# A Flexible, Longitudinal and Surrogate Consent Model: Consent of Infants for Neonatal Secondary-use research (CoINS) Model

by

Yvonne Choi

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

**Master of Health Sciences in Health Informatics**

Faculty of Health Sciences

University of Ontario Institute of Technology (Ontario Tech University)

Oshawa, Ontario, Canada
April 2020

## THESIS EXAMINATION INFORMATION

Submitted by: **Yvonne Choi**

**Master of Health Sciences** in **Health Informatics**

| Thesis title: | A Flexible, Longitudinal and Surrogate Consent Model: Consent of Infants for Neonatal Secondary-use research (CoINS) Model |
|---|---|

An oral defense of this thesis took place on April 29, 2020 in front of the following examining committee:

**Examining Committee:**

| | |
|---|---|
| Chair of Examining Committee | Dr. Manon Lemonde |
| Research Supervisor | Dr. Carolyn McGregor |
| Examining Committee Member | Dr. Jennifer Heath |
| Thesis Examiner | Dr. Amandeep Sidhu |

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

## ABSTRACT

Documenting healthcare, along with technology enabling capture of streaming patient telemetry, can deliver large datasets offering opportunities to discover new insights primarily identified through retrospective secondary use research. Research involving health data requires consent of the subject patient or someone with the power to speak on that patient's behalf. Flexible consent models that capture consent preferences while allowing updates as preferences change are needed. This research proposes and demonstrates one solution in a case study collecting surrogate consent from parents for the physiological data of infant inpatients in the Neonatal Intensive Care Unit (NICU) and attaching this consent as a wrapper controlling access to their data. 145 parents were approached and 134 provided consent: with 78 percent of infants consented during their first week of life. This research supports the contention that using a flexible consent approach enhances willingness to consent use of infant's health data for secondary research purposes.



**Keywords:** surrogate consent, secondary use, physiological data, neonatal research

## CO-AUTHORSHIP STATEMENT

**Choi, Y.**, McGregor, C., 2017, “A Flexible Parental Engaged Consent Model for the Secondary Use of Their Infant’s Physiological Data in the Neonatal Intensive Care Context”, IEEE International Conference on Healthcare Informatics 2017 (ICHI 2017), pp 502-507

McGregor, C., Heath, J. and **Choi, Y**., 2015, “Streaming Physiological Data: General Public Perceptions of Secondary Use and Application to Research in Neonatal Intensive Care”, The 15th World Congress on Medical and Health Informatics (MedInfo, 2015), Sao Paulo, Brazil, pp 453-7.

**Choi, Y.**, McGregor, C., (2014) "The secondary use of identifiable physiological data for health research purposes”, Paper presented at the 7th Canadian Student Conference on Biomedical Computing and Engineering & IEEE EMBS International Student Conference, (CSCBCE IEEE EMBS ISC), Oshawa, Ontario.

## AUTHOR'S DECLARATION

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

The research work in this thesis that was performed in compliance with the regulations of Ontario Tech's Research Ethics Board under #14736 for the Artemis Cloud data collection study and #15536 for the Late Onset Neonatal Sepsis study.

| | |
|---|---|
| | Yvonne Choi |

## STATEMENT OF CONTRIBUTIONS

The data collection described in Chapter 5 was performed at McMaster Children's Hospital in Hamilton, Ontario by Geoff Travis. I was responsible for analysis of the collected data.

Part of the work described in Chapter 2, 3, 5 and 6 has been published as:

**Choi, Y.**, McGregor, C., 2017, "A Flexible Parental Engaged Consent Model for the Secondary Use of Their Infant's Physiological Data in the Neonatal Intensive Care Context", IEEE International Conference on Healthcare Informatics 2017 (ICHI 2017), pp 502-507

Part of the work described in Chapter 2 and 3 has been published as:

McGregor, C., Heath, J. and **Choi, Y**., 2015, "Streaming Physiological Data: General Public Perceptions of Secondary Use and Application to Research in Neonatal Intensive Care", The 15th World Congress on Medical and Health Informatics (MedInfo, 2015), Sao Paulo, Brazil, pp 453-7.

## DEDICATION

To my immediate family – Dad, Mom and Michelle. Thank you for always being there for me, bringing life to me and bringing me back to life.

To Scott – thank you for adding colour to my life and showing me a whole new world.

To my guardian angels – biological, adoptive, friends and my beloved Kiwi! Your presence is missed every day, but your legacies live on. Until I see you all on the other side of the stars, I hope you know that you each played a role in inspiring me to never give up.

To my friends – those who had some idea what I was doing and to those who don't – you can finally stop asking me what my thesis is about because you can finally read it for yourself. Special shout out to: Nadja Johnson Bressan, Yvonne Lam and Samantha Lee for your mentorship and friendship over the years. As Paulo Coelho said, "Friendship isn't a big thing – it's a million little things".

To all the good people I have been so blessed to have in my life – thank you from the bottom of my heart for all of your never-ending encouragement, support, kindness and love. Thank you for taking a chance on me by giving me the time, resources and opportunities I needed to discover my passions. Thank you for being my teachers, mentors, friends and for making me laugh, but most importantly for reaching out to me when I least expected anyone to understand and needed someone to turn to the most.

I hope that one day I can repay all of your kindness and pay it forward by inspiring someone else to go for their dreams.

## ACKNOWLEDGEMENTS

I want and need to express my deepest gratitude to my supervisor, Dr. Carolyn McGregor, without whose invaluable guidance, encouragement and support, this academic journey would never have started – or come to fruition. Thank you for the all that you have done and are still doing for me.

Dr. Jennifer Heath, for your time and efforts in making the thesis content much clearer and more coherent and for all the additional references you have linked me to. This work would not have been possible without building on the foundation of your work.

Dr. Scott McLachlan, for your your extreme patience when it came to my academic journey. Whether it was teaching me how to organize content for a literature review, sending me interesting journal articles, proofreading, for making that fancy Excel feature (that was supposed to make life easier) function the way it was meant to, seeking your lab mates' particular expertise that I needed in writing this thesis, helping me prepare for the oral defense or just listening to my incoherent thoughts, you were always there for me.

To the Health Informatics Research (HIR) team – Catherine Inibhunu, Anastasiia Prysyazhnyuk, Jennifer Yeung and Tobias Cibis – thank you for your encouragement, proofreading assistance and help in forming the initial structure of some of the middle chapters and/or presentation feedback. It's not very often that I have a live audience on demand!

Yvonne Lam, Samantha Lee, Melissa Koh and Pavidra Ambiganithy for all those proofreading sessions and/or impromptu Excel crash courses when we were supposed to be just "hanging out". Even though some of you do not have any (health) science backgrounds nor did you know what you signed up for when you agreed to help me revise my 100+ page document, all of you provided important feedback from your perspectives and understanding. Yvonne and Sam, your

accounting backgrounds (i.e. attention to detail) were beneficial to me in the formation of this thesis. I hope that I have not turned any of you off from the possibility of pursuing further studies in your professions.

