here!” were also coded as insights identified in the visual display.

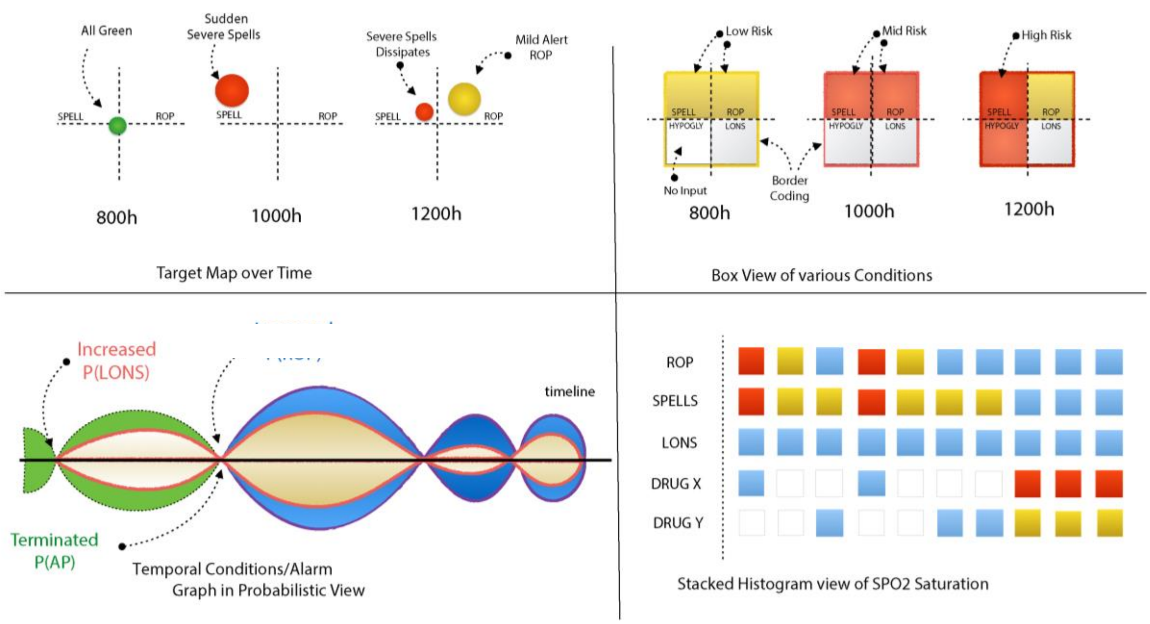


Figure 39: An illustration of a paper prototype presented in the secondary phase. Each prototype was evaluated using think-aloud sessions.

The think-aloud of the evaluations were also coded using codes that emerged in the initial phase and concepts were compared to initial findings, the list of codes are presented in §5.1.4.2. The visual designs presented to the participants include traditional method (line graphs), and more novel metaphoric representations that utilized asymmetric shapes and colour shades to highlight abnormal ranges or events. The visual designs were selected due to their ability to communicate insight in the literature. The aim of each session was to gain qualitative feedback from domain experts about their ability to understand and verbalize clinical information across several visual designs.

5.1.4 Data Collection and Analysis

The literature review of visual representations using physiologic data informed the coding strategy of this study. A hierarchy of tasks was identified in §4.4.4, that hierarchy was then used to code the qualitative data in the initial and second phase of the qualitative study. While these six codes are not comprehensive, they provide early knowledge about current practises in the neonatal unit. Definition of the codes as used in the study are described in Table 3.

Table 3: Description of Codes

Code	*Description*
Data Gathering	Directly interfacing technology in order to retrieve facts, such as monitoring vital signs on the patient monitor
Data Cognition	Reflecting on the significance (or mumbling) of a fact collected from a technology source, or collaboratively brainstorming possible reasons for the retrieved fact with another person in the group
Information Retrieval	Actions where the participant actively seeks previously synthesized facts (abnormal results or nursing notes) from a technology source
Information dissemination	Explicitly sharing synthesized facts with another group member or collectively discussing the results of a particular information retrieval task using technology sources
Knowledge gathering	Directed browsing activities in search of an actionable explicit or implicit knowledge, such as a confirmed report, or a plan of action
Knowledge Dissemination	Sharing actionable explicit facts retrieved from technology sources with other members of the group

5.1.4.1 Methods used in the initial and second phases

The initial phase collected direct transcriptions and used codes in the analysis. Additional codes emerged during the coding of the transcript, but they were not used in the secondary phase, because they were identified to be too context specific. Themes were clustered using an informal affinity diagram method, using Microsoft Excel spreadsheets and presented in Appendix 4. Codes were organized into groups with similar themes. For instance, in Table A4-1 (page 270), five theme groups were identified. In the second phase, data on general processes in information analysis involving technology were studied using clinical artefacts; these artefacts included worksheets, and interaction with the clinical information management system. Worksheets were included because they were computer-generated reports that were printed each morning and made available for each clinician to use.

Traditional methods in group scenarios use video and audio records to capture verbal statements [147], however in the observed setting, video and audio records were not permitted due to policies surrounding the protection of patient privacy. In the absence of video-audio recorded transcripts, summaries of directed interactions with technology were hand recorded as activities. Furthermore, while comprehensive verbal statements were hand recorded in the initial phase, the second phase recorded only comprehensive statements about the interactions involved between clinicians and technology (computers, patient monitors, printed worksheets, and telephones). This was performed to enhance the focus of the field notes on general processes involved in the information analysis pipeline.

Codes were clustered using an informal affinity diagram method to generate themes. The affinity diagram was created using Microsoft Excel spreadsheets. Affinity diagram topics for the first phase included groups such as: Analysis—Exploration Activities; Consumption—monitoring; Temporal Parameter—Trajectory; Temporal Parameter—Frequency; and Temporal Parameter—Duration. Examples of the classification is presented in Table 4.

Table 4: Examples of Themes

Comprehensive Note	*Code*	*Theme*
Fellow identifies three patients with similar fluid output values after exploring many patients using CIMs, and thinks aloud about potential correlations.	knowledge gathering; data cognition	Analysis—Exploration Activities
In room three, one of the patient monitors begins to alarm and then stops, nurse quickly glances at the screen from across the room.	data gathering; data cognition;	Consumption—monitoring
Resident describes issues they've had with the care of an infant, shows circled values on the printed worksheet, remarks values getting worse.	information retrieval; information dissemination	Temporal Parameter--Trajectory
Staff asks resident to find out how many spells the nurse has observed over the evening shift, resident looks over the nursing notes on CIMs.	information retrieval; knowledge gathering	Temporal Parameter--Frequency
Fellow observes from the patient monitor that an infant has been persistently desaturating for several minutes, and mumbles out loud if there is need for re-intubation due to the extended length of desaturation.	data gathering; data cognition;	Temporal Parameter—Duration

5.1.4.2 Methods used in the third phase

While the initial and second phases provided insight about information analysis processes that occur in the neonatal intensive care unit, additional efforts were required to gain knowledge about heuristics for visual designs that are most effective for that specialized population. Therefore the third phase explores effective visual designs using think-aloud sessions followed by semi-structured interviews. In the third phase, visualizations were presented to the participant, therefore, in addition to the codes that emerged during the initial and second phase observations, additional codes were used from the literature to measure insights generated from the scenarios presented via those Microsoft PowerPoint illustrations. Moreover, in order to move the focus from group interactions to individual observations, codes generated from the initial and secondary phases were modified as presented in Table 5. The characteristics of insights were measured using the coding themes contributed by Saraiya et al. [142]. The codes were clustered using an informal affinity diagram method, and emerging themes were identified.

5.1.4.3 Validation of the coding

Two researchers performed independent coding in the initial phase, and a third researcher was included when differences in coding occurred. Data from the second and third phase was coded by a single researcher (author of the dissertation), and the coding method was peer-reviewed by an independent team member who was not part of the research [266].

Table 5: List of codes used in the third phase

Code	*Description*
Data gathering	An activity where the user actively searches for data, i.e. a piece of information that can support their understanding of the current scenario
Data cognition	A state in which the user processes the information through a series of cognitive tasks and classifies that data as a potentially useful information
Information retrieval	An activity where the user uses an information displayed on the visualization to make a decision
Information dissemination	An activity where the user transfers information that can be helpful for others to make a decision
Knowledge gathering	User performs task to find meaning and relevance in the information presented
Knowledge dissemination	An activity where the user shares their insight and thoughts with others who can potentially benefit
Trajectory	User performs task to identify a trend or trajectory of an event
Frequency	User performs task to count of events of interest
Duration	User performs task to identify the temporal duration of an event of interest
Insight generation	A realization that the knowledge available from the information presented can be useful and meaningful
Hypothesis	User generates new hypothesises about a scenario

5.1.5 Limitations

Although directed observation can be used to rapidly gather different types of information, there are limitations [267]–[269]. Due to the limited sample size in the observations, and the convenience sampling method used in this study, the transcription that was collected may be relevant to that particular subset. Transcription errors are concerns in paper recorded field notes, however, due to patient privacy concerns audio-video records were not possible to obtain [148], [268]. Previous interaction with the researcher may have influenced the verbal statements made during observations. Due to the limited time available to complete the study, salient observations may have been missed from the transcription record. Due to a non-continuous observation timeline, the days selected by the researcher for observation may have been atypical and skewed the results.

Further, the observational studies and the interviews were conducted on one clinical site. Hence there are limitations on how far it can be extended across other hospitals both domestically and internationally. Clinicians at the teaching hospital represent a segment of the population who may not be privy to modern clinical information systems, or monitoring facilities. However, the benefit of using a highly specialized and international centre as The Hospital for Sick Children, Toronto, is the presence of a great amount of diversity in experiences and background found across staff and visiting fellows in the unit. Finally, some clinicians were asked to evaluate visual designs in the form of paper prototypes at the end of their 12 hour shift. This may have influence their alertness and ability to extract information from the presented displays. This may have also influenced their ability to provide detailed response on its usability. That being said, the effects may indeed provide valuable knowledge as it is routine for clinicians to spend up to 24 hours on the unit during their rotations. Moreover, due to memory recall limitations, interviews performed at close proximity to the event can be advantageous for recall from episodic memory [263], [270]. In that case the immediate proximity to their clinical service may have allowed participants to provide rich episodic anecdotes during the think-aloud sessions.

5.2 Results

This section presents results from the analysis of all three phases of study. This section begins with results obtained from the observations performed in the initial and second phases. A theoretical continuum is presented arising from the analysis of themes from those phases in §5.2.2. Finally results from the third phase are presented in §5.2.3.

5.2.1 Observational Results from Initial and Second Phases

A total of 158 codes appeared in the transcript collected in the initial phase, of which 122 represented codes identified in the literature, 36 codes were classified as other. In the second phase a total of 1313 codes were identified, of which 1055 were represented by the codes from the literature and 258 codes belonged to the other category. These other codes, from both phases, represented activities identified as supportive tasks, and included: "Gets Number of Episodes", "Makes Future Projections", and "Asks Time-based Queries". While these codes are important for analysis of the temporal tri-event parameter, in this section we present results relating to the six codes identified in the literature (Table 3).

5.2.1.1 Codes from the Literature

In the initial phase, 46 codes appeared in the transcripts collected during the bed-side rounds. 35 codes appeared during the evening rounds transcription, 22 codes appeared in the morning round transcripts and 19 appeared during the neurologic rounds. The remaining codes are presented in the next subsection. Among the codes identified in the literature, the most common code was knowledge dissemination, however this code was significantly more likely to appear during the evening rounds. The second and third most codes were information retrieval, and information dissemination.

The second phase produced similar distribution of codes, with 1055 codes identified over the 185 hour duration. While it was information retrieval and information dissemination that were second most frequently recorded in the initial phase, in the second phase, data gathering was observed more frequently. This may be due to an increased focus on the transcription of the second phase on directed interventions that involved technology in the unit. An illustration of the distribution is presented in Figure 40. Knowledge dissemination activities dominated the conversation throughout evening rounds, where the emphasis is on transferring actionable insight about patients to the clinician responsible for the evening shift (Figure 41). Data gathering and data cognition tasks were however more dominant during the bed-side rounds. This may be due to the fact that the bed-side rounds often included a computer trolley that accompanied clinicians as they saw each patient in the ward.

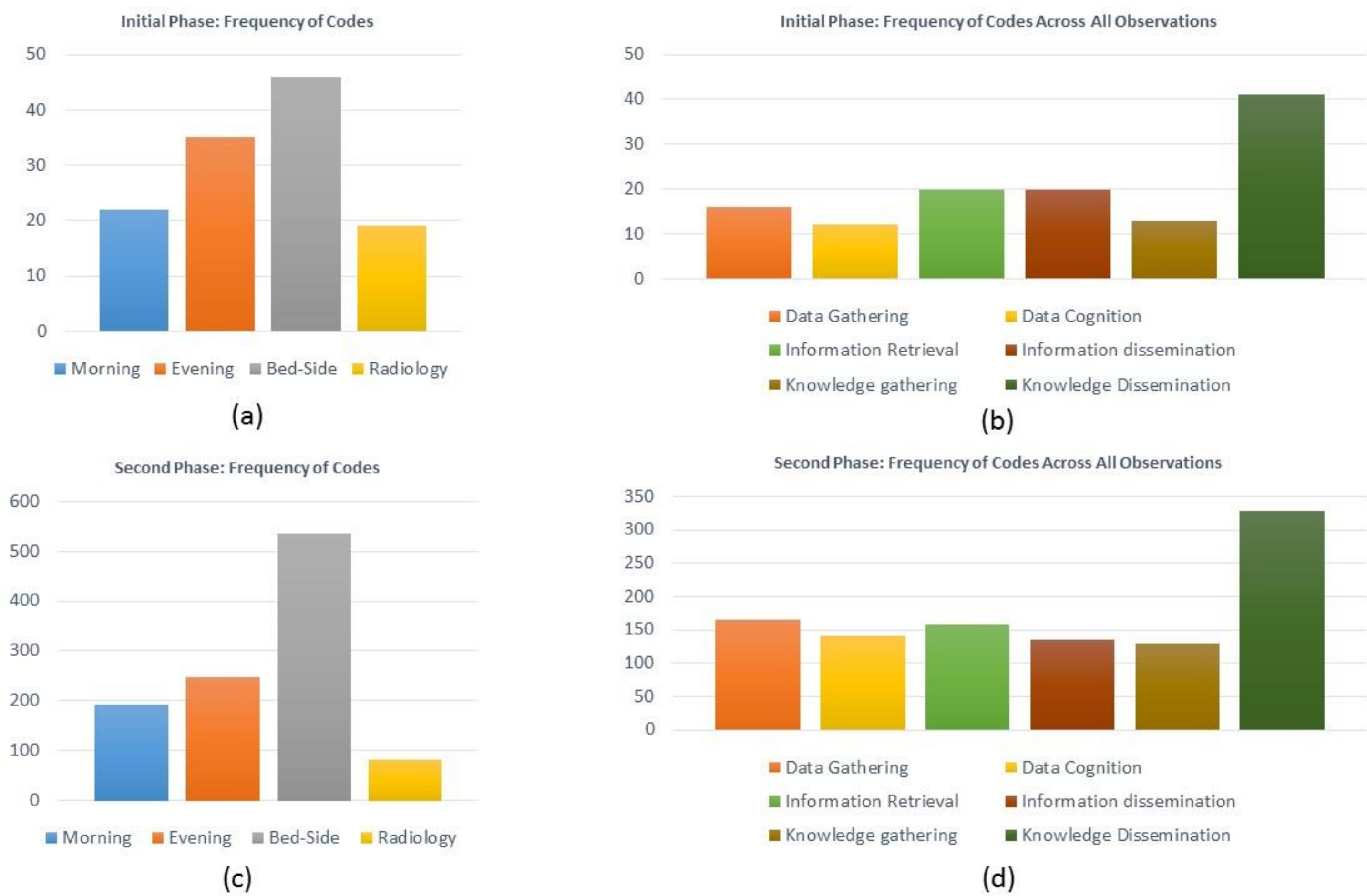
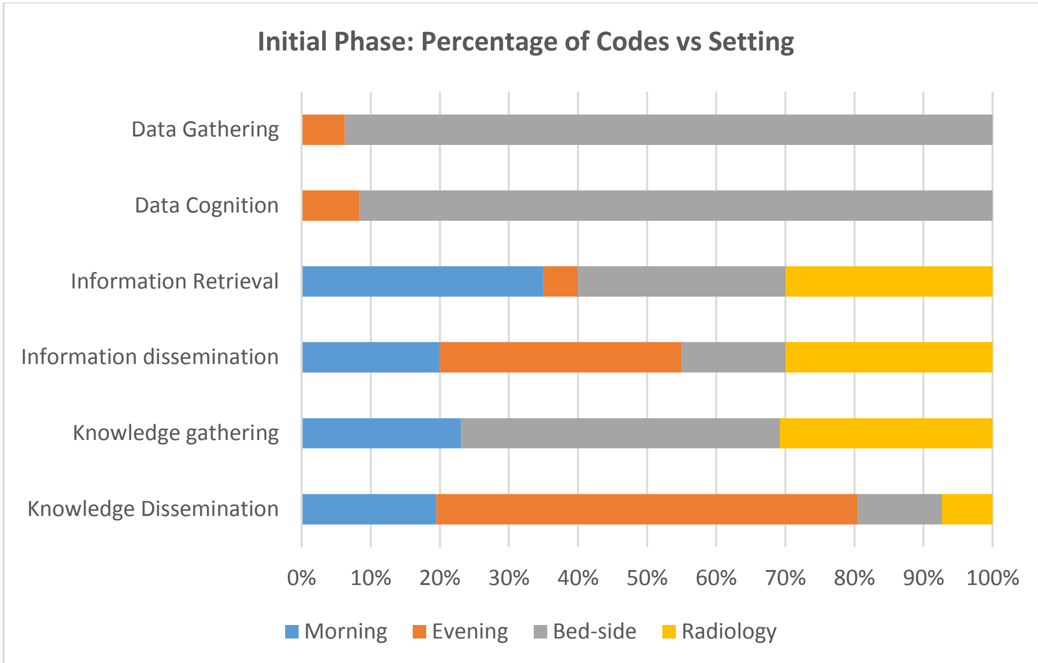
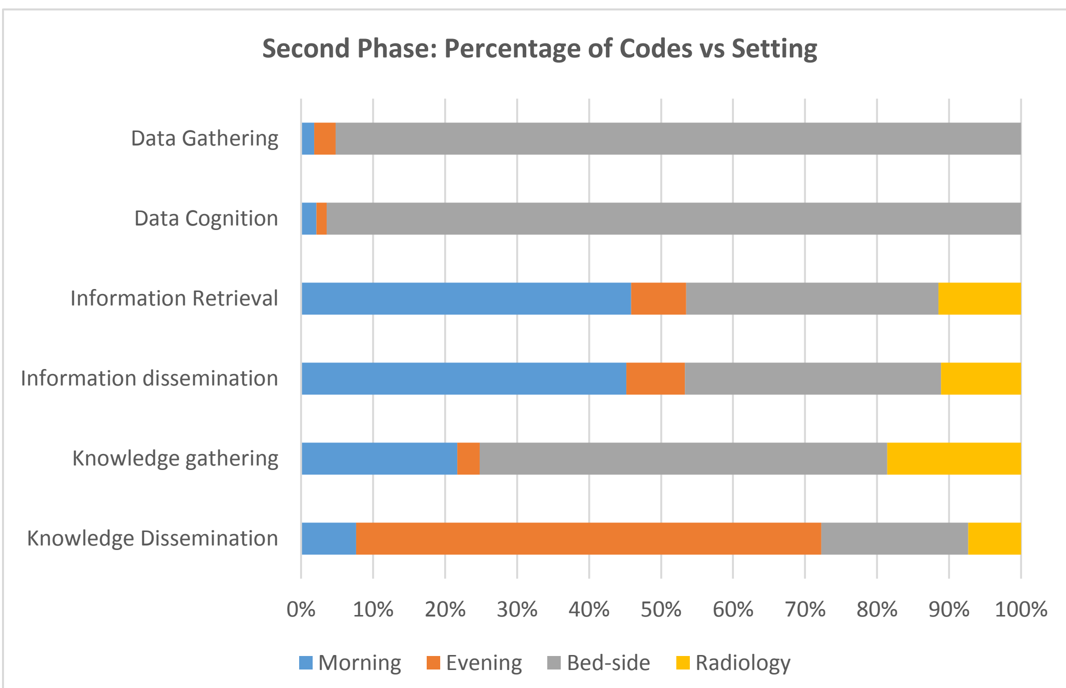


Figure 40: A series of graphs illustrating the distribution of codes collected using observational data gathered in the initial and second phase of the qualitative study. Graphs (a) and (c) illustrate the number of codes identified across each observation setting in both phases, while graphs (b) and (d) describe the absolute frequency distribution in both phases.



(a)



(b)

Figure 41: Graph (a) illustrates the distribution of codes across each observed setting in the initial phase, while (b) illustrates the same distribution in the second phase.

5.2.1.2 Codes Relating to Clinical Event Parameters

In addition to codes from the literature, three additional codes emerged. These codes were: "Gets Number of Episodes", "Makes Future Projections", and "Asks Time-based Queries". The first code appeared a total of 58% in the first phase and 78% in the second phase. The "Makes Future Projections" code appeared 19% in the first phase, and 24% in the second phase. The "Asks Time-based Questions" appeared 22% in the first phase and 30% in the second phase.

5.2.1.3 Themes from clustering both groups of codes

The coded transcripts were clustered using an informal affinity diagram method and five emerging themes were identified (Appendix 4). The first theme is trajectory, participants contemplated on the current trajectory of the patient. Questions such as "I feel like this patient is getting worse", highlighted the "Makes Future Projections" based comments that were observed. The second theme is frequency, participants would think-aloud about the distribution and frequency of clinical events, and was coded as "Gets Number of Episodes". This metric was important when patients were considered to be 'very sick'. Duration was the third theme that was identified arising primarily from the code "Asks Time-based Queries", participants requested clarification on the duration of events. Two participants expressed that due to the lack of tracking duration of events using current technology, they were not able to provide an answer. The final two themes were consumption of facts and exploration tasks for knowledge gathering. Additional themes that emerged during the qualitative analysis were challenges in data gathering and cognition, and the lack of technology to support information and knowledge generation at the point of care.

Clinicians who wished to express behaviours occurring in a span of seconds faced the most difficulty communicating with their peers, i.e. data gathering and data cognition. This was performed with minimal aid from clinical information systems as: (1) the data was unavailable for the time the event occurred or (2) there was a lack of analytic facility. Information was easily verbalized but the verbalization of specific clinical statuses pertaining to a patient was often ambiguous and required additional follow-up between clinicians and bed-side nurses.

After the analysis of artefacts and tasks, it was recognized that very few systems exist to support the clinician through analytic activities performed using retrospectively stored physiologic data. However, when discussions were initiated about physiologic behaviour of a patient, the discussion was often times delivered as anecdotes at the bedside. Common phrases such as “I am not concerned”, or “The baby looks sick” were often used to convey clinical knowledge that may affect a patient’s management strategy in the absence of quantitative means for identifying apnoea or sepsis.

These results pointed to an early theory involving the diverse tasks that can be supported in the neonatal intensive care unit. A second analysis of the codes revealed a tendency to shift towards monitoring and exploration activities. This lead to the construction of a continuum, which is presented in the following subsection.

5.2.2 Exploration-Consumption Continuum

During the observation, it was noted that clinicians frequently fluctuate between different knowledge seeking dimensions. This observation, along with the semi-structured interviews with clinicians and clinical researchers resulted in the creation of the Exploration-Consumption continuum illustrated in Figure 42.

Two prominent themes emerged during the analysis of the qualitative data, they were: Analysis—Exploration Activities and Consumption—monitoring (Table 4). Transcripts from both themes were studied and the six codes used for analysis of the transcripts were grouped into Exploration and Consumption categories. Within each of those groups, several activities emerged. These activities ranged from hypothesis generation (exploratory research), to active perceptual data gathering (focal monitoring). It was also noted that these activities differed in the context of the urgency of the situation and the engagement level of the person involved. These emergent activities were compared with the findings of Sanderson's study on focal and peripheral attention of anesthesiologists [28]. The focal and peripheral attention activities showed similar behaviour to the activities identified in the analysis of the qualitative data from this study. For instance, a responsible physician within a highly urgent situation was likely to exist in the focal monitoring phase, whereas a clinical observer (without involvement in the care management) in the same situation with less engagement may be involved in peripheral monitoring, in which data is observed through peripheral attention [28]. These findings were integrated within a theoretical model that forms the basis of the Exploration-Consumption continuum.

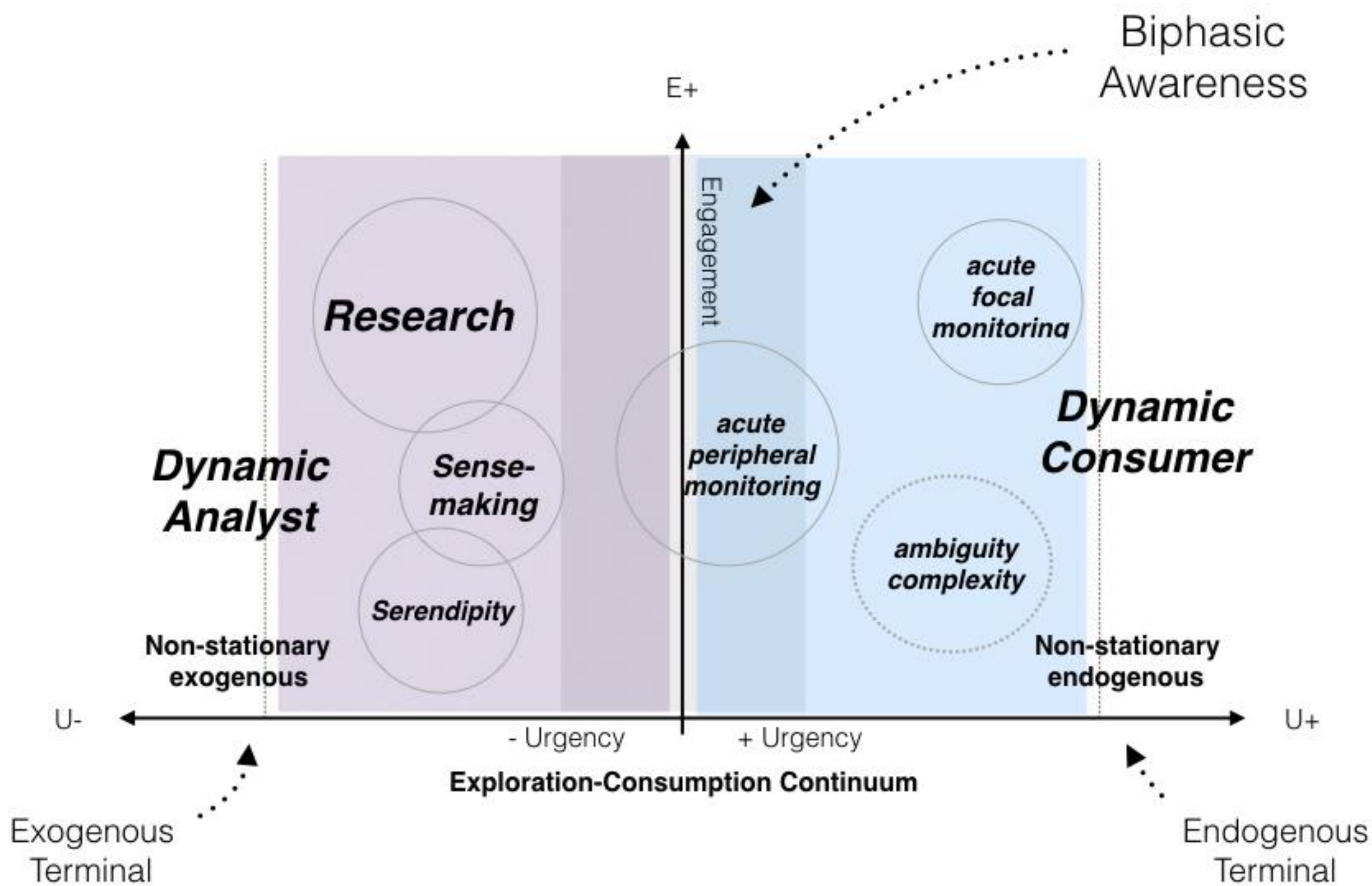


Figure 42: Knowledge exploration and consumption continuum.

The first dimension is the knowledge consumption domain illustrated in the right side of Figure 42. In this dimension the user is predominantly interested in focal monitoring of clinical events. This includes events such as breathing patterns, skin tone, or data gained from the clinical information system. In other complex domains, a dynamic consumer may be interested in rapidly identifying events such as network intrusions, or power-system failure. Therefore the consumption dimension exists for those who require situational analysis within very short durations.

The second dimension is the knowledge exploration dimension illustrated in the left side of Figure 42. In this dimension the clinician is often asking questions about conditions that may be hidden or untapped due to lack of information. As noted in the results obtained from the think-aloud sessions, clinician in the explorative dimension spent upwards of four minutes on each display, carefully identifying elements and predicting future trends without verbal cues from the researcher. In similar complex domains, dynamic analysts existing in this dimension perform detailed exploratory tasks to uncover trends and patterns that may indicate an abnormal condition. The dynamic analyst is not constraint by time and therefore can perform deep insight activities such as, research, sense-making, and spend time to arrive at knowledge via serendipity.

The axes of the continuum represents the urgency of the environment and the engagement capacity of the human. As the urgency (horizontal axis) and engagement (vertical axis) increased, the participant pivots towards the acute focal monitoring (top right, Figure 42), a phase in which the user discerns critical temporal events. Conversely, when the participant operated in a less urgent (left, Figure 42) environment, such as their office, and their engagement levels are increased, there were increased instances of exploratory research activities. The continuum identifies other activities that can occur as both axes are controlled.

The presence of the endogenous and exogenous terminals are relevant for the maximum amount of information that can be processed at any point in time. It was identified in the observation, that significant transfer of data was discouraged in the consumption oriented task, while encouraged in the exploratory activities. Participants expressed desires for synthesis, using statements such as “Is the infant unwell?” to request summarization of events. When tasks such as browsing, and data gathering were the prime focus, statements such as “How does this value compare over to that over the weekend?” and “Can you bring up all the results on screen?” were noted. The actual extent of this terminal may well be influenced by the individual, hence further research is needed to understand specifically which tasks can be considered bordering the terminal on the continuum.

Due to the lack of information systems that consider both dimensions identified in the Exploration-Consumption continuum, existing clinical systems were seen to provide minimal facilities and complicate previously trivial tasks, based on the feedback received throughout the study. Moreover, the participants’ feedback suggests that existing displays also were identified to have poor consumption facilities. For instance, to retrieve information about blood glucose over the past 24 hours, a series of menus needed to be access along with multiple selections. Participants verbalized anecdotal evidence to suggest this challenge is seen in numerous instances.

5.2.3 Results from the third phase

The third phase exposed clinicians to novel visual designs that were presented as paper prototypes (Appendix 3, Figures A3-1 to A3-6). Clinicians were asked to perform think-aloud sessions to elicit knowledge about effective displays. 127 codes were collected from the analysis of the transcript of which 11 were excluded because they were too context specific. Information Retrieval (n=27), Knowledge Gathering (n=15), Trajectory (n=13), Data Cognition (n=12) and Insight Generation (n=11) were dominate codes that appeared in the transcript (Figure 43). Codes were then grouped using the informal affinity diagram method, and this process involved the author grouping each code item into five affinity groups using Microsoft Excel Spreadsheet. Those emerging themes include: reactions to existing systems, reactions to novel representations, engagement levels, perception of usefulness, and analysis of cohorts. These themes are discussed in the remaining parts of this section.

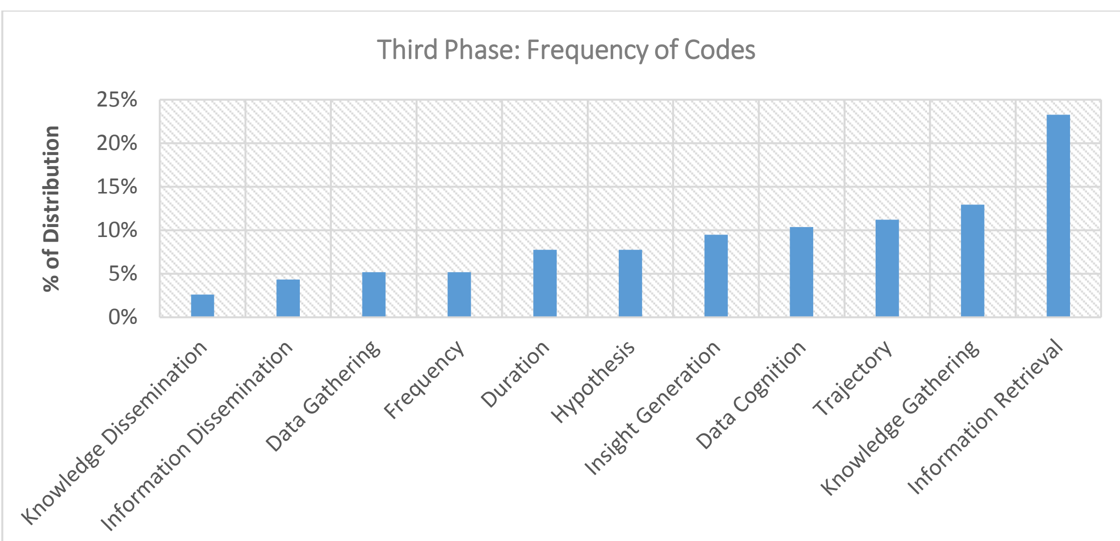


Figure 43: Distribution of codes from the third phase.

During the think-aloud sessions, clinicians associated current clinical information systems with negative emotions, and complained of their limitations when analysis workflows were required. It is to be noted however, that a distinct group of clinicians with highly observant skills (clustered by time spent on each case, cursor movements, and number of questions asked) expressed greater dissatisfaction with the current information systems and their abilities to expose more data for clinician-driven analysis. These clinicians were also clinical researchers and hence, routinely performed in-depth analysis of physiological data.

Generally clinicians responded with interest in the novelty of the metaphoric displays, but preferred standard line charts. They stated familiarity as the reason for choosing tables and line charts to visualize clinical information. The heatmap representation (Figure A3-1, page 264) was identified to be the most revealing visual representation, and was found to the easiest to understand. Novel visual representations were identified with greater interest among department fellows and medical residents. Specifically the ability to interact with visuals on the screen to reveal additional details on demand were associated with positive emotions (Figure A3-3, page 266).

Within participants recruited for this study, there existed a set of clinicians who were much more involved in understanding the details of the visual display. These participants were found to have more frequent analysis related codes (data gathering, data cognition, information gathering, information dissemination, hypothesis generation, and insight characteristics) than their peers. This set included one fourth year medical undergraduate student, one neonatal program fellow and two neonatal department fellows. These 'observant' clinicians utilized the paper prototypes to predict the next cycle of events and spent generally a greater amount of time with each display. The average time spent by the 'observant' clinicians before making initial verbal comments exceeded 4 minutes per display, while the overall average for all eight participants was under 1 minute. These observant clinicians, as well as the general population of clinicians eluded to a new category of 'information consumers' that have limited exposure in the literature. This discovery leads to the creation of an exploration-consumption continuum which is discussed in the next section.

The novel metaphoric displays were verbalized as useful and usable with that set of clinicians. Specifically, the DxRadar (Figure A3-2, page 265) representation and the star-plot (Figure A3-4, page 267) was identified to be very useful for monitoring clinical conditions by four participants. They had positively responded to using these displays in their workflow to improve their understanding of complex physiological data at the bedside.

Clinical researchers while progressing through the think-aloud session frequently voiced intent for extracting features from a single patient and evaluating that across a patient population. Although the tree-map (Figure A3-5, page 268) was not intended to be used for such, but one participant found it to be a potentially useful representation to support analysis of patient cohorts. However, explicitly, this function was not available in the paper prototypes and presented an opportunity that was not anticipated prior to participant exposure.

5.3 Discussion

Analysis of observations conducted in a dynamic complex environment show that experts exist along a continuum that spans strictly exploratory analysis to strict analytic consumption.

The observation results provide insight into the day-to-day activities of those complex domain experts existing in both dynamic dimensions. In particular, the initial and secondary phase feedback and anecdotal evidences, shows that there is a need to enhance currently available clinical information systems with dynamic components that support each of the requirements for both the consumer and the researcher. For instance, results from the secondary phase revealed that clinical information systems do not provide consumers with the ability integrate heterogeneous data streams with events of interest.

Moreover, those clinical systems (electronic medical records, order-entry, and clinical narratives) as observed in each of the qualitative studies were found to lack even high-levels of analysis of physiological data streams. Much of the physiological displays were summaries of data samples abstracted at a predefined interval. It was also found that during bed-side rounds for instance, clinicians spent a larger portion of their time gathering facts, and fewer time was dedicated to knowledge dissemination activities.

The initial and secondary phases provided results that were used in a requirement analysis. Four functional requirements were elicited from limitations observed in the setting:

R.1. **Dynamic visual analytic**: Applying visual analytics to physiologic event stream processing algorithms that produce complex association and classifications from multi-stream analysis of raw data. This requirement supports the need for user-driven explorative and explanatory tasks.

R.2. **Monitoring relatively aligned temporal events**: Using temporal relative alignment techniques to visually expose for analysis real-time and retrospective events across multiple data streams to support exploratory research.

R.3. **Coordinated streams visual analytics:** Supporting dynamic coordinated interaction on multi-dimensional temporal data streams to allow the analyst to elicit information on event features and classifications.

R.4. **Analysing across cohorts**: Hypothesis testing performed on one subset, needs to be applied across cohorts or populations to observe similar patterns has not been demonstrated in real-time visual analytic systems.

Research is warranted in the space identified above in R.1, where system plays a proactive role by performing low-level data processing tasks, and generates features for interactive exploration by the analyst at the visual analytic layer. Novel relative alignment of temporal streams, complex associations and conditions can be exposed for further pattern analysis, this approach addresses requirements stated in R.2. R.3 further implements on requirements established in R.2, with the addition of coordinated interactive analysis views. These coordinated views support the analyst in performing exploratory and explanatory research. Finally, the need to support case-controlled studies for extending results of a single analysis to other unique systems is stated by R.4. Applying hypothesis generated from a single case, to similar populations enables the domain expert to generate new knowledge.

