Quantitative Analysis of Vital Signs of Premature Infants

A thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Science degree in Physics from the College of William and Mary

by

Sarah Maddox

Advisor: John Delos

Senior Research Coordinator: Henry Krakauer

Date: April 28, 2016

Contents

1 Introduction **1**

2 Literature Review **2**

2.1 Background 2

2.2 NICU Measurements 2

2.3 Apnea of Prematurity 3

2.4 Periodic Breathing 5

2.5 Desaturations and Hypersaturations 5

2.6 Retinopathy of Prematurity 5

3 Data and Methods **7**

3.1 NICU Signal Data 7

3.2 Apnea and Periodic Breathing Database 8

3.3 ROP Database 8

4 Results and Discussion **9**

4.1 Query User Data Interface for Apneas and Periodic Breathing Events 9

4.2 SpO_2 Statistics 12

4.3 ROP Analysis 15

5 Conclusion **20**

Abstract

Due to their underdeveloped systems, low birth weight, and immaturity of control, premature infants (born before 37 weeks gestation) experience high risk of pathology and an increased mortality rate. Critical medical conditions are often predicated by aberration of vital signs, such as heart rate and respiration, and thus monitoring of these signals is essential to decreasing morbidity associated with prematurity. Two tools were developed to better organize and assess the multi-terabyte data sets available from monitor signals of patients in the University of Virginia Medical Center Neonatal Intensive Care Unit (NICU). The first of these projects consisted of development of a Query User Data Interface (QUDI). This type of graphical user interface allows both researchers and clinicians alike to generate tables of recorded apnea events or epochs of periodic breathing according to specified characteristics, such as patient age or duration of the event. An analysis of SpO_2 signal data was used to generate information on desaturation and hypersaturation events. A second tool, a database categorized by bed and day in William and Mary's Sciclone Computing Complex, organized these events with associated characteristics, such as duration or %-time area of a desaturation below a given threshold (or likewise, of a hypersaturation above a given threshold). When associated with patients, this database can easily be used in future research to correlate aberrant oxygenation with pathology or mortality. Lastly, the hypothesis that aberrant oxygenation correlates with retinopathy of prematurity (ROP) was tested. Using a Kolmogorov-Smirnov Statistical Test, it was found that SpO_2 was significantly higher for infants that did not develop ROP in weeks 34 to 39 postmenstrual age at p=.05 significance level. However, after controlling for gestational age, as extremely preterm infants are more likely to develop ROP, no significant difference between the outcomes could be found. Thus SpO_2 signals, controlling for gestational age, were not able to make a significant prediction of ROP.

1 Introduction

Every year, about ten percent of infants are born prematurely, or before 37 weeks gestation [6]. These preterm infants may be very low birth weight (VLBW; less than 1.5 kilograms) or extremely low birth weight (ELBW; less than 1 kilogram), and are susceptible to increased risk of several clinical conditions due to their underdeveloped systems. Because of this, those of extremely and very low birth weight have high mortality and morbidity rates. In the United States, over 55,000 VLBW infants are born annually, and many of these infants experience critically adverse outcomes and conditions [10]. Though the problems associated with infant prematurity are well known, much remains unknown about optimal conditions for those born preterm. Eradicating the negative outcomes premature infants experience requires not only medical researchers and physicians, but also quantitative researchers who can analyze past data and make sense of it for the future.

Our group at William and Mary, headed by Dr. John Delos, works closely with a 45-bed University of Virginia Neonatal Intensive Care Unit (NICU), a quaternary care facility that has more than 500 admissions per year. At William and Mary, we utilize the data from this NICU, a quarter of whose admittance are VLBW, in order to quantitatively develop algorithms for better monitoring of vital signs. In particular, the goal is predictive monitoring: in order words, using current and recent data from the NICU patient with the quantitative algorithms of our lab to predict the future state of the infant. This collaboration gives our team of physicists the ability to assist physicians' on a fundamental issue, the high mortality of preterm infants, given that a medical doctor's time is extremely valuable and requires complete dedication to their patients. The goal for our medical physics research group at William and Mary is to better understand how different factors affect the health of premature infants in NICUs, and in doing so, we may better predict critical clinical events before they occur. Understanding the vital sign response to clinical events allows us to develop more beneficial monitoring techniques, prevention strategies, and therapies for medically suffering preterm infants. Thus, analysis and predictive monitoring of clinically significant events can help to save and improve the lives of preterm infants.

2 Literature Review

2.1 Background

As mentioned above, our knowledge of the optimal environment for NICU patients is still limited. That being said, the commonality of medical distress for those born preterm ensures, unfortunately, that we have a great amount of data on the topic. Premature infants may experience critical clinical events including sepsis, periodic breathing, very long apneas, retinopathy of prematurity, and necrotizing enterocolitis (a bacterial disease that destroys the intestine of the infant). Lack of respiratory and heart rate control, and insufficient development of the immune system, visual system, and other systems contributes greatly to the adverse outcomes preterm infants experience, which can increase health care costs in and beyond the NICU [12]. In order to record these clinical events, it is necessary to analyze vital sign characteristics and provide feedback to the medical practitioner. Vital sign characteristics used by our research group include electrocardiogram signals and chest impedance to determine heart rate and respiration rate, respectively, and peripheral capillary oxygen saturation, or SpO_2, an estimate of the amount of oxygen in the blood. In earlier work of the UVA and William and Mary groups, heart rate characteristics were used to develop a predictive monitoring system known as the HeRO System. My research has focused on the respiration rate, which can determine when the baby is experiencing an apnea, as well as the SpO_2 measurement, used to recognize critical events in the infant's oxygen level.

2.2 NICU Measurements

Chest impedance is the easiest way to monitor respiration. Two electrodes are placed across the chest, which measure a combination of basic impedance (giving measurements in hundreds of ohms, and consisting of impedance from muscles, tissue, blood, wires, and electrode-skin transitions), heart activity (contributing about .5 ohms) and respiration, which contributes about 2 ohms. Air has poor conductivity, so as the lungs inspire and expire air, the chest impedance signal rises and falls, and breathing can be monitored. A cardiac artifact is due to blood pumping out of the thorax with each heartbeat. Blood is more conductive than air, so this rhythmic action also causes the impedance across the chest to rise and fall. Luckily, this artifact of about .5 ohms can be filtered out, together with movement artifacts [11]. After removing these disrupting signals, the chest impedance gives a cleaner signal, which is better for discerning critical medical events related to breathing.

The SpO_2 level is a measurement of the amount of oxygen in the blood, defined as the percentage of oxygenated hemoglobin (oxyhemoglobin) compared to the total amount of hemoglobin in the blood (both oxygenated and deoxygenated). Hemoglobin is the protein molecule in red blood cells that carries oxygen from the lungs to cells throughout the body's tissue, and returns waste carbon dioxide from these tissues back to the lungs to be expired. To measure SpO_2, pulse oximetry is used as a noninvasive measure of blood oxygen levels. A clip-like device is placed on a body part, most commonly a finger or ear lobe, and the probe uses two or more wavelengths of light to determine the oxyhemoglobin saturation at this peripheral location from the heart [7].