My thanks to all of you for your dedication, efforts, time and for not chucking a textbook or a laptop at me even when I really deserved it. None of this would not have been possible without all of your encouragement and support.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# LIST OF ABBREVIATIONS AND SYMBOLS

CoINS – Consent of Infants for Neonatal Secondary-use research
CRISP-DM - CRoss Industry Standard Process for Data Mining
DC – Dynamic Consent
DM – Data Mining
EBM – Evidence-based medicine
ECG – Electrocardiogram
EEG – Electroencephalogram
HR – Heart Rate
EHR – Electronic Health Records
IT – Information Technology
IRB – Institutional Review Board (see REB)
LHS – Learning Health Systems
MC – Meta Consent
NICU – Neonatal Intensive Care Unit
PCI - Participant-Centered Initiative
PHI – Patient/Personal health information
PF4SUMD – Privacy Framework for Secondary Use of Medical Data
RA – Research Assistant
REB – Research Ethics Board (see IRB)
SDM – Surrogate decision maker
$STDM^{n}_{0}$ – Service based Multidimensional Temporal Data Mining Framework
TA – Temporal Analysis
TDM – Temporal Data Mining

# 1. Introduction

Florence Nightingale helped to transform healthcare through the introduction of formal and descriptive clinical documentation that would become a vital component of effective care delivery (Iyer & Camp, 1995). Clinical documentation serves many purposes, with the most significant being to ensure continuity and quality of care for the subject patient (Fischbach, 1991; Iyer & Camp, 1995).

Modern practice incorporates technological devices into the care process that are capable of displaying continuous real-time data streams on various aspects of the patient's physiological state. These devices are capable of recording these observations with much higher fidelity than could be documented by the care team, who may only do so twice hourly, hourly, or four hourly (Considine, Trotter, & Currey, 2016; Sriwatanakul et al., 1983). Variations in measured physiological parameters are common in the critically ill, but observations documented too infrequently may not make obvious the onset of potentially dangerous disease processes. While most medical devices are capable of reporting and storing the data as it is acquired, even many times per second, hospitals generally do not or are unable to capture the data and it is ultimately lost. Thus, the data is unavailable for further analysis because the monitoring device is not connected to an appropriate network or data storage systems. If healthcare professionals had access to clinical decision support systems (CDSS) that could source datasets of high-frequency observational data for analysis, there is greater potential for earlier recognition of the deteriorating patient. This recognition is vital for increasing patient safety and reducing the significant harm that results from potentially serious adverse health events (Cardona-Morrell et al., 2016; Lister, Bryan, & Tracy, 2000; McGregor, James, et al., 2013; van Veen, 2008)

Recent studies on neonatal health conditions have demonstrated the potential of continuous real-time physiological data streams, with one study, *Late Onset Neonatal Sepsis* (LONS), providing an example for how large collections of physiological data could be used in predictive modelling (Fairchild, 2013; McGregor, Catley, & James, 2012). Use of patient data to perform research is

always a *secondary use* of that data with the primary use of the data for the care of the patient (McLachlan et al., 2018).

Secondary use of medical records is an important means of scientific investigation, supporting a wide range of research including epidemiology, public health, drug surveillance and the design of decision support systems (Miller, 2008; Neamatullah et al., 2008; Weiskopf, Hripcsak, Swaminathan, & Weng, 2013).

Research often relies on secondary use of patient data which, unless it has been anonymized with disclosure permitted by a research ethics board waiver, requires consent from every individual subject (Willison et al., 2007). Informed consent is a communication process between the patient/participant and the clinician/clinical researcher. Competent adults and Gillick competent minors, who are determined under medical law to be able to consent on their own behalf, are advised of the clinical test or treatment recommendation and related risks and benefits, and are given an opportunity to ask questions prior to providing agreement or permission for those tests or treatments to occur (American Cancer Society, 2019).

In the current privacy-concerned age patients prefer to be involved in the decision to undertake research involving the use of their personal medical records (Lunshof, Chadwick, Vorhaus, & Church, 2008; Nair, Willison, Holbrook, & Keshavjee, 2004; Robling et al., 2004; Sanderson et al., 2017; Willison et al., 2009). Patients often want to know when and why their medical records are being accessed, and to have the opportunity to consent or decline participation in such research (Nair et al., 2004; Robling et al., 2004).

Conventional consent for research participation is most often achieved through the researcher explaining the purpose and degree of participation required for a particular study. Different consent models have been described and may be in clinical use, and while they are meant to reduce the consenter's stress and decision-making burden, in situations involving children, and in particular infants, many do not consider parental autonomy and, in some cases, may not result in valid informed consent (Golec, Gibbins, Dunn, & Hebert, 2004). In some cases, if a clinician believes in the benefit to the infant patient, it is not uncommon for them

to engage the child in treatment or research without seeking consent and potentially against the wishes of the parents (McHaffie, Laing, Parker, & McMillan, 2001) or limit parental decision-making authority if the parents' decision is not in the infant's "best interest" although this would primarily be based on the healthcare professional's judgment of "best interest" (Albersheim, Lavoie, & Keidar, 2010). One study that offered hypothetical scenarios found that neonatologists would limit parental decision-making authority in situations where there was a clear treatment benefit (i.e. greater than 50% chance of intact survival or survival without severe handicap), but only up to 18% of them would limit parental decision-making authority despite a very poor chance of intact survival for a preterm neonate (Albersheim et al., 2010). When treatment is either clearly beneficial or clearly of no benefit, parental decision-making typically follows the medical estimate of the baby's best interest, based on the neonate's threshold of viability even though the neonatologists' self-rated respect for parental decision-making authority was 8/10 (Albersheim et al., 2010). Healthcare professionals may even approach internal ethics review boards and family courts, seeking a higher authority's approval in overriding the parent's views (Kopelman & Kopelman, 2007). The use of a less pressuring, more open, and informative approach to obtaining consent may make it easier to approach parents.

Obtaining informed consent presents ethical and legal difficulties with certain groups of people or in situations requiring proxy consent (Burgess, Singhal, Amin, McMillan, & Devrome, 2003; Canadian Institutes of Health Research, 2005; Mason, 1997). This includes conducting neonatal research in the neonatal intensive care unit (NICU) environment with premature or ill term infants. The law acknowledges that young children and neonates are not competent to communicate their opinions concerning medical treatment and research participation, and allows proxy or surrogate consent from their parents or caregivers as a way of ensuring family values, preferences and parental discretion are all considered (Burgess et al., 2003; Canadian Institutes of Health Research, 2005; Committee on Bioethics, 1995; Cooke, 2005; Golec et al., 2004; Mason, 1997; Stenson, Becher, & McIntosh, 2004).