5.4 Research Directions in Novel Physiological Displays

This study exposes constraints in technology developed for the neonatal intensive care unit, specifically with respect to the representation of physiological data. It was noted in the analysis of the qualitative study that physiological displays can be designed and developed using three temporal metrics identified as the tri-event temporal parameters namely, trajectory, frequency, and duration of salient events.

It can be further noted that clinicians, as well as general complex and dynamic domain analysts, can exist across either the exploration or the consumption dimension with respect to the urgency of the situation and their particular level of engagement, this is referred to as the Exploration-Consumption continuum [42] and was introduced in this chapter. The tri-event parameters and the exploration-consumption continuum encapsulate critical temporal properties along with higher-level processes that may be critical in supporting practices essential for evidence-informed care [207], [42].

In the systematic survey presented in the previous chapter, trajectory, among the tri-event parameters, was found to be the most popularly expressed. 31 studies were found to incorporate some form of trajectory information. However, longitudinal trajectory was found in only nine studies, moreover seen rare among displays that were found in anesthesiology but more common in critical care. Displays that incorporated an aspect of the tri-event temporal parameters exclusively adopted trajectory. Nine visual representations were found to have included the duration and frequency metrics.

Most of the representations that included duration and frequency used glyphs (n=5) or text (n=4) to communicate episodic information. Glyphs among one of the most popular methods in information visualization to communicate episodes and their properties [249]. Hence, it is not surprising to see many authors exploit the discrete property of glyphs to communicate information such as duration and frequency. Text also remains a popular method for communicating discrete events. Law et al., found text to be superior to waveform and numeric displays when communicating clinical episodes, even while clinicians reported a preference for graphical displays [271].

While the exposure of tri-event parameters remain limited, they have become increasingly prevalent in visual representations developed over the past decade. Gorges et al., for instance, introduce novel glyph based far-view display that discretizes events into hourly summaries [129], resulting a 26% improvement in decision accuracy. However the utilization of all three tri-event parameter have seldom appeared in design considerations of the papers analysed. Further research is required to investigate the influence of these tri-event parameters in clinical workflows and ultimately their potential impact on situational awareness and cognitive errors. Where multiple views were presented, only one representations utilized interactive coordination between independent views [225]. Finally, it was noted that none of the visualizations that were surveyed supported the function to identify case-control groups.

The results presented in this and the previous chapter demonstrate the need for further research in novel visual representations developed for the complex medical domain. The visual representation should at minimum support the tri-event temporal parameters, and enable new tasks such as hypothesis generation and testing. The contribution of this thesis is centered on addressing some of these limitations through the introduction of a framework for developing consumption, and interactive exploratory dynamic visual analytic tools.

5.5 Threats to Validity

This section details threats to validity-- including external, internal and conclusion validities for the grounded qualitative study that was presented in this chapter. Due to the use of naturalistic observation design, there is a high degree of external validity presented in this research. However, due to the use of a single site, even when it is composed of a highly diverse participant set, the naturalistic observation may not translate to all similar hospital environments. The use of the semi-structured interviews, and the think-aloud sessions strengthen the internal validity of the research.

Conversely, due to the influence of the 'Hawthorne Effect', participants may have changed their internal disposition within the context of the research study. Finally, due to the limited set of visualizations that were shown to the user, there may be threats to conclusions. The conclusions of this research may pertain only to the types of visualizations (density, metaphoric, and traditional) that were used.

5.6 Chapter Summary

The chapter begins with a detail of the methodology used in the observation study. The qualitative results are presented in §5.2. Two major findings were introduced, the first finding exposed the tri-event temporal parameters as critical elements of discussion between clinicians in the intensive care environment. It also motivates the need to investigate the extent to which health visualizations support these parameters. The second finding revealed two dimensions, exploration and consumption existing along a common 'urgency' axis. A secondary axis, 'engagement' was also observed. These two axis forms the 2-dimensional plane called the Exploration-Consumption continuum.

The next chapter presents detailed walkthrough of the traditional data warehouse architecture, followed by introducing the components of the TDVA framework, along with the TDVA methodology. That chapter concludes with the presentation of an architectural representation of the TDVA framework, as instantiated to support clinical research activities.

6. Tri-event Parameter Dynamic Visual Analytic Framework

This chapter introduces the Tri-event Parameter Dynamic Visual Analytic (TDVA) framework, TDVA methodology and TDVA platform design. The literature review as detailed in chapter three demonstrated open research opportunities in the broad areas of dynamic visual analytics, and data warehouse frameworks to support dynamic environments. The traditional data warehouse is presented, then the TDVA framework, followed by details of the TDVA methodology and TDVA platform design. This chapter concludes with examples of two prototype visual analytic tools developed using the TDVA framework called the Heart Rate Variability Graph and the SeqEvent graph in §6.5.

6.1 Traditional Data Warehouse Design

The traditional data warehouse architecture (Figure 4, §3.3) [91] loads data from operational data systems which are typically databases that support organization operation systems. The limitations of that approach include: the use of restrictive OLAP cubes for analysing fixed arrays of data, overemphasis of longitudinal normalized data, unidirectional data flow paradigm, and the lack of continuous feedback between the OLAP engine and front-end user interactions [94], [95], [105], [110].

To address the limitations presented above, a new approach is needed that integrates multi-dimensional temporal features while supporting interactive dynamic visual analysis for analysts in complex domains. This separation of the presentation layer, therefore, presents limitations to enable dynamic visual analytics as discussed in § 3.3.

6.2 TDVA Framework

There are numerous reasons for which sensors are deployed, including to support fault detection and avoid costly interventions. Sensor networks can be found in a variety of environments including, flight recorders on airplanes [272], network traffic monitoring [273], and also in the intensive care unit [37]. The proposed Tri-event Dynamic Visual Analytic Framework (TDVA) incorporates real-time sensor network processing to enable dynamic visual analytics.

The key insight introduced by the TDVA framework is the instantiation of a dynamic visual analytics publishing interface between the Data Presentation Area and the Data Access Tool of the traditional data warehouse architecture [91]. Figure 44 illustrates the Dynamic Visual Analytics (DVA) Mart Publisher (mPub) engine that receives data from the data warehouse. After performing a series of pre-processing steps, the mPub engine produces independent and loosely coupled Dynamic Visual Analytic Marts (DVAM). These DVAMs exist as a toolkit of visual interfaces that are available to support dynamic requirements of the analyst. DVAMs can be instantiated as exploration or consumption instances. Exploration instances support analysis tasks such as explanatory and exploratory research. Consumption instances present dynamic content by assigning visual representations that contain limited interactions.

The TDVA framework, use components from the traditional data warehouse architecture, and extends components where interactive exploratory analysis is unsupported. The introduced concepts extend beyond the traditional data warehouse data mart by intrinsically coupling unique consumption or exploration requirements with an appropriate visual analytics approach to adaptably explore the data. The user is able to customize their analytic space by submitting requests to mPub, which creates a personalized DVAM with unique views that are interactive and independent. Each instantiated DVAM allows the user to interactively explore the analytic space. The user performs interactive controls on visual toolkits made available by the mPub engine to perform a series of basic direct manipulation functions such as zooming in on lower levels of information, and hovering over visual objects to get details on demand. Furthermore, an interconnected Event Stream Processor (ESP) replaces the silo processing component of the data staging area presented by Kimball and Ross in (Figure 4, page 27). This allows deployed DVAMs to have rich data processing routines that are required on-demand by dynamic and responsive visual interfaces. Hence, in contrast to the unidirectional ingestion of data via traditional data warehouse architectures, DVAMs are tightly integrated bi-directionally with the ESP. The presentation area is replaced by the mPub engine, and the Data Access Tools are replaced by DVAMs. Details of each component of the framework are described in §6.2.1, and §6.3 presents details of a novel methodology that makes instantiation of DVAMs feasible in complex environments.

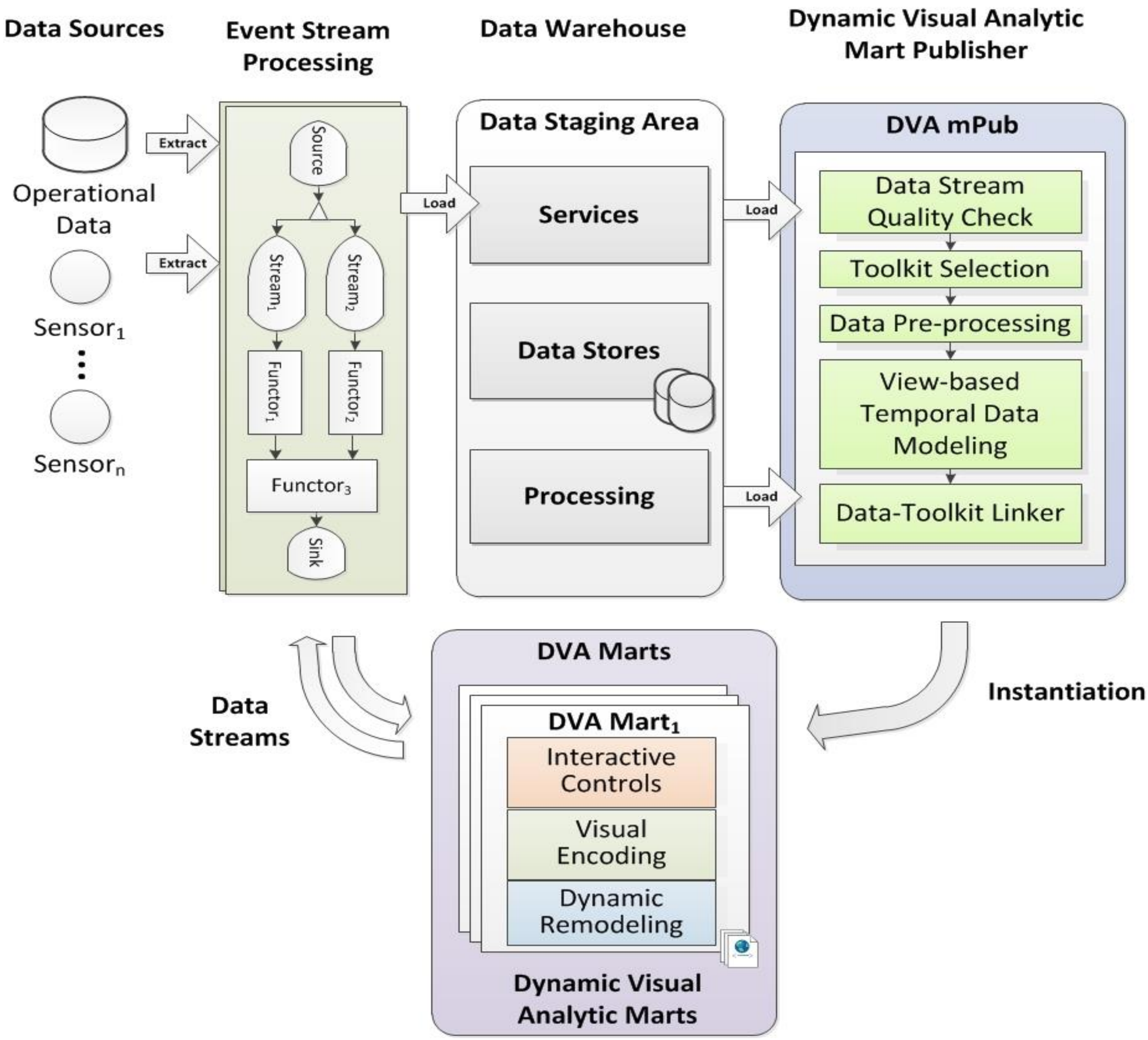


Figure 44: Extensions to the data warehouse architecture adapted from Kimball & Ross, 2011 [91]

6.2.1 TDVA Framework Components

There are four key components of the TDVA framework that is illustrated in Figure 45; they are the (1) Event Stream Processor (ESP), (2) Persistent Data Store, (3) Dynamic Visual Analytic Mart Publisher (mPub), and finally the (4) Dynamic Visual Analytic Marts (DVAM).

(1) **Event Stream Processing (ESP)**: The ESP is the first component of the framework, and sources real-time and retrospective data streams from sensors and event stream databases. In this component multiple secondary data streams are generated belonging to one of four classes. Namely the output streams are: (a) event features, (b) variability, (c) trending, and (d) event classifications.

ESP.a. *Event Feature streams* are the most basic of the four complex streams generated by the ESP; these streams provide rapid asynchronous view of the dynamic environment relating to the state of sensors and the health of the monitored system.

ESP.b. *Variability streams* are generated for synchronous data streams that exhibit measurable changes in variability of the underlying data stream. Variability streams are synchronous and independent to the trending and classification streams. Variability is calculated as a delta that is measured against a moving average and recalibrated upon incremental or rapid stabilization of the source signal.

ESP.c. *Trending streams* are generated as a trajectory metric between two features within a single stream or between one or more streams that exhibit coordinated behaviour. Trending streams are transmitted as synchronously bursts even before the event can be classified.

ESP.d. *Event Classifications* are generated as ordinal outputs that are transmitted after events have been classified as belonging to one or several conditions. These streams are asynchronously generated based on conditions established in algorithms running in the ESP engine.

(2) **Persistent Data Store**: In this component, data streams from the source, as well as the secondary streams generated by the ESP engine are stored persistently. The data store also stores system generated actions, such as the tertiary data generated by the mPub engine. Analyst generated streams, such as interactions, click streams, and analysis session data are stored for providing analytic provenance, such as the ability to revert to a previous analysis pathway.

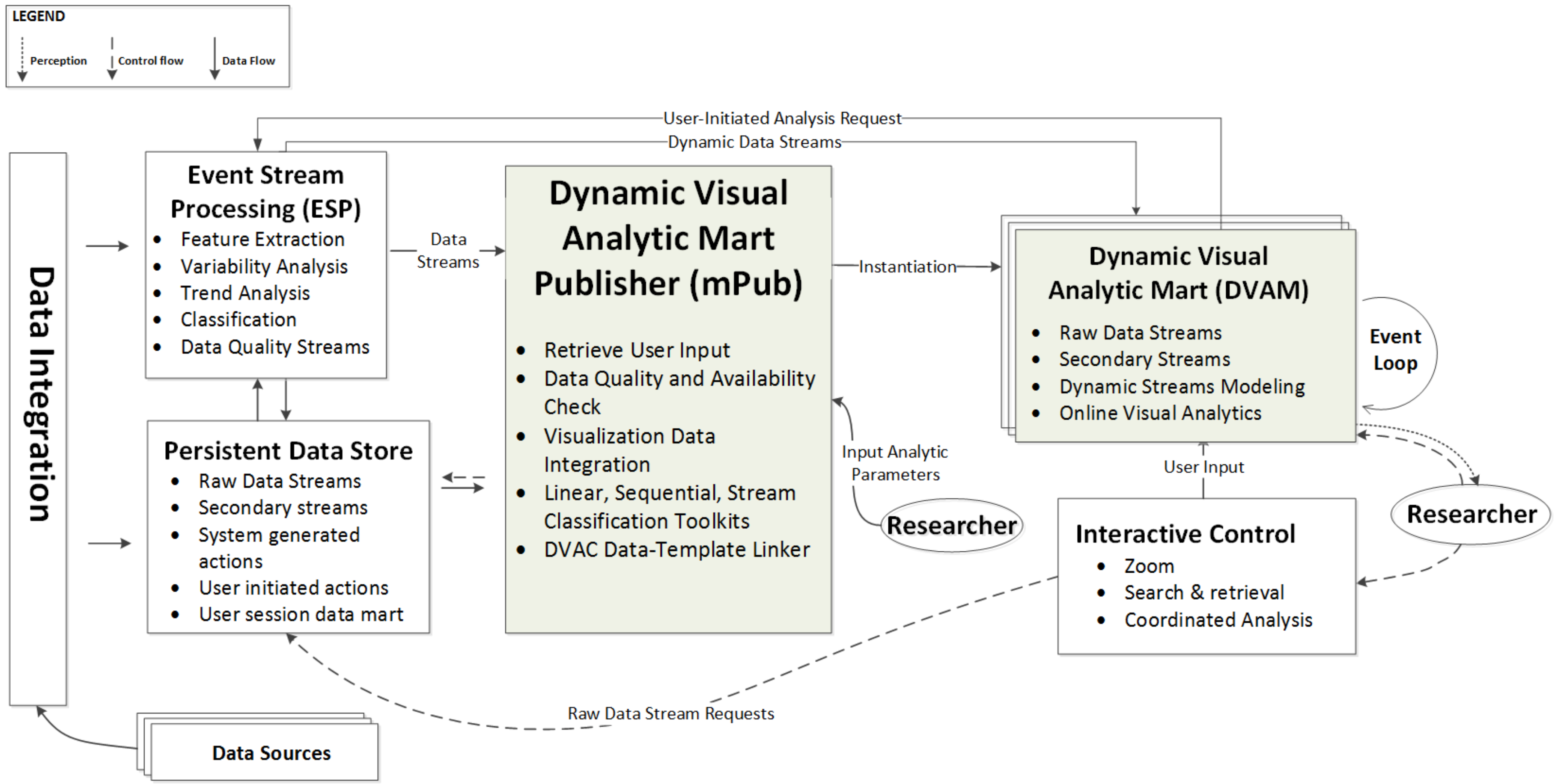


Figure 45: The Tri-event Parameter Dynamic Visual Analytic Framework

(3) **Dynamic Visual Analytic Mart Publisher (mPub)**: This component serves as the backbone for enabling dynamic visual analysis in the critical and complex environment. It receives input from the dynamic analyst, ESP engine and the persistent data store components. The end product is an instantiation of one of many visual analytic toolkits existing in this interface.

The mPub also allows the researcher to balance exploratory and explanatory research. By using retrospective analysis the dynamic analyst may perform iterative exploratory research by exploring various views provided by the mPub component. Once the dynamic analyst is satisfied with a view, they can replicate these instantiations for new data streams, without querying retrospective data. Hence, dynamic consumers may be able to utilize the view with minimal configuration to perform tasks existing in the consumption dimension.

Initially, the dynamic analyst submits a request for one or more visual analysis mart(s), using necessary parameters such as, the data identification, analysis epoch, relative alignment of other events, and the specific temporal views. To arrive at the dynamic visual analytic mart, the mPub component must first check the source stream for sufficient data quality to identify missing data, or void regions. Once sufficient data quality has been met, the mPub component creates a wireframe mart with all components and socket addresses of the respective source streams. Using information gathered from the analyst, the mPub component creates DVAMs using one of six customizable visual toolkits

(VT). Additional VTs can be developed in future work to expand the collection of VTs available for use. Each VT supports the representation of one or all of the temporal tri-event parameters.

VT.1 **Linear Plot Graph**: Plotting of event feature streams in linear time against duration, one of the key tri-event parameter. This graph also exposes frequency by controlling opacity and trajectory using time along the x axis.

VT.2 **SeqEvent Parallel Coordinates**: An extension of parallel coordinates [274] designed for highlighting sequence patterns in temporal datasets.

VT.3 **Sequence Graph:** Plotting of classification events by aggregating frequency over an epoch. The default view of this graph shows Day x Hour. In which, hourly intervals are expressed across the x-axis, and day across the y-axis. If the analyst has submitted the Hour x Minute request, the x-axis would contain one minute interval and the y-axis would contain one-hour intervals.

VT.4 **Streams Graph**: A stacked area plotting of frequency of events calculated at one minute intervals against frequency. The area is determined by the frequency of the event. The x-axis is measured in one minute intervals.

VT.5 **Temporal Intensity Map**: A unique heatmap representing frequency and duration of events along time in the x-axis. Frequency is controlled by stacking and duration is plotted along the vertical axis. Trajectory is implicitly expressed by change over time.

VT.6 **Cohort Relative Alignment Map**: A collection of heatmaps representing a common event feature stream. The event feature trajectory, frequency and duration can be identified for a population.

VT.7 **Contextual Bar**: A small heatmap that can be associated to a linear graph (VT.1) or a Cohort Relative Alignment Map (VT.6). This contextual bar can be made visible on demand using a selection interaction.

Once all of the above mentioned processes are completed, and necessary data quality prerequisites are met, the mPub instantiates the DVAM. Once instantiated, the DVAM exists independent to the mPub component until the session is terminated, after which the mart is deactivated.

(4) **Dynamic Visual Analytics Mart (DVAM)**: The instantiation is activated by the mPub component using parameters provided by a designer, and immediately upon activation the mart ingests real-time data streams. Hereafter the dynamic visual analytic mart independently sources, and remodels the data as the analyst begins to interact directly with the interactive component of the mart. The independent mart can then request for raw or secondary data from the event stream processing engine, and if required access historical sensor data from the data store.

Event Loop: A unique feature of the mart, as opposed to traditional marts in data warehousing architecture, is the dynamic nature by which it accepts interactive controls and submits instructions to the event stream processor for retrospective analysis.

In this subsection, the components of the TDVA framework were detailed. The subsequent subsection presents the methodology used to generate DVAM using the TDVA framework.

6.3 TDVA Methodology

The instantiation of the TDVA framework follows a series of phases collectively identified as the TDVA Methodology, and begins with phase one: finding the right person, phase two: categorization and data modelling, phase three: deployment of DVAMs, and phase four: evaluation of the DVAM. The TDVA Methodology is illustrated in Figure 46.

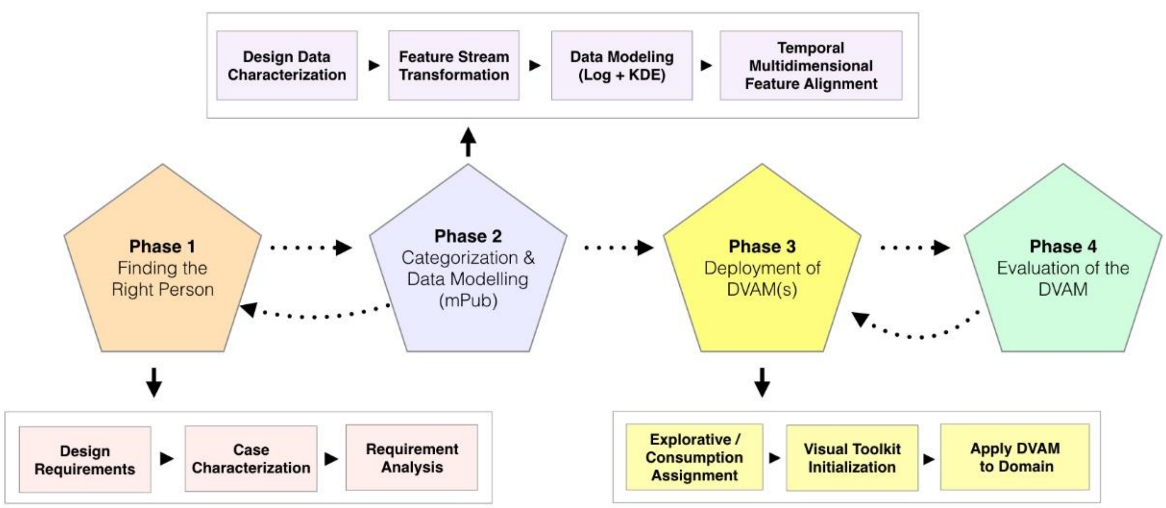


Figure 46: Tri-event Parameter Dynamic Visual Analytic Methodology

6.3.1 Finding the Right Person

Apart from instances where knowledge about the domain is well documented, there are cases in which the domain needs to be better understood. To support the generalization of the TDVA framework, the TDVA methodology includes an important initial step called ‘Finding the Right Person’. In this first phase of the methodology, important information about the context and the user are isolated prior to the instantiation of DVAMs. To that end, a series of requirement gathering methods are available, including ethnographic observations and surveys such as the one performed in the previous chapter relating to clinical handover. The users are primarily observed to identify the extent to which the user requires exploratory or consumption displays. These observations help inform task requirements, as well as design goals that are required modules for the dynamic visual analytic marts.

Following that step is the case characterization, in which the domain itself is carefully studied, and their related workflows observed. Completing both steps result in a list of design requirements. Information requirements are then analysed from prior observations. The produced list of information requirement is made available for subsequent visual encoding that will transform information into visual representations. Moreover, observations may also reveal individual user preferences, which can provide additional knowledge about deploying the appropriate VTs effectively for analysts with varying needs for detail.

6.3.2 Categorization and Data Modeling

The second phase is to extract appropriate data as defined in the case characterization to input as design requirements to the mPub engine. In domains where data streams are analysed in real-time, this step may also involves making available secondary-data from existing online algorithms. These algorithms have been modified as to extract key secondary-data regarding the temporal parameters relating to the identification and detection of features. This secondary-data is made available to the mPub engine so the appropriate feature scaling density, and data transformation algorithms can be executed to match the selected VT.

Secondary-data from the algorithms also provides unique trajectory information. For instance, every occurrences of an event produced by the online algorithm can be used to determine potential trajectories. This involves identifying data that populates the tri-event parameters. Specific data transformation and statistical analysis techniques such as kernel density estimation (KDE) and non-linear non-parametric estimation are used by the mPub engine to strengthen and enhance the dynamic visual analytic mart’s ability to communicate inferences on existing and potential trajectories.

Moreover, depending on the requirement analysis conducted in the first phase, the user may require an exploratory or explanatory analytic environment. Before the DVAM is deployed, the mPub engine produces event alignment in data. Generally, event alignment is an important aspect in explanatory research, as the user intends to study sequences, and patterns exhibited by low-level event features that precede certain high-level contextual events.

The final component of this phase will be to reveal temporal relative alignment of raw, variability, event and classifications. In data intensive sensor networks, this means temporal relative alignment of events must exist across multiple data streams. The temporal domain of data can exist over multiple hierarchies, such as, detection of spells in seconds, calculation of heart rate variability in hours, and gestational age in days and weeks. Event alignment and depth of overview as required in the operation of the dynamic visual analytic mart may result in creative development of visual displays, and also result in novel method for conveying sequences of events using coordinated visual displays. In cases where frequency and duration require emphasis, the use of novel methods such as combinatory logistic and KDE modeling can be used to expose kernel centres in data points and produce anchor points for the relative alignment. Following the completion of this phase, a DVAM is created and ready for deployment.

6.3.3 Deployment of the DVAM

Once both phases are complete, the DVAM can be deployed as an exploratory or consumption tool. The DVAM is generated through design goals and criteria that are gathered and documented by the designer in the first phase. The difference between the exploration and consumption instances lies in the degree of exploration and manipulation the user can perform. The resulting DVAM allows the researcher to partake in knowledge discovery, and to modify the underlying data model to produce new views. This phase of the TDVA methodology may require several iterations until the optimal DVAM is produced. An optimal DVAM is met when a DVAM satisfies all requirements established in the first phase of the methodology. This phase of the methodology provides flexibility on part of the designer to develop multiple instances of the DVAM to evaluate each instantiation based on criteria developed in the first phase. For instance, if among the criteria of the first phase was improving detection time, then the DVAMs will be iterated until an instance has been developed which improves detection time against tools used in the existing workflow.

6.3.4 Evaluation of the DVAM

The final phase of the TDVA methodology is the evaluation of the DVAM utilising domain experts to validate its utility with a complex domain problem. This phase is typically performed when the DVAM is developed and instantiated in the domain by a usability researcher. Task requirements identified in the first phase server as the evaluation criteria. Using the generated criteria, a list of evaluation tasks can be devised to identify key outcome metrics as appropriate to the environment where the DVAM will be used. These outcome metrics can be used to identify limitations in the DVAM. Should limitations prevent the successful completion of the task, the results obtained in this phase can be used as input in a subsequent iteration of the DVAM creation (arrow to phase 3). Subsequent iterations result in a DVAM that addresses the requirements of the complex domain expert.

The four phase methodology presented in this section provides a means of instantiating DVAMs within the context of a specific use case. The next subsection describes components of a platform design that can be used to support the DVAM instantiation.

6.4 TDVA Platform

This section details an instantiation of the TDVA framework as a platform known as the TDVA platform illustrated in Figure 47. The entry point of new data within the TDVA platform begins with integrating data streams collected by sensors, and parallel multi-dimensional feature extraction and classification algorithms deployed in the ESP engine (second from left, Figure 47) and this function aligns with the ESP engine of the TDVA Framework.

Once features are generated, the Dynamic Visual Analytic mPub engine (middle, Figure 47), subsumes the responsibility of the mPub engine described in the TDVA framework. This engine runs at compile time for every unique instance of a DVAM created by the user (right, Figure 47). A user can create multiple instances and use them for hypothesis generation or for hypothesis testing between the same or different system. In this way the platform allows the user to perform both exploratory and explanatory research. The DVAM is tightly bound to the ESP engine, hence, if a user requests to see retrospective analysis of data streams the instruction is sent to the ESP engine to dynamically remodel the dataset using new parameters determined by the analyst. Each component is described in detail below.

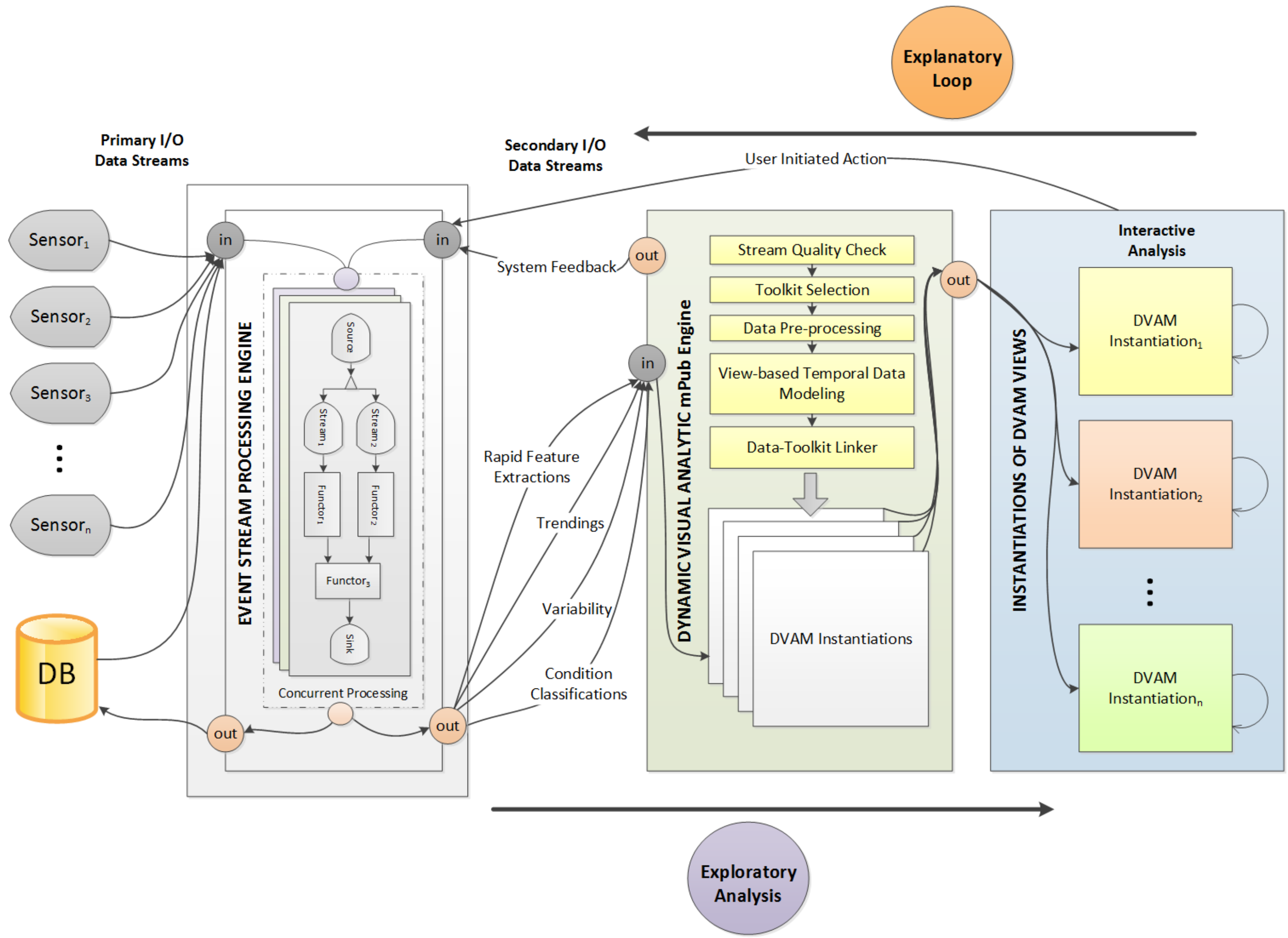


Figure 47: TDVA Application Platform

6.4.1 Event Stream Processing Engine

The first container labelled "Event Stream Processing Engine" found on the left in Figure 47 performs basic data analytic and event stream processing functions thereby, producing one of four essential algorithmic outputs. These algorithmic outputs are determined by the underlying event stream processing algorithm and align with the output of the ESP engine in the TDVA framework. The outputs generated by the ESP engine in the platform are:

O.1. ***Rapid Event Feature Extraction***: Consisting of high velocity asynchronous output streams that describe low-level characteristics of the streaming data. Examples include, data quality streams measuring the quality of data received from sensors, intermediate event streams produced before a classification can be determined, and finally system environment parameters such as the health of the engine.

O.2. ***Trending***: Consisting of synchronous streams describing temporal trends that occur within or between multi-stream classifications, such as a single data stream describing a real-world system that has breached absolute or relative thresholds thereby signalling abnormal conditions.

O.3. ***Variability***: Consisting of synchronous streams that describe the degree of variability that exists within single or multiple series of data streams of critical systems.

O.4. ***Event Classifications***: Consisting of asynchronous streams that output classifications pertaining to real-world conditions through real-time event stream processing. The output is often delivered as a single ordinal value, such as the name of condition detected.

The engine also accepts two inputs.

I.1. ***System Feedback***: The dynamic visual analytic engine, the next component in the architecture can provide requests for the execution of algorithms required for the generation of specific visual analytic templates.

I.2. ***User Initiated Action***: The instantiated dynamic visual analytic mart can independently request the event stream processing engine for re-execution of algorithms with user defined parameters. The output from the engine is delivered directly to the respective template. This supports the 'Explanatory Loop' of the TDVA framework. The explanatory loop, and the user initiated action are both used to support continuous hypothesis generation tasks.

6.4.2 Dynamic Visual Analytic mPub Engine

Once the ESP has started generating features, the mPub engine, (Figure 47, middle) begins to prepare a DVAM instantiation. Each mart is allowed to be flexible in order to accommodate dynamic data ranges, baselines, and frequency of events. For instance, if the frequency of one class of events overwhelms all other events then a non-parametric normalization is performed to enable the analyst to see those events. In addition to this custom data dimension modelling, the mPub component also:

P.1. Prepares the mart for instantiation by customizing the visual encoding of the visual analytic unit, such as accommodating for users screen width and height, thus preventing obstructed views.

P.2. Performs independent instantiation and establishing uninterrupted session with the ESP for the mart to receive data streams and enable the user to perform dynamic visual analytics.

6.4.3 Dynamic Visual Analytic Mart

Each dynamic visual analytic mart instantiation (Figure 47, right), once it has been deployed, has direct access to the ESP from where all sensor data and retrospective data are retrieved. The mart is an independent instance, which means that the mart exists within its own local scope, therefore the analyst can dynamically interact with the mart to elicit novel features and visually analyse the data without imparting changes on other active sessions. This feature enables the user to perform dynamic hypothesis generation actively across one or more instances of the same analysis.

6.5 Sample Applications: Prototype Instantiations of the TDVA Framework within the TDVA Platform

Using the TDVA platform, two early prototypes of visual analytic tools were developed and integrated to the platform. These prototypes include the Heart Rate Variability Graph [275], and the Neonatal Spells Explorer [276]. These prototypes at minimum, received data from the ESP engine or data warehouse, provided interactive functionality to the end user, supported expression of the tri-event parameters, applied degree of interaction based on the requirements for exploration, and supported automated data modelling tasks. The sample applications are described with a walk-through of the framework. These prototypes served as a proof-of-concept demonstration for the full implementation that resulted in PhysioEx and Co-Rad which are detailed further in the following two chapters.

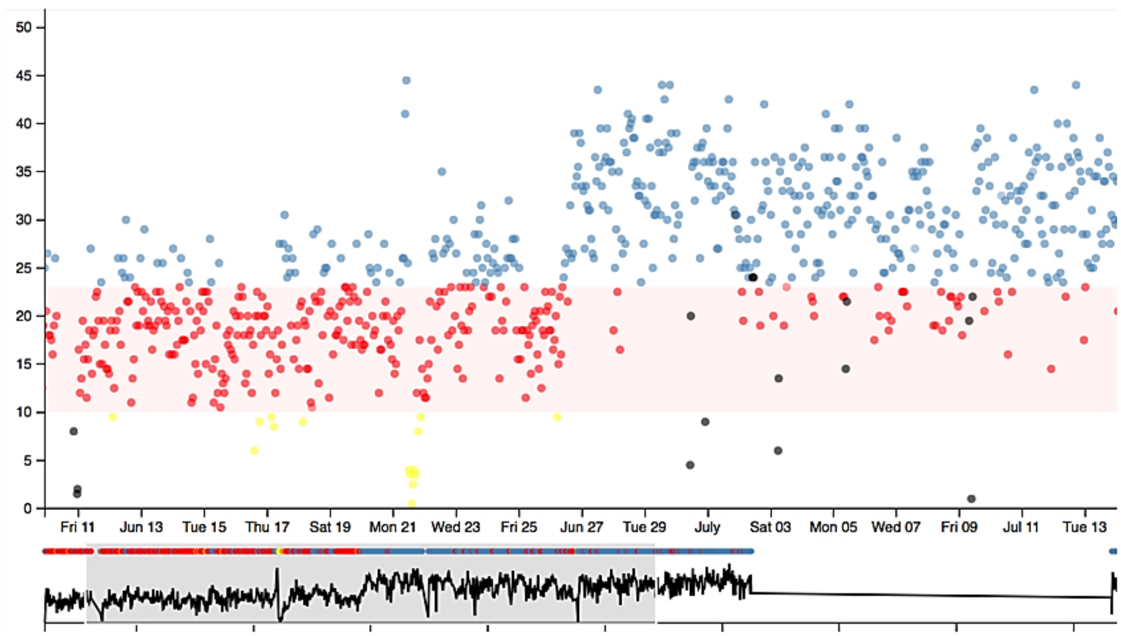


Figure 48: The Heart Rate Variability Graph- real-time visual analytics for the consumer

6.5.1 Heart Rate Variability Graph

The first prototype DVAM that was instantiated using the mPub engine of the TDVA platform, was the heart rate variability graph [275]. In order to visualize heart rate variability previously, the user had to perform numerous manual processes, including retrieving the data from the data warehouse, populating it to Microsoft Excel or IBM SPSS before a visual can be made available. The goal of this instantiation is to make heart rate variability available in real-time to support the workflow of a consumer. The prototype DVAM, illustrated in Figure 48, was developed to support a consumer with real-time knowledge requirements. Each heart rate variability score is represented by a circle and visually encoded using colours. A yellow colour signifies low scores, while a red colour denotes acceptable regions, and finally the blue circles represent consistent variability.