When a finger is placed between the light source and the light detector within the pulse oximeter, part of the light will be absorbed so that only a fraction of the light passes through to the detector. The amount of light absorbed depends on the oxyhemoglobin and deoxyhemoglobin in the finger. Hemoglobin absorbs light proportional to the concentration in the blood vessel and the thickness of the finger through which the light has to travel. In order to determine the amount of oxyhemoglobin compared to deoxyhemoglobin, thereby giving the SpO_2, the oximeter takes advantage of the absorption spectrum of the two proteins. Oxyhemoglobin absorbs more infrared light (about 950 nm) than red light (about 650 nm), while deoxyhemoglobin absorbs more red light than infrared. The pulse oximeter is able to compute the oxygen saturation by comparing how much the blood absorbs this red light and infrared light [9]. Though the optimal level of oxygenation is still unknown, it is theorized that "a major factor contributing to death and long-term disability in these infants is aberrant oxygenation" [10]. Thus monitoring both of these vital signs gives us clues as to the optimal levels of oxygen needed, and notifies the physicians of any abnormal breathing that may anticipate medical distress.

2.3 Apnea of Prematurity

Apnea of prematurity (AOP) is a very common problem for premature infants, and its criticality in predicting distressing and even fatal clinical events indicates that more attention for AOP is needed. Apnea is clinically defined as either I. cessation of breathing for longer than 20 seconds, or II. cessation of breathing for longer than 10 seconds, together with bradycardia, where the heart rate dips below 100 beats per minute, and oxygen desaturation, where the SpO_2 levels go below 80%. In II, these three events together are called an ABD event [11]. AOP is a result of immaturity of the brainstem and peripheral chemoreceptors in underdeveloped preterm infants, and may lead

to abnormal development, physical and cognitive impairment, and sometimes, death [12]. These apneas may also delay discharge from the hospital, as well as indicate a warning of many other illnesses. Chronic or acute pathologies may also impact the severity and duration of apneas in premature infants. Medical distress of the central nervous system, abnormal oxygenation, and acute inflammatory processes such as infection are all associated with dysregulation of chemoreceptor function and thus can increase the duration or severity of apnea [8]. In NICUs, over half of all VLBW infants and almost all ELBW infants experience apnea [11]. The number of occurrences and duration of AOP events decreases with gestational age (GA), or duration of the pregnancy, and postmenstrual age (PMA), or the number of weeks since conception, and thus extremely preterm infants are most likely to develop AOP.

VLBW and ELBW babies have a high risk of neonatal sepsis, or infection, which can be caused by a bacterial infection of the bloodstream, leading to an immune response as the body tries to rid itself of the foreign toxins [13]. Out of approximately 4 million births per year, 56,000 are VLBW, and their risk of sepsis is around 25-40%. The diagnosis of sepsis is difficult and furthermore there is a high rate of false negatives. Because of this, bedside physicians administer antibiotics early and often. However, there is evidence that apneas can predict sepsis, and thus predictive monitoring of vital signs of preterm infants may give an early warning of sepsis. Connections between apnea, decelerations of the heart rate, and sepsis have been found. Therefore, accurate detection of neonatal apnea is imperative in order to distinguish which apneas are clinically significant, and may predict other clinical events.

One of the current questions our lab is investigating is whether we can further develop earning warning signs of clinically significant apneas. Currently, nurses record only about one third of apneas. Alarms are set to go off in the NICU if the infant is experiencing apnea, bradycardia, desaturation, or another failure of a vital sign. However, the apnea alarm fails in about 14% of events [11]. Even more significantly, apnea alarms have been found to activate in only about two thirds of the events for very long apneas, where the duration is longer than 60 seconds, and on average of 28 seconds after cessation of breathing [12]. Bradycardia alarms activated on average of over a minute late [12]. It also appears that NICU bedside physicians and nurses responded inconsistently to apnea alarms. Most events are self-resolved or terminate when the nurse stimulates the infant in response to an alarm, so reliable apnea detection is necessary for improved patient outcomes. Thus, to develop early warning signs of sepsis and increase efficiency and accuracy of predictive monitoring, we must

advance our techniques for consolidating and analyzing the AOP data received from the UVA NICUs.

2.4 Periodic Breathing

Another aberrance of respiratory rate is termed periodic breathing, or periodic apneas. Similar to the apnea events described above, periodic breathing is recurring apnea events at a regular interval; in other words, an abnormal repetitive pattern of fluctuation in the chest impedance signal. Periodic breathing is an artifact of the delayed respiratory control system. Because the control of the system has a time delay, the preterm infant may over-correct slowing or quickening breaths, and thus go into oscillation [12]. Periodic breathing appears to be correlated with clinically significant events as well, and in some cases are correlated with periodic decelerations of the heart.

2.5 Desaturations and Hypersaturations

My research has also focused on a third category of clinical events: desaturations. As mentioned, desaturations may occur as a part of an apnea event (an AD or ABD event), but they also can be present on their own. A desaturation is clinically defined as an event in which the SpO_2 level dips below 80%, and my research also considers events in which the SpO_2 drops below some defined threshold, including 75% and 85%. A hypersaturation may also cause or correlate with pathology in preterm infants and thus the oxygen level can be analyzed for these events as well, including events where the SpO_2 level goes above the thresholds of 95%, 96%, 97%, and so on. The optimal oxygenation level that balances the risk of death from low oxygen and the risk of other pathology from high oxygen is still unknown, yet a good range is suspected to be about 90-95% SpO_2. According to Zupancic, "aberrant oxygenation can lead to multi-organ damage, with hypoxia linked to brain injury" [15]. Thus our ability to recognize changes in SpO_2 and correlate these abnormalities with clinical events is critical for improving the lives of preterm infants.

2.6 Retinopathy of Prematurity

One such pathology that is correlated with aberrance in SpO_2 levels is known as retinopathy of prematurity, or ROP, and was first discovered to be a critical problem for premature infants in the 1940's. Though understanding of this disease is more complete than it was 70 years ago, and monitored oxygen use and improved conditions in NICUs better the outcomes for preterm infants,

the incidence of ROP has actually increased. This trajectory is most likely due to the increase in survival of preterm infants, but suggests that ROP is a condition that must be seriously researched and reduced or eradicated. In the 1940s, premature infants were often placed in closed incubators with supplemental oxygen (often at a level of 100%) to increase the chance of survival by decreasing the chance of major desaturations. However, these infants were seen to experience blindness, as the high level of oxygen causes pathological vascularization of the retina, which may lead to complete detachment of the retina [6].

Later termed ROP, the babies were experiencing pathology of immature retina vascularization, a process that begins in the fourth month of pregnancy and only develops completely just before a healthy birth. For a premature infant, the early birth changes the conditions under which the retina develops, and the infants have underdeveloped retinas with stunted vascularization, leading to a peripheral avascular zone. The risks factors associated with ROP were determined to be the use of high oxygen levels in NICUs and decreased gestation period (extremely preterm infants are more likely to suffer from ROP), but the condition was not more thoroughly understood for years [3]. ROP occurs in two phases dependent on the oxygen of the environment. As explained by Chen, "The first phase begins with delayed retinal vascular growth after birth and partial regression of existing vessels, followed by a second phase of hypoxia-induced pathological vessel growth" [3].