## 1.1 Background and Motivation

Health records are created to record and inform clinical care; this is termed the *primary use* (Sandhu, Weinstein, McKethan, & Jain, 2012). When health records are used for purposes unrelated to care delivery, including research, analysis, quality and safety measurement, payment, provider accreditation or for other commercial activities; these are termed a *secondary use* (Safran et al., 2007). Secondary use of health records plays an essential role in expanding current knowledge and understanding of treatment and health care delivery (Black, 2003; Safran et al., 2007). A body of research reports the general public's feeling of support for secondary use of their anonymized health data as long as they are: consulted as part of the process; provided with information about the research being undertaken; and given the option to approve their own participation (Damschroder et al., 2007; Kim, Joseph, & Ohno-Machado, 2015; Nair et al., 2004; Sanderson et al., 2017; Vermeulen, Schmidt, Aaronson, Kuenen, & van Leeuwen, 2009; Willison et al., 2009). This is exemplified in the outcome for secondary use of Guthrie cards collected as part of the newborn screening (NBS) program. Card reuse received negative public attention and was the subject of a successful lawsuit after unconsented release for secondary research by a public health department (Cunningham, O'Doherty, Senecal, Secko, & Avard, 2015).

We were unable to identify any prior investigations specifically assessing public perceptions for physiological data captured from infants for secondary use. This data has the potential to reveal important health knowledge about an individual, and prevent unnecessary death (Blount et al., 2010; McGregor, 2013; Pinsky & Dubrawski, 2014; Saria, Rajani, Gould, Koller, & Penn, 2013). Unless protected through vigilance and adherence, it is possible that any patient record, whether a screening blood sample or a collection of neonatal physiological data, could be exploited later in the subject infant's life (Cunningham et al., 2015; Ramshaw, 2010).

The primary motivation for this research was to address multiple limitations in existing research recruitment and consent models in areas relating to longitudinal consent. Longitudinal research requires ongoing engagement with

participants, yet relies on traditional one-time *blanket consent* to encompass potential and unknown future uses (Kaye et al., 2012; Kaye et al., 2015). We contend that blanket consent models should not be considered as true and meaningful informed consent as it is impossible to inform the participant of a future use that is unknown at the time consent was given, and users of blanket consent approaches may be reticent to approach participants a second time for fear that the person may withdraw consent (Kaye et al., 2012; Kaye et al., 2015). Information technology (IT) based participant-centered initiatives (PCI) that assist in capturing the informed consent event, with ongoing support for participant engagement and communication is an area of open research. The recently updated Australian National Health and Medical Research Council (NHMRC*) National Statement of Ethical Conduct in Human Research* identifies three types of consent around the use of data: 1) specific, 2) extended or 3) unspecified. 'Unspecified' is defined as given for the use in any future research. Unspecified needs to include agreement for data to be included in a 'databank' for future use (The National Health and Medical Research Council, The Australian Research Council and Universities Australia, & Commonwealth of Australia, 2007 (Updated 2018)).

Systemic approaches towards secondary use of medical data could include health informatics infrastructure and associated data mining tools, such as Artemis and Service based Multidimensional Temporal Data Mining Framework ($STDM^n_0$). McGregor et al. (2011) proposed a clinical decision support system that enables multi-dimensional temporal abstraction and data mining. This system is known as the $STDM^n_0$ (McGregor, Catley, & James, 2011; McGregor et al., 2012; McGregor, Smith, & Dhanoa, 2013). $STDM^n_0$ has been instantiated within the knowledge discovery component of Artemis. Artemis is a high frequency, multisource, real time, online health analytics platform developed through a collaboration between the University of Ontario Institute of Technology (Ontario Tech University) and the IBM T.J Watson Research Center (Blount et al., 2010; McGregor, 2017). This data intensive platform allows for concurrent multi-patient, multi-diagnosis and multi-stream temporal analysis in near real-time for the purpose of clinical decision

support and ongoing research (McGregor, 2013). Artemis obtains high frequency physiological data from neonatal monitors that are located at the infant's bedside along with other clinical information. The information can be simultaneously processed and returned, or stored for later analysis (McGregor, 2013, 2017; Pugh & McGregor, n.d.). Several research studies have been performed to date utilizing Artemis. Examples of neonatal research using Artemis include late onset neonatal sepsis (LONS) (McGregor, Catley, Padbury, & James, 2013), apnoea of prematurity (Thommandram, Eklund, & McGregor, 2013; Thommandram, Eklund, McGregor, Pugh, & James, 2014), anemia of prematurity (Pugh, Keir, McGregor, & James, 2013), premature infant pain (Naik, Bressan, James, & McGregor, 2013) and sleep wake cycle detection (Eklund et al., 2014).

At present, the $STDM^n_0$ framework and its instantiation within Artemis does not support flexible or ongoing consent, nor does it encourage longitudinal parent and caregiver engagement with the research process. Few research IT approaches encourage individual participation in research, or support ongoing engagement and communication between participants and researchers (Kaye et al., 2012; Kaye et al., 2015). Developments in participant-centered initiatives (PCI) and REB requirements for *patient and public involvement* (PPI) are changing this.

Two models described in the literature that seek to meet the need for flexible and ongoing consent are *dynamic consent* (DC) (Kaye et al., 2015) and *meta consent* (MC) models (Ploug & Holm, 2015a, 2016). The DC model is a PCI approach that supports participant engagement (Kaye et al., 2015). MC is any flexible consent model providing individuals an ability to choose how and when they consent to future secondary use of their retrospective or prospective biological samples or health information. PCIs can address some of the ethical and legal challenges that result from continual advancement and innovation, while ensuring the protection of research participants (Kaye et al., 2012).

### 1.2 Research Problem

The efficient use of valuable patient data requires a consent model that enables flexible, initial and longitudinal proxy consent for the secondary use of

health data. Limited research attention has investigated approaches for surrogate ongoing flexible consent for the collection of real-time physiometric data that can subsequently be used for secondary research. This thesis investigates existing and potential consent models seeking a solution that is capable of offering the functionality necessary to enable flexible and longitudinal surrogate consent.

## 1.3 Research Aim and Objectives

*Research Aim:*

This thesis investigates consent models applicable to circumstances using streaming physiological data from neonatal patients for secondary use. A secondary aim relates to the instantiation of a suitable consent model within a NICU context.

*Research Objectives:*

The research objectives are to:

1) Create a flexible multi-level participant consent model that enables changing consent preferences longitudinally and the ability for surrogate consent;
2) Enable that consent model to be integrated within database infrastructures that enable secondary use of data research studies that analyzes streams of data collected from sensors in relation to patient care;
3) Instantiate that model within a research database;
4) Demonstrate the use of that model within an ethically approved secondary use of data research clinical case study where surrogate consent is required.