The entire technical architecture, including the software utilized in the prototype is illustrated in Figure 49. IBM InfoSphere Streams was used as the ESP, and physiologic data was streamed into the ESP from the patient monitor using CapsuleTech's DataCaptor software. The ESP would then stream data to a node.js application that served as the mPub engine. The mPub engine could also retrieve retrospective data from the persistent data store using the ibm_db library. The mPub engine was coded in JavaScript. The mPub engine would then generate a visual representation that was coded using the d3.js library [277].

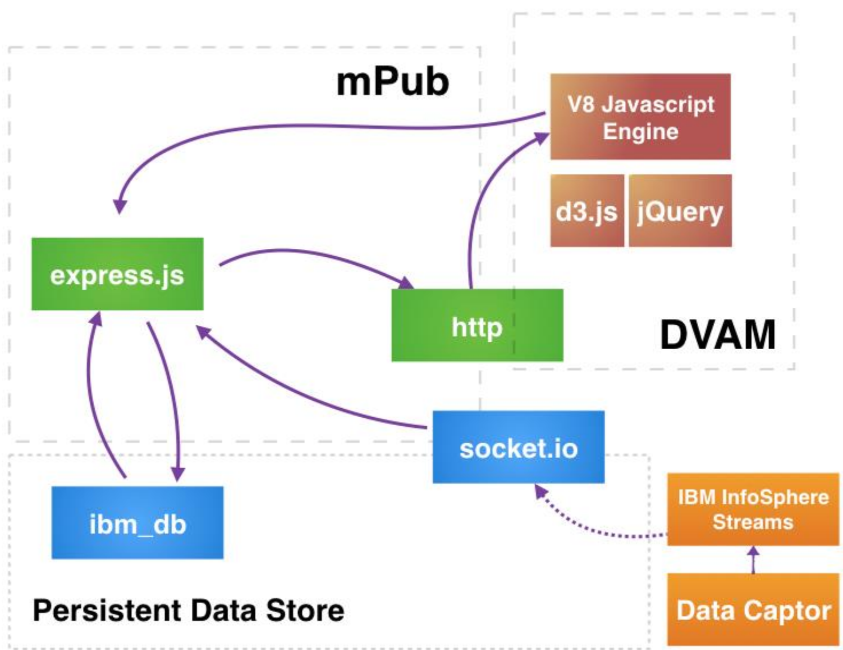


Figure 49: Components of the TDVA platform used to provide real-time analysis feeds.

Once the requirements were established, and the supporting platform identified, the process of identifying the correct visual toolkit for the DVAM that would best support visual consumption of heart rate variability trends was initiated. Using the TDVA methodology, the design processes were organized into four phases. The sequence of steps followed are described below.

P1: Finding the right person

The first phase involves processes soliciting design requirement that identified specific needs of the end-user. Two domain experts were identified with a minimum of 10 years of neonatology expertise as the subject of study. The tasks they performed were then noted, which included (1) performing heart-rate variability temporal abstractions on raw heart rate data; (2) extracting the temporal abstractions to a file; (3) importing the file to Microsoft Excel or IBM SPSS to visualize the data.

P2: Categorization & Data Modelling

The second phase of the TDVA methodology involves categorization and data modelling. In this phase the data modelling was performed by an event stream processing algorithm enacted in the ESP engine. That algorithm consumed real time streams of heart rate values and generated a variability score between 0 and 60, every hour. The score represented the actual number of minutes of high variability. The zero value represented no minute in that hour that had periods of high variability, while 60 represented full variability.

P3: Deployment of DVAM(s)

In this step, the user is initially presented with a prompt that involves selection of a consumption or exploration display type, here the user selects consumption. In a subsequent prompt presenting a list of data sources, the user selects heart rate variability. Following the enactment of the heart rate variability algorithm in the ESP engine, the mPub uses the meta-data retrieved from the user prompt to instantiate a DVAM called the heart rate variability graph. The heart rate variability graph contains several interactive options, including the ability to click on a circle to reveal raw data. Since consumption display type was selected, only the selection and zoom interactions were enabled.

The Linear Plot Graph (VT.1, page 140) was selected as the visual toolkit in this deployment. Four additional visual encoding strategies were applied to the visual toolkit to support the specific tasks identified in phase one. The first strategy included the ordinal categorization of heart rate variability abstraction scores based three levels of risk: low, medium, or high [278]. Scores that belonged to the medium risk were marked with a red fill. Scores above that mark were filled blue, and low risk, filled yellow. The medium threshold was double encoded as a ‘safe zone’ by appending a red rectangle to the canvas. Circles appearing in that region were marked as safe, and hence noncritical. An absolute opacity score of 0.8 to highlight regions where multiple circles were overlapping was applied. Finally, all invalid abstractions, were marked i.e. without at least 90% of the data with a black fill.

P4: Evaluation of the DVAM

The final phase of the TDVA methodology was the evaluation of the DVAM. The mPub engine, after receiving instructions from the user generates the visual displays and directs the user to the d3.js render. The heart rate variability graph was verified as a prototype using two mutually exclusive domain experts with at least 10 years of experience in neonatology. Participants provided anecdotal evidence, including improved satisfaction and reduced task completion times with the display over alternatives. However the prototype had limitations, due to the slow changing nature of the visualization, it may be difficult for consumers to identify any slow developing abnormal reading. Moreover, the yellow colour for low heart rate variability can be difficult to discern against a white background. Hence, using the knowledge gained in the first deployment, a secondary deployment can be pursued to correct those limitations.

6.5.2 Neonatal Spells: Sequence of Events

The TDVA methodology was further utilised to instantiate a second DVAM, called Sequence of Events (SeqEvent), for exploring primitive event sequences found in multidimensional neonatal spells events. The TDVA platform components used in SeqEvent are illustrated in Figure 50. Components in this figure demonstrate much greater communication with the front-end tool. In contrast to the Heart Rate Variability Graph, the SeqEvent graph used Redis [279], an in-memory NoSQL database for rapid recovery of short term data. The communication is supported by a web socket connection from express.js, part of the mPub engine, to the visual analytics tool for dynamic and responsive interaction functionality. The processes completed as defined by the TDVA methodology are described below.

P1: Finding the right person

Three clinical researchers, with at least five years of experience using physiologic data were consulted through semi-structured interviews. The researchers identified difficulties in analysing large volumes of cardiorespiratory sequences that was produced by an algorithm developed to analyse neonatal spells throughout an infant's care [32]. The analysis had to be performed manually, often using Microsoft Excel spreadsheets. Identifying specific sequences were difficult, as the manual process often meant key data was mis-entered or mis-analysed.

An opportunity was identified to instantiate a DVAM that can address the exploratory needs of the researchers.

P2: Categorization & Data Modelling

The second step involved passing parameters to the mPub engine to initiate the categorization and data modelling for a prototype DVAM, with rich interactive support. A large part of the categorization and data modelling phase involved connecting to the Persistent Data Storage (lower, Figure 50), to collect retrospective data as required for each patient recruited in the analysis. The dataset selected for this display was the neonatal spells - event sequences dataset.

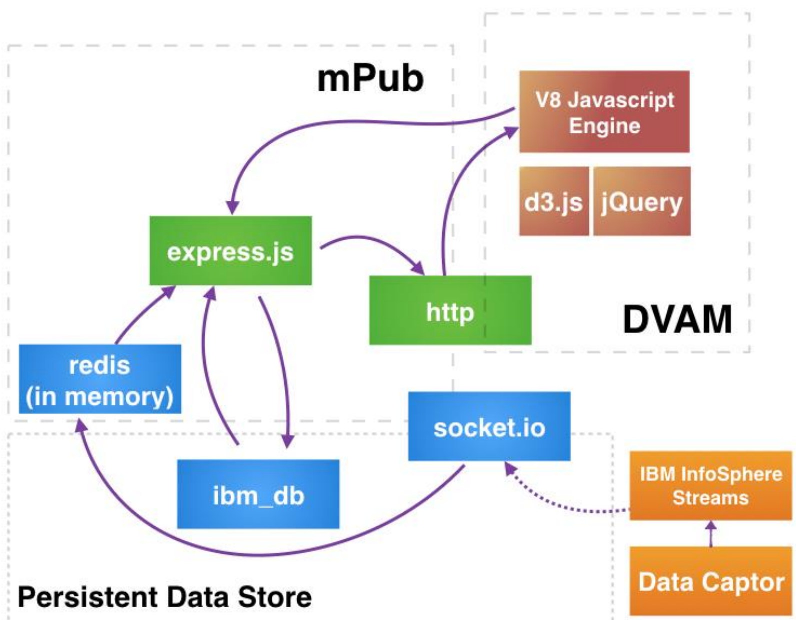


Figure 50: The TDVA platform components supporting the SeqEvent visual analytic tool

The mPub engine after processing the request from the user, generates an instance of SeqEvent and makes it available for the user. The SeqEvent tool allowed clinical researchers to identify all potential pathways that was detected by the algorithm. The algorithm produced up to six multidimensional event classifications (O.4, §6.4.1): central, vagal, obstructive, obstructive-central, central-obstructive, and unclassified; and two primitive trending streams (O.2, §6.4.1): bradycardia and desaturation.

P3: Deployment of DVAM(s)

The SeqEvent DVAM uses the SeqEvent visual toolkit (VT.2, page 140), which extends concepts from a visual analytic representation called Parallel Coordinates, where each axis becomes an selection object that can be used to highlight a subset of the dataset. The Neonatal Spells algorithm was executed in Artemis, which produced a collection of event classifications [280]. These event classifications describe a sequence of primitive events. However, the algorithm outputs only a string of the multidimensional event classification. Analysts are able to access the primitive event sequences, however, this data is unavailable without accessing meta-data, stored as text. The SeqEvent DVAM, illustrated in Figure 51, was instantiated to simply these tasks and to support exploration of primitive sequences in multidimensional events detected by the neonatal spells algorithm. Each output of the algorithm contained up to three sequences that were observed across various physiologic streams. The coordinate stream1 in (middle, Figure 51) represents the first primitive event, followed by up to two additional sequences across different physiologic streams. A duration coordinate is provided to give additional context about the duration tri-event parameter. Opacity is controlled to highlight frequency. Trajectory is implicitly observed using the "`starttime`" coordinate and duration together.

The NA coordinate acts as a null event, hence informing the researcher that the sequence had terminated.

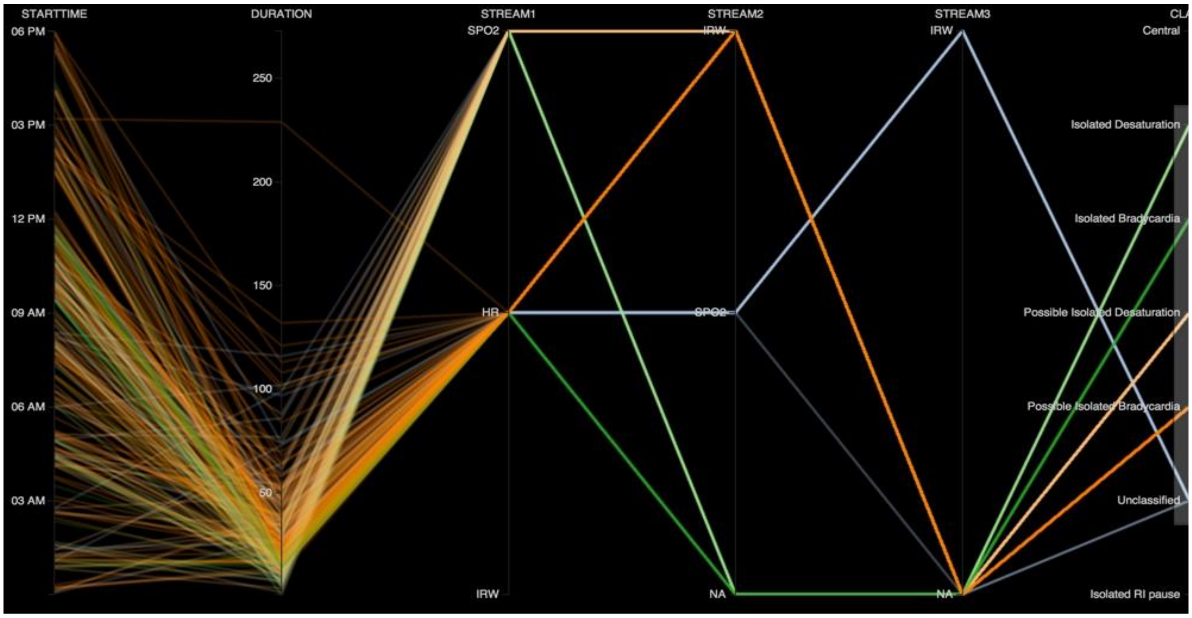


Figure 51: SeqEvent- analysis of primitive sequences in multidimensional events

P4: Evaluation of the DVAM

Once the DVAM was instantiated, the final phase of the TDVA methodology commenced. In the verification of the prototype, a clinical researcher used the SeqEvent DVAM to analyse 7828 primitive events that were extracted from 3248 instances of the six potential multidimensional event classifications. While the sequence of all events were generally well understood, the unclassified output was much more ambiguous. Hence, SeqEvent provided the researcher a novel method of exploring unknown sequence pathways in the unclassified output. Figure 52 illustrates a clinical researcher using SeqEvent to explore all potential sequence paths that led to an unclassified output. The researcher performs this simply by brushing over the '`Unclassified`' coordinate in the '*CLASS*' axis (Figure 52- far right). Immediately all other classifications are removed from the canvas and three unique sequence pathways emerge. The researcher refers to the `starttime` and `duration` coordinates, and finds that longer duration events (>100 seconds) reflect a two stream sequence, with primitive events are observed in the oxygen saturation stream (SpO_2) before breathing (IRW).

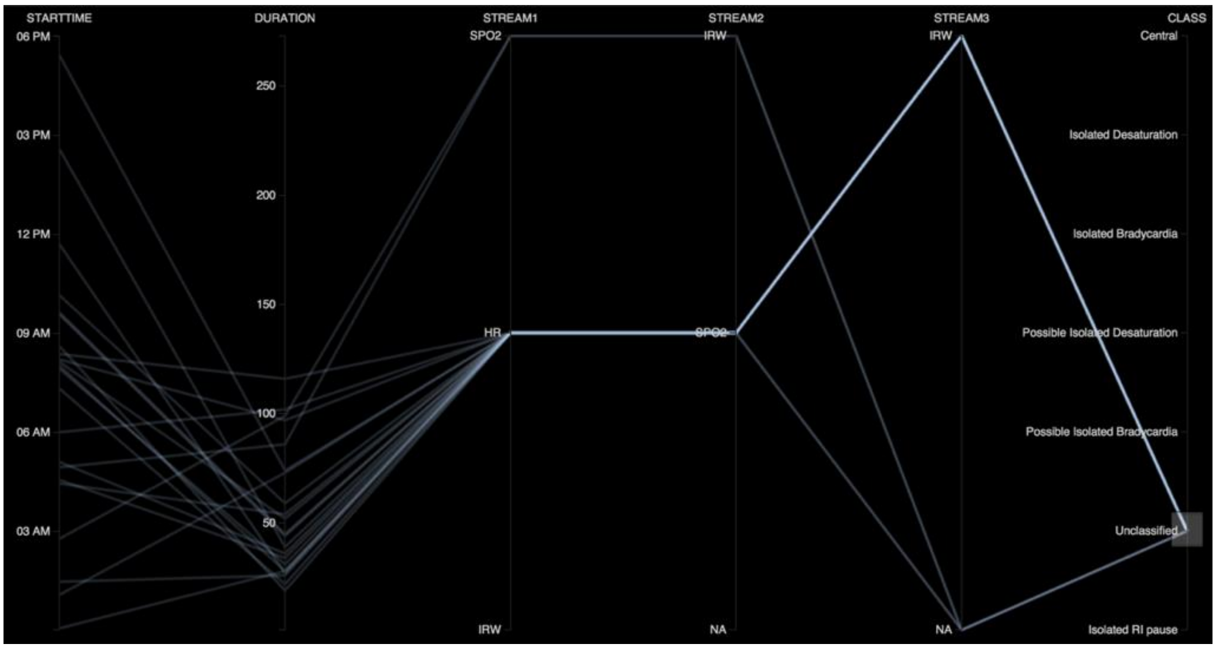


Figure 52: SeqEvent- a researcher exploring three unique sequences in unclassified.

SeqEvent DVAM presents a unique use case of the parallel coordinate visual representation that was applied to sequences in data streams. The visual analytic tool is shown to support the tri-event parameters, and deliver interactive exploratory functionalities to a researcher while employing the TDVA framework.

6.6 Chapter Summary

This chapter outlined the limitations of the data warehouse paradigm in §6.2 and then presented as alternatives, the TDVA framework (§6.2), methodology (§6.3), and a platform design (§6.4). The TDVA framework offers an alternate means of creating interactive visual analytic toolkits, referred to as visual marts. The TDVA methodology presented a four phased approach to instantiate dynamic visual analytic toolkits for the researcher or the consumer. The TDVA platform presented an architecture that supports both a real-time implementation and a retrospective implementation of the TDVA framework. The chapter ends with a demonstration of two prototypes that were developed through the application of the TDVA framework and methodology.

The subsequent chapters will further expand the designs introduced in this chapter through full implementation and evaluation of visual analytics solutions addressing the clinical challenges identified earlier in this thesis (§2.1).

7. PhysioEx: A coordinated visual analytic tool for physiologic data streams

This chapter presents material interpolated from a publication that was accepted and in press [44]. The publication was co-authored by Christopher Collins, Carolyn McGregor and Andrew James. The author of this dissertation designed, developed, and evaluated PhysioEx. Christopher Collins, Carolyn McGregor, and Andrew James provided input on the design of PhysioEx and write-up of the publication. An additional content, §7.3, does not appear in the publication, and is presented in this chapter to aid the reader in details involving the instantiation of PhysioEx.

7.1 Introduction

This chapter presents the creation and evaluation of PhysioEx which is an instantiation of a TDVA mart to support hypothesis generation. The creation and evaluation of PhysioEx followed the TDVA framework and methodology. Before a hypothesis can be generated, clinical researchers elicit knowledge from multidimensional streams of physiologic data by isolating features and analysing behaviours that may predict the early onset of clinical conditions. Conducting a hypothesis generation task in large volumes of physiologic data is a complex undertaking and require manual siphoning of raw waveform data. To address some of these challenges, a novel visualization technique called the Temporal Intensity Map (TIM) was developed. TIM reveals critical information about the frequency, duration and trajectory of streaming events generated by real-time event stream algorithms. A event-stream algorithm was developed by Thommandram et al. [32] that produces event features and classifications in real-time. The visualizations utilize these output to highlight salient temporal features that may assist the user in generating hypotheses about physiologic behaviour. A unique representation of the bubble chart, named the Sequence Graph for identifying high level periodic patterns is also contributed. Finally, methods are presented to highlight three salient temporal properties called the temporal tri-event parameters that include frequency, duration, and trajectory.

In a preliminary study of domain experts using PhysioEx, participants detected correlations between low-level event features and high-level event classifications, identified salient features in the physiologic data streams that illustrate the infant's cardiorespiratory health, and deliberated over the presence of infection by carefully studying physiological trends. These findings are valuable in the face of a current lack of tools available to perform deep insight analysis of physiological data as identified in chapter 4. The study was approved by the Research Ethics Board at our institutional REB, and all patient data was de-identified. One of the goals of that study was to determine whether neonatal sepsis was present at the time of suspicion of infection at the bedside.

The research contributions of PhysioEx are as follows:

I. *The Temporal Intensity Map (TIM) visualization technique for frequency, duration and trajectory of events.*

II. *The PhysioEx dashboard of coordinated views including TIMs, sequence graph, linear graph, and streams graph.*

III. *A case study of PhysioEx with NICU clinical researchers.*

In the remainder of this chapter, the background of the specific problem domain will be presented, followed by related work, design requirements, design of PhysioEx, preliminary user study, discussions and ending with conclusion.

7.2 Problem Characterization

Neonatal sepsis, a form of nosocomial infection, is a life threatening condition that is difficult to detect and for which early detection significantly improves mortality [281]. Apnoea is condition that is defined as a pause in breathing for 20 seconds or more [282]. The term neonatal spells is commonly used in NICUs for cardiorespiratory events that may include pauses in breathing, fall in heart rate, or fall in blood oxygen saturation [32]. An increase in frequency of spells may be associated with neonatal sepsis. A research study by Moorman et al. [283] reported a potential association between reduced heart rate variability and increased bradycardia in the hours prior to the clinical suspicion of neonatal sepsis. Other studies have also linked the presence of sepsis with heart rate characteristics, especially reduced heart rate variability and bradycardia [35], [58].

PhysioEx is contributed as a tool enabling the end-user to explore neonatal spells event classifications produced by the real-time data stream algorithm around the time of suspicion of neonatal sepsis. By exposing novel neonatal spells event classification information, juxtaposed with the relatively aligned time of suspicion of neonatal sepsis, we provide clinical researchers with an expressive tool to support their analysis and hypothesis generation.

7.3 Instantiation of the TDVA Framework

The PhysioEx DVAM serves as an instantiation of the TDVA framework applied to Artemis. Artemis, an online analytic platform for physiologic data streams, produces numerous physiologic events (PEs) in real-time, a PEs can include one or more primitive, complex and multi-dimensional events. Each PE describes one of several features, such as instantaneous drifts from a baseline, or in the case of multidimensional PE a temporal sequence across different physiologic streams [32]. PhysioEx primarily supports the researcher operating in a high engagement, and low urgency environment within the Exploration-Consumption Continuum (Figure 42). The researcher uses a coordinated dashboard, along with a backend database system to interactively analyse PE details on demand.

The instantiation of TDVA framework begins with the output of multi-dimensional physiological event classifications generated by Artemis to the Dynamic Visual Analytic Mart Publisher (mPub) engine (Figure 53). The mPub menu in Figure 53 illustrates an options list based on a request submitted by a researcher. The researcher has identified the requirement for an interactive view, this is denoted by the filled box next to the 'Researcher' option. This satisfies the first requirement of the TDVA methodology, "Finding the Right Person". Once the researcher option has been selected, the mPub component is made aware of the need for interactive functions, such as selection, filtration, and detail on demand to be made available to the user. In Figure 53, the researcher has selected the options for a coordinated view, and to generate views that contain exploratory functions, including the ability to select and zoom into areas of interest, highlight, and expose details on demand. The second step in the TDVA methodology is categorization and data modelling. To initiate this task, the mPub prompts an additional menu that solicits details about the required data source, visual representations (TIMs, {sequence, linear, streams} graphs, standard heatmap, and context bar) for the analysis task.

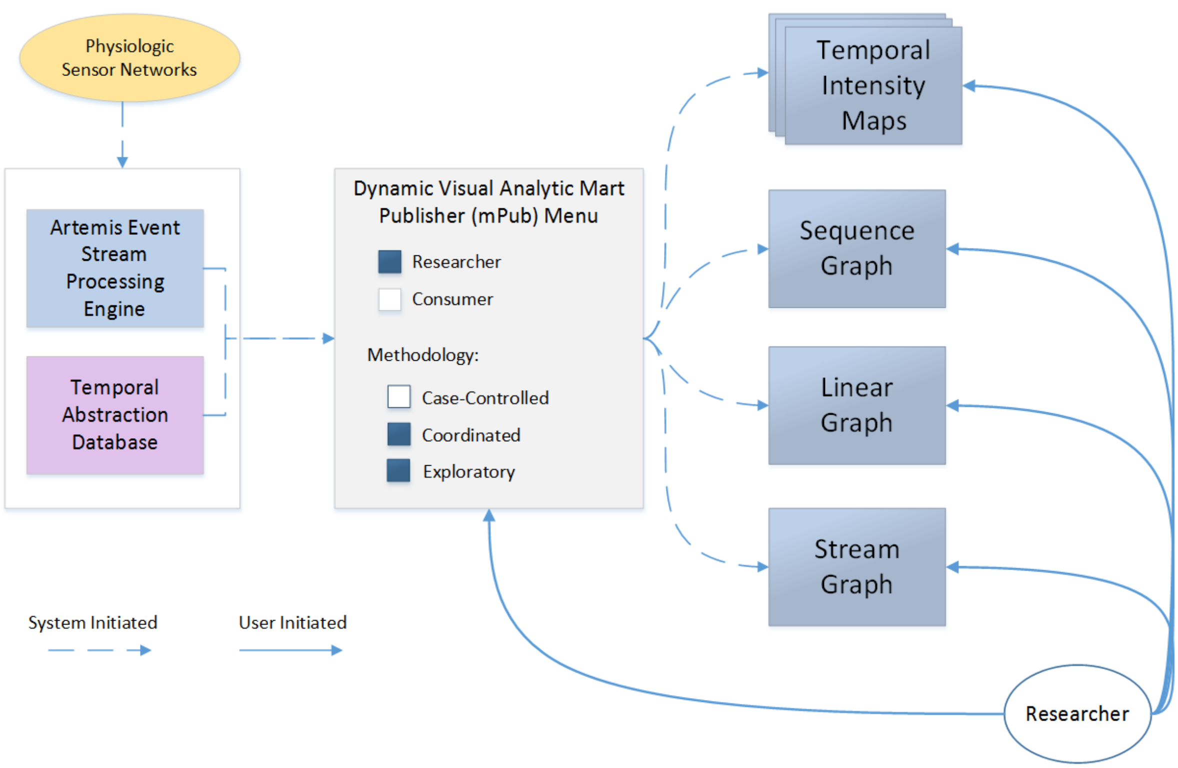


Figure 53: Instantiation of PhysioEx using the Dynamic Visual Analytic Mart Publisher (mPub)

Once the request has been submitted by the researcher, mPub assumes the responsibility of performing necessary checks on the data for sufficient data quality and performs data modelling to create relative temporal alignments. In this component additional PE features are generated to support each of the TIM and event classifications views. Once the data processing stage has been completed, an array of views, including three TIM views, and three event classification views are generated, satisfying the final step of the TDVA methodology “Deployment of the DVAM”.

The neonatal spells algorithm was executed against the retrospectively stored raw data for a total of 47 patients who had sufficient data quality and clinical data, which generated PEs that were saved to a database in real-time and used in this work. All output physiological event features and event classifications were saved to a database in real-time. All algorithm generated secondary data were stored to a common table and labelled by the type of abstraction. All instances of each abstraction type included duration of a sequence; duration was measured as the total time since the onset of abnormal shift in the first signal until the recovery or stabilization of all signals in that sequence.

7.4 Task Analysis

We asked three domain experts to describe specific tasks they currently perform to predict physiological behaviours prior to the point of suspicion of infection (PSI). The common tasks were:

T.1. *Identify the PSI*: The researcher uses the PSI as an anchor for subsequent analysis.

T.2. *Identify PEs in the respiratory physiologic signal before PSI*: PEs having breathing pauses greater than 20 seconds were noted and associated with neighbouring clusters.

T.3. *Analyse PEs across heart rate and SpO_2 data streams*: Heart rate signals and blood oxygen saturation signals are analysed to determine downwards shifts before the PSI.

T.4. *Identify abnormal PEs*: Abnormal PEs are flagged and sometimes investigated to verify algorithm accuracy.

T.5. *Create mental temporal map of underlying physiology*: Information gathered from all previous steps were used determine a hypothesis about the presence of infection.

Supporting these tasks was the primary design goal of PhysioEx.

7.5 Design of PhysioEx

PhysioEx is illustrated in Figure 54, and consists of three groups of views: three TIM views; the sequence graph, linear graph, and streams graph; and three raw data views. The interface was developed using D_3 [277]. In this section we explain each component in detail.

The first group of views, namely the Respiratory Pause TIM, Heart Rate Flux TIM, and the SpO_2 TIM provide the user with the ability to rapidly analyse behaviours in event features stream. The second group of displays assist with analysing event classification data. A third view, when activated, provides the user with deeper contextualization by providing raw data that would be observed at the bedside. We mark the canvas with a red cross. This red-cross indicates that a blood result was obtained after a physician suspected the infant of having infection. We do not show whether it was positive or negative to allow the researcher to use this position marker to conduct explanatory research for generating hypothesis about the onset of infection.

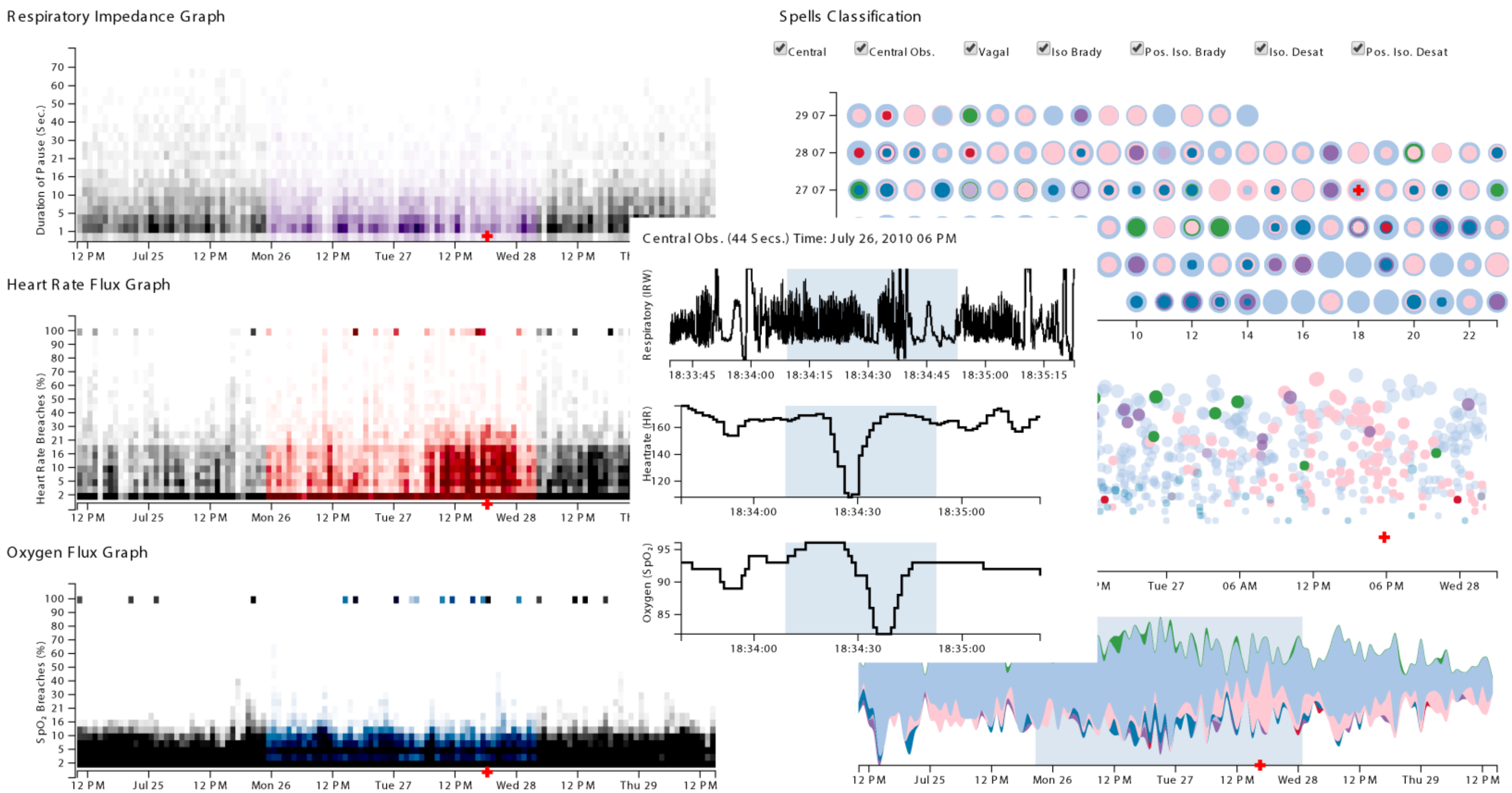


Figure 54: PhysioEx is a visual analysis tool for event stream analysis of multiple streams. Several Temporal Intensity Maps (left), in the coordinated dashboard reveal the duration, frequency, and intensity of physiologic data over time, alongside a selected raw data display (middle), and three visualizations (right, top to bottom): a sequence, linear, and stream graph.

7.5.1 Temporal Intensity Map View

Each TIM provides users the ability to rapidly discern subtle behaviour in streaming data. We employ a novel use of the heatmap visual encoding, where positions along the vertical axis represents an aspect of an event's nonlinear *critical distance interval*, such as duration of breathing pause. It is termed a critical distance interval, because it helps determine the PE's severity. PEs are aggregated into critical distance interval bins as determined by the density estimation function. Hence, durations with smaller values are represented at the bottom of the graph while larger durations appear near the top. The horizontal axis represents temporal range of the dataset. A red cross is placed where a nominal clinical event (e.g. PSI) exists, to support task T1.

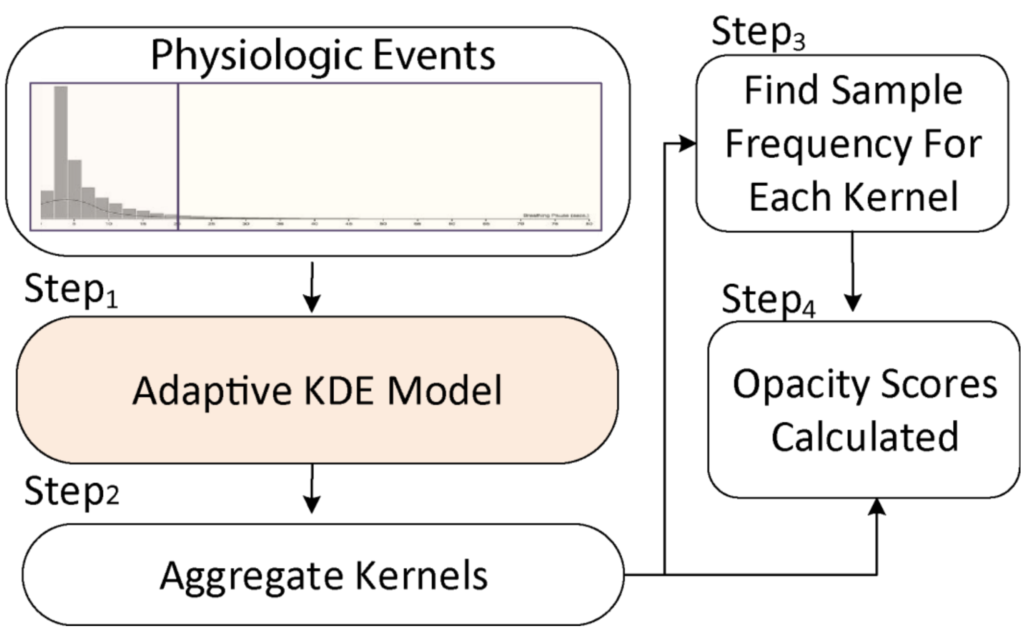
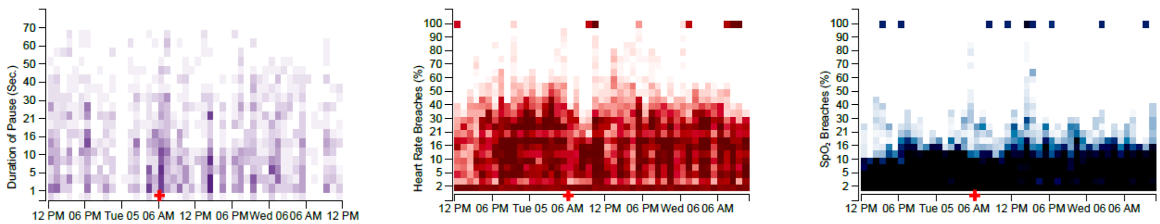


Figure 55: The four-step method of constructing the Temporal Intensity Map beginning with identifying kernels.

In order to support task T2 and T3, we contribute a combined adaptive bandwidth method of vertical binning, using the KDE generated probability density function (pdf) as illustrated in Figure 55. We began the process by calculating the KDE pdf for the entire dataset (Step 1). We utilized the `scikit-learn` to implement the density estimation [284]. The top-hat kernel form, an alternative to Gaussian, was selected as this kernel form involved less smoothing which produced more kernels. The width was also made narrow, and set to a value of 0.2. These two modelling decisions increased the likelihood of kernels identified in the heavy-tail of the distribution. All PEs were then aggregated into hourly sets (Step 2) and reduced to produce sample frequencies for each kernel (Step 3). The binning produces a two-dimensional array of PE critical distance interval sums ranging from 0 to N, where N is the furthest critical distance interval. The value of each element in the array are used to encode opacity.

The visual encoding of the TIM is a heatmap controlled for hue and opacity. The hue indicates the PE classification and is metaphoric: red for heart rate which evokes the colour of blood, and blue for desaturation of oxygen, due to blue-like colour of the skin when oxygen levels fall. The hue selection supports T2--T4, in which one must rapidly associate PE type. The opacity is controlled by the frequency value. The width is controlled by available space of the canvas, divided by the temporal range.