The uterine environment is hypoxic, so the high-oxygen environments in which premature infants may be placed can cause a change in the growth hormones of the infant. Specifically, hyperoxia causes a reduction in oxygen-regulated growth factors such as vascular endothelial growth factor (VEGF) and insulin-like growth factor 1 (IGF-1). Lower levels of VEGF slow the growth of the blood vessels in the retina, and cause the loss of some that have already been formed [6]. IGF-1 helps to regulate VEGF, and thus aberrant levels of the growth hormone correlate with ROP by allowing VEGF to maximally affect the amount of vascularization [3]. Over time, the retina becomes more active, and the pathology begins to transition from phase I to phase II. Since it cannot receive adequate oxygen from the blood due to poor vascularization, the tissue becomes hypoxic, initiating phase two. Hypoxia increases the levels of IGF-1 and VEGF, contributing to new blood vessel formation in the retina. These new blood vessels are poorly formed and leaky, which can lead to retinal detachment and thus blindness [6]. The transition between phases is dependent on postmenstrual age and can be modified by oxygenation in the NICU, though controlled oxygenation reduced but does not necessarily eliminate ROP [6]. Since ROP has this biphasic oxygen susceptibility pattern,

hyperoxia early in life and hypoxia later may trigger the pathology, or exacerbate a preexisting case of ROP for the premature infant [1]. Though it has been seen for over 70 years, critical barriers have retarded the growth of knowledge on the condition. A lack of high-speed computing and data storage capacities no longer limit the research on ROP, nor apnea and other preterm clinical events.

3 Data and Methods

3.1 NICU Signal Data

Our group at William and Mary works closely with the 45-bed University of Virginia NICU, and we utilize the signal data from this NICU, a quarter of whose admittance are VLBW, in order to quantitatively develop algorithms for better monitoring of vital signs. The NICU has generated and shared over five years of data, or over 200 baby-years of data. Bedside monitor data includes chest impedance waveforms (at a frequency of 60 Hz), electrocardiogram waveforms (a tripartite signal, with each lead at 240 Hz), and pulse oximetry signals, delivering oxygen saturation data every 2 seconds (and using 8 second averaging). These electronic signals equate to approximately 10 terabyte of data, making the criticality of high-speed computing and large data storage easily evident. The clinical information related to this data can be found at the UVA Medical Center, while the signal data at William and Mary is housed in the SciClone Computing Complex. This data is in a PostgreSQL format, and is then converted to Matlab files to be used by other algorithms. A sample of the signal data can be found in Figure 1 [11].

Figure 1(a) shows 30 seconds of filtered chest impedance signal with the original data for heart rate, EKG, and chest impedance. In 1(b), we see these signals over a 4-minute from which part (a) was taken. This graph also included the SpO_2 data from the pulse oximetry signal, and a probability of apnea marker on which and ABD event has been labeled. It is clear from the chest impedance measure that the infant has stopped breathing for this time period. Immediately following the period of apnea is a bradycardia, in which the heart rate signal dips below 100%. A desaturation can also be seen, where the SpO_2 dips below 80%. Each of these thresholds are shown on the graph as a dashed line.

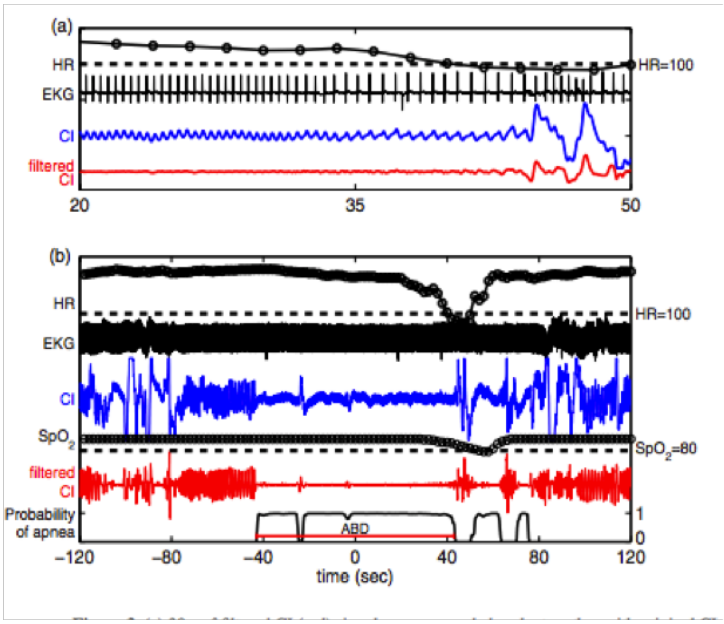


Figure 1: (a) 30 seconds of signal data for a baby experiencing an ABD event. Signals shown for heart rate, EKG, chest impedance, and filter chest impedance data. (b) The same event, expanded to show a 4-minute segment of data. Heart rate, EKG, chest impedance, filtered chest impedance, SpO_2, and probability of apnea (calculated after the event) are shown.[11]

3.2 Apnea and Periodic Breathing Database

Mary Mohr, a recent graduate student from our lab at William and Mary, has created a second database of clinical events. The respiratory data from the SpO_2 and chest impedance signals were analyzed for apnea events and periodic breathing events. Continuous bedside monitor chest impedance and EKG waveforms, as well as oxygen saturation data, from 2009-2014 was used in order to define the apnea events. This resulted in a database of 78,857 apnea events. My goal was then to consolidate the data we have on apnea events, as well as periodic breathing tags, in an easy-to-use format for the clinicians at UVA. The 10 terabyte data set is huge and unwieldy, and researchers, especially clinicians, need systems for easy and efficient access.

3.3 ROP Database

Another data set was created that details the SpO_2 levels for 56 days (8 weeks) for 241 infants that were tested for ROP. Infants are generally tested for ROP around 30-32 weeks, though much of the data continues after this time. It is assumed that if the baby has not developed ROP by 32 weeks, the pathology will not form because the retina has progressed properly by this age. The data set

includes a variable organized by baby and day that describes the total number of readings, as well as the total number of readings for each baby for each day at a certain oxygen level between 60% and 100%. Thus, to find the percentage of time that the infant's SpO_2 is 80%, for example, we would divide the number of readings at 80% for that baby for each day by the total number of readings for that baby for each day. This data set also includes information about the patient, including the PMA in days, the GA in weeks, and the outcome of the ROP test, which may be one of three results: 'ROP,' 'no ROP,' or 'died.' Though there are 8 weeks of data for each baby, much of the data is missing. Some of this data is specified in a variable detailing the number of readings for a particular day and infant for which O2 equals -1, and these readings can easily be removed from the total. However, there is also a large portion of the data for which the reading is simply missing rather than an error. This may be due to moving infants between beds, accidentally unplugging of the monitors, or another human error.

4 Results and Discussion

I have carried out three distinct projects. The first project was the development of a Query User-Database Interface (QUDI) for apnea events and periodic breathing epochs. The second consisted of construction of a database of desaturation and hypersaturation events for all infants at all times for which data is available. Both of these projects are tools that will make the data sets from the UVA Medical Center NICU easier to use for both quantitative researchers and clinicians alike. The third project focused on a research question, testing the hypothesis that statistics on oxygen saturation, given by SpO_2, can predict retinopathy of prematurity.