## 1.4 Research Method

Constructive research is the method of choice for the research presented in this thesis. The constructive research method is a systematic approach that allows for the meaningful creation of methods, modules, techniques and tools, that have the potential to be applied in other areas well beyond the specific scenario for

which their creation was originally intended for (McGregor, 2018). This research approach is widely used in technical science, mathematics, operations analysis, and clinical medicine (Kasanen, Lukka, & Siitonen, 1993). The method seeks to produce constructions based on the use of existing knowledge in novel ways, raising the possibility of incorporation of missing links in a manner that, as represented in Figure 1, bridges knowledge gaps between academia and practice (Crnkovic, 2010; Kasanen et al., 1993; Lehtiranta, Junnonen, Kärnä, & Pekuri, 2015). The aim of constructive research is to solve a practical problem such that knowledge can be generated about how the problem can be solved, understood, explained or modeled in principle while producing an academically appreciated theoretical contribution within a specific domain (Crnkovic, 2010; Lehtiranta et al., 2015). The research should address several related knowledge problems, with regards to feasibility, improvement and innovation (Crnkovic, 2010).

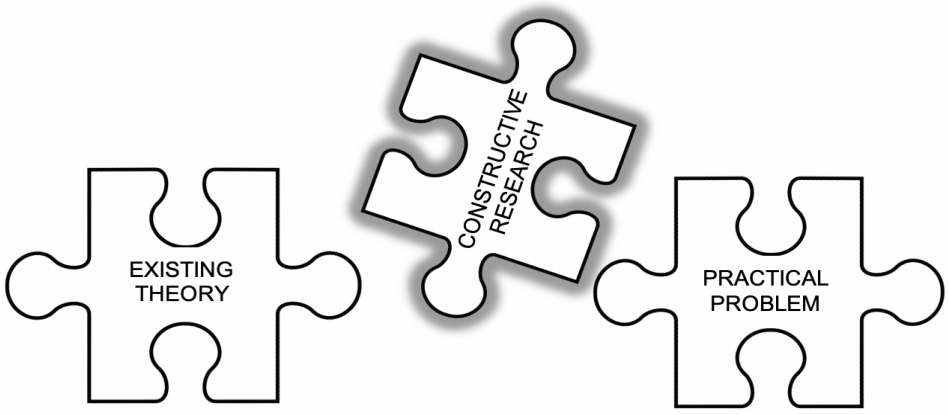


*Figure 1: Potential for a constructive research contribution.*
*Adapted from: (Lehtiranta et al., 2015).*

Constructs generally refer to entities which develop solutions to explicit problems (Kasanen et al., 1993). A construct may be practical, theoretical or both practical and theoretical in nature (Crnkovic, 2010). Examples of constructs include processes, practices, tools or artifacts such as models, diagrams, organization charts, plans, algorithms and artificial languages, system designs and software development methods (Crnkovic, 2010; Lehtiranta et al., 2015). The theoretical significance of the construct should be highlighted (Crnkovic, 2010). The development of a construction tends to contribute to the creation of a new reality

particularly in situations where it differs significantly from anything that existed prior to its creation (Kasanen et al., 1993). This allows for comparison of the new reality against the pre-existing one (Crnkovic, 2010). The constructive approach divides the research process into the following steps (Kasanen et al., 1993; Lehtiranta et al., 2015):

1. The selection of a practically relevant problem, which also has research potential
2. Obtain a general and comprehensive understanding of the topic
3. Innovate (i.e. construct one or more applicable solutions to the problem)
4. Demonstrate the solution's feasibility
5. Show the theoretical connections and the research contribution of the solution concept (i.e. linking the results back to the theory and demonstrating their practical contribution)
6. Examine the scope of applicability of the solution (i.e. examine whether the results can be generalized)

A summary of the Constructive Research Phases as they relate to research informing the *Consent of Infants for Neonatal Secondary-use research* (CoINS) model is presented in Table 1.

| Constructive Research Phases | Consent of Infants for Neonatal Secondary-use (CoINS) Constructive Research |
|---|---|
| 1) Selection of a practically relevant problem which also has research potential | Provide a new consent approach/model that offers participants and surrogate consenters adaptable and flexible consent, both initially, and longitudinally, in retrospective and prospective clinical research studies that involve the secondary use of physiological data |
| 2) Obtain a general and comprehensive understanding of the topic | • Understanding the current state of consent models used for retrospective and prospective clinical research<br>• Understanding of the current state of surrogate consent models used for retrospective and prospective medical research in the NICU context |

| | |
|---|---|
| | ● Understand the current state of secondary use of physiological data methods, systems and tools from a perspective of their integration of consent models |
| 3) Innovate | ● A flexible, multi-level participant consent model that allows for the dictation of consent preferences regarding if and how health data can be used for clinical research is proposed. In addition, this model needs to support changing consent preferences longitudinally and the ability for surrogate consent<br>● Integrate that consent model within a database infrastructure that enables secondary use of data research studies that analyzes streams of data collected from sensors in relation to patient care<br>● Instantiate that model within a research database<br>● Demonstrate the use of that model within an ethically approved secondary use of data research study where surrogate consent is required. |
| 4) Demonstrate that the solution works | ● This consent model is an extension of the current $STDM^n_0$ data model, providing surrogate consenters with the ability to dictate their consent preferences regarding if and how their infant's collected physiological data can be used for clinical research.<br>● The extension of the current $STDM^n_0$ data model has been instantiated within the Artemis platform database<br>● This new consent model and extended data model that stores the consent information has been used within a retrospective LONS case study using an instantiation of the $STDM^n_0$ data model within the Artemis database |
| 5) Show the theoretical connections and the research contribution of the solution concept | ● Contributions to health informatics in the areas of consent, and (temporal) data mining by implementing a flexible consenting construct within the current $STDM^n_0$ data model<br>● Contributions to medicine in the area of consent, medical research and (surrogate) patient engagement by providing research participants the ability to offer flexible consent<br>● Contribution to medicine through the use of this consent model in the LONS research study |
| 6) Examine the scope of | The consent model can be applicable to activities that require consent outside of the NICU context and health care domain |

| | |
|---|---|
| applicability of the solution | |

*Table 1: Constructive Research Phases in relation to the creation of the Consent of Infants for Neonatal Secondary-use research (CoINS) model*
Adapted from (McGregor, 2018)

## 1.5 Thesis Structure

This thesis is structured as follows:

- Chapter 2 presents a literature review scoped to include the secondary use of medical data with a focus on physiological data, public perspectives of such secondary use, and consent models that are currently used in medical and research practice.
- Chapter 3 introduces the NICU environment, a discussion of the value of and parental attitudes towards neonatal research, and the concept of delayed consent followed by a background of Late Onset Neonatal Sepsis (LONS), a common and serious condition affecting preterm infants.
- Chapter 4 presents a description of the present state $STDM^n_0$ framework and the extension of $STDM^n_0$ to enable flexible, longitudinal and surrogate consent as well as the ability to tie consent data with research data. This is followed by an introduction to the Artemis platform and the instantiation of the $STDM^n_0$ framework within that. Finally, the extension of the Artemis to enable flexible collection of surrogate consent is presented.
- Chapter 5 identifies the data elements that need to be captured for consent by identifying the functional requirements. This is accomplished via the inclusion of additional tables as an extension to the current $STDM^n_0$.
- Chapter 6 contains a case study that examines how the flexible, longitudinal and surrogate consent model will work for a retrospective LONS research study in the NICU environment.
- Chapter 7 presents a discussion arising from the results in chapter 6 including how specified functional requirements were addressed and met. The chapter also reviews changes to and limitations of the study as well as limitations regarding the data collection study.