Where there are significant number of samples found in a particular kernel, the opacity score of each is reduced, and where the frequency is low the opacity is increased (Step 4). Thereby, events appearing in low-frequency kernels, such as in the heavy-tailed portions, are represented with increased visibility. These heavy-tail events, such as an extended breathing



(a) Respiratory Pause (b) Heart rate flux (c) Oxygen saturation (SpO_2) flux

Figure 56: Temporal Intensity Maps, compact visualizations for gaining rapid situational awareness of low-level behaviours in data streams.

pause lasting several minutes are clinically significant and warrant increased visibility. Moreover, this method effectively addresses the requirement of highlighting outliers involved in the task T2 and T3, where a constant opacity score would have otherwise excluded them from view. The temporal trajectory of the health status is visually elicited from observations made on each distinct view generated by the encoding. As rectangles with varying hue are appended along the horizontal temporal axis, the user is able to visually glean information about ongoing changes in the physiologic signal. Finally, we considered the use of bar graphs as an alternative design, due to their familiarity. However, that encoding was not appropriate for illustrating all three temporal properties without creating visual clutter. Due to the nature of our dataset, the TIM encoding was more appropriate for identifying both frequency (dense areas) and duration (vertical dimension) along a temporal axis.

Figure 56 illustrates three uses of TIM, beginning with the respiratory pause map (Figure 56a), displaying data in the form of duration of breathing pauses between 0--80 seconds. In this dataset intermittent clusters of breathing pauses are seen throughout the entire duration.

Breathing pause durations are also seen extending to pathological ranges above 21 seconds. The heart rate flux (Figure 56b) illustrates a measure between zero variability (0%) to high variability (100%) in heartrate. A sliding window sampling approach is used to compare the instantaneous heart rate every second against the average of the previous 30 seconds. The percent change is calculated and a block is added to the TIM at the appropriate height, if the heart rate reduced (bradycardia). In this chart, clinical researchers would be looking for repeated occurrences of severe bradycardia (high percentage change), or periods of low overall variability (high density low on the TIM). Both are indicative of pathological status.

Figure 56b shows a region of reduced variability (three columns of lighter blocks) after 12 p.m. on Monday, and then a period of high variability with more density (darker red blocks) from 3 p.m. There is high oscillatory behaviour observed in this patient, potentially due to the influence of drugs or other systemic influences. Finally Figure 56c illustrates the oxygen flux. The data for this visualization is measured using the same metric as heart rate flux, however oxygen flux data is gathered each time a desaturation occurs in the SpO_2 signal. Observing Figure 56c, one sees a period of low variability initially, followed by a region of higher variability between 12 p.m. on the Tuesday and lasting 24 hours. Blocks at the 100% level in the flux TIMs likely indicate data errors (such as when a sensor disconnected) but are left in the chart as they may be clinically relevant and should be investigated. To differentiate zero data from missing data requires further research and improvements in data collection.

The researcher can use the interactive brushing functionality to highlight a region on any one of the TIM views, all other views are immediately updated to highlight that section. Figure

54 illustrates how each of the TIM views appear when a region is brushed. Here the researcher is interested in 48 hours prior and 24 hours post an infection event. Highlighting this region also triggers coordinated updates across the linear and the streams graph for more detailed analysis of event classifications.

7.5.2 Physiologic Event Classification Views

We developed three coordinated views to show PE classifications, coming from Artemis, including the sequence graph, linear graph and streams graph. We use similar hues with varying saturation to highlight complementary PE classifications of varying severity. For instance, an isolated bradycardia receives a more saturated pink than a possible isolated bradycardia. Oxygen desaturation events are blue.

7.5.2.1 Sequence Graph View

The first PE classification view found on the top right of Figure 54 is the sequence graph (highlighted in Figure 57). This view supports T5, in which the user requires a rapid means of understanding temporal discontinuous event data. The advantage of this representation is that it reveals events occurring during the same hour across multiple days. This can be useful in associating the influence of routine events, such as bed-side interventions to changes in physiologic data. Each vertical position represents the same hour over multiple days. Specifically, the horizontal x-axis shows progression over 24 hours, and the vertical y-axis shows progression of events over days of the month. The axes can be configured to express seconds (x) by minutes (y), or days (x) by months (y), each producing a periodic view of high-level event classifications.

In order to control the size of circle in this view, we calculate the sample frequency for every hourly epoch. Less significant PE classifications receive a lower opacity, while more significant PE classifications have higher opacity. This allows the user to visually discern areas where greatest clinically significant PEs exist. The radius encodes for the log transform of the total duration in the hour (default view, Figure 57). The transformed values are then sorted in descending order and painted largest to smallest, producing a layered view. The fill hue is determined by the event classification type. The user can hover over the circle to reveal details of each inner circle. An alternative to this design was to use a stacked bar representation, which summarized the frequency of each event over the hour. However that representation does not convey periodic events that occurred over the same time-period spanning multiple days.

Figure 57 illustrates a vagal PE (green) at 1 a.m., followed by central apnoea PE (purple), at 3 a.m., 4 a.m., and 6 a.m. (horizontal) on 29th day of April (vertical). Possible isolated bradycardia (pink) and possible isolated bradycardia (light blue) are sustained over the next several hours. The researcher notices that a red cross, denoting a PSI, is visible at 9 a.m. that day. The researcher notes that till that period, the salient and clinically relevant PEs have become more prevalent by integrating the observed frequencies of vagal, central and possible isolated bradycardia and desaturation PEs.

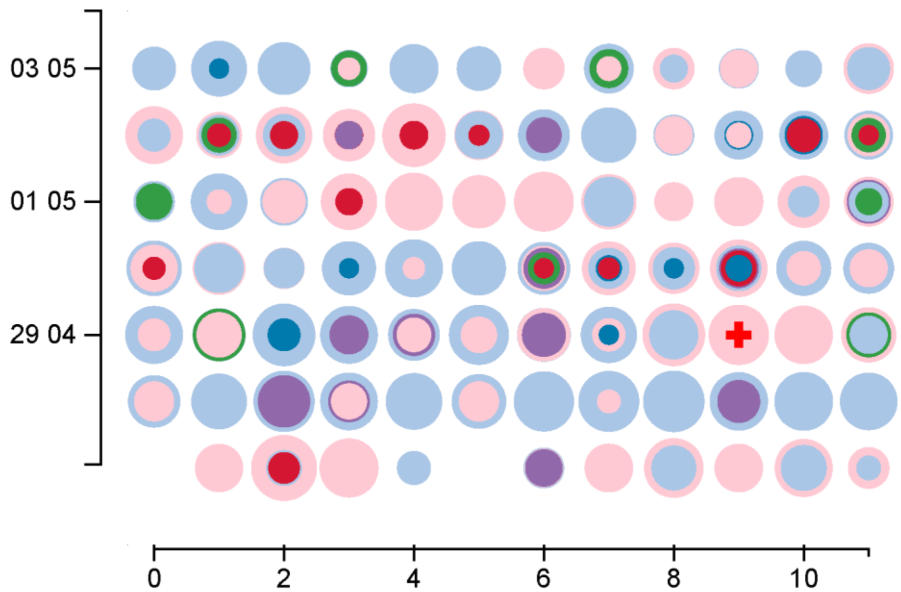


Figure 57: The Sequence Graph, illustrating a matrix of hours by days (truncated to 10 hours). Each bubble's radius encodes the total duration of episodes within that hour, and smaller bubbles are drawn on top.

7.5.2.2 Linear Graph View

The linear graph, Figure 58, supports T4 and T5, in which the user identifies abnormal PEs as well as requiring detailed temporal view of all PEs over time. The y-axis represents a log transform of PE duration and x-axis the linear timeline view. PEs are plotted as circles where hue is determined by the classification type. The radius is double encoded with the log transform of the duration value. Reduced opacity is applied to PEs that are less important, while PEs with higher clinical significance maintain full opacity. Smaller bubbles are of low durations, while high duration events are larger and have more prominence at the top of the graph. A tooltip is available for additional information about each PE. Selecting a PE launches an overlay view of the associated raw data graphs. Figure 58 illustrates several prominent vagal apnoea (green) PE appearing before 12 p.m. and continuing till 6 a.m. the following day. Intermittent central apnoea events (purple), along with possible isolated desaturations (pink) and possible isolated bradycardias (light blue), are observed throughout the night. Event classifications are rendered according to their frequency, and severity. Low severity events like possible isolated bradycardia and desaturation are rendered first, followed by the more significant PEs.

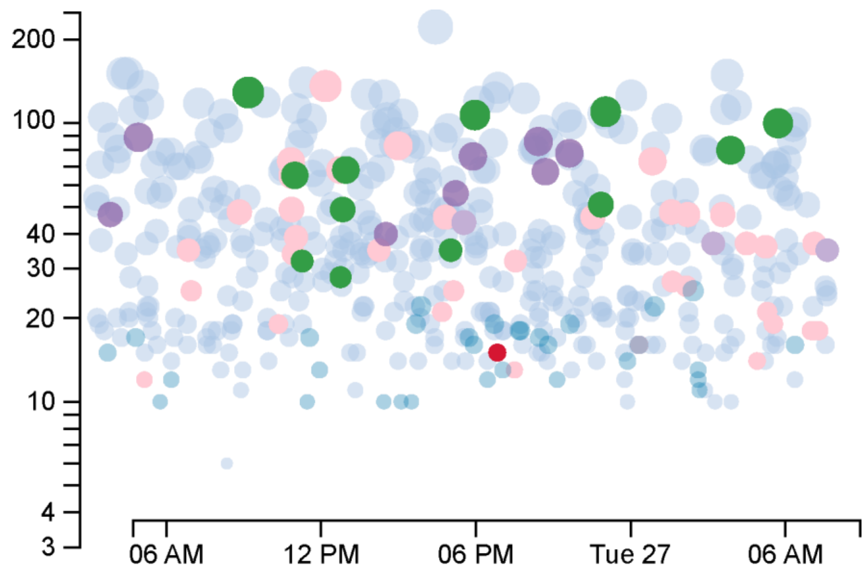


Figure 58: The Linear Graph shows a log-transformed duration of each event classification in a linear temporal view.

7.5.2.3 Streams Graph View

The third event classification view, illustrated in Figure 59, is the streams graph, revealing continuous event classification frequency over time, with the data summed to a count per hour and supports tasks T3--T5. Each stack is coloured with the event classification hues shared across all event classification views. A tooltip is available to explore details about the event classification. Brushing a stack causes all other stacks to fade, giving visual prominence to the hovered stack and reducing clutter. Figure 59 illustrates relatively high frequencies of possible isolated bradycardia (pink) lasting from the 12 p.m. mark, along with possible isolated desaturations (light blue), until 12 a.m.. Following that, possible isolated desaturation events diminish, only to return again in the late afternoon of the 6th. Between this range, there are also several other PE classifications identified, such as intermittent central apnoea episodes (purple), and vagal apnoea (green). An alternative to this design was to use line graphs, while commonly utilized in electronic medical records, the line graph encoding fared poorly when compared to the streams graph. The streams graph, through the use of filled area, allowed the user to rapidly elicit information about the most frequent event within one or more time windows.

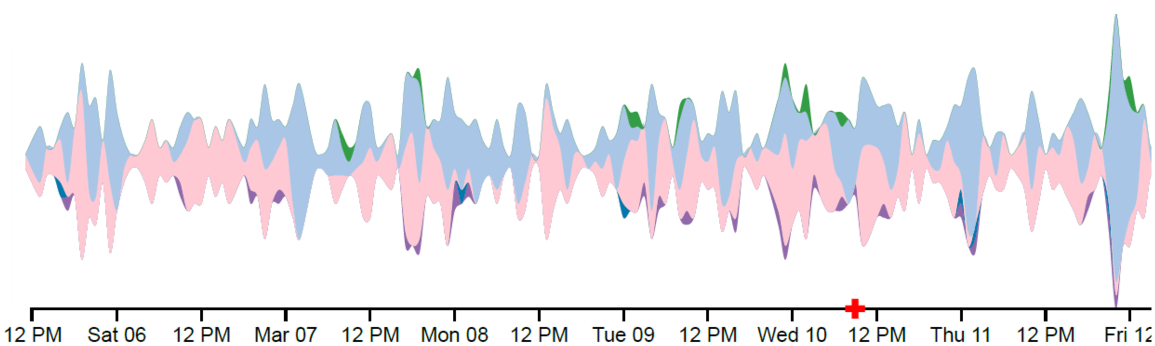


Figure 59: The Stream Graph illustrates the flow of event classification frequency over the analysis duration.

7.5.2.4 Raw Data View

The final user interface component, designed to primarily support T5, which serves as a critical step in confirming whether a patient is believed to be positive for sepsis, is the raw data display that illustrated in Figure 60. In this view the respiratory impedance graph is displayed at the top, followed by the heart rate trace, and finally the oxygen saturation graph at the bottom.

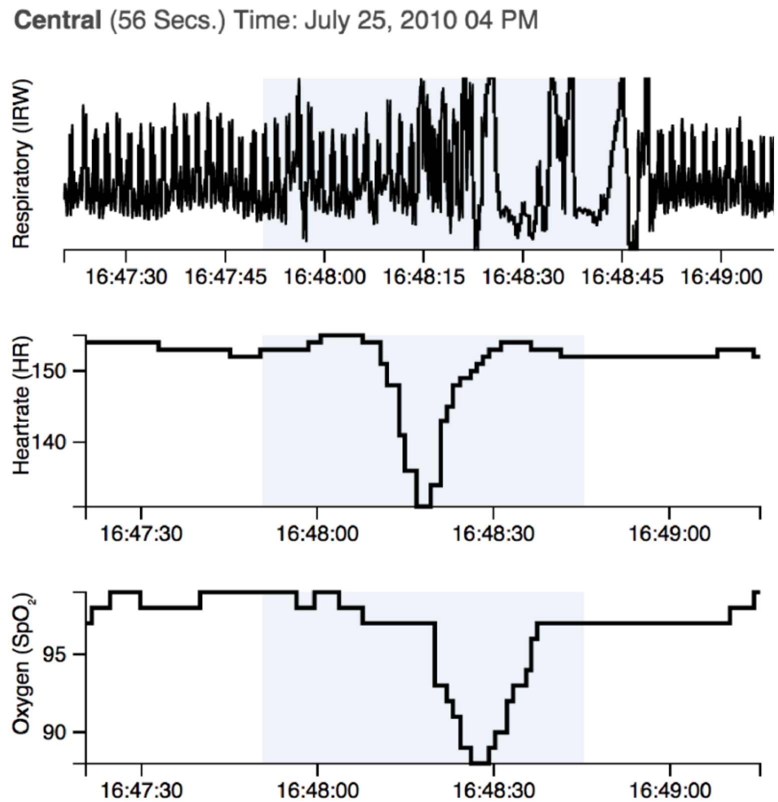


Figure 60: The Raw Data View displays sensor data using 3 line graphs. The highlighted region corresponds to a PE classification, and the white region is a 30 second buffer.

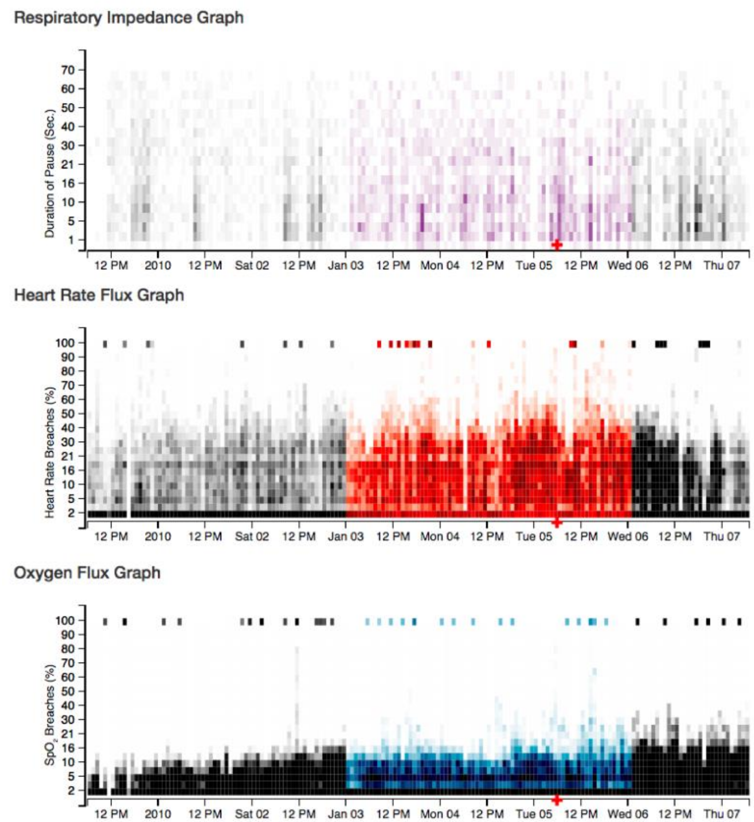


Figure 61: PhysioEx provides coordinated interactive focus for enabling analysis of segments of the timeline across multiple temporal views. In this figure a central obstructive event classification was selected by the user to examine underlying physiological signals.

This view is activated when the user performs a selection on one of the PE classifications in the linear graph view. In this view the analyst can immediately access low-level sensor data that lead up to the PE classification. This line-graph method is a familiar design for displaying sensor data. A background band is appended to the chart, representing the actual duration for the event classification.

7.5.2.5 Coordinated Analysis

When an analyst selects a portion of a graph using interactive brushing, all other graphs immediately update to highlight that section. For instance, in Figure 61, the highlighted region appears prominent in colour on each of the TIM displays, and is also highlighted the streams chart. The linear chart is zoomed in to show the selected time period in detail, from 6 a.m. of the 26th to 6 a.m. of the 28th. Due to its design as a summary graph, the sequence graph maintains its view to provide high level details.

7.6 Expert Evaluation

We conducted an expert evaluation to gain a better understanding of the utility of PhysioEx for clinical researchers. The primary condition in this study was the visualization technique, with two levels. PhysioEx was compared to a stacked bar view (illustrated in the next section) that is currently used to perform clinical research of neonatal spells behaviour [38]. Due to the difficulty in recruiting a large number of highly specialized domain experts, we adopt a primarily qualitative evaluation approach, engaging the available experts in real analysis tasks and both observing their experience and requesting their feedback, to build a holistic understanding of the potential for PhysioEx.

7.6.1 Methodology

Participants: We engaged four domain experts with experience working with neonatal physiologic data on a day-to-day basis ranging from five to 35 years. Three of the experts were males and one was a female. All four experts report using the computer at least once a work day for analytic purposes. Both visualization techniques used in the study were unknown to all participants.

Dataset: The study dataset consists of 29 patients who were suspected of infection and for whom we had truth data about the presence of infection. Suspicion of infection was defined by the presence a blood draw for a laboratory test for infection. The results of the laboratory test provided the truth data for this study. The apnoea event classification algorithm was run over seven days' worth of data for each patient: 120 hours before and 48 hours after the time the blood culture results were received. Prior research suggests that neonatal sepsis may be detected in physiological data several days before current practices suspect it at the bedside. We decided to use this case study as it provides an exploratory means by which the domain experts can investigate and potentially arrive at novel findings.

Task and measures: The task of the domain experts was to use each visualization to determine whether the patient has an infection (sepsis) and state any additional insights they had about the data. We measured the accuracy of determination of infection and the time taken in analysis. In addition, we engaged participants in a semi-structured interview about their analysis process, preferences, and usability issues which arose. The transcript was coded using open-coding, and themes were generated by clustering codes into their respective categories.

The themes were analysed and reported as research findings. The codes and clustering method was validated by another researcher. Screen and voice recording was used to allow for detailed analysis as well as easy transcription of the collected data.

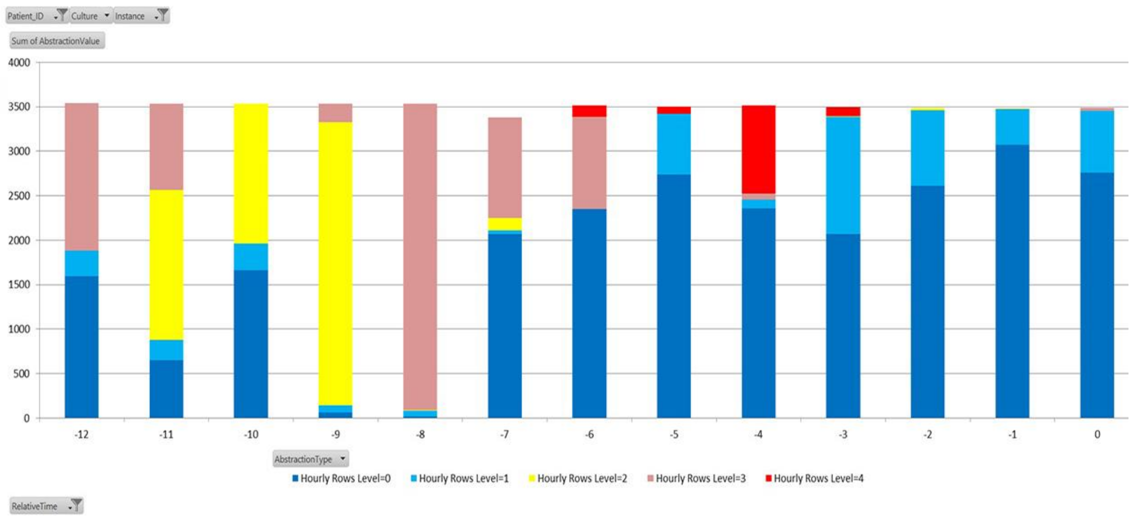


Figure 62: Stacked Bar Chart used as control against PhysioEx

Control: We compare PhysioEx against a stacked bar graph, illustrated in Figure 62, which has been used by clinical researchers to understand trends in neonatal spells (PEs relevant to the neonatal domain) preceding a point of suspicion of infection. This view provides a higher level and non-interactive view on the physiologic data by counting event classifications every second and summarizing them on an hourly basis. In the stacked bar graph the levels refer to event classifications (in order): all normal, heart rate variance changes, respiratory pauses, oxygen desaturation, and blood pressure drop. The stacked bar view is aligned with the time of the suspicion of infection (red cross on PhysioEx) at the zeroth hour, then all the preceding events sorted by hour to the left. An analyst would start at the zeroth hour to analyse the spells behaviour. Figure 62 shows that beginning at the -12th hour up to the -4th hour there are sustained fluctuations in the infant's cardiorespiratory status. The infant seems to be improving as it approaches the zeroth mark (more times classified as normal). Note that there is missing data in hour -7, perhaps due to infant movement or sensor malfunction.

Procedure: Each session began with a brief introduction to the study and a semi-structured interview to assess participant prior knowledge about the domain of neonatal care and spells. This was followed by a series of 7 training trials, 11 timed experimental trials, and a brief questionnaire eliciting feedback on the interface design, repeated for each visualization technique. Due to data availability, the same dataset of 7 samples (in random order) were used for the training trials in both conditions. For the experimental trials, two different datasets of 11 samples each was used, one for each condition. The ordering of technique and experimental datasets was counterbalanced. There was a total 2 techniques × 11 trials × 4 participants = 88 trials. The analysis task was repeated for each training and experimental trial. Feedback about the correctness of determination of infection was provided for training trials only. Participants were asked describe the spells activity around the point of suspicion of infection and to state whether an infection was present. At the end of the experiment, a brief questionnaire was administered to collect participant preference between techniques. Experimental sessions lasted two hours and participants were able to take breaks as needed.

7.7 Results

In this section we report the results of the study comparing PhysioEx and the stacked bar representation of physiologic data. The accuracy of determination of sepsis was uniformly distributed and below 50%, for the dataset containing 7 sepsis and 22 non-sepsis patients for both conditions, thus we did not investigate further. We instead focus our analysis on the quality and depth of insights expressed during the analysis process, and the subjective expert feedback.

7.7.1 Identification of Physiological Behaviours

Experts provided a range of comments the stacked bar method of representation when trying to elicit physiological behaviour. Although it was a simple way of seeing how much of the hour was attributed to one physiological measure, it did not provide additional and more salient information about what was going on in the hour. Experts found it difficult to decide whether a bulk of events occurred uniformly in the hour, the breadth of their intensity, and durations. The time to analysis was rapid, with a majority of the analysis being completed within ten seconds.

Meanwhile PhysioEx allowed them to rapidly elicit physiological behaviour, frequency within an hour, the duration of all event classifications aggregated in an hour in addition to duration of single classifications. When asked to describe the physiological status of the infant, experts often spent several minutes describing the intricate behaviour, frequency, duration and sequences of events seen in TIMs and also on the sequence and streams graph. This was seen consistently, with analysis time ranging from 2-10 minutes per patient. One expert comments about the Respiratory TIM: "I see a burst of activity here, on this Friday starting before 11 p.m., and going through to about noon, then I see a trivial amount of activity about 24 hours later, and then I see another burst of activity starting about midnight starting about the 28th, which seems to be of the same intensity as the first burst I observed but has a longer duration. In the middle, I see very little variation.'' The stream graph was also noted to be a unique tool in the domain of physiological research. Experts had not encountered this representation and therefore required some time to adapt to it. One expert found that he was relying on it as a final 'truth' indicator, after having analysed all other representations.

7.7.2 Hypothesis generation

Using the stacked bar view, experts found it difficult to generate hypotheses unless there was a clear and distinguishable trend. Where events occurred without any clear trend, all experts stated difficulty with determining whether these events had any relationship with the point of interest at the zeroth hour. All experts described the colour scheme to be very favourable when determining patterns and trends. One expert mentioned “I'm looking for the stacks with a lot of yellow, the red is distracting for me, but the yellow is interesting''. Another expert physician stated that “[the stacked bar] is too simple, it doesn't work for me”.

Experts described PhysioEx as useful, and powerful when generating hypotheses, they also mentioned that the coordinated interactive brushing was most useful when they wanted to reaffirm incremental patterns. They found the coordinated brushing and highlighting across all TIMs provided the most benefit in terms of closely analysing neonatal spells preceding the infection suspicion point. The ability to select the event classifications to reveal low-level sensor data was appreciated by all experts and heavily utilized by one expert. Two experts were able to derive bed-side intervention information from the patterns exposed on the Respiratory Pause TIM. They revealed information about potential respiration modality of the infant. Some quotes received from experts include:

"Oh wow look at that… look at that… this is a baby that got intubated… a fully manually intubated baby. Well this child cannot apnoea… if you look at the respiratory pauses they are all so uniform."

"Look at the heart rate variability, it swings everywhere and then it comes back. […] It looks like they had a ventilator mode change, maybe to biphasic, but they've also taken a culture at the same time, this is an odd practise, we tend to do things one at a time."

7.7.3 Satisfaction of Use

Domain experts who used both the stacked bar view and PhysioEx reported greater satisfaction with the simplicity of the former, but expressed concern over excessive simplicity and hiding of potentially useful data. When analysing trends on the stacked bar view experts found that while they were able to verbalize trends of high-level event classifications more easily, they were unable to provide detailed descriptions. Moreover, the stacked bar also provided the domain experts with a familiar format. This familiarity factor was seen as favourable for immediate use without much training.

Experts were positive about the additional detail available for analysis in PhysioEx. In particular, the TIM representations were favourably received by all. They paid keen attention to the behaviours expressed in heart rate and SpO_2 TIMs, and stated that it was helpful when conceptualizing the infant's status over many hours. The sequence graph was used by three experts for determining sequences of events prior to the suspicion of infection, the fourth expert did not use the display at all. On the simplicity of PhysioEx, the responses were mixed. While domain experts greatly appreciated the increased level of detail, it also proved to be cognitively demanding task, requiring learning new interaction methods for selecting, filtering, and retrieving information about physiological signals. The experts attributed the cognitive load due to the overwhelming number of possible events that had prominence in almost all patients. Moreover, experts also noted the usefulness and utility of PhysioEx could be even further improved with the addition of contextual information, such as the infant's gestational age, gender, method of respiration, and other comorbidities.

7.7.4 General Comments

Experts provided numerous comments on the usability and potential utility of PhysioEx. Two experts, also physicians, mentioned that TIMs may contribute additional means of gaining insight on subtle physiological behaviours of the infant that are currently unavailable for bedside decision makers. Six coordinated views, as currently instantiated in PhysioEx, were found to be useful for research but likely too complex for use at the bedside. All experts using the TIMs representation were immediately cognizant of the data quality available for analysis. Data quality is an ongoing challenge in the neonatal intensive care environment. However, obtaining consistent and continuous data samples is very difficult, due to the dynamic nature of the environment where the sensors attached to the infant often disconnect or have to be removed for transport.

7.8 Discussion and Future Work

We used an expert evaluation consisting of four domain experts analysing neonatal spells behaviour in an attempt to predict the likelihood of neonatal sepsis. Although the results of determination of sepsis in our dataset was inconclusive, our study revealed that PhysioEx deeply involved clinical researchers in the analytic pipeline. Experts using PhysioEx were able to verbalize subtle physiologic behaviour spanning numerous days and for numerous patients. Many of the insights discovered with PhysioEx were hidden by the standard stacked bar representation. While the time for using PhysioEx was much longer, this may be explained by the richer interface, interactivity, and novelty of the visualization. Rapid analysis is needed in bedside situations, but for retrospective research, such as analysing the relationship of physiologic measures, spells, and neonatal sepsis, depth of insight is more important than speed.

The analytic tool gave experts the first opportunity to interactively explore physiological event features and event classifications. To our knowledge, there are currently no other tools that provide interactive exploration of detailed physiological changes of neonatal spells. However, introducing such a novel tool does have limitations. Some experts experienced fatigue after enduring a long training phase and then analysing a total of 18 patients on PhysioEx. Contributing to the fatigue was the significant cognitive load imposed by using novel tool in detail to perform a difficult task. The TIM views provided experts with a simple and rapid method of appreciating physiologic behaviour. Most experts relied on the TIMs to base their decisions on whether the infant was experiencing normal or abnormal changes in physiology. Dense and low density regions were rapidly identified by all experts. This information was then augmented by the event classification display. Experts, especially practitioners, also used the TIMs to characterise the data quality for that particular patient. Since this is a commonly faced issue in NICUs, the ability to see drops in data quality gave more insight about the infant and their management.

The sequence graph was heavily utilized by some to track incremental hourly changes leading up to the point of suspicion. One expert commented that the bubble matrix provided a unique ability to recognize patterns that commonly occur at various times of the day. Events such as blood draw occurring in the afternoon, loss of data for short durations, and transfer of the infant to other units, were speculated. While this information was provided to the experts, the ability for the experts to augment clinical expertise provides an opportunity as future work for automated annotation capabilities for PhysioEx. The automated annotation of events would further supplement researchers with much needed context to explore the event space in more detail.

PhysioEx was found to provide a greater advantage to explain neonatal spells behaviour than the alternative. One expert physician with extensive involvement in neonatal spells research, had mentioned that they are now inclined to invest a day in training a neonatal fellow so they would be better able to describe physiological behaviour of spells.

There are however, limitations with PhysioEx and our preliminary study. We only tested PhysioEx with four expert participants drawn from the larger clinical researcher population. Moreover, there are no established clinical links yet between neonatal spells and infection. Therefore, the experts participating in the study were not looking for known associations. Many experts noted that lack of contextual information (patient metadata) as a limitation of both techniques. We had developed PhysioEx to cater to exploring physiological data, however in future work incorporating clinical information would certainly be highly advantageous for supporting analytic activities. To address the cognitive overload from analysing several patients independently, in the next chapter, we present a visual analytic tool called CoRAD that assist in analysing population cohorts in a single view.

7.9 Chapter Summary

In this chapter, a novel visualization technique, the Temporal Intensity Map was presented, and PhysioEx, a visual analytic tool for complex multidimensional sensor data exploration was also introduced. We present a task analysis for designing visualization displays for the complex and heterogeneous sensor network environment in neonatal care and draw on this analysis to inspire design. Our preliminary study supports further investigation into PhysioEx as an important addition to the tools available for clinical researchers. In future work we aim to deploy PhysioEx to support additional use cases, such as exploring physiological behaviours for other clinical conditions. Moreover, we aim to integrate more contextual information such as clinical histories into PhysioEx for the development a more tightly integrated physiological clinical research system.

8. CoRAD: Cohort Relatively Aligned Dashboard

This chapter presents material from a publication currently under review [45]. The publication was co-authored by Christopher Collins, Carolyn McGregor and Andrew James. The author of this dissertation designed, developed, and evaluated CoRAD. Christopher Collins, Carolyn McGregor, and Andrew James provided input for the design of CoRAD and write-up of the publication. An additional content, §8.4, does not appear in the publication, and is presented in this chapter to aid the reader in details involving the instantiation of CoRAD.

8.1 Introduction

This chapter presents the creation and evaluation of CoRAD, which is an instantiation of a TDVA mart to support hypothesis testing. The creation and evaluation of CoRAD followed the TDVA framework and methodology.

Case-control studies are among the most utilized research methodologies in clinical research [5]. A case-control study involves isolating retrospective data for patients with a condition of interest, and comparing those features of interest to a sample of individuals without the condition [133], [134]. The goal is to explore correlations across relevant clinical variables. In most cases, cohorts must be relatively aligned to an epoch. The alignment may be a time period when a test result was received, such as a blood result confirming or rejecting a possible infection. The relatively alignment process typically involves a large number of manual data cleansing and data preparation activities to align clinical data of each patient to a single and representative time. Most case-controlled studies use clinical data stored in databases and electronic medical records. Furthermore, performing case-controlled studies using physiologic data is a challenging task. Physiologic data is often collected at a consistent sample frequency, and appear in their raw form, as a list of values. This is in contrast to a limited set of discrete clinical variables, such as lab reports, or physical observations.

This chapter introduces a novel dynamic visual analytic tool called the Cohort Relative Aligned Dashboard (CoRAD). The CoRAD tool represents an instantiation of the dynamic visual analytic publisher component of a larger framework called the temporal tri-event parameter based Dynamic Visual Analytic (TDVA) framework. CoRAD supports the integration of relatively aligned algorithm-generated output to visual interface, to automate and enhance the analysis workflow. In addition, CoRAD allows the user to drill through multiple hierarchies of data, from quality of signals, to abstractions and ultimately classifications of clinically relevant events.

To validate the effectiveness of CoRAD against a separate visualization, an expert evaluation was conducted at Neonatal Intensive Care Unit at The Hospital for Sick Children, Toronto. The subsequent sections details related works, problem characterization, task analysis, CoRAD design, evaluation methods and the results of the evaluation.

8.2 Related Works

A case-control study involves retrospective analysis that separates patients based on the presence of a condition [134]. Case-control studies, among many observational research methods, remain an important aspect of clinical research [133]. Differences are studied and hypotheses are generated based on the analysis, to motivate deeper investigation and more rigorous research. However visualizations that support these efforts in physiologic data remain elusive.

A. Artemis Platform

Artemis is an online analytic platform that was developed to source, analyze, and perform real-time feature detection on multiple physiological data streams, for multiple conditions in multiple patients [37]. Artemis supports the deployment of real-time event stream processing algorithms. In this research, we use data generated by an algorithm running in the Artemis platform for neonatal sepsis that was executed to detect and classify Heart Rate Variability (HRV) scores between 0 and 60, where zero signifies no variability and 60 demonstrated that the patient's heart rate varied consistently in the hour. The details of the neonatal sepsis algorithm have been previously published [57]. Results from the analysis are then sent to a database and also available for real-time streaming for visualization. The output are then processed and sent to a platform that was developed using the TDVA framework. That platform produces instantiations called dynamic visual analytic marts, such as the CoRAD.

B. Cohort health visual representations

In the general space of health-based cohort analytics, some recent work has resulted in high fidelity visualizations with a time component. TimeSpan [135] provides an interactive dashboard for identifying door-to-needle time for stroke patients at a large tertiary hospital. LifeLines presents graphical summaries of patient journey [136]. The Cohort Comparison (CoCo) tool, provides a simple interface for exploring statistical correlations across multiple clinical datasets [137]. DecisionFlow presents graphical summaries of patients who developed heart failure relative to a population [138]. VISITORS is a dashboard for analysing clinical temporal abstractions in oncology patients [139]. EventFlow presents a method to simplify event sequence information to rapidly identify abnormalities [140].

While all of these visualizations introduce cohort analysis of patients using clinical information, there is a need for research in representing temporal abstractions of physiologic data across cohorts, and supporting automated temporal relative alignment, while allowing the user to gain contextual awareness using low and higher-level summarizations of data.

C. Visual analytics of temporal data

Domain specific dynamic visual analytic tools have been shown to perform well in communicating anomalies to the end user. The VisAlert system [183], for example, provides situational awareness for network security analysts. Another system in the same domain is LiveRAC [16], which supports additional exploratory features such as semantic zoom to search through the data set, and allows for side-by-side comparisons between different clusters. However, this system presents a complicated user interface with potential for visual clutter.

Director [285] is a visual analytic tool for computer network simulations. It provides a heatmap-based timeline visualization to identify the health of multiple nodes, along with a temporal view of their health deterioration. CloudLines [185] introduces an incremental event visual analytic tool using kernel density estimation (KDE) to amplify signals from highly dense areas and minimize low density areas. The technique is applied to online news stream analytics, and multiple time-series data are used to highlight topic emergence, and when the topic is no longer emerging, a visual decay function is applied to emphasize more popular topics.

While most visual displays are temporally aligned to the most recent epoch, in this research we present a novel visual analytic tool that uses relative alignment to a real-world independent event. Two heatmap timelines are presented in the main display to allow clinical researchers the ability to visually explore patterns in HRV across multiple patients.

8.3 Problem Contextualization

Sepsis is a form of hospital acquired infection, and remains a serious health problem requiring antibiotic therapy [35]. Currently it is very difficult to detect using non-invasive methods, such as by bed-side monitoring. Clinicians rely on qualitative observational methods for identifying signs on this illness. When sepsis is suspected, blood samples are drawn and required to confirm any diagnosis. However, neither method has been found to be reliable [56]. There is growing body of evidence that shows new pathophysiologic behaviours can be identified earlier using physiologic data. One such case involves the study of reduced HRV as a potential indicator of sepsis [286], [287]. In addition, Flower et al [58], present results that indicate periodic cycles of heart rate decelerations, or bradycardias, are common and seen to be clinically correlated with sepsis in addition to reduced HRV and they propose heart rate characteristics as a means to correlate the occurrence of the two together.

McGregor et al., developed an algorithm that produces real-time HRV scoring for neonatal infants [57]. This scoring can be used to identify temporal areas where there is reduced HRV that indicates some sign of illness. A dataset containing HRV information and algorithm-generated classifications of bradycardia as part of McGregor's neonatal spell research are available from a prior study [44]. Data from a total of 47 patients are available in that dataset, of which 33 patients have sufficient data quality. The goal of this study is to investigate the hypothesis exposed in Flower et al. [58] that periodic cycles of heart rate decelerations together with reduced HRV are common and clinically correlated with neonatal sepsis. This information is presented in CoRAD and we performed an evaluation to test participants' ability to determine sepsis based on Flower's hypothesis. The study was approved by the Research Ethics Boards at The Hospital for Sick Children and at UOIT.