4.1 Query User Data Interface for Apneas and Periodic Breathing Events

I first developed a Query User Data Interface. This QUDI, a type of graphical user interface (GUI), provides lists of either ABD events or epochs of periodic breathing, with other criteria specified by the user. In order to create this QUDI, I consolidated the data we have in several different matrices containing to the bed number, time of the event, and the patient ID. This information allowed me to correspond the information about the patient (organized by patient ID), such as age, with information about each event (organized by bed number and time), such as length of the event. The user interface is searchable by patient ID, gestational age (the number of weeks of pregnancy before

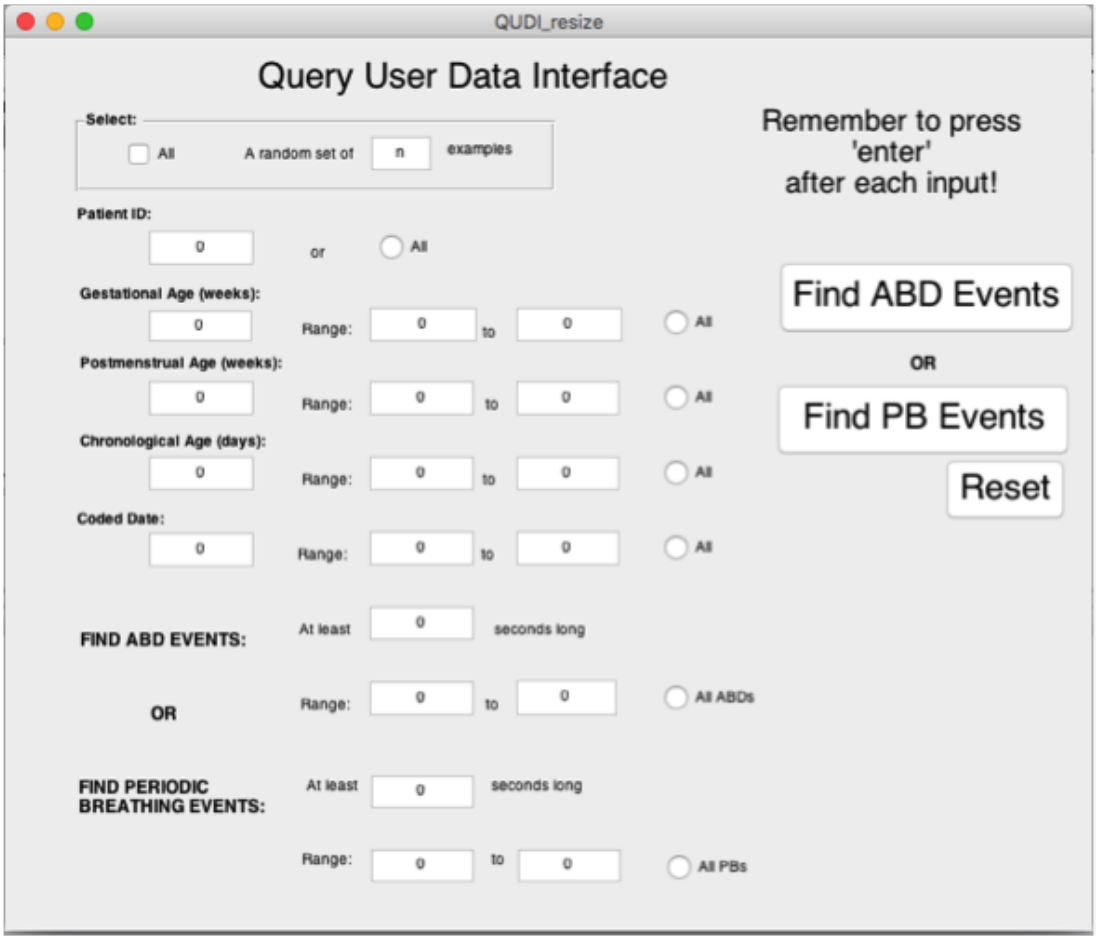


Figure 2: Query User Data Interface.

birth), chronological age (the number of days from birth), post-menstrual age (the number of weeks from last menstrual period to date of event; i.e. gestational age plus chronological age), coded date of event (given in seconds), and the length of the apnea. If a single value is chosen for any of the time inputs (i.e. an age or the coded date), the algorithm returns all events within one unit of time of that event. For example, if all events for infants of chronological age of 20 days are chosen, any infant greater than or equal to 20 days old and less than 21 days old at the time of the event will be returned. It is also possible to return all events that meet the given criteria, or to return only a random subset of n events, determined by the user. I then updated the QUDI to include periodic breathing events, searchable by the same criteria. A sample of the QUDI is shown in figure 2.

An example of the QUDI output is shown in figure 3. For this example, the user chose a subset of 30 random examples from apnea events longer than 20 seconds for infants between 30 and 40

(a)

	1 Patient_ID	2 Bed_Number	3 GA_weeks	4 PA_weeks	5 CA_days	6 Coded_date	7 Duration_of_event
1		25	28	32.0980	28.6863	117834200	23.5000
2		16	32	36.3517	30.4619	19456686	22.2500
3		39	31	34.0414	21.2900	47303078	20.5000
4		25	28	31.9051	27.3355	117717483	25
5		39	32	32.2253	1.5772	50039287	29.7500
6		38	25	31.6596	46.6174	142280040	22
7		36	31	35.6273	32.3912	151784541	33.7500
8		38	25	30.9260	41.4823	141836367	35.2500
9		25	28	31.9607	27.7249	117751133	30.7500
10		9	34	34.0979	0.6854	143004182	23.2500
11		37	25	32.6591	53.6136	141119256	24.2500
12		16	32	34.9728	20.8095	18622722	33.7500
13		38	25	30.9791	41.8538	141868465	24.2500
14		25	28	32.0641	28.4490	117813696	30.2500
15		32	29	31.4087	16.8607	136066581	29.5000
16		7	35	35.2755	1.9284	140191136	25.2500
17		7	36	36.2754	1.9276	138694844	27.5000
18		25	28	32.0292	28.2047	117792584	28.5000
19		9	38	38.2589	1.8126	144707529	29
20		11	31	37.4225	44.9572	77981724	37.2500

(b)

Figure 3: Sample output of QUDI, displaying (a) the number of events found and number of random events selected and (b) the first 20 results. The table is saved as a .mat file including information about the events such as the patient ID, bed number, age of patient, and duration of event.

weeks PMA (postmenstrual age). Upon hitting go, the user will see a box such as Figure 3 giving the summary of events, including the total number of events and the number of random events (if a random set of n examples is chosen). This subset of data of events that meet the criteria is then saved in a Matlab file as a table with information on each relevant event. The output is shown in this report, but the patient ID column is removed due to the sensitivity of the data. This QUDI currently uses a database of 78,857 apnea events and approximately 894,700 periodic breathing events, and it can easily be updated to read in another set of data as we continue to receive monitoring data from the UVA NICU.

4.2 SpO_2 Statistics

After this project was complete, I began to analyze the SpO_2 signal. I used the data housed in the SciClone Computing Complex for this analysis, focusing on the SpO_2 measurement in order to determine the desaturation events experienced by each infant in our database. Desaturations are defined as an event in which the SpO_2 level drops below 80%, and I used this threshold along with 75% and 85%. I also considered the hypersaturations, which are correlated with clinical problems such as ROP. While the optimal level of oxygen is unknown, as explained, hypersaturations can cause imperfect retinal development and even blindness in preterm infants. Thus I cataloged events in which the SpO_2 levels rose above 95%, 96%, 97%, and so on. Each subset of data is categorized by bed and day; thus one Matlab file of data could correspond to different infants, depending on when each infant is born, moves beds, or leaves the NICU. Thus the desaturations and hypersaturations must be correlated with infant IDs in order to be used for future research. This can be done using the database that connects the bed number and time to a specific infant (via the patient ID).