- Chapter 8 reiterates the thesis contributions, identifies areas for future work and concludes the thesis.

# 2. Literature Review

This chapter presents a literature review that examines informed consent with a focus on how consent applies to the secondary use of health data, and public perceptions of such uses. Secondary use of health data requires valid informed consent. The primary focus is on medical data, specifically streaming physiological data as captured from medical devices. The insights that might be revealed in physiological data remain largely unknown, however with increased interest in the real-time analysis of physiological data there are biometrical and privacy concerns that need addressing. Background information regarding the history and value of informed consent is introduced, followed by the concept of proxy and longitudinal consent, information technology consent tools to support the need for a flexible multi-level participant consent model that enables changing consent preferences longitudinally and the ability for surrogate consent and for that consent model to be integrated within database infrastructures that enable secondary use of data research studies that analysis streams of data collected from sensors in relation to patient care is introduced. This is then followed by a comparison of existing consent models that are used in research and medical practice.

## 2.1 Identification of Literature Review

The literature review followed a three-phase approach as shown in Figure 2:

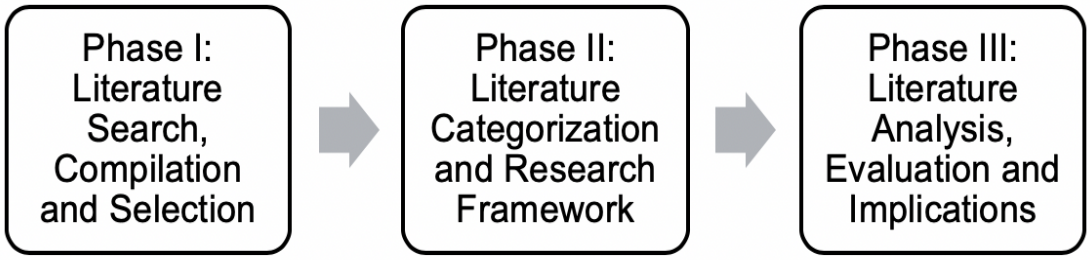


*Figure 2: Literature review method.*
*Adapted from (Yao, Chu, & Li, 2010)*

### I) Literature Search, Compilation and Selection

A search using PubMed, IEEE Xplore Digital Library, JSTOR, ScienceDirect, Springer and EBSCOhost was performed in seeking literature discussing the

issues of informed consent. First, as it applied generally to secondary use of health information for research purposes, and second, more specifically in situations of proxy, or surrogate, consent. Search terms included the following, combinations and their synonyms: consent (e.g. “proxy consent”, “surrogate consent”), consent models, health data, identifiable data, medical data, patient data, physiological data, health records, neonatal research, health research, neonatal intensive care (unit), secondary use (e.g. “secondary use of health data”) and privacy. Papers were constrained to English publications. No date constraint was applied due to the fact that many articles cited seminal literature published during at least the past century.

### II) Literature Categorization and Literature Review Framework

Literature was separated into categories at each step of the workflow process shown in Figure 3. The focus and key ethical issue of this thesis relates to the need for a flexible and longitudinal consent model that can be easily applied. Papers discussing the potential benefits and challenges relating to the implementation of novel consent models were also identified.

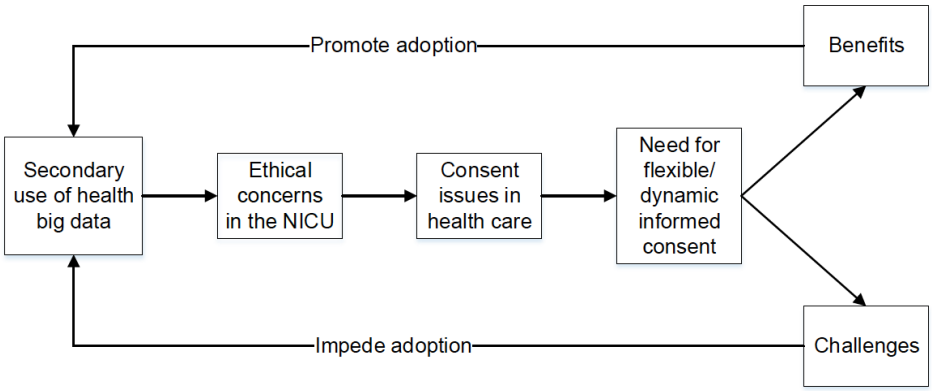


*Figure 3: Literature Review framework*
*Adapted from (Yao et al., 2010)*

### III) Literature Analysis, Evaluation and Implications

The literature review framework was applied to the literature. Categorizing topics in the previous step helped to guide identification of issues that may arise from the secondary use of patient data. It assisted in limiting the literature being reviewed to those discussing surrogate consent by caregivers for the participation

or treatment of very young infants, especially those in the NICU environment. Suggestions and implications relating to the proposed implementation of consent models along with areas of open research were also examined.

The central themes, issues and concerns identified during this analysis are illuminated and discussed in this chapter.

### 2.2 Secondary use of Medical Data

The former Article 29 Working Party (which has since been replaced by the European Data Protection Board upon enactment of the General Data Protection Regulation [GDPR]) defined personal data as *health data* when:

1. The data are inherently/clearly medical data
2. The data are raw sensor data that can be used by itself or in combination with other data to draw a conclusion about the actual health status or health risk of a person
3. Conclusions are drawn about a person's health status or health risk (irrespective of whether these conclusions are accurate, legitimate illegitimate, or otherwise adequate or inadequate)

(Article 29 Data Protection Working Party, 2015)

When raw data should be considered health data is also a matter of scale: a pedometer tracking and storing how many steps one has taken for a few days may not be "health data" but the combination of several years' worth of extensive quantified-self records of an individual (e.g. tracking sleep and exercise habits, detailed records of diet, and other vital statistics) will be considered health data. In this latter case, the conclusions and inferences, and the raw data will be considered health data (Article 29 Data Protection Working Party, 2015; Covington & Burling LLP, 2015).