8.4 Instantiation of the TDVA Framework

CoRAD is the second instantiation of the TDVA visual mart within the TDVA platform which has been created by extending the Artemis Platform [37]. The instantiation process is illustrated in Figure 63. Bradycardia primitive events were gathered through the execution of the neonatal spells algorithm [280]. Heart-rate variability was calculated using a scoring system identified by McGregor et al. [278]. CoRAD supports the researcher existing in a high engagement, and low urgency environment typical of retrospective clinical research case-control studies.

P1: Finding the Right Person

The first step of the TDVA methodology is a requirement gathering phase. A task analysis was performed using domain experts. The details of the task analysis are described in §8.5. In this instance, the researcher requests mPub to generate views using the "Researcher" context (middle, Figure 63). This informs the mPub engine to generate views that are interactive, including the support for selections, filtration, and detail on demand functions.

P2: Categorization and Data Modelling

The second step of the TDVA methodology requires information about the data source and modelling requirements. To gain that information the mPub engine prompts the user with a secondary menu that solicits information about the required data source and visual toolkits

(VT.1-7, page 140). Once the user has selected the appropriate data source and visual toolkits (in this case VT.6 and VT.7), the mPub engine performs the necessary data modelling activities. Details of the design is described in §8.5. Since the user has selected the intent for a case-controlled methodology, data is pulled from its appropriate source and relatively aligned to the comparator variable. In this instance, the user selects "blood culture" to be the comparator variable. Hence, all patients having that data are particularly identified and divided into positive and negative cases. All cases are aligned to the comparator's time, in this instance the timestamp associated to the blood draw is used.

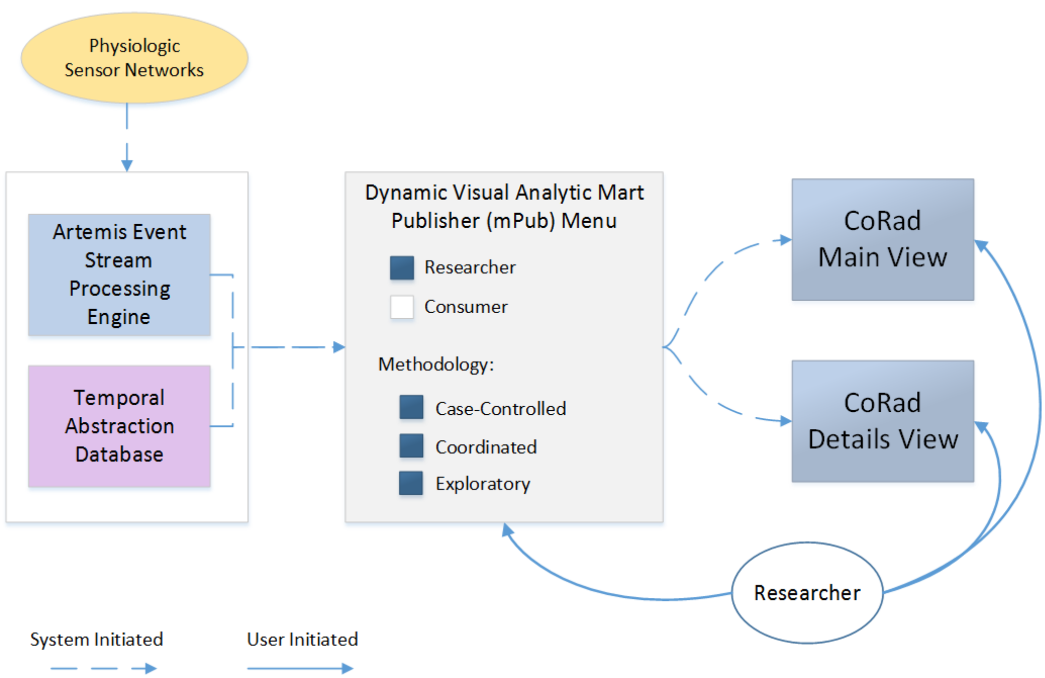


Figure 63: Instantiation of CoRAD using the Dynamic Visual Analytic Mart Publisher (mPub)

P3: Deployment of DVAM(s)

The third step of the TDVA methodology, requiring the instantiation of the DVAM is initiated. The publisher generates a visual mart that contains three unique visual toolkits including a heatmap (VT.6), context bar (VT.7), and a detail view (VT.1). The specific set of VT's that are selected is determined by the user. The DVAM serves as an independent analytic sandbox, the DVAM is independent of other instantiations and allows the analyst to perform directed exploration. The researcher then interactively analyses the dataset to extract correlations between case cohorts. The CoRAD DVAM allows the researcher to scan for areas of reduced heart rate variability using an interactive brushing tool. The brush tool highlights temporal regions that can be viewed in the details view (illustrated in §8.8.4).

P4: Evaluation of the DVAM

The final step to complete the TDVA methodology is to enlist five domain experts to evaluate the DVAM and assess its overall utility. This evaluation is described in §8.6.

8.5 Task Analysis

Two domain experts, who were representative of the population, were asked to describe specific tasks they currently perform to conduct hypothesis testing using physiologic data across a cohort of patients. The common tasks were:

T1 Relatively aligned temporal abstractions: Relevant HRV values are filtered and manually aligned to an anchor point. The relative alignment performed manually, can introduces errors, and can be time consuming.

T2 Import relatively aligned abstractions to a spreadsheet: Each HRV value is then sorted by the relative aligned time and imported to a spreadsheet manually, this also introduces scope for potential error.

T3 Graph abstractions*:* Once the HRV values were imported into the spreadsheet, line charts and stacked bar graphs were frequently used to visualize the data.

T4 Identify correlations: The domain expert would find associations by comparing HRVs before and after the anchor point. Further, the domain expert might highlight multiple patients of interest and investigate patterns between the selections.

These tasks were performed manually, and was stated to be time-consuming and error prone. These tasks informed the design of CoRAD and serve as a guide for future research in similar application domains.

8.6 Design of CoRAD

We describe CoRAD with its design goals that were informed from the observations and task analysis with domain experts.

DG1 Integrate heterogeneous data: The first task, the relative alignment of physiologic data to clinical data, can involve a mix of numeric, continuous, or ordinal data types. Our design goal is to unify the representation of these data types for extendibility of CoRAD.

DG2 Single holistic view: Currently most of the current tasks performed are manual, however, the ultimate goal is to collect all important disparate data into a single environment. Patient clinical data is closely associated with the patient's physiology, which is correlated to the device measuring that data. Therefore the goal is to provide an integrated view of all direct and indirect patient data.

DG3 Details on demand: The user requires access to details, however current tasks limit the degree of data that can be accessed in a timely manner. Moreover, access to details can be useful in determining the salience of an observation. Our goal is to provide the user convenient access to details on demand.

DG4 Access to statistical tools: Many of the activities performed are by nature, statistical. So our goal is to provide the user with a simple statistical view of the data to assist potential discovery of salient features.

CoRAD is illustrated in Figure 64, and consists of four components: the main view (Figure 64a), detail view (Figure 64b), properties view (Figure 64c), and the context bar (Figure 65). The interface was developed using D^3 [277]. In this section each component is described in detail.

8.6.1 Main View

The main view, illustrated in Figure 64a, consists of several patient bars that utilize an opaque controlled colour scale to present heart rate variability (HRV) information to the user. The darker bars reflect higher HRV and the lighter shades denote lower scores. Each patient bar is painted from left to right, where the left most region shows -120 hours, about five days prior for 48 hours after the aligned pivot. The zeroth hour is marked by a grid line that extends from the top of the main view and repeat every 20 hours. This method of relative alignment supports tasks T1 – T3 and DG1 and DG2. Each patient is stacked from bottom up, with the bottom being the population bar. This vertical arrangement provides a convenient means of comparing HRV patterns within their respective relatively aligned epoch. An anonymized patient identification is appended to the left vertical axis.

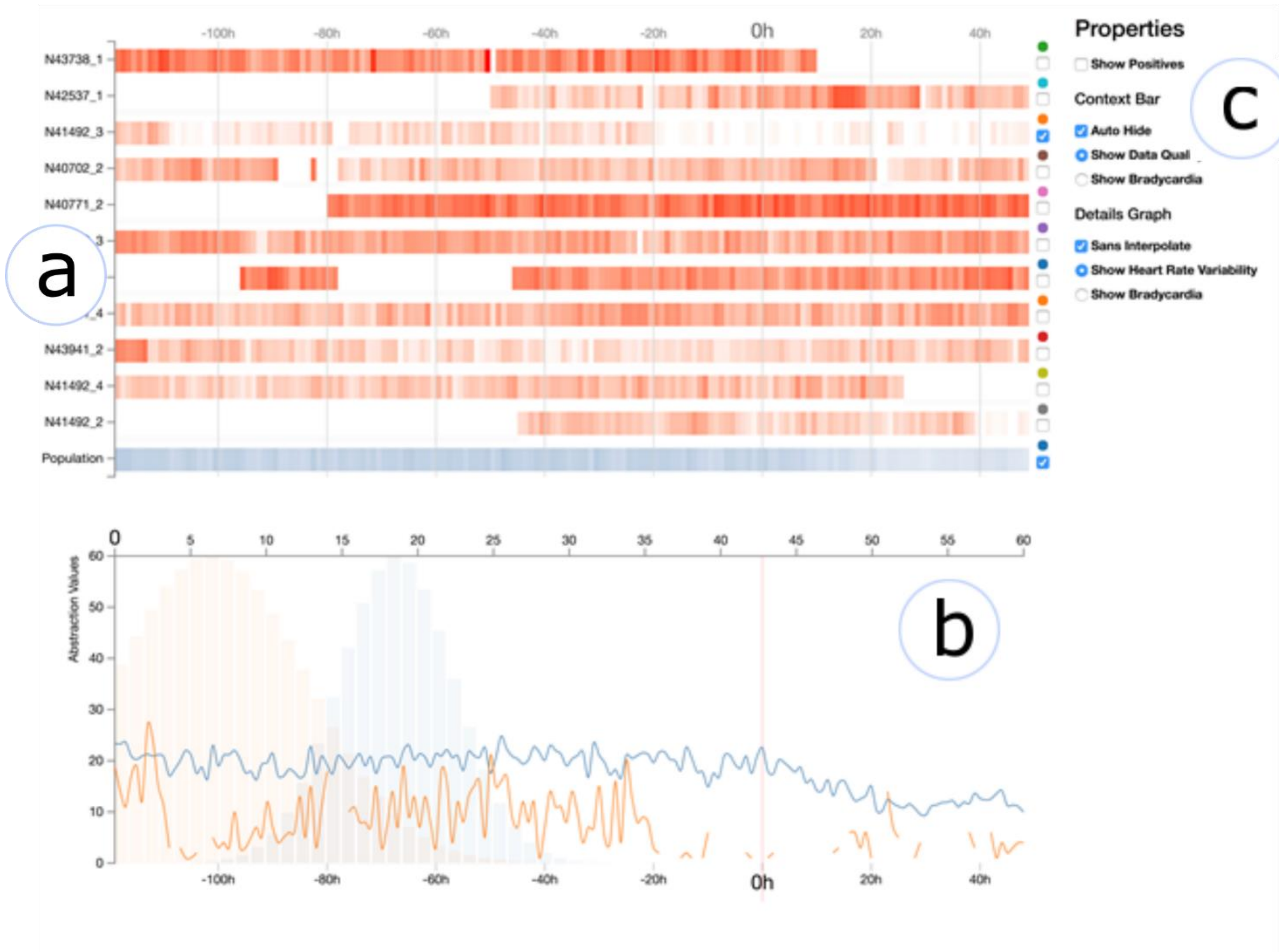


Figure 64: CoRAD provides interactive focus supporting analysis related to event relatively aligned at the zero hour (0h) mark. (a) In this figure all patients are aligned to the y-axis, and the relative-time is marked across the top horizontal position. All patients are coloured using a red scale (lighter means reduced HRV, darker means more variable heart rate), unless the 'Show Positive' control is active. The normalization of all results were used to produce the population map coloured in blue. The detailed view on the bottom (b) provides a line-chart view of details including the raw-data, heart rate variability for selected datasets, or high-level classifications. A histogram is also available and highlights the distribution of HRVs over the entire duration. (c) Provides a view of the properties control, functions are provided to manipulate the dashboard view interactively.

8.6.2 Detail View

The detail view provides an alternative view for selected data from either of the other two views. It consists of a line graph and a histogram. The line graph is a plot of HRV values for an interval selection in the main view. A line graph was previously used to display HRV values [278]. If there are no selections in the main view, the line graph displays HRV values for the entire duration. The user is also able to display the line plot of the average HRV of the population. Having access to this raw data can be helpful in associating discrete values to observations. The line graph supports DG2. For instance, Figure 64b, shows the HRV line graph for patient N41492_3 and the population pinned to the same canvas, while all other lines are set to be transparent. The line graph can be configured to show interpolation, should missing data be present in the dataset. The default option is to avoid interpolation, and make the line transparent when there are missing data.

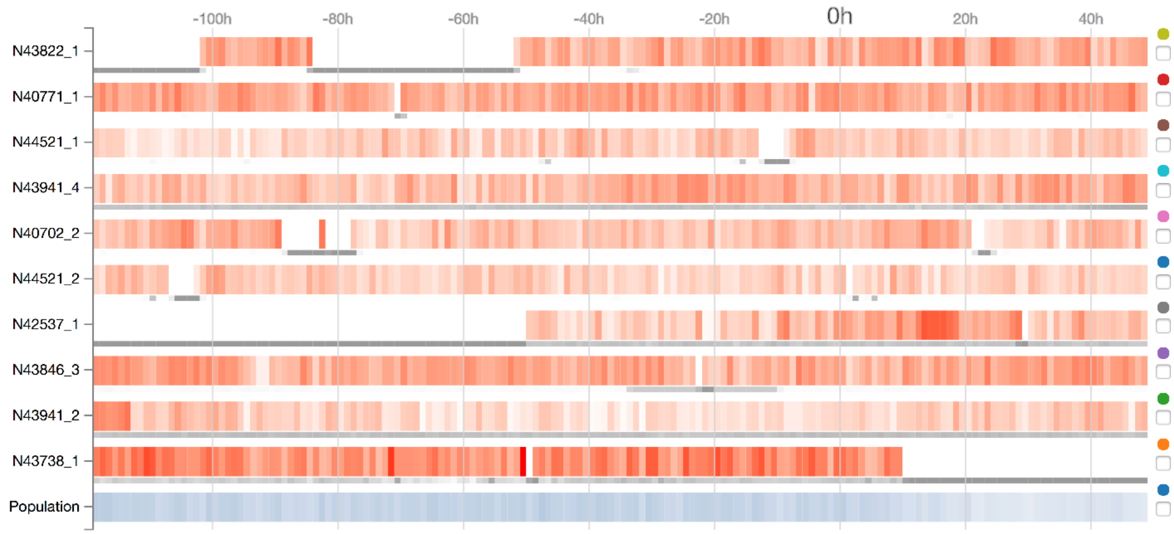
The detail view also contains a histogram that displays the distribution of HRV values for each selection in the main view. The distribution is a Gaussian plot derived from the mean, and standard deviation of the HRV data for each sample. Should the user select the population, a population mean and standard deviations of HRV's are used based on the values of all 33 patients in the dataset. The availability of the histogram fulfils DG4.The detail view can be altered to higher-level classifications, such as the temporal presence of bradycardia. This view also exposes details about the HRV value and the associated patient when the user selects a single line on the screen. The detail view more specifically supports T4, as it allows the user to directly compare two or more patients within a window of time. The interactive details tooltip allows CoRAD to provide the domain expert details on demand, thus supporting DG3.

8.6.3 Properties View

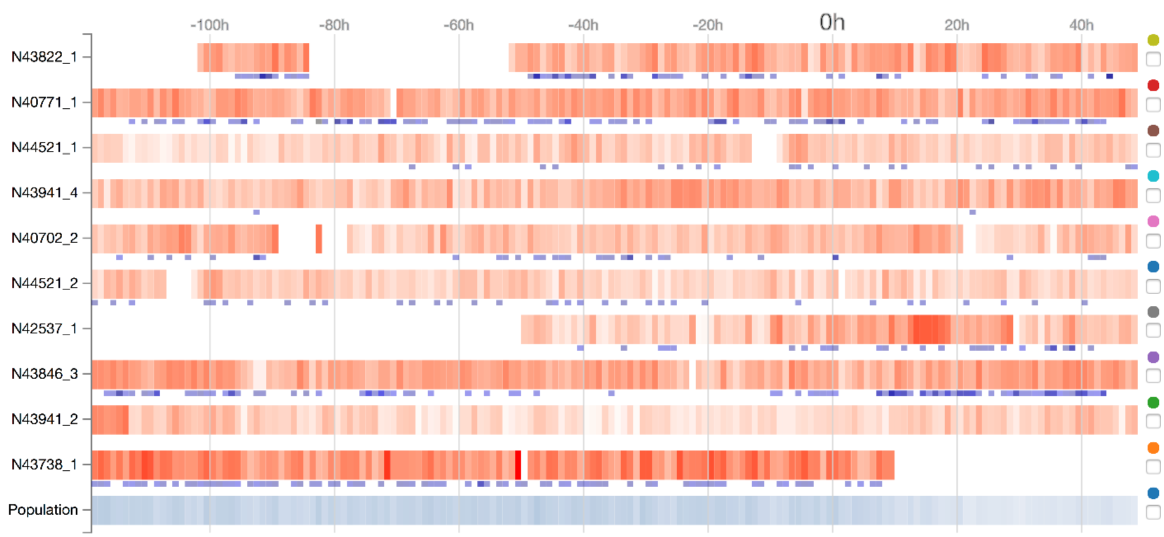
The particular methods by which information is presented in the main and detail views are controlled by the properties view presented in Figure 64c. The first checkbox allows the user to highlight patients that were tested positive, and alternatively to turn the highlighting off if the user did not want to make positive cases explicit. The subsequent selection buttons are grouped according to the views they manipulate. The data quality and classifications button in the context bar group control the data being represented in that view. The raw data, abstraction and classification selection buttons controls the information visible in the details view.

8.6.4 Context Bar View

The context bar resides immediately under the patient bar and can represent one of two types of information. Figure 65a, shows the context bar illustrating regions of poor data quality. For instance, patient N43738_1 is shown to have compromised data quality just before the 20th hour and continues until the 48th hour. Meanwhile, N43941_2 is shown to have comparatively better quality throughout the entire duration. The second type of data the context bar can represent is classifications data. Figure 65b, illustrates the presence of bradycardia episodes during an hour by affixing a green box under the appropriate relative time period. The user can interactively control the data represented in this layer, hence, providing information on demand.



(a)



(b)

Figure 65: Context Bar View, allows the user to select one of two potential data being represented, (a) shows the data quality, the darker lines being times when the data quality was compromised, and (b) representing bradycardia events.

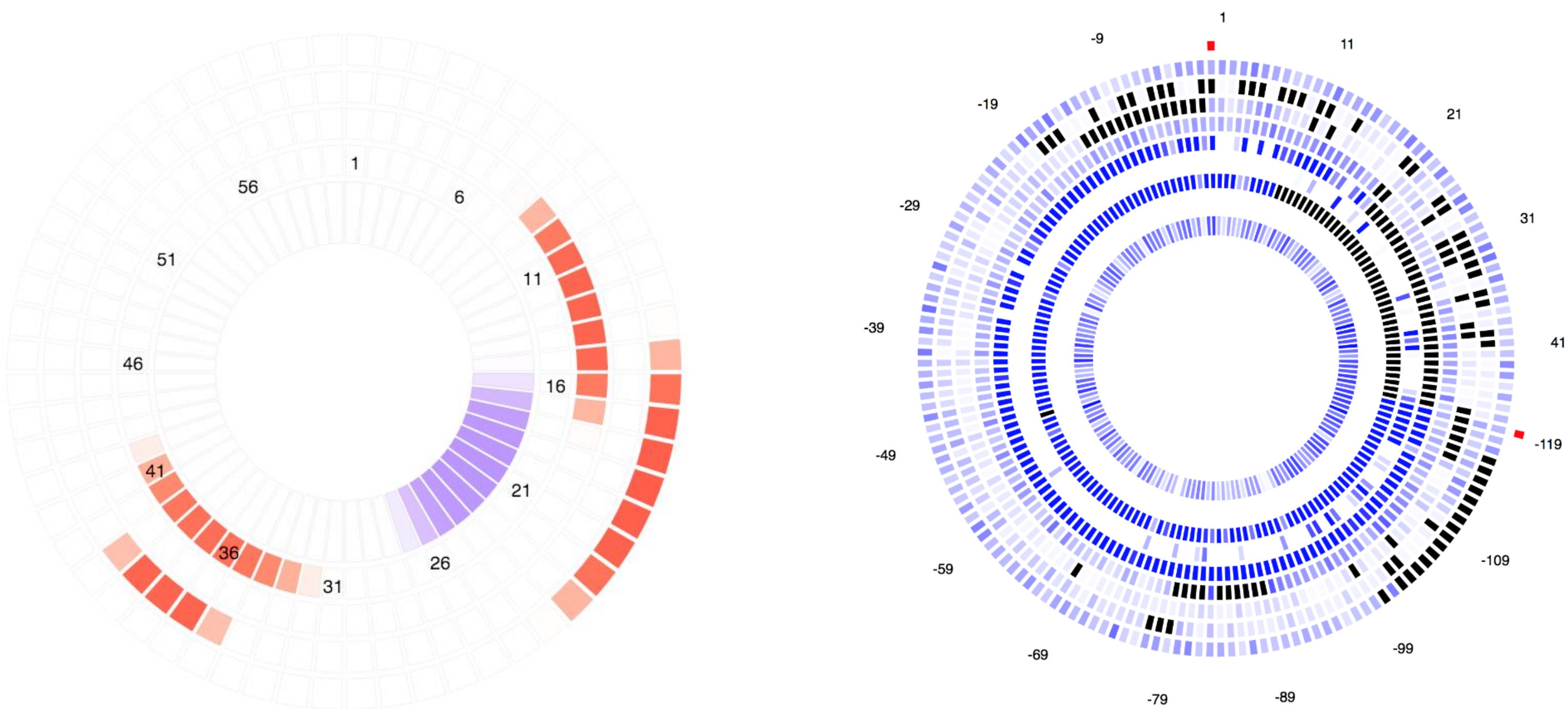
8.6.5 Design Alternatives

Prior to finalizing the visual components of CoRAD, two alternatives were investigated. Among the most prominent alternatives was a radial graph that consisted of two views: a distribution and temporal view. The distribution view illustrated in Figure 66a, consists of a central arc that describes the average distribution of heart rate variability scores for the population, and the each ring representing a separate patient. The arc begins as zero at the top of the ring and extends to the 60th mark. Zero represents no variability, while 60 represents variability in each minute of the hour. For the distribution illustrated in Figure 66a four patients are compared to the average of the population. The average of the population has a mean around the 21 mark. However for the patients the first and third ring, a mean for the distribution is observed around 36 mark. Significantly, these patients have had a higher than average heart rate variability scoring recorded during the monitored period.

A temporal radial graph was also constructed to support the identification of abnormal trajectories of heart rate variability values in cohort populations using an average of the population as a baseline. The temporal radial graph illustrated in Figure 66b presents seven patients who are aligned to population average as separated rings at fixed radii from the centre. Opacity is controlled to show regions of higher and lower heart rate variability values. For instance, the first and third patient from the population are seen to have very dark blue rings, signifying higher heart rate variability scores. While the patients in the outer ring have lighter blue rings, signifying reduced heart rate variability.

While there has been many forms of radial graphs produced [288], there have been some concerns that have emerged about the interpretation of radial graphs [289], [290]. However, other instances of radial graphs were shown to be successful in identifying trends [291]. The radial visual representations were evaluated in a preliminary involving two clinical researchers. Both displays required longer training time to understand, and, the temporal radial graph presented a challenge when interpreting the tail-ends of the monitoring duration. Evaluators had a difficult time observing patterns only in the -120th hour without being influenced by the +48th hour that was within its immediate vicinity.

For these reasons the radial graphs were not selected for the full evaluation. While these challenges show that radial graphs may involve more training, more research needs to be done to further enhance the visual representation to address those shortcomings. In future work, both radial graphs will be evaluated using similar multidimensional datasets.



(a) (b)

Figure 66: Alternative designs for a cohort based relatively aligned dashboard. (a) A radial graph representing the distribution of heart rate variability scores over 120 hours for each patient. The population is represented as the large arc in the centre of the circle, while each patient is a ring extending at fixed radius from the population. (b) A radial graph representing the temporal trajectory of heart rate variability scores for patients. Similar to the former representation, the temporal radial graph has the population at the centre and each patient as separated rings. A red mark is annotated to determine the zeroth hour, as well as the 48th hour.

8.7 Expert Evaluation

To determine the usability and usefulness of CoRAD, we conducted an expert evaluation. An expert is defined as an individual with at least five years of experience in neonatology and physiologic data. Two key quantitative values that were measured were accuracy of the verbal statements and task completion.

8.7.1 Methodology

The evaluation of CoRAD was conducted with five experts including, clinicians and clinical researchers. A single factor, technique, was varied, with two levels: CoRAD (Figure 64), and stacked bar display (Figure 67). The stacked bar representation is inspired from an alternate design used in the neonatal spells research, however this research involves only the bradycardia episodes [292]. Seven key measures were collected including, demographic information, completion rate, accuracy of response, usability problems verbalized, errors made during the evaluation, posture, and the subjective satisfaction.

The experimental task was to determine and verbalize suspicion of infection for a single patient (a row in CoRAD, a bar in stacked bars). When the participant began the new task they were asked to state “I’m moving to the next patient”, this statement served to mark the end of the former task and the start of a new task. Following exposure to a technique, they were asked to provide feedback on the usability and acceptability of the user interface. The participants were directed to provide their honest opinion of the presented display and to participate in a post-session subjective questionnaire involving a 5 point Likert scale. All verbal discussions, as well as the cursor movements were recorded and transcribed.

Participants received an overview of CoRAD and the stacked bar graph at the start of the experiment, along with the test procedure, and equipment. There was one training scenario consisting of 10 patient datasets. Training consisted of the experimenter reading aloud interpretations of three patient datasets, taking 5 – 10 minutes. Then the participant was provided time to explore the interface and familiarize themselves with the functionality. The 10 patients used in the training set were not included in the evaluation set.

Each *evaluation scenario* consisted of 10 *tasks*. Two evaluation scenarios were carried out for each technique, and repeated for the other technique (data order was randomized). Due to data availability, the same datasets (in random order) were used for the training tasks in both techniques across all participants. The ordering of technique was counterbalanced to limit learning effects. In summary, from the original 33 datasets, 10 were used for training, and of the remaining 23, 20 were randomly selected and used in evaluation scenarios.

Nine personnel with research interests in physiologic data were initially identified. Five were identified to meet all conditions of the inclusion criteria, which included at least five years of expertise in neonatology. The sample was chosen purposefully to represent the local demographics with respect to age, sex, years of experience, and involvement in physiologic research. Trainees and fellows were excluded from this study. There were a total of 5 (participants) x 2 (evaluation scenarios) x 2 (techniques) x 10 (datasets) = 200 evaluation tasks. Study sessions lasted an average of 45 minutes.

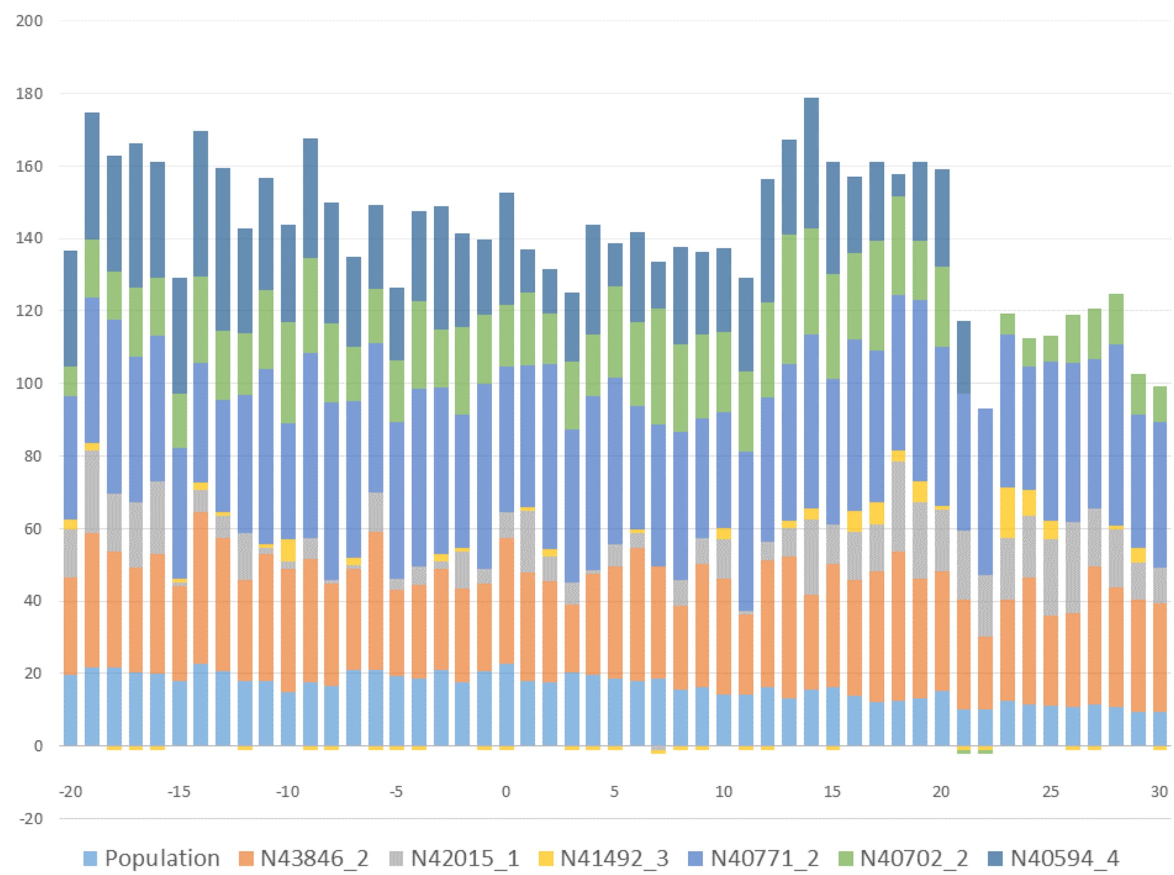


Figure 67: Stacked bar representation used to stack all patients above a population average (bottom). The zeroth mark represents the point of suspicion of infection, and negative numbers illustrate heart rate variability scores in each preceding hours, while positive numbers signify heart rate variability scores in the hours after the event.

3.7.2 Procedure

A laptop computer with Web site/Web application and supporting software was used in a typical office environment. The participant's completion of the task was video recorded for aiding transcription and analysis of time to completion. The evaluation was initiated with a brief description of the CoRAD application, and the participant was made aware that the facilitator would be evaluating the application, rather than the diagnostic abilities of the participant. Participants were then prompted to sign an informed consent sheet that acknowledges: the participation is voluntary, that participation can cease at any time, and that the session will be videotaped but their privacy of identification will be safeguarded.

The participant was then asked to complete a demographic and background questionnaire. Once the demographic questionnaire was completed, the participant was introduced to one of the two techniques. In both the training and experiment phases, the participant was frequently asked to think aloud, describing their analysis process. The participant body posture was observed and entries were made to the observation diary. After each the second exposure to each technique, the participant was asked to complete the post-task questionnaire and elaborate on the task session with the facilitator. After all evaluation scenarios were attempted, the participant completed the post-test satisfaction questionnaire.

3.7.3 Analysis

Each session was video recorded and transcribed (with field notes). Using the open-coding method, sentences were classified using the cluster method of similar codes to generate themes. Themes were identified using informal affinity diagrams. These themes are discussed further in §8.7.4. This iterative approach of integrating more sentence derived themes to adjust the clusters was ongoing throughout the fieldwork to allow emergent themes to be included into the data collection process. The associated themes and distinctions formed the basis of the coding strategy. Review of the evolving themes contributed to the data synthesis and interpretation. To analyse the accuracy of detection the sensitivity-specificity binary classification method was used. This method is a popular clinical measure for determining the efficacy of an intervention [293]. In addition, this method was chosen, as detailed in section

3.4.3 of the literature review because of its depth of use in visual analytics. Average timing was manually determined from the video recording and rounded to the nearest second. Bias was mitigated through independent peer-review of the coding.

8.8 Results

The study yielded data from a total of 200 tasks performed across both conditions (10 datasets × 4 evaluation scenarios × 5 participants). This section highlights the main differences in demographics, accuracy of detection of sepsis, task completion, and subjective feedback received from expert participants.

8.8.1 Demographic Differences

Five clinical researcher participants were recruited in the study and all participants completed each component to completion. All participants had at least ten years of practise in critical care medicine. Two females and three males were recruited. The average age of the sample was 40 – 50 years of age. The average length of total clinical experience was 18 years. All but one subject reported using the computer multiple times a day for analysis purposes. All participants had at least 15 years of experience working with physiologic data. The average reported score of participants' familiarity with physiologic data was 4 out of 5, where 1 represented minimal familiarity and 5 represented expert proficiency. On the same scale, participants reported their familiarity with heart rate variability as 2.5 out of 5 and knowledge of neonatal sepsis as 3.5 out of 5. Two of the five participants were aware of the hypothesis exploring the link between heart rate variability and neonatal sepsis. The years of experience also did statistically differ in the clinical researcher's familiarity with the relationship between heart rate variability and neonatal sepsis.

8.8.2 Accuracy of Detection

Table 6 summarizes the results of the display condition, true positive, true negative, false positive, false negative, and sensitivity and specificity for all tasks performed. True positive refers to the number of true sepsis patients that were correctly identified to be septic. True negative to the correct identification of negative cases as non-septic. False positive refers to the number of patients who were incorrectly identified as positive, and false negative the number of patients who were incorrectly identified as negative. The sensitivity and specificity scores were collected for each condition and an average specificity and sensitivity score was generated.

8.8.3 Task Completion

Table 7 summarizes results of the tasks successfully completed, errors, average time in seconds, as well as the standard deviation in seconds. Non-crucial errors occurred in the CoRAD condition that did not obstruct task completion. The error was a result of using an external monitor that did not reproduce colour saturations, hence the normal distribution histograms were less visible. This error was fixed after the first trial by reverting to the laptop monitor.

8.8.4 Subjective feedback

As a result of the utilisation of the open-coding method for this research, three themes emerged namely: the usefulness and utility; visual encoding; and ultimate choice of preferred visualisation. Each of these themes are further detailed below.

Clinical researchers provided rich subject feedback about the usefulness and utility of both conditions. On the stacked bar representations, clinical researchers noted that as they progressed through each it became progressively difficult to analyse the patient's HRV scoring due to the non-aligned vertical height. The stacked representation was seen to lack the ability to allow the expert to compare a certain temporal range against the rest of the data set. Clinical researchers also noted that using the stacked bar representation required manual scrolling to get a perception of the entire duration of the dataset. The lack of contextual information was noted to be a significant negative of the stacked bar display.

When compared to the stacked bar, CoRAD was perceptually simpler and easier for the experts to use, from expert feedback gained in the post-test survey. The heatmap representation was unanimously noted as being very helpful for analysis. All clinical researchers appreciated having a single view of the dataset. One of the clinical researchers expressed having been confused with the red colour coding, they identified the darker red regions as being more severe. Interactive zooming was frequently used and noted as a positive component. While many experts found the detail view important to their analysis, two experts voiced having options to have the normal distribution appearing as a histogram on a separate display.

CoRAD generated more thoughtful responses, such as: "I like that I can select a region of interest without the rest of the graph obstructing my view. This function makes me focus my attention on the task I want to complete." The stacked bar display drew negative emotions on that topic, one researcher mentioned: "I am not able to see the patient I want to look at. This makes completing the task challenging".