Each event has associated characteristics that are generated by my algorithms in Matlab. SpO_2 is recorded in the data set every two seconds, giving a discrete signal, and sometimes there is missing data. Therefore, definitions of characteristics of events involve some nontrivial issues, and the definitions decided upon required discussions with and approval by our clinical colleagues at UVA. T_{up} is a variable that marks the beginning of a desaturation event, and T_down is one that marks the end of a desaturation event. For a hypersaturation, these markers are switched; in other words, T_{up} delineates the beginning of the event, and T_{down} the end of the event. The duration of the event is simply the subtraction of these two values. How these variables are defined with regards to the discrete data set is explained as follows, and thus the definition of an event is dependent upon the definitions for T_{up} and T_{down}.

Considering desaturations, if the data goes below the threshold, stays below for a time, and later goes above, then the event is easy to define. T_{down} is defined as the first value below a given threshold; i.e., when the data goes from above a threshold to below it, the first data point less than 85% is considered T_{down}. Likewise, T_{up}, the end of the event, is the last point below the threshold before returning above the threshold. If the data just touches the threshold and then goes back above it, we do not count this as a desaturation event. However, we determined that clinically it is logical to consider that touching the threshold should not necessarily end the event. In other words,

if the signal starts below the threshold, touches the threshold, and then goes back below (in other words, touching the threshold in the middle of an event), it is considered to be one desaturation event. T_{down} in this case is still the first point below the threshold, and T_{up} is still the last point below it. Clinically, it was important to treat this as one event in case the discrete signal fluctuates near the threshold. To find the minimum or maximum SpO_2 for each of these events, I calculated the lowest or highest value within each event. For the duration of each event, I subtracted T_{up} from T_{down}.

There were some more complicated definitions due to the missing data. If SpO_2 is above the threshold before and after the missing data, it is ignored. If the signal is above the threshold prior to the missing data and below when the data resumes, then T_{down} is considered to be the first point below the threshold after the missing data. Conversely, if the signal is below the threshold prior to missing data and above when the data resumes, the last point before the missing data is marked as T_{up}. If during a desaturation event there is missing data (i.e. before and after the missing data the signal is below the threshold), we must consider two cases. If the missing data lasts less than or each to six seconds (three data points or less), then it is ignored in the definition of T_{down} and T_{up} and interpolated as a part of a single event. If the missing data lasts longer than six seconds, then the last point before the missing data is considered T_{up} of the previous event, and the first point after the missing data is considered T_{down} of the following event.

Calculating the %-time area was also complex due to missing data and the discrete nature of the sum. Because each signal was taken at two-second intervals, the area is calculated as the sum of the difference between the threshold and the signal value, multiplied by two (for the time interval). However, due to some missing data, some time intervals are longer than two seconds. There was missing data for unknown reasons related to problems with recording or monitoring the signals in the NICU, such as an infant becoming unplugged from the monitor. If there is missing data for more than six seconds (three data points), and the next data point is still below the threshold, the data is counted as two different desaturation events. The first event ends at the last point before the missing data, and the second one begins and the first point after the missing data. The missing data is not included in an area or duration. If the missing data lasts six seconds or less (three or fewer points are missing), and the next data point is below the threshold, it is considered one event. The area is then interpolated by multiplying the duration of missing data by the average of the first point after and the last point before the missing data. We include the missing data time in the

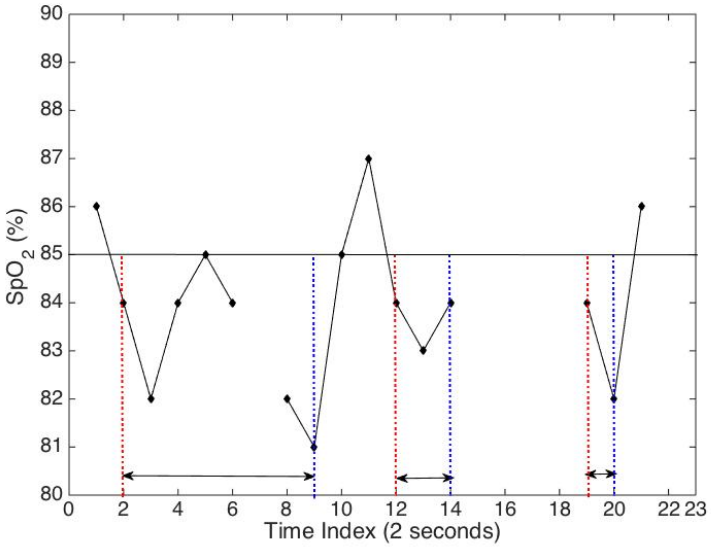
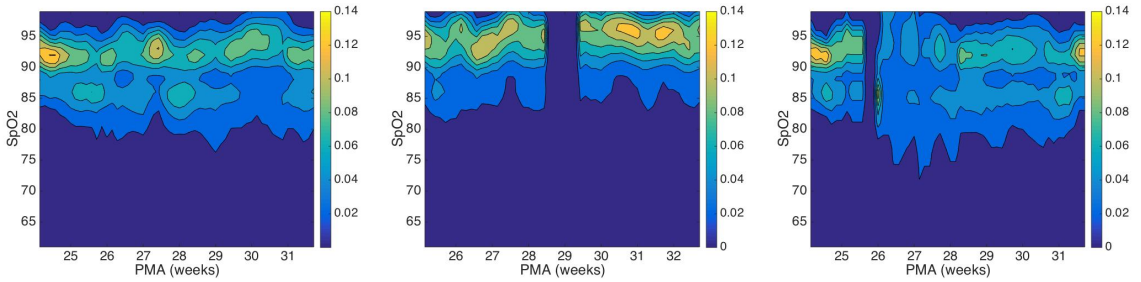


Figure 4: An example of a (fabricated) SpO_2 signal for 42 seconds (2 second intervals) displaying nontrivial issues in defining events. Three 85% desaturation events, with T_{down} and T_{up} shown by red and blue dotted lines, respectively. In the first event, missing data ($\leq$ 6 seconds) is ignored, and interpolation is used to calculate the %-time area of this event. Between the second and third events, missing data ($>$ 6 seconds) causes the desaturation to be labeled as two separate events.

duration of the event. A sample of these aberrations is shown in figure 4 with an explanation of how each event is defined.

The generated files are arranged in the Sciclone Computing Complex by bed (folder) and day (file name) to correlate with the original Matlab files. There are 45 beds, and each bed contains a number of days in the order of 10^3 to 10^4. For each bed and day, thousands of desaturation and hypersaturation events are analyzed, and thus the number of analysis-generated events is on the order of 10^7 to 10^8. Each set of files by bed takes approximately 12 hours to analyze. Once it is evident which patient is in a particular bed at a specified time, as described previously (using the bed number database), the user can locate a particular infant's desaturation and hypersaturation events in the SciClone Computing Complex. The Matlab files for each bed and day contain all of the events from the available data arranged by variable and correlate with one another by index (e.g. $T_{up_{85}}$ is the variable of T_{up} values for the threshold of 85%, and $T_{down_{85}}$ is the variable correlated with these values).



(a) Baby 6; Outcome: ROP (b) Baby 98; Outcome: No ROP (c) Baby 209; Outcome: Died

Figure 5: Sample contour plots of SpO_2 levels for (a) an infant with an outcome of ROP, (b) an infant with an outcome of no ROP, and (c) an infant that died within the 56 days of monitoring.