A category of information which is uniformly considered as health data is the category of medical data. Health data (or all data pertaining to the health status of a data subject) is a broader term than the term "medical" (Article 29 Data Protection Working Party, 2015). Medical data is data about the patient collected at point-of-

care for the *primary use* of informing the subject's immediate and ongoing medical care (Sandhu et al., 2012). This includes all data related to patient contact, their diagnosis and/or treatment, as well as any related information on diseases, disabilities, medical history and clinical treatment that are generated within the healthcare context.

When that data is used for a purpose beyond or unrelated to the original care event, this is termed *secondary use*. Secondary use includes activities such as research, analysis, quality and safety measurement, accounting and billing, provider accreditation and numerous other non-commercial and commercial activities research and data mining by advertising companies to increase and target pharmaceutical sales (Safran et al., 2007). Personal health data used for many of these secondary purposes plays an essential role in expanding knowledge and understanding health care and service delivery via knowledge translation.

The Canadian Institutes of Health Research (CIHR) defines knowledge translation as "a dynamic and iterative process that includes synthesis, dissemination, exchange, and ethically sound application of knowledge to improve the health of Canadians, provide more effective health services and products, and strengthen the healthcare system" (Graham, 2012). The following clinical knowledge translation life cycle (Figure 4) demonstrates how aggregated patient personal health data is translated into patient health information (PHI) when the health data is given context. Once a problem is defined, the health information can be analyzed to produce knowledge related to the problem. Decisions such as treatment options are made based on available information. The outcomes of the decision(s) made are then assessed and evaluated and the cycle repeats itself. The clinical knowledge translation life cycle is depicted as a continuous life cycle as improvements are made to current understanding of medical conditions and new tools are developed to help analyze the raw patient data.

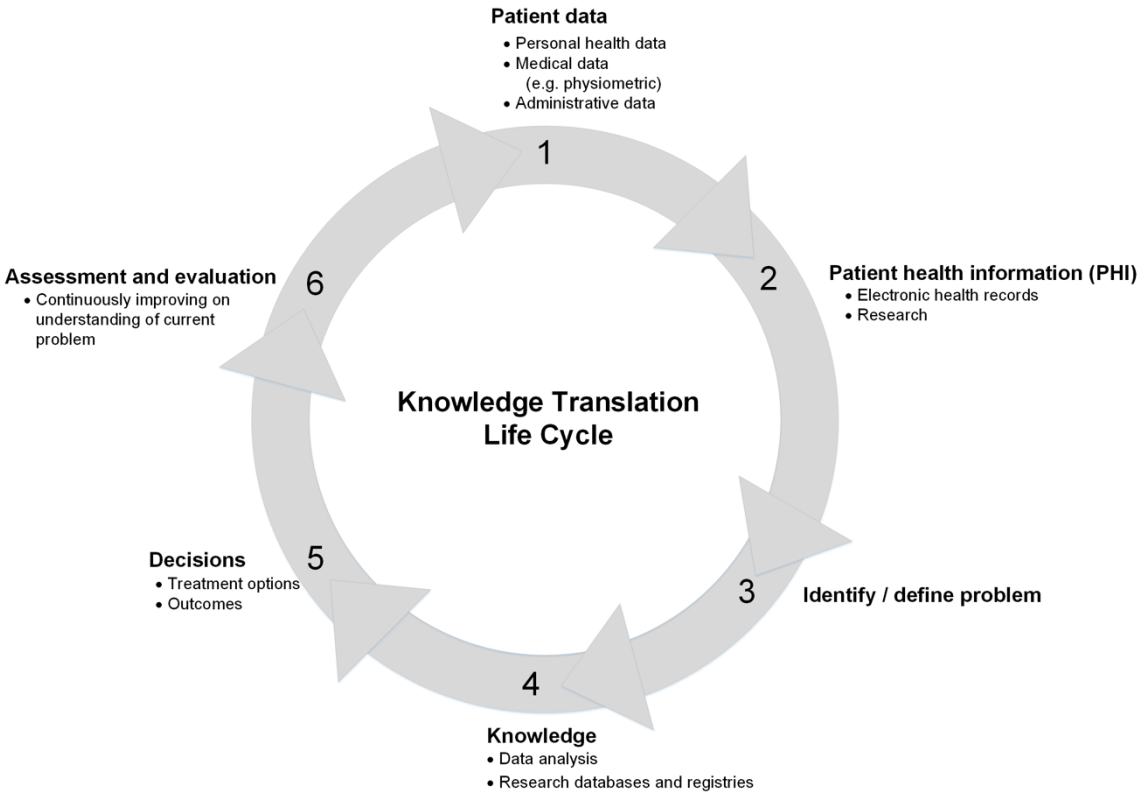


*Figure 4: Clinical knowledge translation life cycle*
*Adapted from (Deeny & Steventon, 2015; Foley & Fairmichael, 2015)*

Increased adoption of electronic health records (EHR) and health information systems (HIS) has also increased the availability and accessibility of PHI in electronic form (Grande, Mitra, Shah, Wan, & Asch, 2014; Jones, Shipman, Plaut, & Selden, 2010). PHI has substantial value in a wide range of secondary uses (Grande et al., 2014; Weiskopf et al., 2013). Some secondary uses such as disease progression research are utilitarian in nature, while others are potentially undesirable. Examples of the latter include where EHRs, some still including personal identifiers, have been made available to law enforcement for undisclosed and warrantless use in immigration and drug investigations, or debt collectors and credit reporting agencies being provided treatment records for medical indebtedness, and more recently where hackers and criminals have created false identities based on the contents of stolen EHRs for use in medical billing and pharmaceutical fraud (Humer & Finkle, 2014; Tanner, 2017).

Many accept utilitarian secondary use of EHR in research that benefits the community (Hill, Turner, Martin, & Donovan, 2013). Ploug and Holm (2017) found that very few people do not want their data used for any kind of research – in fact, many are willing to permit their data be used without specific consent, in particularly for public research. However, the majority consistently across data and research types want some control over the use of their data (Ploug & Holm, 2017). Another 2009 study examined the consent preferences of a group of breast cancer survivors (Vermeulen et al., 2009). 70% of respondents said they would appreciate the opportunity to decide whether they wish to participate in research. Genetic research was also perceived to be valuable and unproblematic and, similar to many other consent-related studies, that adult respondents support research that could benefit future patients in some way (Vermeulen et al., 2009). A dichotomy exists between what the general public considers acceptable research deserving of altruistic data sharing, and less acceptable research conducted by commercial enterprise (Hill et al., 2013; Vermeulen et al., 2009; Willison et al., 2007). Rather than objecting to issues arising from the ethics or study design, the key division is seen to arise when the commercial entity profits from the research, patenting the knowledge gained and limiting its availability for wider use (Hill et al., 2013; Vermeulen et al., 2009). Patients generally disapproved and felt greater restrictions should apply where their health information was to be used for commercial research (Hill et al., 2013), believing that profit motivation was contrary to the common good (Vermeulen et al., 2009). With regard to the secondary use of their data for commercial research and international research in comparison to public research, approximately twice as many people want to be approached for specific consent (Ploug & Holm, 2017). University researchers often exist across a spectrum where acceptability of their research depends on the funding source and level of control that funder may exert over the research conduct, results and any eventual publication (Angell, 2005; Goldacre, 2014; Hill et al., 2013; Mirowski & Van Horn, 2005).