Table 6: Sensitivity and Specificity of both conditions

Participant	Condition	True Positive	True Negative	False Positive	False Negative	Sensitivity	Specificity
1	CoRAD	2	9	4	5	29%	69%
1	Stacked	3	7	6	4	43%	54%
2	CoRAD	2	11	3	4	33%	79%
2	Stacked	0	15	1	4	0%	94%
3	CoRAD	1	13	4	2	33%	76%
3	Stacked	0	13	3	4	0%	81%
4	CoRAD	0	12	5	3	0%	71%
4	Stacked	2	12	2	4	33%	86%
5	CoRAD	2	13	3	2	50%	81%
5	Stacked	1	9	6	4	20%	60%
Average	**CoRAD**	-	-	-	-	**29%**	**75%**
Average	**Stacked**	-	-	-	-	**19%**	**75%**

Table 7: Task Completion measures for both conditions

Participant	Condition	Successfully Completed	Errors	Average Time (seconds)	Standard Deviation (seconds)
1	CoRAD	20	3	25	12
1	Stacked	16	0	23	16
2	CoRAD	20	1	9	11
2	Stacked	17	0	5	5
3	CoRAD	20	0	20	7
3	Stacked	19	0	17	6
4	CoRAD	20	0	16	17
4	Stacked	20	0	15	15
5	CoRAD	20	1	15	8
5	Stacked	18	0	27	19
Average	**CoRAD**	**20**	**1**	**17**	**11**
Average	**Stacked**	**18**	**0**	**18**	**12**

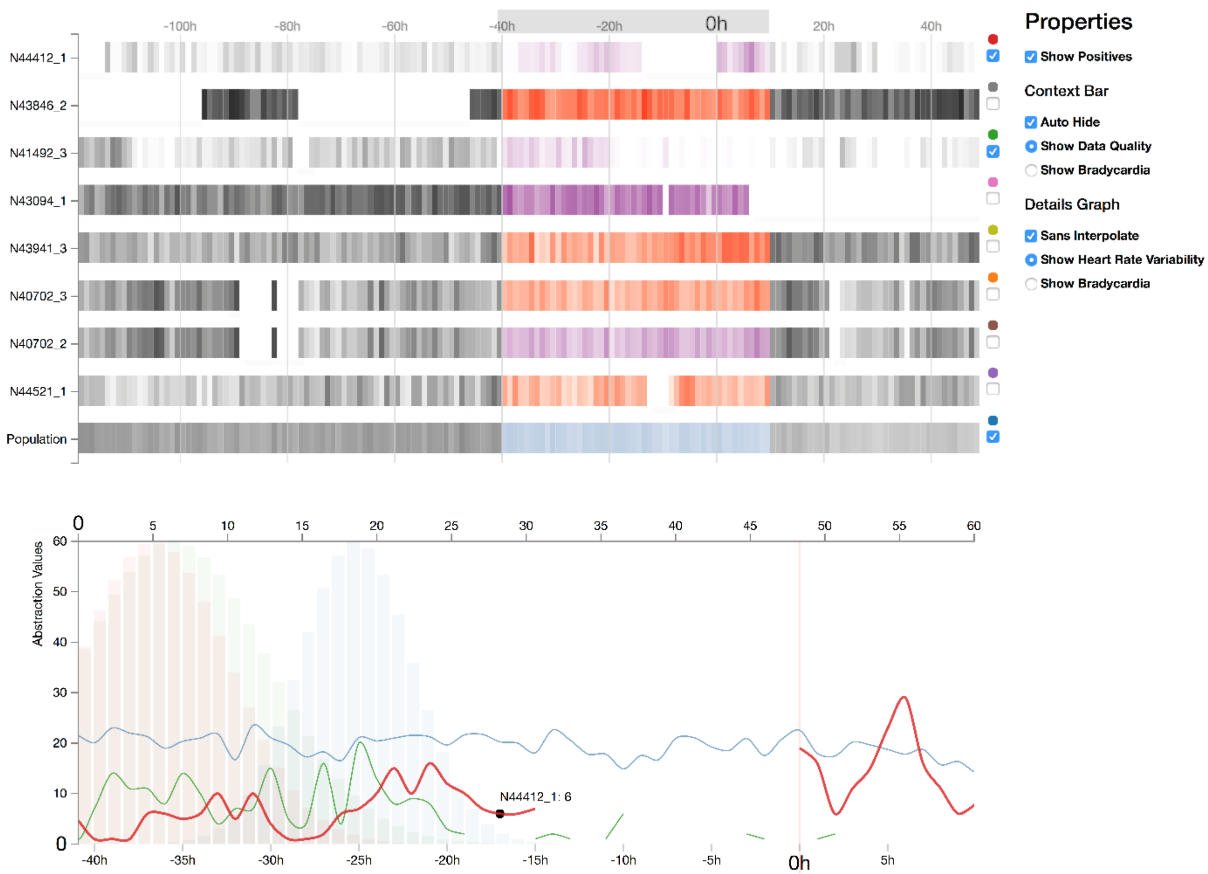


Figure 68: Interactive selection and filtering functions on CoRAD allow the clinical researchers to isolate patients of interest. In this figure, the 'Show Positives' function is selected, which filters patients based on a positive clinical result for neonatal sepsis. The clinical researcher is shown here highlighting -40 hour to +10 hour two positive cases N44412_1 and N41492_3 in the detail view. The average of the population is also highlighted (in blue), to assist the clinical researcher in identifying potential correlations.

The contextual bar was heavily utilized, however three of the five clinical researchers requested to see both bradycardia and data quality at the same time. One clinical researcher found the CoRAD display too cluttered and overwhelming, however that clinician did not use any of the interactive selection and filtering functions. Moreover, that clinical researcher preferred to see a summary graph showing only the most deviant patient. Other clinicians reported high satisfaction with the availability of the interactive selection and filter functions, and stated it helped to reduce excess information. When interactive selections were used, most clinical researchers also utilized the filter to display key patients of interest in the detail view. A typical workflow is illustrated in Figure 68, where two patients of interest are compared to the population mean in the detail view. In the main view, the user has highlighted an interval of interest. All clinical researchers stated the highlight function to be useful for determining changes in HRV across multiple patients at the same time, within salient temporal windows. One clinical researcher started the analysis by immediately highlighting a temporal window, and maintained that same window throughout the entire duration of the analysis. That researcher stated that they did not view data in other durations to be relevant.

One clinical researcher stated a desire to see distributions over only a fixed temporal range. That clinical researcher found the display of the average distribution across the entire duration not significantly helpful for completing their task. Researchers used the detail view to confirm their visual suspicions, one subject verbalized: “I am not sure (whether I am correct) visually about these subsets of patients, I want to see them statistically using the detail view. Ah, I see that my visual interpretations were correct”.

After both conditions were tested, clinical researchers were asked to state their preference for one display. All experts preferred CoRAD over the stacked bar display. All clinical researchers stated they would utilize CoRAD as one of the applications in their analytic toolkit. Three clinical researchers with significant bed-side research interests expressed an inclination to use CoRAD as a tool as part of their bed-side rounds. One clinical researcher mentioned that after some suggested modifications, such as including a dynamic histogram for the normal distribution, they would see themselves actively using CoRAD.

8.9 Discussions and Future Work

An expert evaluation consisting of five domain experts analysing HRV and bradycardia events was conducted in an attempt to predict the infant's neonatal sepsis status. Results from the expert evaluation revealed several key insights. The demographic differences in this study reveal broad coverage in age, sex, and years of experience. Based on the results that were observed, there seems to be little differences between age, gender, and years of experience to both the accuracy and task completion ($p > 0.05$). The relative low score attributed to familiarity of HRV is significant as this measure has yet to be established as a routine clinical indicator in practice [294]. One clinical researcher mentioned that, while they did not use HRV actively, they had knowledge of its potential relevance.

Accuracy of sepsis detection was reported with sensitivity appearing below 50% for both conditions (Table 1). CoRAD allowed for a 10% increase in sensitivity, however it was not statistically significant ($p > 0.05$). With respect to the specificity, both the stacked bar and CoRAD displays indicate an identical score at 75% ($p > 0.05$). The low sensitivity score across both displays may support the notion of a weak link between HRV, bradycardia and neonatal sepsis, thereby providing counter evidence against the initial hypothesis for the dataset used by this evaluation [58], [286], [287]. Since the commencement of this research another independent study has also reported low accuracy results for the detection of late onset neonatal sepsis using these two physiological behaviours as part of the heart rate characteristics approach in a three year observational study [295].

Task completion (Table 7) was significantly higher on CoRAD than on the stacked bar display ($p < 0.05$). All instances of unsuccessful task completion occurred when these clinical researchers failed to analyse one of the required patients in the display. The omitted tasks were not subsequently identified by the clinical researcher in most cases (8 out of 10), in one instance the researcher spoke aloud to confirm whether they may have missed a patient in their analysis. Analysis of the video screen recording, revealed that most of the omitted tasks appear as patients stacked in the middle or upper region of the representation.

Non-crucial errors were seen early in the evaluation with CoRAD, in particular with colour accuracy with the external display utilized in a single experiment. The CoRAD display was subsequently shown on another display which produced accurate colour representation. An additional error were encountered with subject 3 and 7 where the database communication was temporarily timed-out. A refresh of the web page allowed the evaluation to continue. The average time for task completion was not statistically significant between the two conditions (17 vs 18 seconds). Even with the additional number of interactive manipulations that were performed by clinical researchers, CoRAD still allowed the user to perform their task in the same amount of time. The greater functionality afforded by CoRAD, along with the general interest in the tool was not seen to have contributed to longer task completion times.

The general subjective feedback shows greater interest in the CoRAD display. A unanimous agreement was present on the integration of CoRAD as an informatics tool that should be deployed as a tool in the hospital analytics suite. In particular, clinical researchers found having the ability to interactively select, filter, and expose details on demand to be helpful to their analysis workflow. Some researchers report using the tool, however with other forms of data, such as electroencephalogram, or an oxygen saturation dataset. The clinical researchers also suggested two major areas for future work. Including having the option to manually change the colour scheme, allow the context bar to represent both data quality and bradycardia at the same time, and separate the histogram view from the details graph. Future work with CoRAD will address the identified limitations.

CoRAD has shown positive effects in supporting clinical researchers explore patterns across multiple modes of physiologic data using an interactive cohort based visual analytic tool. The CoRAD display was tested in the context of an application by conducting an expert evaluation and experimentation against a control stacked bar display. Exposure to CoRAD within this limited case study, resulted in interest on the part of the clinical researchers to use this tool in other scenarios, such as electrocardiography and oxygen saturation variability. The relatively aligned heatmap allowed each researcher to rapidly identify event details, which was more difficult on the control display. However, open challenges remain in studying alternative visualizations that can be used to display multiple features, such as data quality, and bradycardia without producing visual clutter.

8.10 Limitations

There are limitations with the presented research. The use of a single site to conduct the research has limited external generalizability. However, the use of randomization within the pool of experienced participants allowed us to mitigate external validity. Moreover, as exploratory research, results produced in this study allow for subsequent research to study generalizations to other sites. Secondly, due to the limited dataset that was available, the training dataset was shared by all participants, this may have exposed participants to a limited set of features that were only visible within the training dataset. The questionnaire relied in several self-reported quantitative values (years of experience, expertise, etc.). There may have been biases introduced from the self-reporting of those measures [296].

8.11 Chapter Summary

This chapter introduces CoRAD a visual analytic tool for exploring patterns across cohort populations, and to conduct case-controlled studies. CoRAD is demonstrated as a DVAM instance within the TDVA framework that has been created in the TDVA platform and was developed using the TDVA methodology. The CoRAD display was evaluated by clinical researchers and results were presented in §8.7. Future versions of CoRAD display will address limitations identified in this chapter, including the need to make multiple contextual information visible without contributing to visual clutter.

9. Conclusion and Future Work

This thesis has presented the TDVA framework for supporting dynamic visual analytics in complex environments. The key insight of the TDVA framework is a novel publisher software called mPub, which allows users in complex domains to generate multiple instances of DVAMs. These DVAMs are developed following the TDVA methodology proposed in this thesis, and using the TDVA platform as a physical environment supporting the deployment. Multiple DVAMs were instantiated, including two as a prototype application, and two others as full instantiations in chapters seven and eight. This chapter concludes the dissertation. In §9.1 a summary of the major work and contributions presented in this thesis are outlined. Future work is discussed in §9.2, followed by concluding remarks in §9.3.

9.1 Summary and Contributions

Sensors have become cheaper to manufacture, and the storage required to capture all generated data continues to expand. Organizations ranging from agriculture to intelligence, now actively seek insight from their large collections of data. One of the methods to providing that insight, has been to create visual representations of historical data. However, the limitation with that approach, is that those visuals were often static and query-based. This implied a silo approach to hypothesis generation.

Modern approaches now adapt interactive functionalities applied to visual interfaces to support visual analysis, the methods supporting that workflow termed as visual analytics. A specialized, but growing field of interest is in visual analytics applied to the temporal domain, called dynamic visual analytics. To that end, this dissertation presents a novel framework, called the Tri-event parameter based Dynamic Visual Analytics.

The focus of this dissertation has been on designing a framework, establishing a methodology to support the use of the framework for the creation of a physical platform that effectively allows users existing in the dynamic domain to access highly interactive visual interfaces to perform analysis tasks. Two analytic tasks were addressed: hypothesis generation and hypothesis testing using case-controlled studies. In order to arrive at requirements for the framework, a systematic review was conducted and presented in chapter four. This was followed by a qualitative study presented in chapter five, that elicited two novel concepts namely the tri-event parameters (§5.2), and the Exploration-Consumption continuum (§5.2.1). Both tasks resulted in the development of the TDVA framework (§6.2), which motivated the development of a TDVA methodology (§6.3), and a template TDVA platform (§6.4) that can be deployed in a real-time environment through an extension to the Artemis platform.

The feasibility of the proposed framework was examined by providing two prototype DVAMs instantiations within the TDVA platform, namely the Heart Rate Variability Graph (§6.5.1) and SeqEvent (§6.5.2). Heart Rate Variability Graph addressed the consumption requirements of complex users, while SeqEvent DVAM addressed the need for exploring temporal sequence of primitive events for hypothesis generation purposes. These instantiations were followed by two full instantiations, the first being PhysioEx presented in chapter seven and the second is CoRAD in chapter eight. PhysioEx also addressed the hypothesis generation requirement, while CoRAD addressed the requirement for effective physiologic visual analytic tools supporting case-controlled hypothesis testing.

The key insight in the proposed framework, is that by employing temporal tri-event parameters, along with considerations for the unique requirements of the end-user, we can develop effective visual mediums to support analysis of complex data. The key component in that framework that supports such tasks is the Dynamic Visual Analytic Mart Publisher. The publisher serves as an intelligent agent, providing the user with tools that directly support analysis tasks. In this dissertation, the publisher was presented supporting researcher-oriented tasks across three DVAMs including SeqEvent, PhysioEx, and CoRAD. A DVAM was illustrated using the Heart Rate Variability Graph to support consumers.

In summary, the framework, methodology, platform, and techniques presented in this dissertation present a viable solution to address challenges of supporting dynamic visual analytic workflows in complex data environments.

9.2 Future Work

The contributions proposed in this dissertation are made to specific challenges within a larger open research area. In addition to the future work presented in §7.8 and §8.9, this section presents the following avenues for future work.

9.2.1 Automated Context Awareness

The instantiations of PhysioEx and CoRAD were performed manually. The mPub engine, introduced in §6.2.1 currently makes distinctions between consumer and researcher using input provided directly by the user. This can be a cumbersome and tedious process, moreover, it may be an additional workflow introduced into an existing intricate environment. What is needed, therefore, is a level of intelligence applied to the mPub engine that allows it to automatically detect the context based on events produced in the data stream, or physical location. For instance, if potentially pathologic complex and multidimensional events were observed for a particular patient or system, the event stream processor should alert the mPub engine to modify its context to provide visual representations suitable for the new condition.

Alternatively, if the mPub engine detects that a physical location has changed, for instance, if a clinician is removed from the bed-side, then the visual representations are adapted from consumption displays to support more research tasks. To effectively apply these adaptions, there needs to be extensive post-instantiation studies that are conducted. These studies should evaluate the impact of such awareness functionality on domain experts, and arrive at evidence justifying its adoption.

9.2.2 Instantiation in Real-time

The TDVA framework proposed in this dissertation supports many functionalities that ingest real-time streaming data to support consumers and researchers. The particular instantiations provided in this dissertation, however, directly support the researcher. The effects this framework would have on a consumer in a complex real-time environment where not investigated in this research. Hence that presents a unique opportunity as future work, to develop instantiations for the complex real-time user. Those instantiations however, need to consider the salience of balancing interactivity with the needs of the user as determined by their immediate context and level of engagement.

This Exploration-Consumption continuum proposed in chapter five presents a conceptual template that can be used to advance this work. For instance, by using urgency and engagement as metrics, an automated engine can be developed to support very specific focal monitoring tasks. The framework, methodology, and platform presented in this thesis can be used to advance such a task.

9.2.3 Scaling out to the cloud

The TDVA platform currently assumes all components exist in a local cluster. Moreover, a key aspect of the TDVA framework is to support isolated independent dynamic visual analytic marts. However as number of instantiations increase, and the dataset expands, there is a need to modify the platform to support horizontal scalability. This involves modifying the mPub engine to support parallelized and distributed instantiations.

9.2.4 Extending to non-domain experts

The case study applications demonstrated in this thesis were evaluated using domain experts, who were familiar with concepts that were important in maximizing the utility of the DVAMs. In future studies, the framework will be demonstrated for non-domain experts using visual cues to aid decision making when key concepts are unfamiliar.

9.3 Concluding Statements

Large volumes of sensor data continue to be collected across all major industries, however tools that allow end-users to analyse those datasets are limited. Most dynamic visual analytic tools are developed to support a standalone workflow, where each application follows a unique development pipeline. Besides, they do not support the temporal tri-event parameters: trajectory, frequency, and duration, which are persistently observed in complex domains. I provide an alternate instantiation strategy by introducing a framework, methodology and platform design that integrates the tri-event parameter and used to construct dynamic visual analytic applications. The framework allows flexibility to produce instantiations that support unique workflows of consumers and researchers. The introduction of such framework, not only addresses the challenge of supporting visual analysis of complex data, but also provides a blueprint for future advancement of intelligent context-aware systems.

Reference

[1] C. McGregor, "Big Data in Neonatal Intensive Care," *Computer*, vol. 46, no. 6, pp. 54–59, 2013.

[2] C. C. Aggarwal, *Data streams: models and algorithms*, vol. 31. Springer Science & Business Media, 2007.

[3] S. K. Card, J. D. Mackinlay, and B. Shneiderman, *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[4] S. Blackburn, *Maternal, fetal, & neonatal physiology*. Elsevier Health Sciences, 2014.

[5] N. B. Gabler, N. Duan, S. Vohra, and R. L. Kravitz, "N-of-1 Trials in the Medical Literature: A Systematic Review," *Medical Care*, vol. 49, no. 8, 2011.

[6] P. Craig, P. Dieppe, S. Macintyre, S. Michie, I. Nazareth, and M. Petticrew, "Developing and evaluating complex interventions: the new Medical Research Council guidance," *Bmj*, vol. 337, p. a1655, 2008.

[7] F. Mansmann, F. Fischer, and D. A. Keim, "Dynamic visual analytics—facing the real-time challenge," in *Expanding the Frontiers of Visual Analytics and Visualization*, Springer, 2012, pp. 69–80.

[8] A. Thudt, U. Hinrichs, and S. Carpendale, "The bohemian bookshelf: supporting serendipitous book discoveries through information visualization," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 1461–1470.

[9] A. Foster and N. Ford, "Serendipity and information seeking: an empirical study," *Journal of Documentation*, vol. 59, no. 3, pp. 321–340, 2003.

[10] C. M. Burns, G. J. Skraaning, G. A. Jamieson, N. Lau, J. Kwok, R. Welch, and G. Andresen, "Evaluation of ecological interface design for nuclear process control: situation awareness effects.," *Human factors*, vol. 50, no. 4, pp. 663–679, 2008.

[11] M. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995.

[12] T. H. Davenport and L. Prusak, *Working knowledge: How organizations manage what they know*. Harvard Business Press, 1998.

[13] C. Oppenheim, J. Stenson, and R. M. S. Wilson, "Studies on information as an asset I: definitions," *Journal of Information Science*, vol. 29, no. 3, pp. 159–166, 2003.

[14] P. Checkland and S. Holwell, "Information, systems and information systems: making sense of the field," 1997.

[15] J. J. Thomas and K. A. Cook, *Illuminating the path: The research and development agenda for visual analytics*. IEEE Computer Society Press, 2005.

[16] P. McLachlan, T. Munzner, E. Koutsofios, and S. North, "LiveRAC: interactive visual exploration of system management time-series data," in *Proceedings of the twenty-sixth*

annual SIGCHI conference on Human factors in computing systems. ACM, 2008, pp. 1483–1492.

[17] R. Jaeschke, G. H. Guyatt, J. Keller, and J. Singer, “Ascertaining the meaning of change in quality of life questionnaire score: data from N of 1 randomized control trials,” *Control Clin Trials*, vol. 12, no. Suppl, pp. S226–S233, 1991.

[18] D. H. N. Barlow, M. Hersen, M. D. Barlow, M. Nock, and M. Hersen, *Single case experimental designs: Strategies for studying behavior for change*, no. Sirsi) i9780205474554. 2009.

[19] T. R. Kratochwill, *Single subject research: Strategies for evaluating change*. Academic Press, 2013.

[20] G. H. Guyatt, J. L. Keller, R. Jaeschke, D. Rosenbloom, J. D. Adachi, and M. T. Newhouse, “The n-of-1 randomized controlled trial: clinical usefulness: our three-year experience,” *Annals of Internal Medicine*, vol. 112, no. 4, pp. 293–299, 1990.

[21] E. O. Lillie, B. Patay, J. Diamant, B. Issell, E. J. Topol, and N. J. Schork, “The n-of-1 clinical trial: the ultimate strategy for individualizing medicine?,” *Personalized medicine*, vol. 8, no. 2, pp. 161–173, 2011.

[22] A. E. Kazdin, *Single-case research designs: Methods for clinical and applied settings.* Oxford University Press, 2011.

[23] P. Nishith, D. E. Hearst, K. T. Mueser, and E. B. Foa, “PTSD and major depression: Methodological and treatment considerations in a single case design,” *Behavior Therapy*, vol. 26, no. 2, pp. 319–335, 1995.

[24] K. J. Rothman, S. Greenland, and T. L. Lash, *Modern epidemiology*. Lippincott Williams & Wilkins, 2008.

[25] L. Golab and M. T. Ozsu, “Issues in data stream management,” *ACM Sigmod Record*, vol. 32, no. 2, pp. 5–14, 2003.

[26] F. Wang, S. Liu, P. Liu, and Y. Bai, “Bridging physical and virtual worlds: complex event processing for RFID data streams,” in *Advances in Database Technology-EDBT 2006*, Springer, 2006, pp. 588–607.

[27] C. Yu, Y. Zhong, T. Smith, I. Park, and W. Huang, “Visual data mining of multimedia data for social and behavioral studies,” *Information Visualization*, vol. 8, no. 1, pp. 56–70, Feb. 2009.

[28] P. Sanderson, “The multimodal world of medical monitoring displays.,” *Applied ergonomics*, vol. 37, no. 4, pp. 501–12, Jul. 2006.

[29] A. J. Starmer, N. D. Spector, R. Srivastava, A. D. Allen, C. P. Landrigan, and T. C. Sectish, “I-pass, a mnemonic to standardize verbal handoffs.,” *Pediatrics*, vol. 129, no. 2, pp. 201–4, Feb. 2012.

[30] J. Horsky, L. Gutnik, and V. L. Patel, “Technology for emergency care: cognitive and workflow considerations.,” *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, pp. 344–348, 2006.

[31] K. Orphanou, E. Keravnou, and J. Moutiris, “Integration of Temporal Abstraction and Dynamic Bayesian Networks in Clinical Systems. A preliminary approach,” in *OASIcs-OpenAccess Series in Informatics*, 2012, vol. 28.

[32] A. Thommandram, J. E. Pugh, J. M. E. Smieee, C. M. Smieee, and A. G. James, “Classifying Neonatal Spells Using Real-Time Temporal Analysis of Physiological Data Streams : Algorithm Development,” in *IEEE EMBS Special Topic Conference on Point-of-Care Healthcare Technologies*, 2013, pp. 240 – 243.

[33] D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar, “Big data in health care: using analytics to identify and manage high-risk and high-cost patients,” *Health Affairs*, vol. 33, no. 7, pp. 1123–1131, 2014.

[34] B. Goldstein, “Heart rate characteristics in neonatal sepsis: a promising test that is still premature.,” *Pediatrics*, vol. 115, no. 4, pp. 1070–2, 2005.

[35] M. P. Griffin, T. M. O’Shea, E. A. Bissonette, F. E. Harrell, D. E. Lake, and J. R. Moorman, “Abnormal heart rate characteristics preceding neonatal sepsis and sepsis-like illness,” *Pediatric research*, vol. 53, no. 6, pp. 920–926, 2003.

[36] C. Catley, K. Smith, C. McGregor, A. James, and J. M. Eklund, “A Framework for Multidimensional Real-Time Data Analysis: A Case Study for the Detection of Apnoea of Prematurity,” *International Journal of Computational Models and Algorithms in Medicine (IJCMAM)*, vol. 2, no. 1, pp. 16–37, 2011.

[37] M. Blount, M. R. Ebling, J. M. Eklund, A. G. James, C. McGregor, N. Percival, K. P. Smith, and D. Sow, “Real-Time Analysis for Intensive Care: Development and Deployment of the Artemis Analytic System,” *Engineering in Medicine and Biology Magazine, IEEE*, vol. 29, no. 2, pp. 110–118, 2010.

[38] C. McGregor, A. James, M. Eklund, D. M. Sow, M. Ebling, and M. Blount, “Real-time multidimensional temporal analysis of complex high volume physiological data streams in the neonatal intensive care unit.,” in *MedInfo*, 2013, pp. 362–366.

[39] A. Thommandram, J. E. Pugh, J. M. Eklund, C. McGregor, and A. G. James, “Classifying neonatal spells using real-time temporal analysis of physiological data streams: Algorithm development,” in *Point-of-Care Healthcare Technologies (PHT), 2013 IEEE*, 2013, pp. 240–243.

[40] M. J. Richards, J. R. Edwards, D. H. Culver, and R. P. Gaynes, “Nosocomial infections in pediatric intensive care units in the United States. National Nosocomial Infections Surveillance System.,” *Pediatrics*, vol. 103, no. 4, p. e39, Apr. 1999.

[41] R. Kamaleswaran and C. McGregor, “A Review of Visual Representations of Physiologic Data,” *JMIR Medical Informatics (in review)*.

[42] R. Kamaleswaran and C. McGregor, “A Real-Time Multi-dimensional Visualization Framework for Critical and Complex Environments,” in *Computer-Based Medical Systems (CBMS), 2014 IEEE 27th International Symposium on*, 2014, pp. 325–328.

[43] R. Kamaleswaran, R. Wehbe, J. E. Pugh, L. Nacke, C. Mcgregor, and A. James, “Collaborative Multi-Touch Clinical Handover System for the Neonatal Intensive Care Unit,” *electronic*

Journal of Health Informatics, vol. (In Print), 2015.

[44] R. Kamaleswaran, C. Collins, A. G. James, and C. Mcgregor, "PhysioEx: Visual Analysis of Physiological Event Streams," *Eurographics Conference on Visualization (EuroVis) 2016*, vol. 35, no. 3, 2016.

[45] R. Kamaleswaran, C. Collins, A. James, and C. McGregor, "CoRAD: Visual Analytics for Cohort Analysis," in *IEEE International Conference on Healthcare Informatics 2016 (ICHI 2016) (in review)*.

[46] H. Blencowe, S. Cousens, M. Z. Oestergaard, D. Chou, A.-B. Moller, R. Narwal, A. Adler, C. Vera Garcia, S. Rohde, L. Say, and others, "National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications," *The Lancet*, vol. 379, no. 9832, pp. 2162–2172, 2012.

[47] H. Klocke, S. Trispel, G. Rau, U. Hatzky, and D. Daub, "An anesthesia information system for monitoring and record keeping during surgical anesthesia.," *Journal of clinical monitoring*, vol. 2, no. 4, pp. 246–261, 1986.

[48] F. C. Inc., "Fifty Years of Physiologic Monitors," 2006. [Online]. Available: http://femtosimclinical.com/History of Physiologic Monitors.htm.

[49] F. A. Drews, "Patient monitors in critical care: Lessons for improvement," *Advances in Patient Safety: New Directions and Alternative Approaches. Rockville, MD: Agency for Healthcare Research and Quality*, 2008.

[50] M. M. Shabot, "Software for computers and calculators in critical care medicine.," *Software in healthcare*, vol. 3, no. 1, pp. 26–30, 1984.

[51] M. M. Shabot, P. D. Carlton, S. Sadoff, and L. Nolan-Avila, "Graphical reports and displays for complex ICU data: a new, flexible and configurable method," *Computer Methods and Programs in Biomedicine*, vol. 22, no. 1, pp. 111–116, Mar. 1986.

[52] U. Wenkebach, B. Pollwein, and U. Finsterer, "Visualization of large datasets in intensive care.," *Proceedings / the ... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care*, pp. 18–22, 1992.

[53] C. Schulz, M. Endsley, and E. Kochs, "Situation awareness in anesthesia: concept and research," *Anesthesiology*, 2013.

[54] R. J. Martin, J. M. Abu-Shaweesh, and T. M. Baird, "Apnoea of prematurity," *Paediatric respiratory reviews*, vol. 5, pp. S377–S382, 2004.

[55] M. P. Griffin, D. E. Lake, T. M. O'Shea, and J. R. Moorman, "Heart rate characteristics and clinical signs in neonatal sepsis," *Pediatric research*, vol. 61, no. 2, pp. 222–227, 2007.

[56] M. R. Hammerschlag, J. O. Klein, M. Herschel, F. C. Chen, and R. Fermin, "Patterns of use of antibiotics in two newborn nurseries.," *The New England journal of medicine*, vol. 296, no. 22, pp. 1268–1269, 1977.

[57] C. McGregor, C. Catley, and A. James, "A process mining driven framework for clinical guideline improvement in critical care," *A process mining driven framework for clinical guideline improvement in critical care*, vol. 765, 2012.

[58] A. A. Flower, J. R. Moorman, D. E. Lake, and J. B. Delos, “Periodic heart rate decelerations in premature infants,” *Experimental Biology and Medicine*, vol. 235, no. 4, pp. 531–538, 2010.

[59] T. G. Buchman, “Nonlinear dynamics, complex systems, and the pathobiology of critical illness.,” *Current opinion in critical care*, vol. 10, no. 5, p. 378, 2004.

[60] H. A. Simon, “The New Science of Management Decision (revised edition) Prentice-Hall,” *Englewood Cliffs, NJ*, 1977.

[61] C.-Y. Chong and S. P. Kumar, “Sensor networks: evolution, opportunities, and challenges,” *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[62] A. Labrinidis and H. Jagadish, “Challenges and opportunities with big data,” *Proceedings of the VLDB Endowment*, vol. 5, no. 12, pp. 2032–2033, 2012.

[63] R. Kitchin, “The real-time city? Big data and smart urbanism,” *GeoJournal*, vol. 79, no. 1, pp. 1–14, 2014.

[64] L. Cohen, G. Avrahamibakish, M. Last, a Kandel, and O. Kipersztok, “Real-time data mining of non-stationary data streams from sensor networks☆,” *Information Fusion*, vol. 9, no. 3, pp. 344–353, Jul. 2008.

[65] E. Hoke, J. Sun, J. Strunk, G. Ganger, and C. Faloutsos, “Intemon: continuous mining of sensor data in large-scale self-infrastructures,” *ACM SIGOPS Operating Systems Review*, vol. 40, no. 3, pp. 38–44, 2006.

[66] G. Andrienko, N. Andrienko, and S. Wrobel, “Visual analytics tools for analysis of movement data,” *ACM SIGKDD Explorations Newsletter*, vol. 9, no. 2, pp. 38–46, 2007.

[67] J. Bernard, N. Wilhelm, B. Krüger, T. May, T. Schreck, and J. Kohlhammer, “MotionExplorer: exploratory search in human motion capture data based on hierarchical aggregation.,” *IEEE transactions on visualization and computer graphics*, vol. 19, no. 12, pp. 2257–66, Dec. 2013.

[68] R. Maciejewski, S. Rudolph, R. Hafen, A. M. Abusalah, M. Yakout, M. Ouzzani, W. S. Cleveland, S. J. Grannis, and D. S. Ebert, “A visual analytics approach to understanding spatiotemporal hotspots.,” *IEEE transactions on visualization and computer graphics*, vol. 16, no. 2, pp. 205–220, 2010.

[69] M. Rios and J. Lin, “Distilling massive amounts of data into simple visualizations: Twitter case studies,” in *Workshop on Social Media Visualization at ICWSM*, 2012.

[70] M. De Choudhury and S. Counts, “Identifying Relevant Social Media Content : Leveraging Information Diversity and User Cognition,” 2011.

[71] D. Fisher, I. Popov, and S. Drucker, “Trust me, i’m partially right: incremental visualization lets analysts explore large datasets faster,” in *ACM annual conference on Human Factors in Computing Systems*, 2012, pp. 1673–1682.

[72] D. Fisher, R. DeLine, M. Czerwinski, and S. Drucker, “Interactions with big data analytics,” *interactions*, vol. 19, no. 3, pp. 50–59, 2012.

[73] T. a Manolio, R. L. Chisholm, B. Ozenberger, D. M. Roden, M. S. Williams, R. Wilson, D. Bick,

E. P. Bottinger, M. H. Brilliant, C. Eng, K. a Frazer, B. Korf, D. H. Ledbetter, J. R. Lupski, C. Marsh, D. Mrazek, M. F. Murray, P. H. O'Donnell, D. J. Rader, M. V Relling, A. R. Shuldiner, D. Valle, R. Weinshilboum, E. D. Green, and G. S. Ginsburg, "Implementing genomic medicine in the clinic: the future is here.," *Genetics in medicine : official journal of the American College of Medical Genetics*, vol. 15, no. 4, pp. 258–67, Apr. 2013.

[74] P. A. Whittaker, "What is the relevance of bioinformatics to pharmacology?," *Trends in pharmacological sciences*, vol. 24, no. 8, pp. 434–439, 2003.

[75] I. I. Gueorguieva, I. A. Nestorov, and M. Rowland, "Fuzzy Simulation of Pharmacokinetic Models : Case Study of Whole Body Physiologically Based Model of Diazepam," vol. 31, no. 3, pp. 185–213, 2004.

[76] Y. Zhang, F. a. Drews, D. R. Westenskow, S. Foresti, J. Agutter, J. C. Bermudez, G. Blike, and R. Loeb, "Effects of Integrated Graphical Displays on Situation Awareness in Anaesthesiology," *Cognition, Technology & Work*, vol. 4, no. 2, pp. 82–90, Jun. 2002.

[77] J. Kehrer and H. Hauser, "Visualization and visual analysis of multifaceted scientific data: a survey.," *IEEE transactions on visualization and computer graphics*, vol. 19, no. 3. pp. 495–513, Mar-2013.

[78] T. Von Landesberger, A. Kuijper, T. Schreck, J. Kohlhammer, J. J. Van Wijk, J. Fekete, and D. W. Fellner, "Visual Analysis of Large Graphs : State-of-the-Art and Future Research Challenges," vol. xx, pp. 1–28, 2011.

[79] M. Burns, M. Haidacher, and W. Wein, "Feature emphasis and contextual cutaways for multimodal medical visualization," in *Proceedings of the 9th Joint Eurographics/IEEE VGTC conference on Visualization*, 2007, pp. 275–282.

[80] M. C. de Oliveira and H. Levkowitz, "From visual data exploration to visual data mining: A survey," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 9, no. 3, pp. 378–394, 2003.

[81] T. Pedersen and C. Jensen, "Multidimensional data modeling for complex data," *Data Engineering, 1999. …*, 1999.

[82] D. a. Keim, "Information visualization and visual data mining," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 1, pp. 1–8, 2002.

[83] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski, "Visual methods for analyzing time-oriented data.," *IEEE transactions on visualization and computer graphics*, vol. 14, no. 1, pp. 47–60, 2008.

[84] M. C. Lohrenz, J. G. Trafton, R. M. Beck, and M. L. Gendron, "A model of clutter for complex, multivariate geospatial displays.," *Human factors*, vol. 51, no. 1, pp. 90–101, 2009.

[85] A. Cuzzocrea, I.-Y. Song, and K. C. Davis, "Analytics over large-scale multidimensional data: the big data revolution!," in *Proceedings of the ACM 14th international workshop on Data Warehousing and OLAP*, 2011, pp. 101–104.

[86] J. Gray, D. T. Liu, M. Nieto-santisteban, A. S. Szalay, D. Dewitt, and G. Heber, "Scientific Data Management in the Coming Decade One Microsoft Way Scientific Data Management

in the Coming Decade," *ACM SIGMOD Record*, vol. 34, no. 4, pp. 34–41, 2005.

[87] M. Marinelli, V. Positano, S. G. Nekolla, P. Marcheschi, G. Todiere, N. Esposito, S. Puzzuoli, G. a L'Abbate, P. Marraccini, and D. Neglia, "Hybrid image visualization tool for 3D integration of CT coronary anatomy and quantitative myocardial perfusion PET.," *International journal of computer assisted radiology and surgery*, vol. 8, no. 2, pp. 221–32, Mar. 2013.

[88] T. Kacmarczyk, P. Waltman, A. Bate, P. Eichenberger, and R. Bonneau, "Comparative microbial modules resource: generation and visualization of multi-species biclusters.," *PLoS computational biology*, vol. 7, no. 12. p. e1002228, Dec-2011.

[89] C. G. Chute, M. Ullman-Cullere, G. M. Wood, S. M. Lin, M. He, and J. Pathak, "Some experiences and opportunities for big data in translational research.," *Genetics in medicine : official journal of the American College of Medical Genetics*, vol. 15, no. 10, pp. 802–9, Oct. 2013.

[90] J. K. Laurila, D. Gatica-Perez, I. Aad, O. Bornet, T.-M.-T. Do, O. Dousse, J. Eberle, M. Miettinen, and others, "The mobile data challenge: Big data for mobile computing research," in *Pervasive Computing*, 2012, no. EPFL-CONF-192489.

[91] R. Kimball and M. Ross, *The data warehouse toolkit: the complete guide to dimensional modeling*. John Wiley & Sons, 2011.

[92] M. Golfarelli and S. Rizzi, *Data Warehouse Design: Modern Principles and Methodologies*, 1st ed. New York, NY, USA: McGraw-Hill, Inc., 2009.

[93] P. Zhao, X. Li, D. Xin, and J. Han, "Graph cube: on warehousing and OLAP multidimensional networks," in *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, 2011, pp. 853–864.

[94] T. Freudenreich, P. Furtado, C. Koncilia, M. Thiele, F. Waas, and R. Wrembel, "An On-Demand ELT Architecture for Real-Time BI," in *Enabling Real-Time Business Intelligence*, Springer, 2013, pp. 50–59.

[95] A. Cuzzocrea, L. Bellatreche, and I.-Y. Song, "Data warehousing and OLAP over big data: current challenges and future research directions," in *Proceedings of the sixteenth international workshop on Data warehousing and OLAP*, 2013, pp. 67–70.

[96] A. Rind, S. Miksch, W. Aigner, T. Turic, and M. Pohl, "VisuExplore: gaining new medical insights from visual exploration," in *Proc. Int. Workshop on Interactive Systems in Healthcare (WISH@ CHI2010)*, 2010, pp. 149–152.

[97] C. Despont-Gros, H. Mueller, and C. Lovis, "Evaluating user interactions with clinical information systems: a model based on human-computer interaction models.," *Journal of biomedical informatics*, vol. 38, no. 3, pp. 244–255, 2005.

[98] R. Grams, "The 'new' America electronic medical record (EMR)—design criteria and challenge," *Journal of medical systems*, vol. 33, no. 6, pp. 409–411, 2009.

[99] J. Richardson, K. Schlegel, B. Hostmann, and N. McMurchy, "Magic quadrant for business intelligence platforms," *Core research note G*, vol. 154227, 2008.