4.3 ROP Analysis

With the completion of this second project, I turned my attention to a third project: correlation of SpO_2 with ROP. As mentioned above, it has been hypothesized that higher oxygen saturation levels at an early age and lower levels at a later age correlate with diagnosis of ROP for premature infants. I tested this hypothesis using the SpO_2 signals from the NICU at the UVA Medical Center, which has the diagnosis for 241 premature infants born after 22 weeks. The infants are characterized by one of three diagnoses: "ROP," "no ROP," or "died." I was able to calculate the percentage of time spent at each SpO_2 level (between 60% and 100%), using data[1] on the number of readings for which the infant was at a given oxygen level and the total number of readings per day. Then, using the diagnosis tags, I compared infants of different subsets. Data was smoothed over 3 days and 3 values of SpO_2. After this averaging, a contour and surface plot was produced for each infant. Sample contours plots are shown in figure 5.

By visually examining these plots for individual infants, one could easily see a few striking patterns. The infants labeled "died" often had substantially low oxygenation prior to death, as extreme desaturation can signal fatal respiratory failure. Surprisingly, many of the infants who developed ROP spent the highest percentage of time in a normal range of 90% to 95% oxygenation. In comparison, many who did not develop ROP spent much of their time in a normal range as well, but mostly during the early weeks of life. Generally, an infant without ROP experienced hypersaturations near the end of the 56 days, as evident in the individual displayed in figure 5 (b). In order to verify the occurrence of these trends, I then looked at the average SpO_2 distributions by outcome ("ROP,"

[1]This data was generated for each infant by Professor Emeritus Stephen Knudson, College of William and Mary.

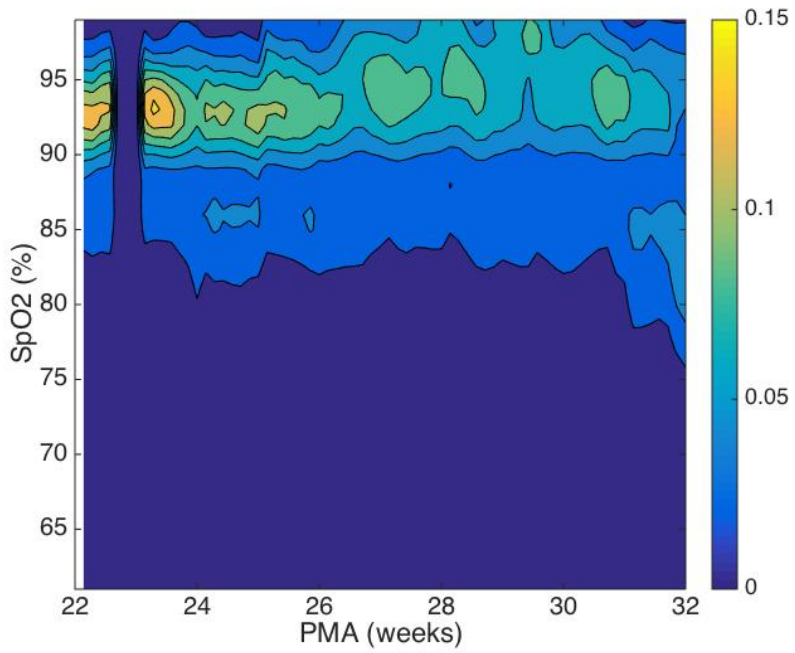
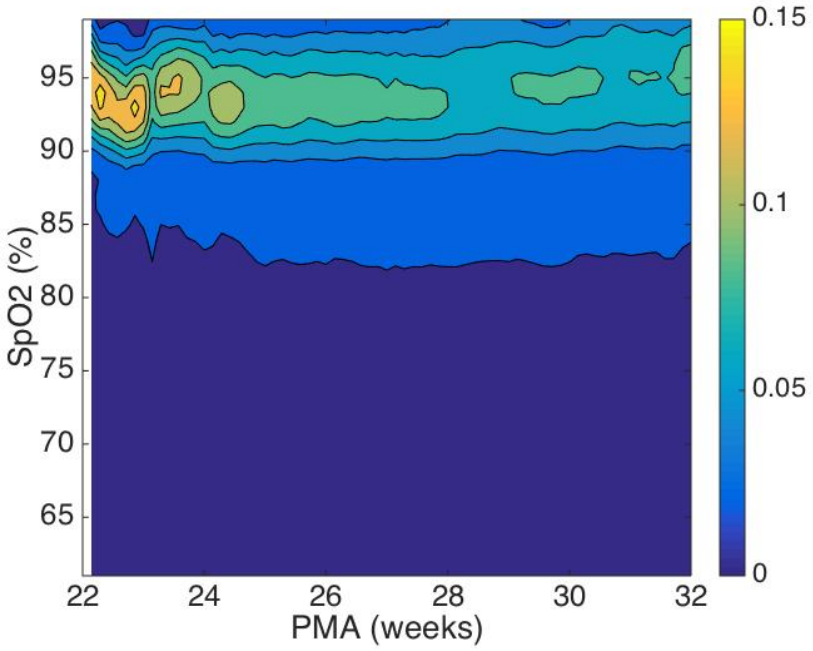


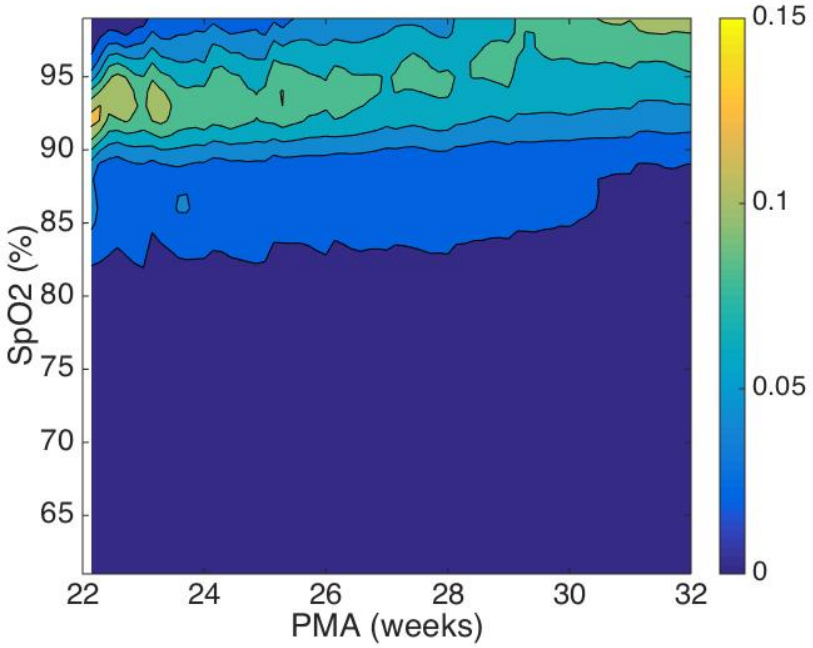
Figure 6: Contour plot of frequency of SpO_2 levels versus age, averaged for infants with an outcome of "died."

"no ROP," and "died"). The data for each category was averaged over the infants so that a mean distribution of oxygen levels was produced for each outcome. These distributions are shown in figures 6 and 7. In figure 6, SpO_2 levels are averaged over the infants with an outcome of "died." The oxygen levels are generally stable early, yet drop the longer the baby is alive, substantiating the results seen in the individual graphs. Data beyond 32 weeks (after the test; not shown in graph) continue to drop, and it becomes clear that this major drop in oxygenation leads to fatality.