When health data custodians receive a request to disclose individual-level health data for secondary purposes, it is their responsibility to decide if the data is

considered to be personal information or not. Should the data be considered personal information, obtaining patient consent may be necessary for disclosure. Alternatively, it may be necessary to de-identify or anonymize the data (El Emam, 2010). As a result, there is substantive evidence that explicit consent for secondary use of personal health data is required.

### 2.3 Public Perspectives Regarding Secondary Use of Medical Data

It is crucial to understand the public opinion of consenting in order to propose an improved consent model that promotes flexibility and longitudinal consent with regards to the secondary use of physiological data. Heath (2012) conducted a survey examining the Australian and Canadian public perspectives regarding secondary use of medical data. Heath noted that there are infinite combinations of citizens with fluctuating levels of 'interest' and 'capacity' when it comes to their engagement with secondary use of their medical data and privacy related matters. The survey examined public opinion concerning privacy aspects surrounding the secondary use of medical data. It was found that citizens have diverse concerns and expectations regarding privacy. It was observed that respondents of the survey can be classified into four broad groups of people described below. Thus, Heath (2012) proposed a Privacy Framework that conceptualizes the different levels of consumer engagement with each group exhibiting different expectations on all matters related to the secondary use of their medical data.

The aforementioned four broad groups were assigned with levels, ranging from levels 0-3 as shown in the consumer engagement diagram in Figure 5. Each of these groups exhibit different expectations on all matters related to the secondary use of their medical data (Heath, 2012). The **first** group (Level 0) involves persons with low or minimal capacity when it comes to engaging in making decisions regarding the secondary use of their medical data. The **second** group (Level 1) includes individuals who have sufficient capacity to participate in secondary matters but low levels of interest in becoming engaged. Over time, there may be people who develop an increased interest in secondary use of their data, if a 'trigger' was involved. This results in a **third** group (Level 2) whose triggers

may be due to a diagnosis of a medical condition in an immediate family member or result from the lingering effect of media reports, such as resale of medical data for commercial purposes. In the first case, a family member's interest level with regards to the secondary use of data may rise should there be a chance to participate in research related to a medical condition that their family member is diagnosed with, whereas in the second scenario, further interest arises as the individual tries to protect their medical data from commercial secondary uses. The **fourth** and final group (Level 3) consist of citizens who have both the capacity and interest in secondary use matters enabling them to fully be involved in the decision-making process when it comes to the secondary uses of their medical data (Heath, 2012).

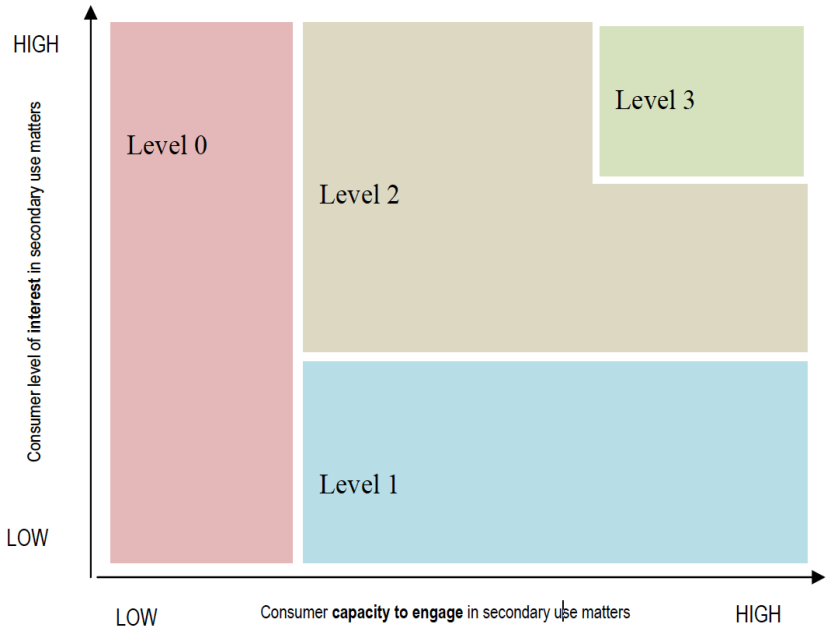


*Figure 5: Consumer engagement for secondary use of medical data (Heath, 2012)*

These conceptual groups can be static or dynamic, depending on the ability and capacity of the individual making the decision regarding the secondary use of data. A capable individual may willingly choose to join the Level 0 group in which limited opportunities would be available for consumer engagement in secondary data use matters. Conversely, if an individual were incapable of making decisions

regarding their levels of engagement, it would not matter how much interest they may express with regards to secondary use matters, as they are unable to make the fundamental decisions. Even if they had a desire to join the Level 3 group, the Privacy Framework would theoretically limit their engagement to Level 0 (Heath, 2012).

The healthcare consumer's capacity to engage determines their engagement levels which is associated with an interest level, which in turn also determines one's consent notification preference when it comes to the secondary use of their medical data. Level 1 provides an individual consumer with an opportunity to have some input into how their data is used for secondary use at a broad level. At this level, the consent is not meant to be at a project specific level but simply a broad "yes" or "no" when it comes to the use of their medical data for any kind of secondary activities. Should an individual choose to decline consent, then their medical data should not be utilized for any secondary purposes. Level 2 is for those consumers who wish to be more engaged in the research contribution process by allowing them to provide project specific consent, such as choosing if they wish for their data to be available for commercial and/or non-commercial. For consumers who voluntarily choose to allow themselves to be identified and available for contact for future secondary use matters should they meet a project's inclusion criteria, belong to Level 3. It is intended for an individual's consent preference to be stored with their electronic medical record (EMR) (Heath, 2012).

In a proof of concept study that examines the use of an electronic implementation of a *Meta Consent (MC)* model via the use of a smartphone application in Danish adult citizens where they had the ability to choose their consent preferences for future secondary research use of health data, it was found that the participants should be offered the opportunity to make MC choices given the significantly different consent preferences (Ploug & Holm, 2017). The study found that the majority of people consistently across data and research types want some control over the use of their data. When it came to the secondary use of their data for commercial research and international research in comparison to public

research, approximately twice as many want to be approached for specific consent (Ploug & Holm, 2017).