[100] F. Fischer, J. Fuchs, F. Mansmann, and D. A. Keim, “BANKSAFE: Visual analytics for big data in large-scale computer networks,” *Information Visualization*, vol. 14, no. 1, pp. 51–61, 2015.

[101] A. Voellmy, H. Kim, and N. Feamster, “Procera: a language for high-level reactive network control,” in *Proceedings of the first workshop on Hot topics in software defined networks*, 2012, pp. 43–48.

[102] G. Alexander and N. Staggers, “A Systematic Review of the Designs of Clinical Technology,” vol. 32, no. 3, pp. 252–279, 2009.

[103] M. Görges and N. Staggers, “Evaluations of physiological monitoring displays: a systematic review.,” *Journal of clinical monitoring and computing*, vol. 22, no. 1, pp. 45–66, Feb. 2008.

[104] A. Rind, T. D. Wang, W. Aigner, S. Miksch, K. Wongsuphasawat, C. Plaisant, and B. Shneiderman, “Interactive information visualization to explore and query electronic health records,” *Foundations and Trends in Human--Computer Interaction*, vol. 5, no. 3, pp. 207–298, 2011.

[105] M. Stacey, C. Mcgregor, and M. Tracy, “An architecture for multi-dimensional temporal abstraction and its application to support neonatal intensive care,” *Management*, pp. 3752–3756, 2007.

[106] D. Ferrucci and A. Lally, “UIMA: an architectural approach to unstructured information processing in the corporate research environment,” *Natural Language Engineering*, vol. 10, no. 3–4, pp. 327–348, 2004.

[107] D. Ferrucci, E. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A. A. Kalyanpur, A. Lally, J. W. Murdock, E. Nyberg, J. Prager, and others, “Building Watson: An overview of the DeepQA project,” *AI magazine*, vol. 31, no. 3, pp. 59–79, 2010.

[108] P. T. Goetz and B. O’Neill, *Storm blueprints: Patterns for distributed real-time computation*. Packt Publishing Ltd, 2014.

[109] R. Ranjan, “Streaming Big Data Processing in Datacenter Clouds,” *Cloud Computing, IEEE*, vol. 1, no. 1, pp. 78–83, 2014.

[110] F. Waas, R. Wrembel, T. Freudenreich, M. Thiele, C. Koncilia, and P. Furtado, “On-demand ELT architecture for right-time BI: Extending the vision,” *International Journal of Data Warehousing and Mining (IJDWM)*, vol. 9, no. 2, pp. 21–38, 2013.

[111] M. Stacey and C. McGregor, “Temporal abstraction in intelligent clinical data analysis: a survey.,” *Artificial intelligence in medicine*, vol. 39, no. 1, pp. 1–24, Jan. 2007.

[112] N. K. Wade and M. Swanston, *Visual perception: An introduction*, 3rd ed. Psychology Press, 2013.

[113] C. Ware, *Information visualization: perception for design*. Elsevier, 2013.

[114] M. Friendly and D. J. Denis, “Milestones in the history of thematic cartography, statistical graphics, and data visualization,” *Seeing Science: Today American Association for the Advancement of Science*, 2008.

[115] A. Vande Moere, “Time-Varying Data Visualization Using Information Flocking Boids,” in

IEEE Symposium on Information Visualization, 2004, pp. 97–104.

[116] C. Tominski, J. Abello, and H. Schumann, “Axes-based visualizations with radial layouts,” in *Proceedings of the 2004 ACM symposium on Applied computing - SAC ’04*, 2004, p. 1242.

[117] S. Havre, B. Hetzler, and L. Nowell, “ThemeRiver: visualizing theme changes over time,” in *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*, 2000, pp. 115–123.

[118] D. A. Keim, J. Schneidewind, and M. Sips, “CircleView: a new approach for visualizing time-related multidimensional data sets,” in *Proceedings of the working conference on Advanced visual interfaces*, 2004, pp. 179–182.

[119] T. Gschwandtner, W. Aigner, K. Kaiser, S. Miksch, and A. Seyfang, “CareCruiser: exploring and visualizing plans, events, and effects interactively,” in *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, 2011, pp. 43–50.

[120] W. Horn, C. Popow, and L. Unterasinger, “Support for fast comprehension of ICU data: visualization using metaphor graphics.,” *Methods of information in medicine*, vol. 40, no. 5, pp. 421–4, Jan. 2001.

[121] R. W. Albert, J. A. Agutter, N. D. Syroid, K. B. Johnson, R. G. Loeb, and D. R. Westenskow, “A simulation-based evaluation of a graphic cardiovascular display.,” *Anesthesia and analgesia*, vol. 105, no. 5, p. 1303–contents, 2007.

[122] S. H. Koch, C. Weir, D. Westenskow, M. Gondan, J. Agutter, M. Haar, D. Liu, M. Görges, and N. Staggers, “Evaluation of the effect of information integration in displays for ICU nurses on situation awareness and task completion time: A prospective randomized controlled study.,” *International journal of medical informatics*, vol. 82, no. 8, pp. 665–75, Aug. 2013.

[123] S. B. Wachter, J. Agutter, N. Syroid, F. Drews, M. B. Weinger, and D. Westenskow, “The Employment of an Iterative Design Process to Develop a Pulmonary Graphical Display,” *Journal of the American Medical Informatics Association* , vol. 10 , no. 4 , pp. 363–372, Jul. 2003.

[124] S. Anders, R. Albert, A. Miller, M. B. Weinger, A. K. Doig, M. Behrens, and J. Agutter, “Evaluation of an integrated graphical display to promote acute change detection in ICU patients.,” *International journal of medical informatics*, vol. 81, no. 12, pp. 842–51, Dec. 2012.

[125] S. B. Higgins, K. Jiang, B. B. Swindell, and G. R. Bernard, “A graphical ICU workstation.,” *Proceedings / the ... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care*, pp. 783–787, 1991.

[126] R. W. Albert, J. a Agutter, N. D. Syroid, K. B. Johnson, R. G. Loeb, and D. R. Westenskow, “A simulation-based evaluation of a graphic cardiovascular display.,” *Anesthesia and analgesia*, vol. 105, no. 5, pp. 1303–11, table of contents, Nov. 2007.

[127] M. Görges, K. Förger, and D. Westenskow, “A Trend Based Decision Support System For Anesthesiologists Improves Diagnosis Speed and Accuracy,” in *Proceedings of the Annual Mountain West Biomedical Engineering Conference*, 2006.

[128] G. T. Blike, S. D. Surgenor, and K. Whalen, "A graphical object display improves anesthesiologists' performance on a simulated diagnostic task.," *Journal of clinical monitoring and computing*, vol. 15, no. 1, pp. 37–44, Jan. 1999.

[129] M. Görges, D. R. Westenskow, and B. a Markewitz, "Evaluation of an integrated intensive care unit monitoring display by critical care fellow physicians.," *Journal of clinical monitoring and computing*, vol. 26, no. 6, pp. 429–36, Dec. 2012.

[130] S. H. Koch, C. Weir, M. Haar, N. Staggers, J. Agutter, M. Görges, and D. Westenskow, "Intensive care unit nurses' information needs and recommendations for integrated displays to improve nurses' situation awareness.," *Journal of the American Medical Informatics Association : JAMIA*, vol. 19, no. 4, pp. 583–90, 2012.

[131] C. M. Burns, "Putting It All Together: Improving Display Integration in Ecological Displays," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 42, no. 2, pp. 226–241, Jun. 2000.

[132] J. A. Effken, R. G. Loeb, Y. Kang, and Z.-C. Lin, "Clinical information displays to improve ICU outcomes.," *International journal of medical informatics*, vol. 77, no. 11, pp. 765–777, Nov. 2008.

[133] K. F. Schulz and D. A. Grimes, "Case-control studies: research in reverse," *The Lancet*, vol. 359, no. 9304, pp. 431–434, 2002.

[134] D. A. Grimes and K. F. Schulz, "Compared to what? Finding controls for case-control studies," *The Lancet*, vol. 365, no. 9468, pp. 1429–1433, 2005.

[135] M. Loorak, C. Perin, N. Kamal, M. Hill, and S. Carpendale, "TimeSpan: Using Visualization to Explore Temporal Multi-dimensional Data of Stroke Patients," 2015.

[136] C. Plaisant, R. Mushlin, A. Snyder, J. Li, D. Heller, and B. Shneiderman, "LifeLines: using visualization to enhance navigation and analysis of patient records.," *Proceedings / AMIA ... Annual Symposium. AMIA Symposium*, pp. 76–80, 1998.

[137] S. Malik, F. Du, M. Monroe, E. Onukwugha, C. Plaisant, and B. Shneiderman, "Cohort comparison of event sequences with balanced integration of visual analytics and statistics," in *Proceedings of the 20th International Conference on Intelligent User Interfaces*, 2015, pp. 38–49.

[138] D. Gotz and H. Stavropoulos, "Decisionflow: Visual analytics for high-dimensional temporal event sequence data," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 20, no. 12, pp. 1783–1792, 2014.

[139] D. Klimov, Y. Shahar, and M. Taieb-Maimon, "Intelligent visualization and exploration of time-oriented data of multiple patients.," *Artificial intelligence in medicine*, vol. 49, no. 1, pp. 11–31, May 2010.

[140] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman, "Temporal event sequence simplification," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 19, no. 12, pp. 2227–2236, 2013.

[141] M. Soegaard and R. F. Dam, "The Encyclopedia of Human-Computer Interaction," *The*

Encyclopedia of Human-Computer Interaction, 2012.

[142] P. Saraiya, C. North, and K. Duca, "An insight-based methodology for evaluating bioinformatics visualizations," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 11, no. 4, pp. 443–456, 2005.

[143] J. Rieman, "A field study of exploratory learning strategies," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 3, no. 3, pp. 189–218, 1996.

[144] C. Plaisant, "The challenge of information visualization evaluation," in *Proceedings of the working conference on Advanced visual interfaces*, 2004, pp. 109–116.

[145] C. Ware, R. Arsenault, M. Plumlee, and D. Wiley, "Visualizing the underwater behavior of humpback whales.," *IEEE computer graphics and applications*, vol. 26, no. 4, pp. 14–18, 2006.

[146] P. Isenberg, T. Zuk, C. Collins, and S. Carpendale, "Grounded evaluation of information visualizations," in *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, 2008, p. 6.

[147] J. W. Creswell, *Qualitative inquiry and research design: Choosing among five approaches*. Sage, 2012.

[148] M. Sedlmair, M. Meyer, and T. Munzner, "Design Study Methodology: Reflections from the Trenches and the Stacks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2431–2440, Dec. 2012.

[149] T. Munzner, *Visualization Analysis and Design*. CRC Press, 2014.

[150] C. Auerbach and L. B. Silverstein, *Qualitative data: An introduction to coding and analysis*. NYU press, 2003.

[151] J. Corbin and A. Strauss, *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications, 2014.

[152] M. Tory and T. Möller, "Evaluating visualizations: do expert reviews work?," *Computer Graphics and Applications, IEEE*, vol. 25, no. 5, pp. 8–11, 2005.

[153] B. Preim and C. P. Botha, *Visual Computing for Medicine: Theory, Algorithms, and Applications*. Newnes, 2013.

[154] Y. B. Shrinivasan and J. J. van Wijk, "Supporting the analytical reasoning process in information visualization," in *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*, 2008, p. 1237.

[155] D. Keim, M. Ankerst, and H. Kriegel, "Recursive pattern: A technique for visualizing very large amounts of data," *… of the 6th conference on Visualization' …*, pp. 279–286, 1995.

[156] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, "Visual analytics: Definition, process, and challenges," in *Information visualization*, Springer Berlin Heidelberg, 2008, pp. 154–175.

[157] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Visual Languages, 1996. Proceedings., IEEE Symposium on*, 1996, pp.

336–343.

[158] D. A. Keim, F. Mansmann, J. Schneidewind, and H. Ziegler, “Challenges in visual data analysis,” in *Information Visualization, 2006. IV 2006. Tenth International Conference on*, 2006, pp. 9–16.

[159] D. A. Keim, F. Mansmann, D. Oelke, and H. Ziegler, “Visual analytics: Combining automated discovery with interactive visualizations,” in *Discovery Science*, 2008, pp. 2–14.

[160] G.-D. Sun, Y.-C. Wu, R.-H. Liang, and S.-X. Liu, “A Survey of Visual Analytics Techniques and Applications: State-of-the-Art Research and Future Challenges,” *Journal of Computer Science and Technology*, vol. 28, no. 5, pp. 852–867, 2013.

[161] P. Chung Wong, S. J. Rose, G. Chin, D. a Frincke, R. May, C. Posse, A. Sanfilippo, and J. Thomas, “Walking the path: a new journey to explore and discover through visual analytics,” *Information Visualization*, vol. 5, no. 4, pp. 237–249, 2006.

[162] T. M. O’Brien, A. M. Ritz, B. J. Raphael, and D. H. Laidlaw, “Gremlin: an interactive visualization model for analyzing genomic rearrangements.,” *IEEE transactions on visualization and computer graphics*, vol. 16, no. 6, pp. 918–926, 2010.

[163] J. Seo and B. Shneiderman, “Knowledge discovery in high-dimensional data: case studies and a user survey for the rank-by-feature framework.,” *IEEE transactions on visualization and computer graphics*, vol. 12, no. 3, pp. 311–322, 2006.

[164] N. Elmqvist and J.-D. Fekete, “Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 16, no. 3, pp. 439–454, 2010.

[165] Z. Xie, M. O. Ward, and E. A. Rundensteiner, “Exploring large scale time-series data using nested timelines,” in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2013, vol. 8654, p. 5.

[166] R. Kosara, F. Bendix, and H. Hauser, “Parallel sets: interactive exploration and visual analysis of categorical data.,” *IEEE transactions on visualization and computer graphics*, vol. 12, no. 4, pp. 558–568, 2006.

[167] J. N. Weinstein, “Biochemistry. A postgenomic visual icon.,” *Science (New York, NY)*, vol. 319, no. 5871, pp. 1772–1773, 2008.

[168] A. Pryke, S. Mostaghim, and A. Nazemi, “Heatmap visualization of population based multi objective algorithms,” in *Evolutionary multi-criterion optimization*, 2007, pp. 361–375.

[169] N. Colaert, K. Helsens, L. Martens, J. Vandekerckhove, and K. Gevaert, “Improved visualization of protein consensus sequences by iceLogo,” *Nature methods*, vol. 6, no. 11, pp. 786–787, 2009.

[170] A. A. Bojko, “Informative or misleading? Heatmaps deconstructed,” in *Human-Computer Interaction. New Trends*, Springer, 2009, pp. 30–39.

[171] L. Wilkinson and M. Friendly, “The history of the cluster heat map,” *The American Statistician*, vol. 63, no. 2, 2009.

[172] R. Atterer and P. Lorenzi, “A heatmap-based visualization for navigation within large web

pages," in *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, 2008, pp. 407–410.

[173] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nat Commun*, vol. 3, p. 745, Mar. 2012.

[174] T. Toyoda, Y. Mochizuki, and A. Konagaya, "GSCope: a clipped fisheye viewer effective for highly complicated biomolecular network graphs," *Bioinformatics*, vol. 19, no. 3, pp. 437–438, 2003.

[175] M. Krstajić, M. Najm-Araghi, F. Mansmann, and D. A. Keim, "Story Tracker: Incremental visual text analytics of news story development," *Information Visualization*, vol. 12, no. 3–4, pp. 308–323, 2013.

[176] G. Sun, Y. Wu, S. Liu, T.-Q. Peng, J. J. H. Zhu, and R. Liang, "EvoRiver: Visual analysis of topic coopetition on social media," 2014.

[177] H. Xu, Z. Li, S. Guo, and K. Chen, "CloudVista: interactive and economical visual cluster analysis for big data in the cloud," *Proceedings of the VLDB Endowment*, vol. 5, no. 12, pp. 1886–1889, 2012.

[178] F. Fischer, F. Mansmann, and D. a. Keim, "Real-time visual analytics for event data streams," in *Proceedings of the 27th Annual ACM Symposium on Applied Computing - SAC '12*, 2012, p. 801.

[179] H. Hochheiser and B. Shneiderman, "Dynamic query tools for time series data sets: Timebox widgets for interactive exploration," *Information Visualization*, vol. 3, no. 1, pp. 1–18, 2004.

[180] J. Zhao, F. Chevalier, E. Pietriga, and R. Balakrishnan, "Exploratory analysis of time-series with chronolenses," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 17, no. 12, pp. 2422–2431, 2011.

[181] T. N. Dang, A. Anand, and L. Wilkinson, "TimeSeer: Scagnostics for high-dimensional time series.," *IEEE transactions on visualization and computer graphics*, vol. 19, no. 3, pp. 470–483, 2013.

[182] C. Shi, W. Cui, S. Liu, P. Xu, W. Chen, and H. Qu, "RankExplorer: Visualization of ranking changes in large time series data," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 18, no. 12, pp. 2669–2678, 2012.

[183] Y. Livnat and J. Agutter, "A visualization paradigm for network intrusion detection," in *IAW'05. Proceedings from the Sixth Annual IEEE SMC. IEEE*, 2005, no. June, pp. 17–19.

[184] F. Reichl and M. Treib, "Visualization of Big SPH Simulations via Compressed Octree Grids," in *2013 IEEE International Conference on BigData*, 2013.

[185] M. Krstajic, E. Bertini, and D. A. Keim, "Cloudlines: Compact display of event episodes in multiple time-series," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 17, no. 12, pp. 2432–2439, 2011.

[186] O. D. Lampe and H. Hauser, "Interactive visualization of streaming data with kernel density

estimation," in *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, 2011, pp. 171–178.

[187] M. Harrison, R. Koppel, and S. Bar-Lev, "Unintended consequences of information technologies in health care—an interactive sociotechnical analysis," *Journal of the American Medical …*, vol. 14, pp. 542–549, 2007.

[188] P. Isenberg and D. Fisher, "Collaborative Brushing and Linking for Co-located Visual Analytics of Document Collections," in *Computer Graphics Forum*, 2009, vol. 28, no. 3, pp. 1031–1038.

[189] J. Stasko, C. Görg, and R. Spence, "Jigsaw: supporting investigative analysis through interactive visualization," *Information Visualization*, vol. 7, no. 2, pp. 118–132, 2008.

[190] G. Klein, B. Moon, and R. Hoffman, "Making sense of sensemaking 1: Alternative perspectives," *Intelligent Systems, IEEE*, vol. 21, no. 4, pp. 70–73, 2006.

[191] C. Chen, J. Zhang, and M. Vogeley, "Making sense of the evolution of a scientific domain: a visual analytic study of the Sloan Digital Sky Survey research," *Scientometrics*, vol. 83, no. 3, pp. 669–688, 2010.

[192] K. Koffka, *Principles of Gestalt psychology*. Hartcourt, New York, 1935.

[193] K. Morton, M. Balazinska, D. Grossman, and J. Mackinlay, "Support the Data Enthusiast: Challenges for Next-Generation Data-Analysis Systems," *Proceedings of the VLDB Endowment*, vol. 7, no. 6, 2014.

[194] J. Alsakran, Y. Chen, Y. Zhao, J. Yang, and D. Luo, "STREAMIT: Dynamic visualization and interactive exploration of text streams," in *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, 2011, pp. 131–138.

[195] C. Rohrdantz, D. Oelke, M. Krstajic, and F. Fischer, *Real-time visualization of streaming text data: tasks and challenges*. Bibliothek der Universit{ä}t Konstanz, 2011.

[196] D. Keim, T. Schreck, M. Krstajic, and C. Rohrdantz, "Real-time Visual Analytics of Text Data Streams," *Computer*, vol. 99, no. 1, p. 1, 2013.

[197] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, *Mastering The Information Age-Solving Problems with Visual Analytics*. Florian Mansmann, 2010.

[198] E. Fioratou, R. Flin, R. Glavin, and R. Patey, "Beyond monitoring: distributed situation awareness in anaesthesia.," *British journal of anaesthesia*, vol. 105, no. 1, pp. 83–90, Jul. 2010.

[199] F. Mansmann, F. Fischer, and D. Keim, "Dynamic visual analytics—facing the real-time challenge," *Expanding the Frontiers of Visual Analytics and Visualization*, pp. 69–80, 2012.

[200] Z. Liu, B. Jiang, and J. Heer, "immens: Real-time visual querying of big data," *Computer Graphics Forum (Proc. EuroVis)*, vol. 32, no. 3, pp. 421–430, Jun. 2013.

[201] M. Krstajić, E. Bertini, F. Mansmann, and D. A. Keim, "Visual analysis of news streams with article threads," in *Proceedings of the First International Workshop on Novel Data Stream Pattern Mining Techniques*, 2010, pp. 39–46.

[202] M. Hao, D. A. Keim, U. Dayal, D. Oelke, and C. Tremblay, "Density displays for data stream

monitoring," in *Computer Graphics Forum*, 2008, vol. 27, no. 3, pp. 895–902.

[203] J. Lin, E. Keogh, and S. Lonardi, "Visualizing and discovering non-trivial patterns in large time series databases," *Information Visualization*, vol. 4, no. 2, pp. 61–82, Apr. 2005.

[204] N. Kumar, V. N. Lolla, E. J. Keogh, S. Lonardi, and C. (Ann) Ratanamahatana, "Time-series Bitmaps: a Practical Visualization Tool for Working with Large Time Series Databases.," in *SDM*, 2005, pp. 531–535.

[205] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, "A symbolic representation of time series, with implications for streaming algorithms," in *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, 2003, pp. 2–11.

[206] P. M. Sanderson, M. O. Watson, and W. J. Russell, "Advanced patient monitoring displays: tools for continuous informing.," *Anesthesia and analgesia*, vol. 101, no. 1, pp. 161–8, table of contents, Jul. 2005.

[207] F. A. Drews and D. R. Westenskow, "The right picture is worth a thousand numbers: data displays in anesthesia," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 48, no. 1, p. 59, 2006.

[208] J. Higgins and S. Green, Eds., *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration, 2011, 2011.

[209] S. H. Koch, C. Weir, D. Westenskow, M. Gondan, J. Agutter, M. Haar, D. Liu, M. Görges, and N. Staggers, "Evaluation of the effect of information integration in displays for ICU nurses on situation awareness and task completion time: A prospective randomized controlled study," *International journal of medical informatics*, vol. 82, no. 8, pp. 665–675, 2013.

[210] D. Margulies, D. McCallie Jr, A. Elkowitz, and R. Ribitzky, "An integrated hospital information system at Children's Hospital," in *Proceedings of the Annual Symposium on Computer Application in Medical Care*, 1990, p. 699.

[211] J. Agutter, F. Drews, N. Syroid, D. Westneskow, R. Albert, D. Strayer, J. Bermudez, and M. B. Weinger, "Evaluation of graphic cardiovascular display in a high-fidelity simulator.," *Anesthesia and analgesia*, vol. 97, no. 5, pp. 1403–1413, Nov. 2003.

[212] A. Miller, C. Scheinkestel, and C. Steele, "The effects of clinical information presentation on physicians' and nurses' decision-making in ICUs.," *Applied ergonomics*, vol. 40, no. 4, pp. 753–61, Jul. 2009.

[213] S. Anders, R. Albert, A. Miller, M. B. Weinger, A. K. Doig, M. Behrens, and J. Agutter, "Evaluation of an integrated graphical display to promote acute change detection in ICU patients.," *International journal of medical informatics*, vol. 81, no. 12, pp. 842–51, Dec. 2012.

[214] H. I. Litt and J. W. Loonsk, "Digital patient records and the medical desktop: an integrated physician workstation for medical informatics training.," *Proceedings / the ... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care*, pp. 555–559, 1992.

[215] D. G. Kilman and D. W. Forslund, "An international collaboratory based on virtual patient

records," *Communications of the ACM*, vol. 40, no. 8, pp. 110–117, 1997.

[216] J. H. Van Bemmel, A. M. Van Ginneken, B. Stam, and E. Van Mulligen, "Virtual electronic patient records for shared care.," *Studies in health technology and informatics*, vol. 52, p. suppl–37, 1997.

[217] H. J. Tange, "The paper-based patient record: Is it really so bad?," *Computer Methods and Programs in Biomedicine*, vol. 48, no. 1, pp. 127–131, 1995.

[218] D. W. Forslund, R. L. Phillips, D. G. Kilman, and J. L. Cook, "Experiences with a distributed virtual patient record system.," *Proceedings : a conference of the American Medical Informatics Association / ... AMIA Annual Fall Symposium. AMIA Fall Symposium*, pp. 483–487, 1996.

[219] E. J. Enison, R. Dayhoff, and R. D. Fletcher, "Graphical electrocardiogram waveforms as part of an integrated hospital system's patient record.," *Proceedings / the ... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care*, pp. 373–375, 1993.

[220] P. M. Kuzmak and R. E. Dayhoff, "The use of digital imaging and communications in medicine (DICOM) in the integration of imaging into the electronic patient record at the Department of Veterans Affairs.," *Journal of digital imaging*, vol. 13, no. 2 Suppl 1, pp. 133–137, 2000.

[221] M. C. Were, C. Shen, M. Bwana, N. Emenyonu, N. Musinguzi, F. Nkuyahaga, A. Kembabazi, and W. M. Tierney, "Creation and evaluation of EMR-based paper clinical summaries to support HIV-care in Uganda, Africa.," *International journal of medical informatics*, vol. 79, no. 2, pp. 90–96, 2010.

[222] Y. Xie, S. J. Redmond, J. Basilakis, and N. H. Lovell, "Effect of ECG quality measures on piecewise-linear trend detection for telehealth decision support systems.," *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 2010, pp. 2877–2880, 2010.

[223] P. R. Norris and B. M. Dawant, "Closing the loop in ICU decision support: physiologic event detection, alerts, and documentation.," *Journal of American Medical Informatics Association*, vol. 9, no. Nov-Dec suppl, pp. S102–S107, 2002.

[224] A. Miller, "Evaluating an information display for clinical decision making in the intensive care unit," pp. 576–580, 2003.

[225] D. Engelman, T. L. Higgins, R. Talati, and J. Grimsman, "Maintaining situational awareness in a cardiac intensive care unit.," *The Journal of thoracic and cardiovascular surgery*, vol. 147, no. 3, pp. 1105–1106, 2014.

[226] N. Stylianides, M. D. Dikaiakos, H. Gjermundrød, G. Panayi, and T. Kyprianou, "Intensive care window: real-time monitoring and analysis in the intensive care environment.," *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 15, no. 1, pp. 26–32, Jan. 2011.

[227] P. Michels, D. Gravenstein, and D. R. Westenskow, "An integrated graphic data display

improves detection and identification of critical events during anesthesia.," *Journal of clinical monitoring*, vol. 13, no. 4, pp. 249–259, Jul. 1997.

[228] U. Gather, M. Imhoff, and R. Fried, "Graphical models for multivariate time series from intensive care monitoring.," *Statistics in medicine*, vol. 21, no. 18, pp. 2685–2701, 2002.

[229] S. Charabati, D. Bracco, P. a Mathieu, and T. M. Hemmerling, "Comparison of four different display designs of a novel anaesthetic monitoring system, the 'integrated monitor of anaesthesia (IMA)'.," *British journal of anaesthesia*, vol. 103, no. 5, pp. 670–7, Nov. 2009.

[230] G. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information. 1956.," *Psychological review*, vol. 101, no. 2, pp. 343–52, Apr. 1956.

[231] M. J. Eppler and J. Mengis, "The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines," *The Information Society*, vol. 20, no. 5, pp. 325–344, Nov. 2004.

[232] D. Bawden and L. Robinson, "The dark side of information: overload, anxiety and other paradoxes and pathologies," *Journal of Information Science*, vol. 35, no. 2, pp. 180–191, Nov. 2008.

[233] C. Speier, J. S. Valacich, and I. Vessey, "The Influence of Task Interruption on Individual Decision Making: An Information Overload Perspective," *Decision Sciences*, vol. 30, no. 2, pp. 337–360, Mar. 1999.

[234] A. A. T. Bui, D. R. Aberle, and H. Kangarloo, "TimeLine: visualizing integrated patient records.," *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 11, no. 4, pp. 462–473, 2007.

[235] R. G. Duncan, D. Saperia, R. Dulbandzhyan, M. M. Shabot, J. X. Polaschek, and D. T. Jones, "Integrated web-based viewing and secure remote access to a clinical data repository and diverse clinical systems.," *Proceedings / AMIA ... Annual Symposium. AMIA Symposium*, pp. 149–153, 2001.

[236] G. H. Kruger and K. K. Tremper, "Advanced integrated real-time clinical displays.," *Anesthesiology clinics*, vol. 29, no. 3, pp. 487–504, 2011.

[237] M. Meyer, W. C. Levine, P. Brzezinski, J. Robbins, F. Lai, G. Spitz, and W. S. Sandberg, "Integration of hospital information systems, operative and peri-operative information systems, and operative equipment into a single information display.," *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, p. 1054, 2005.

[238] A. S. Law, Y. Freer, J. Hunter, R. H. Logie, N. McIntosh, and J. Quinn, "A comparison of graphical and textual presentations of time series data to support medical decision making in the neonatal intensive care unit.," *Journal of clinical monitoring and computing*, vol. 19, no. 3, pp. 183–194, 2005.

[239] A. Ahmed, S. Chandra, V. Herasevich, O. Gajic, and B. W. Pickering, "The effect of two different electronic health record user interfaces on intensive care provider task load, errors of cognition, and performance.," *Critical care medicine*, vol. 39, no. 7, pp. 1626–34, Jul. 2011.

[240] D. Georgopoulos, G. Prinianakis, and E. Kondili, “Bedside waveforms interpretation as a tool to identify patient-ventilator asynchronies.,” *Intensive care medicine*, vol. 32, no. 1, pp. 34–47, 2006.

[241] D. A. Sainsbury, “An object-oriented approach to data display and storage: 3 years experience, 25,000 cases.,” *International journal of clinical monitoring and computing*, vol. 10, no. 4, pp. 225–233, 1993.

[242] K. Zinser and F. Frischenschlager, “Multimedia’s push into power.,” *IEEE Spectrum*, pp. 44–48, 1994.

[243] G. T. Blike, S. D. Surgenor, K. Whalen, and J. Jensen, “Specific elements of a new hemodynamics display improves the performance of anesthesiologists.,” *Journal of clinical monitoring and computing*, vol. 16, no. 7, pp. 485–491, Jan. 2000.

[244] A. Jungk, B. Thull, A. Hoeft, and G. Rau, “Evaluation of two new ecological interface approaches for the anesthesia workplace.,” *Journal of clinical monitoring and computing*, vol. 16, no. 4, pp. 243–58, Jan. 2000.

[245] S. B. Wachter, B. Markewitz, R. Rose, and D. Westenskow, “Evaluation of a pulmonary graphical display in the medical intensive care unit: an observational study.,” *Journal of biomedical informatics*, vol. 38, no. 3, pp. 239–43, Jun. 2005.

[246] J. A. Effken, N. G. Kim, and R. E. Shaw, “Making the constraints visible: testing the ecological approach to interface design.,” *Ergonomics*, vol. 40, no. 1, pp. 1–27, Jan. 1997.

[247] R. R. Kennedy, a F. Merry, G. R. Warman, and C. S. Webster, “The influence of various graphical and numeric trend display formats on the detection of simulated changes.,” *Anaesthesia*, vol. 64, no. 11, pp. 1186–91, Nov. 2009.

[248] J. A. Effken, “Improving Clinical Decision Making Through Ecological Interfaces,” *Ecological Psychology*, vol. 18, no. 4, pp. 283–318, Oct. 2006.

[249] T. Ropinski, S. Oeltze, and B. Preim, “Survey of glyph-based visualization techniques for spatial multivariate medical data,” *Computers & Graphics*, vol. 35, no. 2, pp. 392–401, 2011.

[250] J. Zhao, N. Cao, Z. Wen, Y. Song, Y.-R. Lin, and C. Collins, “#FluxFlow: Visual Analysis of Anomalous Information Spreading on Social Media,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 20, no. 12, pp. 1773–1782, Dec. 2014.

[251] G. Smith, M. Czerwinski, B. Meyers, D. Robbins, G. Robertson, and D. S. Tan, “FacetMap: A scalable search and browse visualization.,” *IEEE transactions on visualization and computer graphics*, vol. 12, no. 5, pp. 797–804, 2006.

[252] W. G. Cole and J. G. Stewart, “Human performance evaluation of a metaphor graphic display for respiratory data.,” *Methods of information in medicine*, vol. 33, no. 4, pp. 390–6, Oct. 1994.

[253] M. Görges, K. Kück, S. H. Koch, J. Agutter, and D. R. Westenskow, “A far-view intensive care unit monitoring display enables faster triage.,” *Dimensions of critical care nursing : DCCN*, vol. 30, no. 4, pp. 206–17.

[254] J. M. Tappan, J. Daniels, B. Slavin, J. Lim, R. Brant, and J. M. Ansermino, “Visual cueing with context relevant information for reducing change blindness.,” *Journal of clinical monitoring and computing*, vol. 23, no. 4, pp. 223–232, 2009.

[255] K. van Amsterdam, F. Cnossen, A. Ballast, and M. M. R. F. Struys, “Visual metaphors on anaesthesia monitors do not improve anaesthetists’ performance in the operating theatre,” *British Journal of Anaesthesia* , Feb. 2013.

[256] W. Horn, “AI in medicine on its way from knowledge-intensive to data-intensive systems.,” *Artificial intelligence in medicine*, vol. 23, no. 1, pp. 5–12, Aug. 2001.

[257] A. Eden, M. Grach, Z. Goldik, I. Shnaider, H. Lazarovici, O. Barnett-Griness, A. Perel, and R. Pizov, “The implementation of an anesthesia information management system,” *European journal of anaesthesiology*, vol. 23, no. 10, pp. 882–889, 2006.

[258] S. Charbonnier, “On line extraction of temporal episodes from ICU high-frequency data: a visual support for signal interpretation.,” *Computer methods and programs in biomedicine*, vol. 78, no. 2, pp. 115–132, 2005.

[259] N. D. Syroid, J. Agutter, F. A. Drews, D. R. Westenskow, R. W. Albert, J. C. Bermudez, D. L. Strayer, H. Prenzel, R. G. Loeb, and M. B. Weinger, “Development and evaluation of a graphical anesthesia drug display.,” *Anesthesiology*, vol. 96, no. 3, pp. 565–575, 2002.

[260] F. A. Drews and D. R. Westenskow, “The right picture is worth a thousand numbers: data displays in anesthesia,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 48, no. 1, pp. 59–71, 2006.

[261] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” *Human mental workload*, vol. 1, no. 3, pp. 139–183, 1988.

[262] D. T. Bauer, S. Guerlain, and P. J. Brown, “The design and evaluation of a graphical display for laboratory data,” *Journal of the American Medical Informatics Association* , vol. 17 , no. 4 , pp. 416–424, Jul. 2010.

[263] J. A. Jacko, *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*. CRC press, 2012.

[264] B. G. Glaser, *Emergence vs forcing: Basics of grounded theory analysis*. Sociology Press, 1992.

[265] L. Knigge and M. Cope, “Grounded visualization: integrating the analysis of qualitative and quantitative data through grounded theory and visualization,” *Environment and Planning A*, vol. 38, no. 11, pp. 2021–2037, 2006.

[266] C. Okoli and K. Schabram, “A guide to conducting a systematic literature review of information systems research,” *Sprouts Work. Pap. Inf. Syst*, vol. 10, p. 26, 2010.

[267] A. E. Blandford, “Semi-structured qualitative studies,” in *The Encyclopedia of Human-Computer Interaction*, 2nd ed., Aarhus, Denmark: Interaction Design Foundation, 2013.

[268] J. A. Smith, *Qualitative psychology: a practical guide to research methods*. Sage, 2015.

[269] S. Attfield and A. Blandford, “Making sense of digital footprints in team-based legal

investigations: The acquisition of focus," *Human--Computer Interaction*, vol. 26, no. 1–2, pp. 38–71, 2011.

[270] E. Tulving, "Elements of episodic memory," 1985.

[271] A. S. Law, Y. Freer, J. Hunter, R. H. Logie, N. McIntosh, and J. Quinn, "A comparison of graphical and textual presentations of time series data to support medical decision making in the neonatal intensive care unit.," *Journal of clinical monitoring and computing*, vol. 19, no. 3, pp. 183–94, Jun. 2005.

[272] B. Zarikoff, D. Martin, and M. Insley, "Lightweight, Low-cost and Flexible Flight Data Monitoring," in *AUTOTESTCON, 2014 IEEE*, 2014, pp. 251–259.

[273] A. Bar, A. Finamore, P. Casas, L. Golab, and M. Mellia, "Large-scale network traffic monitoring with DBStream, a system for rolling big data analysis," in *Big Data (Big Data), 2014 IEEE International Conference on*, 2014, pp. 165–170.

[274] A. Inselberg and B. Dimsdale, "Parallel coordinates," in *Human-Machine Interactive Systems*, Springer, 1991, pp. 199–233.

[275] R. Kamaleswaran and S. McIntyre, "A Real-time Visual Analytic Tool For Heart Rate Variability Analysis," 2013.

[276] R. Kamaleswaran, "Visual Analtic Tool for Exploring Neonatal Spells Classifications," 2014.

[277] M. Bostock, V. Ogievetsky, and J. Heer, "D^3 Data-Driven Documents," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 17, no. 12, pp. 2301–2309, 2011.

[278] C. McGregor, C. Catley, and A. James, "Variability analysis with analytics applied to physiological data streams from the neonatal intensive care unit," in *Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on*, 2012, pp. 1–5.

[279] J. L. Carlson, *Redis in Action*. Manning Publications Co., 2013.

[280] R. Kamaleswaran, J. E. Pugh, A. Thommandram, A. James, and C. Mcgregor, "Visualizing Neonatal Spells: Temporal Visualization of High Frequency Cardiorespiratory Physiological Event Streams," in *Proc. of IEEE VIS 2014 Workshop on Visualization of Electronic Health Records*, 2014.

[281] K. D. Fairchild, "Predictive monitoring for early detection of sepsis in neonatal ICU patients," *Current opinion in pediatrics*, vol. 25, no. 2, pp. 172–179, 2013.

[282] R. J. Martin, M. J. Miller, and W. A. Carlo, "Pathogenesis of apnea in preterm infants," *The Journal of pediatrics*, vol. 109, no. 5, pp. 733–741, 1986.