In figure 7, some general trends appear that were evident in many of the individual graphs, such as those in figure 5 (a) and (b). The initial results we obtained from these graphs were misleading, appearing to show a difference between the ROP and non-ROP infants without regards to age. The infants that develop ROP in 7 (a) had SpO_2 levels that were surprisingly normal. In contrast, the SpO_2 levels for the non-ROP infants were higher, especially near the end of the monitored data. These trends could be confirmed with a Kolmogorov-Smirnov Test. This test is a nonparametric statistical test that compares two samples in order to tell if they possess the same distribution and thus could be from the same population. The test makes no assumption about the type of probability distribution of the variables being assessed. A decision for the two-sample Kolmogorov-Smirnov Test

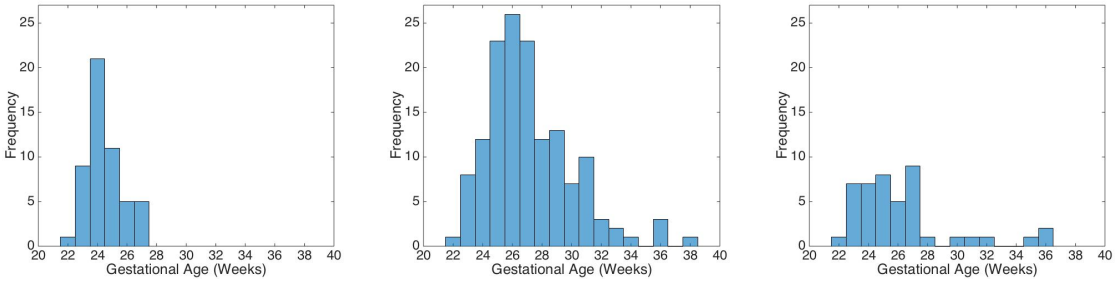


(a) Outcome: ROP, n = 52



(b) Outcome: No ROP, n = 145

Figure 7: Contour plots comparing SpO_2 averaged over infants with an outcome of (a) ROP or (b) no ROP. This figure includes all infants with diagnosis of ROP or no ROP.



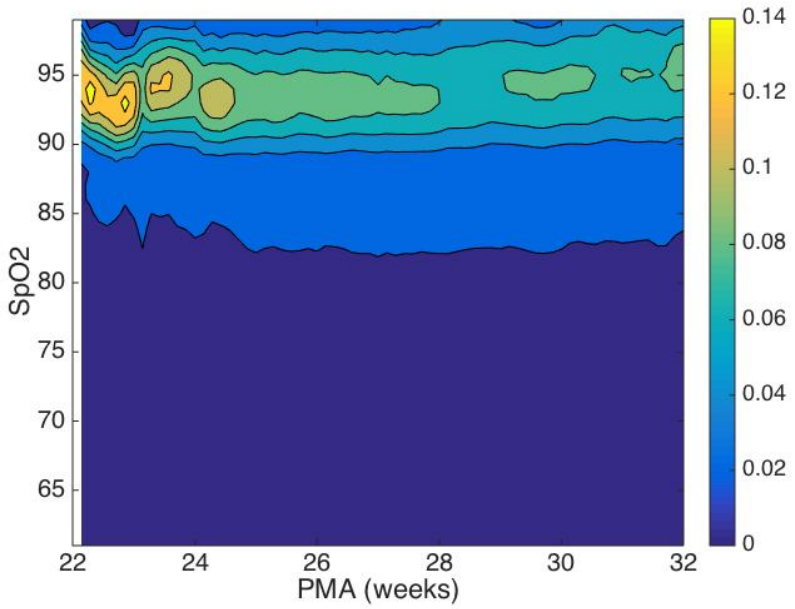
(a) Outcome: ROP, n = 52 (b) Outcome: No ROP, n = 145 (c) Outcome: Died, n = 44

Figure 8: Gestational age distribution for each test outcome.

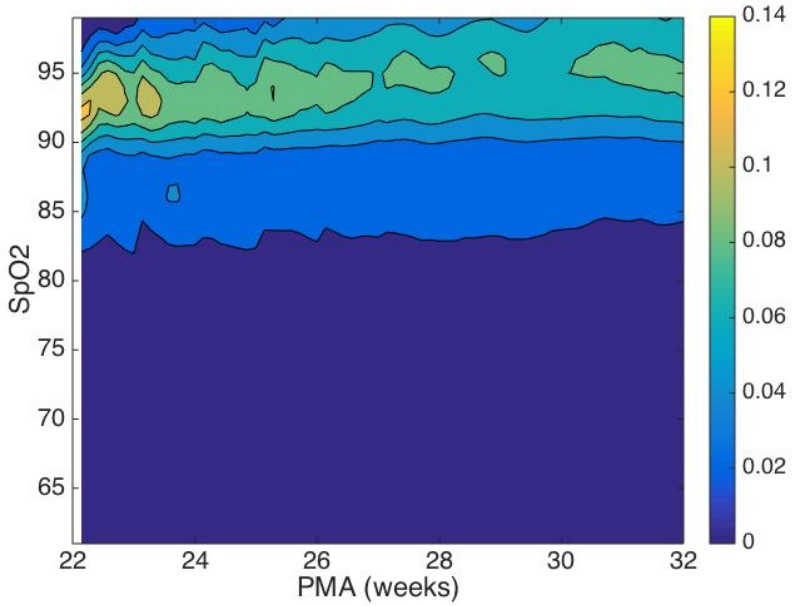
null hypothesis that the data from two samples are from the same continuous distribution is given at the 5% significance level. The test was run on the averaged SpO_2 distribution for each day in order to determine which days had statistically significant differences between the infants who developed ROP and those who did not. As expected, many of the later days (between weeks 34 and 39 postmenstrual age) were statistically significant, where the non-ROP infants had higher saturations. The first two days also had statistically significant differences.

Because the infants were born at different gestational ages, we assumed each infant was born at an integer number of weeks, as recorded in the data (for example, no infants were born at 22 weeks and 3 days). Thus, the averaged SpO_2 data includes a different number of infants in each day's sum. Age distribution data for each outcome are shown in figure 8, and this may clarify the averaging used to generate figures 7 and 6. For example, at the lower bound on the domain in figure 7 (a), only one infant was born at 22 weeks gestational age that eventually developed ROP. Thus, for the first seven days in the graph, only one infant is being "averaged" in this graph.

This led us to believe that age may be a confounding factor in the relationship between diagnosis of ROP and SpO_2 levels. It may be possible, from these graphs, that the higher saturations for later weeks for those infants who did not develop ROP may be attributed only to the infants with a longer gestation; thus, the high oxygenation would have had less impact on the developing retina. The infants that developed ROP, on average, had a much lower gestational age, and thus much less developed retinal vascularization, than those who did not develop ROP. Because no infants who had a "ROP" outcome were born after 27 weeks, the next step was to compare the infants who developed ROP with those who did not develop the disease *and* were born before 28 weeks gestation. This is shown in figure 9.



(a) Outcome: ROP, n = 52



(b) Outcome: No ROP, n = 93

Figure 9: Contour plots comparing SpO_2 averaged over infants with an outcome of (a) ROP or (b) no ROP who were born before 28 weeks gestation.

In comparison to figure 7, where all of the infants are included in the averaging, figure 9 shows a smaller difference between the outcomes. In fact, when a Kolmogorov-Smirnoff Test is used to evaluate the two sample distributions, the only days on which we reject the null hypothesis that "the data from two samples are from the same continuous distribution" are the first two days. There is only one infant in each category born at 22 weeks gestation, so for the first week each distribution has a sample size of one. Since the infants are obviously different, the fact that the first two days of their distributions are different can be disregarded. Thus, when eliminating gestational age as a confounding factor, the distributions for infants that develop ROP and those that do not are not significantly different according to a statistical test with p=.05. Thus, according to this data, once controlling for age, there is no statistically significant difference in the distributions of SpO_2 between ROP and non-ROP infants. Hence, we cannot predict diagnosis of ROP based on SpO_2 statistics alone; age appears to be a great enough factor that, when eliminated, oxygen levels for each group are similar and hold no predictive power.