[283] M. P. Griffin and J. R. Moorman, "Toward the Early Diagnosis of Neonatal Sepsis and Sepsis-Like Illness Using Novel Heart Rate Analysis," *Pediatrics*, vol. 107, no. 1, pp. 97–104, 2001.

[284] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and others, "Scikit-learn: Machine learning in Python," *The Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[285] T. H. Yu, B. W. Fuller, J. H. Bannick, L. M. Rossey, and R. K. Cunningham, "Integrated environment management for information operations testbeds," in *VizSEC 2007*, 2008, pp.

67–83.

[286] M. P. Griffin, D. E. Lake, and J. R. Moorman, “Heart rate characteristics and laboratory tests in neonatal sepsis.,” *Pediatrics*, vol. 115, no. 4, pp. 937–41, 2005.

[287] M. P. Griffin, D. E. Lake, E. A. Bissonette, F. E. Harrell, M. O. Shea, J. R. Moorman, T. M. O. Shea, and A. O. Monitoring, “Heart Rate Characteristics : Novel Physiomarkers to Predict Neonatal Infection and Death,” *Pediatrics*, 2005.

[288] G. M. Draper, Y. Livnat, and R. F. Riesenfeld, “A survey of radial methods for information visualization,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 15, no. 5, pp. 759–776, 2009.

[289] Y. Albo, J. Lanir, P. Bak, and S. Rafaeli, “Off the Radar: Comparative Evaluation of Radial Visualization Solutions for Composite Indicators,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 22, no. 1, pp. 569–578, 2016.

[290] R. Feldman, “Filled radar charts should not be used to compare social indicators,” *Social indicators research*, vol. 111, no. 3, pp. 709–712, 2013.

[291] D. A. Keim, F. Mansmann, J. Schneidewind, and T. Schreck, “Monitoring network traffic with radial traffic analyzer,” in *Visual Analytics Science And Technology, 2006 IEEE Symposium On*, 2006, pp. 123–128.

[292] C. McGregor, E. Pugh, and A. Thommandram, “A Big Data Based Approach for Visualising Neonatal Apnoea and Spells,” 2015.

[293] B. J. McNeil, E. Keeler, and S. J. Adelstein, “Primer on certain elements of medical decision making,” *New England Journal of Medicine*, vol. 293, no. 5, pp. 211–215, 1975.

[294] P. K. Stein, “Challenges of Heart Rate Variability Research in the ICU*,” *Critical care medicine*, vol. 41, no. 2, pp. 666–667, 2013.

[295] S. A. Coggins, J.-H. Weitkamp, L. Grunwald, A. R. Stark, J. Reese, W. Walsh, and J. L. Wynn, “Heart rate characteristic index monitoring for bloodstream infection in an NICU: a 3-year experience,” *Archives of Disease in Childhood-Fetal and Neonatal Edition*, p. fetalneonatal–2015, 2015.

[296] T. D. Cook, D. T. Campbell, and A. Day, *Quasi-experimentation: Design & analysis issues for field settings*, vol. 351. Houghton Mifflin Boston, 1979.

[297] R. R. Kennedy and A. F. Merry, “The effect of a graphical interpretation of a statistic trend indicator (Trigg’s Tracking Variable) on the detection of simulated changes.,” *Anaesthesia and intensive care*, vol. 39, no. 5, pp. 881–886, 2011.

[298] Y. Liu, A. Osvalder, L. Tech, and A. Osvalder, “Usability evaluation of a GUI prototype for a ventilator machine,” *Journal of clinical monitoring and computing*, vol. 18, no. 5–6, pp. 365–372, 2005.

[299] L. Deneault and C. Lewis, “An integrative display for patient monitoring,” *Systems, Man and …*, 1990.

[300] K. Gurushanthaiah, M. Weinger, and C. Englund, “Visual display format affects the ability of anesthesiologists to detect acute physiologic changes. A laboratory study employing a

clinical display simulator.," *Anesthesiology*, vol. 83, no. 6, pp. 1184–1193, 1995.

[301] R. H. Ireland, H. V James, M. Howes, and A. J. Wilson, "Design of a summary screen for an ICU patient data management system.," *Medical & biological engineering & computing*, vol. 35, no. 4, pp. 397–401, 1997.

[302] R. Dayhoff, G. Kirin, S. Pollock, C. Miller, and S. Todd, "Medical data capture and display: the importance of clinicians' workstation design.," *Proceedings / the ... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care*, pp. 541–545, 1994.

[303] P. H. Langner, "The value of high fidelity electrocardiography using the cathode ray oscillograph and an expanded time scale," *Circulation*, vol. 5, no. 2, pp. 249–256, 1952.

[304] A. Burykin, T. Peck, V. Krejci, A. Vannucci, I. Kangrga, and T. G. Buchman, "Toward optimal display of physiologic status in critical care: I. Recreating bedside displays from archived physiologic data.," *Journal of critical care*, vol. 26, no. 1, pp. 105.e1–9, Feb. 2011.

[305] A. Lowe, R. W. Jones, and M. J. Harrison, "The graphical presentation of decision support information in an intelligent anaesthesia monitor," *Artificial Intelligence in medicine*, vol. 22, no. 2, pp. 173–191, 2001.

[306] Medwrench, "Phillips MP70 Medical Monitor," *2014*. [Online]. Available: http://photos.medwrench.com/equipmentPhotos/2000/2354-7114.jpg. [Accessed: 10-Sep-2014].

[307] J. Lin, E. Keogh, S. Lonardi, J. P. Lankford, and D. M. Nystrom, "VizTree: a tool for visually mining and monitoring massive time series databases," in *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*, 2004, pp. 1269–1272.

[308] W. G. Cole and J. G. Stewart, "Metaphor graphics to support integrated decision making with respiratory data.," *International journal of clinical monitoring and computing*, vol. 10, no. 2, pp. 91–100, May 1993.

Appendix 1: Comprehensive Matrix of Design Properties

Paper	Target Users	Year	Clinical Context	Number of Variables	Display Type	Colour Used	Pre-attentive processing	Gestalts	Interactive Controls	Iterative Design
[225]	Intensivists	2014	Yes	>20	WF	Red, Yellow, Orange, Blue, Green	Colour	Continuity	Selection	No
[229]	Anesthesia	2009	No	0-4	WF	Red, Yellow, Blue, Green, Black	Shape, Colour	Continuity	No	No
[211]	Anesthesia	2003	No	11-20	MT	Blue, Red	Colour, Shape, Dimension	Symmetry, Continuity	No	Yes
[124]	Intensivists	2012	Yes	>20	WF, MT	Red, Blue, White	Colour, Shape, Dimension	Symmetry, Continuity	Select, Filter, Overview	Yes

Paper	Target Users	Year	Clinical Context	Number of Variables	Display Type	Colour Used	Pre-attentive processing	Gestalts	Interactive Controls	Iterative Design
[245]	Intensivists	2004	No	5-10	MT	Green, Gray, Yellow, Blue, Pink, Black	Colour, Shape	Continuity, Similarity	No	Yes
[255]	Anesthesia	2013	No	5-10	MT	Green, Red, Yellow, Pink, Blue, White	Shape, Colour, Value, Symmetry, Dimension	Shape, Symmetry, Association, Continuity	No	No
[297]	Anesthesia	2011	No	0-4	OB	Not Specified	Colour, Dimension, Shape	Continuity, Proximity, Similarity	No	No
[298]	Intensivists	2005	No	5-10	MT	Green, Blue, Yellow	Colour, Value, Dimension, Shape	Symmetry, Continuity, Similarity	No	No

Paper	Target Users	Year	Clinical Context	Number of Variables	Display Type	Colour Used	Pre-attentive processing	Gestalts	Interactive Controls	Iterative Design
[128]	Anesthesia	1999	No	11-20	OB	Not Reported	Value, Dimension, Shape	Symmetry, Continuity, Similarity	No	No
[252]	Intensivists	1994	No	5-10	MT	Blue, Yellow, White, Green, Pink	Colour, Dimension, Shape	Symmetry, Continuity, Proximity	No	No
[299]	Anesthesia	1990	No	5-10	MT	Green, Red	Colour, Dimension, Shape	Symmetry, Continuity, Proximity	No	No
[244]	Anesthesia	2000	No	>20	OB, MT	Green, Yellow, Others	Colour, Dimension, Shape	Symmetry, Closure, Continuity, Proximity	No	Yes
[300]	Anesthesia	1995	No	5-10	MT, WF	Not Reported	Colour, Dimension, Shape	Symmetry, Closure, Continuity, Proximity	No	No

Paper	Target Users	Year	Clinical Context	Number of Variables	Display Type	Colour Used	Pre-attentive processing	Gestalts	Interactive Controls	Iterative Design
[301]	Intensivists	1997	Yes	>20	MT	Not Reported	Colour, Value, Dimension, Shape	Symmetry, Closure, Continuity, Proximity	No	Yes
[254]	Anesthesia	2009	No	5-10	MT	Grey, Blue	Colour, Value, Dimension, Size, Shape	Symmetry, Closure, Continuity, Proximity	No	No
[227]	Anesthesia	1997	No	>20	MT	Red, Blue, Green, Yellow	Colour, Value, Dimension, Size, Shape	Symmetry, Closure, Continuity, Proximity	No	No
[246]	Intensivists	1997	No	5-10	OB, MT	Green, Red	Colour, Value, Dimension, Size, Shape	Symmetry, Closure, Continuity, Proximity	No	No
[129]	Intensivists	2012	Yes	11-20	WF, MT	Red, Blue, Green, Beige	Colour, Value, Orientation, Dimension, Shape	Symmetry, Closure, Continuity, Proximity	No	Yes

Paper	Target Users	Year	Clinical Context	Number of Variables	Display Type	Colour Used	Pre-attentive processing	Gestalts	Interactive Controls	Iterative Design
[226]	Intensivists	2011	Yes	>20	WF	Red, Green	Colour, Dimension	Continuity	Selection, Filter, Overview	No
[214]	Intensivists	1992	Yes	>20	WF, MT	Grey	Dimension, Value, Shape	Proximity, Continuity, Symmetry	Selection	No
[218]	Multi	1996	Yes	>20	WF, MT	Not Reported	Value, Dimension, Shape, Size	Proximity, Continuity, Closure	Selection, Filter, Overview	No
[120]	Intensivists	2001	Yes	11-20	MT	Continuity, Similarity, Closure	Continuity, Similarity, Closure	Continuity, Similarity, Closure	No	Yes
[302]	Intensivists	1994	Yes	>20	WF	Not Reported	Colour, Dimensions, Size	Proximity, Continuity	Selection	No

Paper	Target Users	Year	Clinical Context	Number of Variables	Display Type	Colour Used	Pre-attentive processing	Gestalts	Interactive Controls	Iterative Design
[223]	Intensivists	2002	Yes	>20	WF, TB	Blue, Purple, Red, Teal, Green, Orange	Colour, Dimensions	Proximity, Continuity, Closure	Selection, Filter, Overview	No
[303]	Intensivists	1952	No	0-4	WF	Grey	Shape, Size, Dimension	Continuity, Closure	None	No
[304]	Intensivists	2011	Yes	0-4	WF	Green, Red, Yellow, Blue, Teal, White, Purple	Shape, Dimension	Continuity, Closure	None	No
[212]	Intensivists	2009	Yes	>20	WF, TB, OB	Green, Red, Yellow, Blue, Teal, White, Purple	Shape, Dimension	Continuity, Similarity, Closure	Selection, Overview, Filter	Yes

Paper	Target Users	Year	Clinical Context	Number of Variables	Display Type	Colour Used	Pre-attentive processing	Gestalts	Interactive Controls	Iterative Design
[236]	Anesthesia	2011	Yes	11-20	MT	Green, Red, Yellow, Orange	Shape, Dimension	Proximity, Similarity, Closure	Selection, Overview	No
[238]	Intensivists	2004	No	5-10	TB, WF	Not Reported	Shape, Dimension	Similarity, Closure	None	No
[239]	Intensivists	2011	Yes	>20	TB	Black	Dimension	Proximity	Selection, Overview	No
[241]	Anesthesia	1993	No	11-20	OB, WF	Grey	Value, Dimension, Shape, Size	Continuity, Closure	None	No
[76]	Anesthesia	2002	No	5-10	OB, MT	Red, Green	Value, Dimension, Shape, Size	Continuity, Proximity, Similarity	None	No
[247]	Anesthesia	2008	No	0-4	WF	Red, Green, Blue, Black	Dimension	Continuity	None	No

Paper	Target Users	Year	Clinical Context	Number of Variables	Display Type	Colour Used	Pre-attentive processing	Gestalts	Interactive Controls	Iterative Design
[305]	Anesthesia	2001	No	0-4	OB	Red, Blue, Teal, Yellow	Colour, Size, Value, Dimension	Continuity, Similarity	None	No
[258]	Intensivists	2004	No	0-4	TB, WF	Red, Blue, Green	Colour, Dimension	Continuity	None	No
[51]	Anesthesia	1986	No	>20	TB, MT	Grey	Dimension, Size, Value	Continuity, Proximity, Symmetry	None	No
[257]	Anesthesia	2006	Yes	>20	WF, OB, TB	Not Reported	Dimension, Size, Value	Continuity	Selection, Filter	No
[209]	Nurses	2013	Yes	>20	WF, TB, MF	Red, Green, Yellow, Blue, Black	Value, Dimension, Shape, Size	Continuity, Symmetry	Selection	Yes

Acronyms: MT: Metaphoric Display, OB: Object-based display, TB: Tabular display, and WF is waveform display. Multi: Multidisciplinary

Appendix 2: Comprehensive Matrix of Study Results

Paper	Setting	Study Type	Results Reported	Realism	Cognitive Workload	Historic Trends	Visual encoding for Temporal trajectory	Visual encoding for duration	Visual encoding for frequency	Counter-balanced	Clinical Scenario	Case Controlled Functions
[225]	ICU	Eval.	+	Live	NI	C	C	NI	NI	No	Yes	No
[229]	Lab	Exp.	+	Static	-	NI	C	NI	NI	Display	Yes	No
[211]	Lab	Exp.	+	Sim.	-	NI	NI	NI	NI	Display	Yes	No
[124]	ICU	Exp.	Mix	Static	0	C	C	NI	NI	Yes	Yes	No
[245]	ICU	Eval.	+	Live	NI	NI	G	NI	NI	NI	No	No
[255]	Lab	Exp.	-	Static	NI	NI	O	NI	NI	Display	Yes	No
[297]	Lab	Exp.	+	Sim.	NI	NI	O	NI	NI	Display	No	No
[298]	Lab	Exp.	+	Sim.	NI	NI	NI	NI	NI	Yes	Yes	No
[128]	Lab	Exp.	+	Sim.	NI	NI	NI	NI	NI	Scenario	Yes	No

Paper	Setting	Study Type	Results Reported	Realism	Cognitive Workload	Historic Trends Temporal trajectory	Visual encoding for duration	Visual encoding for frequency	Visual encoding for	Counter-balanced	Clinical Scenario Functions	Case Controlled
[252]	Lab	Exp.	+	Static	NI	NI	G	G	G	Yes	Yes	No
[299]	Lab	Exp.	+	Sim.	NI	NI	NI	NI	NI	Yes	Yes	No
[244]	Lab	Exp.	+	Sim.	NI	NI	C	NI	NI	Scenario	Yes	No
[300]	Lab	Exp.	+	Sim.	NI	NI	C	NI	NI	Scenario	Yes	No
[301]	Lab	Eval.	+	Static	NI	NI	C, G	NI	NI	NI	No	No
[254]	Lab	Exp.	+	Sim.	-	NI	C, G	NI	NI	Yes	Yes	No
[227]	Lab	Exp.	+	Sim.	NI	NI	NI	NI	NI	Yes	Yes	No
[246]	Lab	Exp.	+	Sim.	NI	NI	NI	NI	NI	Scenario	Yes	No
[129]	ICU	Exp.	+	Sim.	0	NI	G	G	G	Scenario	Yes	No
[226]	ICU	Eval.	+	Live	NI	C	C	NI	NI	NI	No	No

Paper	Setting	Study Type	Results Reported	Realism	Cognitive Workload	Historic Trends	Temporal trajectory	Visual encoding for duration	Visual encoding for frequency	Visual encoding for Functions	Counter-balanced	Clinical Scenario Case Controlled
[214]	Lab	App.	NA	Static	NI	C	C	G	G	NI	No	No
[218]	Hospital	App.	NA	Live	NI	C	C, G	G	G	NI	No	No
[120]	ICU	Eval.	+	Static	NI	C	G	G	NI	NI	No	No
[302]	ICU	App.	NA	Live	NI	NI	C	NI	NI	NI	No	No
[223]	ICU	App.	+	Live	NI	C	C	NI	NI	NI	No	No
[303]	ICU	Eval.	NA	Static	NI	NI	C	NI	NI	NI	No	No
[304]	ICU	App.	NA	Sim.	NI	NI	C	NI	NI	NI	No	No
[212]	ICU	Exp.	+	Static	NI	NI	C	T	T	Yes	Yes	No
[236]	Surgery	App.	NA	Live	NI	NI	G	T	T	NI	No	No

Paper	Setting	Study Type	Results Reported	Realism	Cognitive Workload	Historic Trends	Temporal trajectory	Visual encoding for duration	Visual encoding for frequency	Visual encoding for Counter-balanced	Clinical Scenario Functions	Case Controlled
[238]	Lab	Exp.	–	Static	NI	NI	C, T	T	T	Yes	Yes	No
[239]	Lab	Exp.	+	Sim.	-	NI	T	T	T	Yes	Yes	No
[241]	Surgery	Eval.	+	Live	NI	NI	C	NI	NI	NI	No	No
[76]	Lab	Exp.	Mix	Sim.	Mix	NI	C, G	G	G	Scenario	Yes	No
[247]	Lab	Exp.	+	Sim.	NI	NI	C	NI	NI	Yes	No	No
[305]	Lab	App.	+	Sim.	NI	NI	C	NI	NI	NI	No	No
[258]	Lab	Design	NA	Sim.	NI	NI	C	NI	NI	NI	No	No
[51]	Lab	Design	NA	Sim.	NI	C	C	NI	NI	NI	No	No

Paper	Setting	Study Type	Results Reported	Realism	Cognitive Workload	Historic Trends Temporal trajectory	Visual encoding for duration	Visual encoding for frequency	Visual encoding for	Counter-balanced	Clinical Scenario Functions	Case Controlled
[257]	Surgery	App.	+	Live	-	C	C	NI	NI	NI	No	No
[209]	ICU	Exp.	+	Sim.	-	C	C	NI	NI	Yes	Yes	No

Acronyms: +: Positive; -: Negative, 0: No Change; Mix: Mixed results; Exp: Experiment; Eval: Evaluation; App: Application; Sim: Simulated; T: Text, O: Object, C: Curves, G: Glyph, NI: None Included;

Appendix 3: Visual Interfaces used in the think-aloud sessions

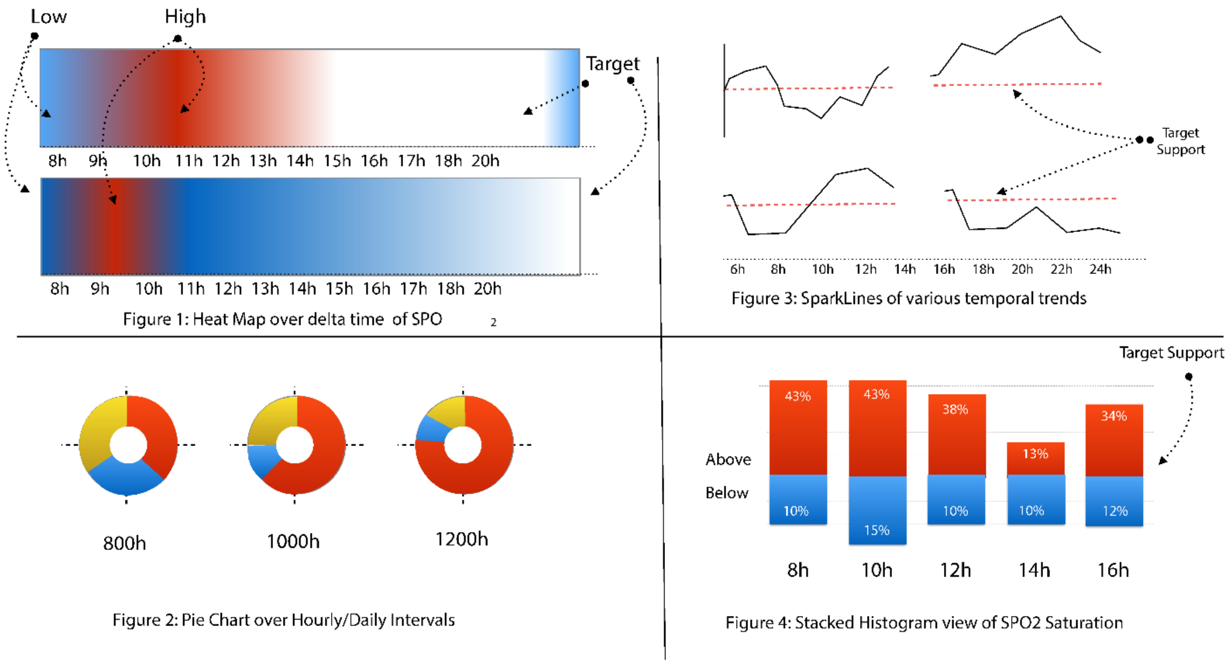


Figure A3-1 Top-left figure presents a heatmap display of oxygen saturations over 24 hours. The bottom-left figure presents hourly distribution of oxygen values that were above target (red), below target (blue) and within target (yellow). The top-right figure presents a Sparkline representation of oxygen values. The bottom-right figure illustrates above and below values of oxygen saturation over several hours.

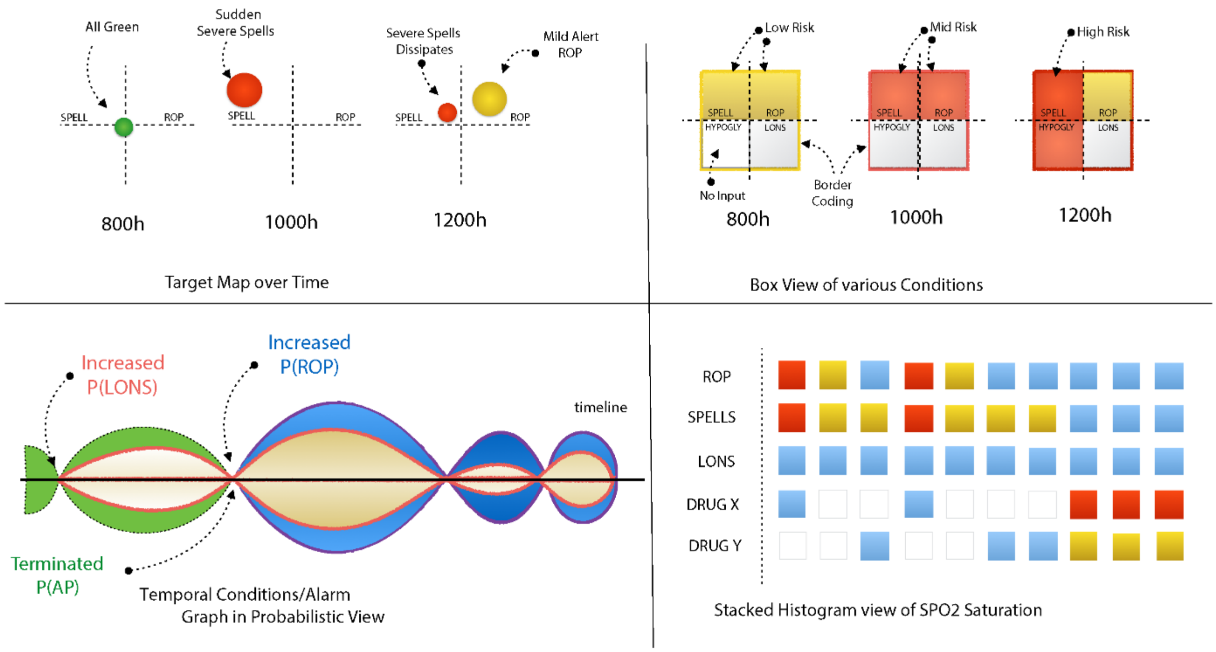


Figure A3-2: The top-left figure illustrates a novel DxRadar representation, where conditions that manifested in each of the hours displayed. Circles, which act as metaphors for physiological events, appear from the middle when they are detected at the bed-side and migrate outwards towards the edges as time progresses. Circles can appear in each of the four quadrants, or monitoring windows. A circle appearing in the Spells window for instance signifies an algorithm detected spell. The circles progressively get larger if the severity during that particular event epoch increases. This view allows the user to perceive conditions as they manifest at the bed-side. The bottom-left figure illustrates a theme-river like metaphor, however, in this graphics, physiologic conditions are constantly re-assessed and a probability score is presented. The top-right figure is an indicator display of four common conditions in the neonatal intensive care unit. The conditions are spells, retinopathy of prematurity, hypoglycaemia, and late onset nosocomial infection (sepsis). Colours change from white (no change), to yellow (mild risk), and to a dark shade of red (high risk), as an algorithm generates risks based on the analysis of physiologic data. The bottom-right display illustrates a stacked histogram view of a patient, who has several conditions that are marked for monitoring. A white colour represents no data, blue represents no change, yellow represents mild risk, and red represents high risk for that condition.

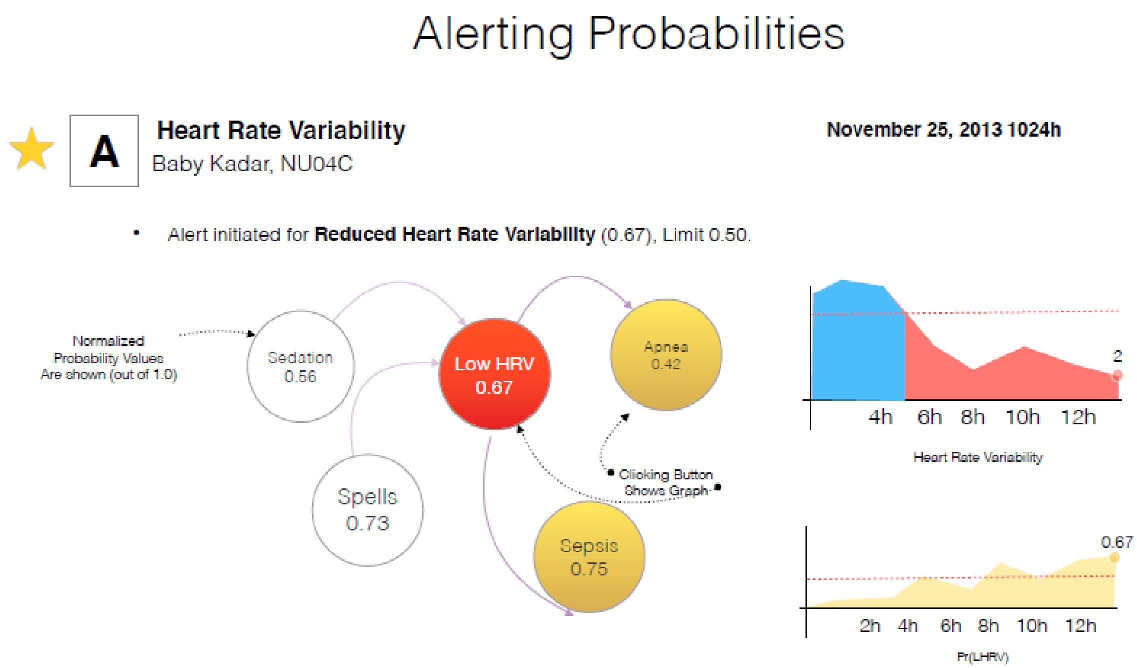


Figure A3-3: This figure illustrates a graph-based visualization of physiologic conditions, along with an area chart (right). Each circle represents a physiologic condition, the colour red signifies a high risk event, while yellow signifies a mild risk and white denotes low risk. In this figure, a high risk physiologic condition 'Low HRV' is shown using a bright red circle. Clicking on the circle buttons reveals their history in form of an area chart (right).

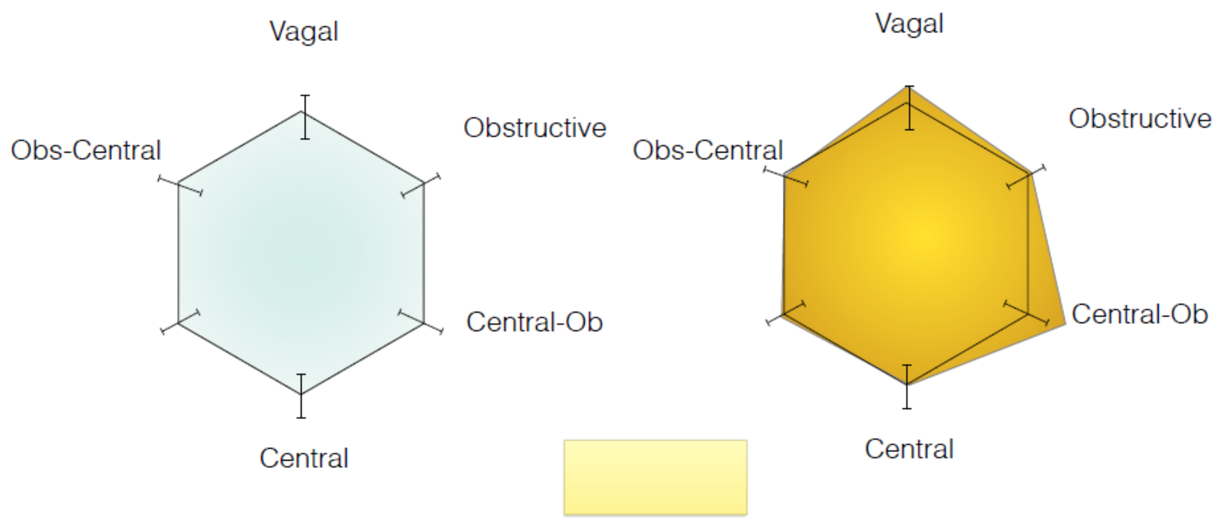


Figure A3-4: This figure illustrates a star-plot representation of five potential neonatal spells. A template star-plot diagram is shown in the left, and begins with an Obs-Central (obstructive then central); Vagal; Obstructive; Central-Ob (central then obstructive); and finally a central spell. When data is populated a yellow highlight appears over each edge, when the highlight protrudes away from the edge, it can be used to convey an increased activity of that particular spell. In this figure for instance, it can be observed that the patient being illustrated had an increased number of vagal and central-obstructive spells.

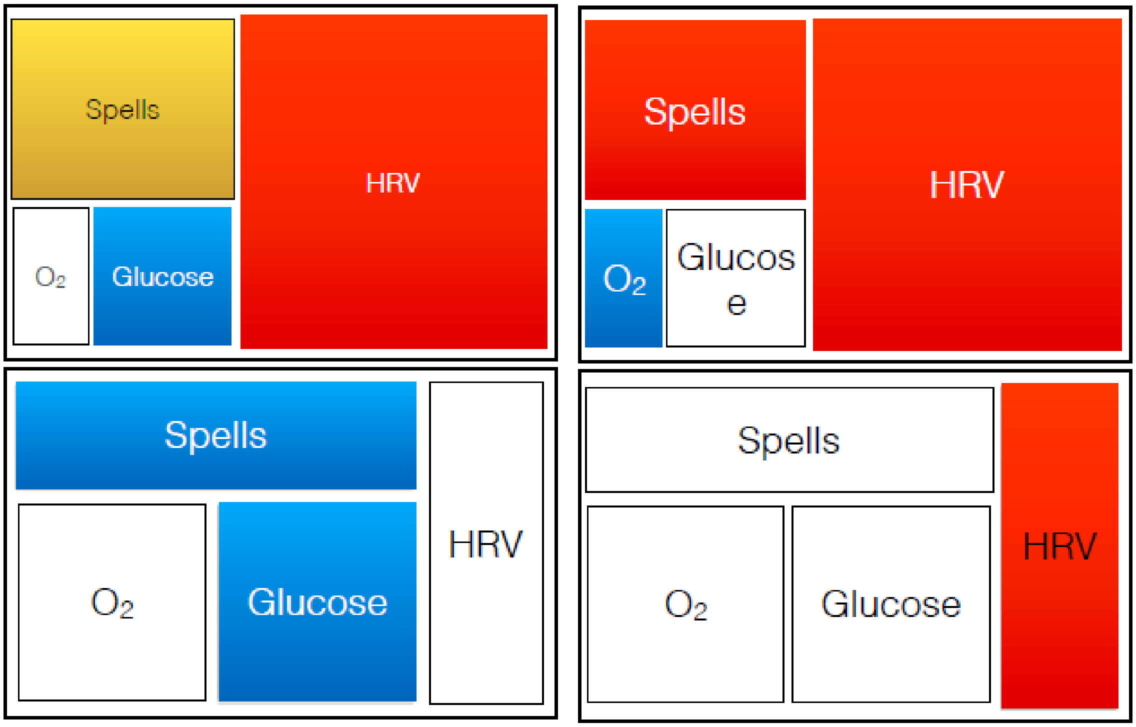


Figure A3-5: This figure illustrates a treemap diagram of various physiologic conditions for a single patient over several hours. Each larger box represents an hour of monitoring. The box represented at the top left shows that during that hour, there was an increased amount of spells, and HRV alerts. Meanwhile in the most recent hour (bottom-right), only HRV is observed, while other conditions are not in alarm state.

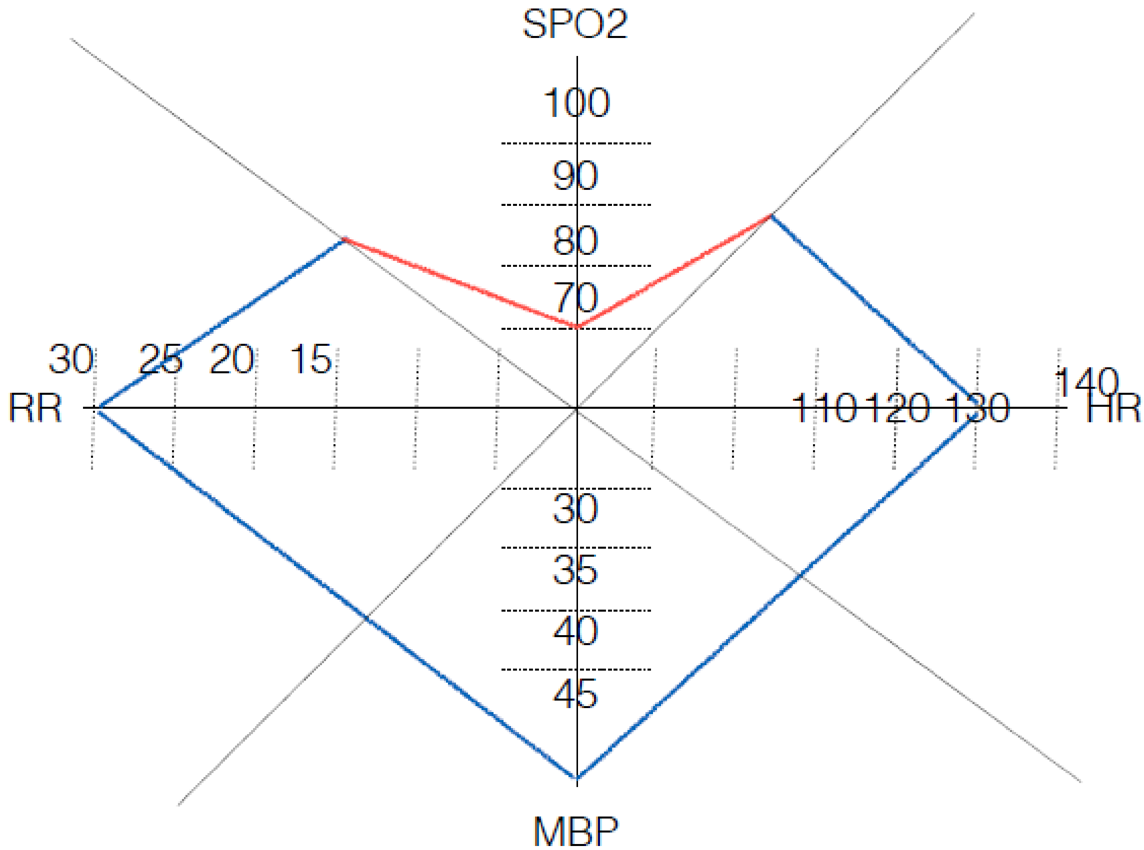


Figure A3-6: This figure illustrates a vector representation of raw physiologic data. Each converging edge of the vector falls along one of four potential axis. Colour is used as an indicator of severity. Blue represents normal ranges, yellow a low risk, and red a high risk. In this figure, SpO_2 reading is shown to be at 70%, which signifies a low oxygen saturation, and hence coloured red.

Appendix 4: Informal Affinity Diagrams for Grouping Themes

Table A4-1: Affinity diagram categories from the initial phase

Trajectory	Frequency	Duration	Consumption	Exploration
Makes Future Projections	Gets Number of Episodes	Asks Time-based Queries	Knowledge Dissemination	Knowledge Gathering
Information Retrieval	Information Retrieval	Information Retrieval	Information Dissemination	Data Gathering
			Information Retrieval	Data Cognition

Table A4-2: **Affinity diagram categories from the second phase**

Trajectory	Frequency	Duration	Consumption	Exploration
Questions about severity	Gets Number of Episodes	Asks Time-based Queries	Knowledge Dissemination	Knowledge Gathering
Makes Future Projections	Data Cognition	Data Cognition	Information Dissemination	Data Gathering
Information Dissemination	Data Gathering	Data Gathering	Information Retrieval	Data Cognition
Information Gathering				

Table A4-3: Affinity diagram categories from the third phase

Reactions to existing systems	Reactions to novel representations	Engagement levels	Perception of usefulness	Analysis of cohorts
Anecdotes on existing systems	Queries about functionality	Directed actions	Positive emotions on utility	Queries about multiple patients
Negative emotions of CIMs	Positive emotions about metaphors	Undirected actions	Things learnt from the representations	Queries about segmenting patients into groups
Limitations of existing systems	Queries about integration with existing systems		Comments on individual views	