5 Conclusion

Pathology associated with premature birth causes a major health care issue around the world, due to the incidence of such births; approximately one tenth of all infants are born prematurely. Due to their low birthweight, immaturity, and undeveloped systems, premature infants have high mortality and morbidity rates. Some types of the pathology that such infants face include those associated with aberrant oxygenation: apnea of prematurity, sepsis, periodic breathing, and retinopathy of prematurity. Though sufficient research has been dedicated to these medical issues, the optimal conditions for premature infants in NICUs still remain unknown, and the monitoring of ill patients still has many faults that increase the cost of such illnesses in both fatalities and monetarily, as health care costs. The goals of this research project have been to develop tools that will make the data we have easier to work with for both researchers and clinicians alike, as well as to test the hypothesis that oxygen saturation statistics can predict diagnosis of retinopathy of prematurity.

My first project required me to develop a Query User Data Interface (QUDI), a type of graphical user interface that outputs a list of apnea events or epochs of periodic breathing. The QUDI takes input as specified by the user that determines the characteristics of the events to be found. For example, the user can identify such characteristics as a range of ages or dates, choose tags such

as the lower bound on the length of the event, or select a particular infant by its patient ID. The QUDI outputs a list of events with characteristics for each event such as the bed number, patient ID, and age associated with each patient, and the date and duration associated with each event. By organizing the data in an easy-to-use format, the QUDI can save valuable time for physicians, and it can help future researchers to better understand the large amount of data on apnea and periodic breathing events available to our group.

I then turned my attention to SpO_2 data in order to better catalog the mass of information we have on past patients. The SpO_2 data used comes from monitor data from the UVA Medical Center NICU, a 45-bed quaternary care facility. We have over 200 baby-years of signals that equates to several terabyte of data. Using the SpO_2 signal, I generated information on desaturation and hypersaturation events, which are known to be associated with several types of pathology in premature infants. Definitions for each of these events required approval of our physician partners at the UVA NICU. For each event, I generated characteristics associated with the desaturation or hypersaturation, including the beginning and end of the event, duration, minimum (or maximum) SpO_2, and %-time area. Due to missing data and the discrete nature of the signal (collected and recorded every two second), there were several nontrivial issues with defining the events that had to be resolved, guided by the advice of our clinical partners. For example, if desaturation points are separated by more than six seconds, they are counted in two separate desaturation events. If the missing data occurred for six seconds or less, it is ignored, only one event is counted, and the %-time area is interpolated. The desaturation and hypersaturation events are organized in the SciClone Computing Complex in folders by bed and in Matlab files by day. For each of the 45 beds, thousands of events were generated.

Lastly, I evaluated a research question regarding the correlation between oxygenation, given by SPO_2, and ROP. ROP, a disease known for more than 70 years, is expected to correlate with high oxygenation in the early weeks of a premature infant's life, and low oxygenation in the later weeks. Using a Kolmogorov-Smirnov Statistical Test, it was found that SpO_2 was significantly different in weeks 34 to 39 postmenstrual age at p=.05 significance level for infants that received a diagnosis of ROP. Age, however, was a confounding factor. Because extremely preterm infants are born before the retina is properly vascularized, they have the highest risk of developing ROP. Thus, it became evident that no infants born after 28 weeks gestation developed ROP. To control for this, I compared only the infants with each outcome born at 28 weeks gestation or prior. This reduced the data set

so that the n_{ROP} was 52 infants and n_{noROP} was 93 infants. No significant difference between the outcomes could be found. Thus SpO_2 signals, controlling for gestational age, were not able to make a significant prediction of ROP. This leaves two hypotheses remaining. There may be no relationship whatsoever between measured oxygenation and ROP. In other words, ROP cannot be predicted from SpO_2 levels at all. A second hypothesis suggests the short-term fluctuations of SpO_2 may be correlated with ROP. Weak evidence favors this hypothesis, suggesting that control of small fluctuations in the oxygenation of VLBW premature infants can decrease the incidence of ROP [14]. Further evidence suggests that a higher incidence of intermittent hypoxia is associated with severe ROP in premature infants [4]. To determine which of these hypotheses can be rejected or supported, more research, and therefore more data, is needed.

References

[1] Biran, V., Verney, C. & Ferriero, D. M. "Perinatal cerebellar injury in human and animal models." *Neurol. Res. Int.* **2012**, 858929 (2012).

[2] Busse, W. W., Erzurum, S. C., Blaisdell, C. J. & Noel, P. "Executive Summary: NHLBI Workshop on the Primary Prevention of Chronic Lung Diseases." *Ann. Am. Thorac. Soc.* **11 Suppl 3**, S123–4 (2014).

[3] Chen, Jing, and Lois EH Smith. "Retinopathy of prematurity." *Angiogenesis* **10.2** (2007): 133-140.

[4] Di Fiore, Juliann M., et al. "A higher incidence of intermittent hypoxemic episodes is associated with severe retinopathy of prematurity." *The Journal of pediatrics* **157.1** (2010): 69-73.

[5] Fairchild, K., and Judy L. Aschner. "HeRO monitoring to reduce mortality in NICU patients." *Research and Reports in Neonatology* **2** (2012): 65-76.

[6] Hellström, Ann, Lois EH Smith, and Olaf Dammann. "Retinopathy of prematurity." *The lancet* **382.9902** (2013): 1445-1457.

[7] "Health Library: Pulse Oximetry." *Johns Hopkins Medicine*. Johns Hopkins University, n.d. Web. 30 March, 2016.

[8] Hofstetter A, Legnevall L, Herlenius E, Katz-Salamon M. "Cardiorespiratory development in extremely preterm infants: Vulnerability to infection and persistence of events beyond term-equivalent age." *Acta Paediatr Int J Paediatr* 2008;97:285–92.

[9] "How Pulse Oximeters Work Explained Simply." *How Equipment Works*. Avada, n.d. Web. 28 March, 2016.

[10] Islam, J. Y., Keller, R. L., Aschner, J. L., Hartert, T. V & Moore, P. E. "Understanding the Short- and Long-Term Respiratory Outcomes of Prematurity and Bronchopulmonary Dysplasia." *Am. J. Respir. Crit. Care Med.* **192**, 134–56 (2015).

[11] Lee, Hoshik, et al. "A new algorithm for detecting central apnea in neonates." *Physiological measurement* **33.1** (2012): 1.

[12] Mohr, Mary A., et al. "Very long apnea events in preterm infants." *Journal of Applied Physiology* **118.5** (2015): 558-568.

[13] Moorman, Joseph Randall, et al. "Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial." *The Journal of pediatrics* **159.6** (2011): 900-906.

[14] York, Jackie R., et al. "Arterial oxygen fluctuation and retinopathy of prematurity in very-low-birth-weight infants." *Journal of perinatology* **24.2** (2004): 82-87.

[15] Zupancic, J. A. F., Richardson, D. K., O'Brien, B. J., Eichenwald, E. C. & Weinstein, M. C. "Cost- Effectiveness Analysis of Predischarge Monitoring for Apnea of Prematurity." *Pediatrics* **111**, 146–152 (2003).