### 2.4 Physiological Data in Research

There is increased interest in the analysis of physiological data, particularly in real-time (Sahoo et al., 2014). Medical devices that continuously monitor some aspect of the patient represent the largest single sources of physiological datasets (Herland, Khoshgoftaar, & Wald, 2014; McGregor, 2013). These datasets may record many hundreds or even thousands of data points per minute (McGregor, 2013). It is already possible to capture electrocardiograms (ECGs), electroencephalograms (EEGs) from the scalp and implantable intracranial electrodes, and pulse oximetry ($SpO_2$) from medical devices (Sahoo et al., 2014). Multiple physiological variables can be computed from a single monitoring device, such as when the heart rate (HR), heart rate variability (HRV), respiration rate (RR) and chest impedance is recorded from the same ECG signal (McGregor, 2013; Pinsky & Dubrawski, 2014).

Predictive monitoring involves analysis of physiological data to detect patterns associated with critical illnesses (Fairchild, 2013). Use of physiological data for early and reliable detection has been demonstrated in neonatal populations for late onset neonatal sepsis (McGregor et al., 2012), pneumothorax (McIntosh, 2000), intraventricular haemorrhage (Fabres, Carlo, Phillips, Howard, & Ambalavanan, 2007; Tuzcu, Nas, Ulusar, Ugur, & Kaiser, 2009) and periventricular leukomalacia (Shankaran, Langer, Kazzi, Laptook, & Walsh, 2006). While healthcare professionals regularly consider physiological variables when making diagnostic decisions, at times this has become a routine activity in the absence of complete comprehension of their determinants and associations relating to the pathophysiology of many conditions. The fact that these variables may result from the interaction of numerous complicated and interrelated processes is sometimes overlooked, albeit unintentionally. Secondary use research using physiological data could investigate and deliver ways to improve detection of health conditions, advancing medical education and our

understanding of disease progression. While progression of research in the analysis of physiological data through secondary use of health data has great potential to improve health outcomes, medical data such as physiological data may reveal more about the source individual than intended by both the consenter (i.e. the patient or their proxy consenter) and the researcher(s).

### 2.5 Big Data Driven Clinical Decision Support Using Physiological Data

One of the primary contributors of Big Data in health care is the result of increased adoption of electronic health records (EHRs) (Murdoch & Detsky, 2013). Majority of EHRs capture quantitative data (e.g. laboratory values), qualitative data (e.g. text-based documents) as well as transactional data (e.g. documentation of medication delivery) (Murdoch & Detsky, 2013). Another contributor of Big Data in healthcare, comes from medical devices that continuously intake new data for purposes of monitoring a patient's current health status in real-time (Herland et al., 2014). The majority of the collected data contributes to a patient's diagnosis, prognosis and treatment based on a healthcare professional's observations and interpretation of the data available (Cios & Moore, 2002).

The use of Big Data for predictive modelling for real-time clinical decision making is increasingly recognized as an approach to achieve what the Institute for Healthcare Improvement (IHI) refers to as the Triple Aim – namely improving outcomes, enhancing patients' experiences and reducing health care costs (Amarasingham, Patzer, Huesch, Nguyen, & Xie, 2014; Institute for Healthcare Improvement, n.d.). Furthermore, the use of Big Data and its associated technologies can support the development of preventative care and personalized medicine (Murdoch & Detsky, 2013; Sahoo et al., 2014)); empower patients by delivering information directly to them, thus encouraging them to play an active role in their own health care (Murdoch & Detsky, 2013); and contribute to knowledge dissemination that help guide best practices within the literature and/or prevent information overload. To solve this issue, data from the EHR can be analyzed and the derived information can be used to design a dashboard to enable easy access to all relevant evidence and guidelines to guide clinical decision

support such that suitable treatment can be delivered to patients especially those living with multiple, complicated and chronic health conditions (Murdoch & Detsky, 2013). Six practical uses of predictive systems that have the greatest opportunities to reduce healthcare costs via the use of Big Data include dealing with high-risk and high-cost patients (along with the identification of low-risk patients), readmissions, triage, decomposition, adverse events and treatment optimization for diseases that affect multiple organ systems (e.g. chronic and/or systemic conditions) which are the costliest conditions to manage (Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014).

Akin to the concept of business intelligence, which results from the prompt interpretation of large volumes of data for actionable information, there is a growing necessity for the health care sector to adopt a similar model of “health care intelligence” in real-time (Sahoo et al., 2014) as one of the limitations impacting the effectiveness and efficiency of healthcare analytics is dependent on the time required to deliver predictions to health care providers and enable action (Bates et al., 2014). In comparison to traditional (clinical) decision support tools which relies exclusively on the use of rule-based decision trees, taking the Big Data approach means that the predictions and suggestions that contribute to clinical decisions are made from real-time data analysis (Murdoch & Detsky, 2013). Simply put, the analysis of Big Data will contribute to what the Institute of Medicine (IOM) refers to as a “learning health system” where the healthcare sector continuously strives to improve on its present-day practices and adopting new approaches such that quality care is delivered at lower costs (Adler-Milstein & Jha, 2013). A key component of this learning is the discovery of new knowledge via the secondary use of health Big Data. It is paramount that patient consent is tied with this knowledge discovery task.

### 2.6 Identifiability of Health Data

PHI datasets consist of information that can be categorized as *direct identifiers* and *indirect identifiers* (or *quasi-identifiers*). Direct identifiers are those fields that directly identify the individual patient, including names and addresses.

Indirect identifiers are those fields that can be used to infer the source patient, including dates, locations, and socio-economic information. It is possible for the source patient to be re-identified from either type, or a combination of both (El Emam, 2010). Quasi-identifiers are data elements in individual-specific data that in combination associate uniquely, or almost uniquely, to that individual and can therefore directly or indirectly re-identify the specific source individual (El Emam, 2010). Even where other direct patient identifiers (e.g. name and street address) were removed, anonymized pharmaceutical marketing data can re-identify the source individual using a relatively small number of innocuous demographics such as a postal code and the patient's age (Sweeney, 2000).

Advantages can arise from the use of patient data for health research purposes, however the secondary use of physiological data for research comes with privacy and confidentiality concerns relating to the potential biometric exposure and identifiability of the individual patient. Biometric data is personal information derived from an individual to determine or verify one's identity (International Biometric Group, 2010). Of note within the context of this thesis is that physiological biometrics are derived from direct measurements of the functioning of systems and organs within the body (International Biometric Group, 2010). The human body would not normally be considered a transmission source of privacy-exposing data (Fairclough, 2014). However, with contemporary electronic physiological measurement devices, the conversion of physiological metrics into computer recordable data has become a triviality (Fairclough, 2014). For example, research studies have demonstrated the ability to identify an individual from features in their electrocardiogram (Israel, Irvine, Cheng, Wiederhold, & Wiederhold, 2005).

Heath (2012) conducted a survey examining the Australian and Canadian public perspectives regarding secondary use of medical data. Almost all respondents (n=90.9%) either agreed or strongly agreed that they would be more likely to consent reuse of their information for research purposes if they could be assured of complete anonymity. As existing anonymization techniques have

repeatedly been drawn into question, further research into privacy-protecting techniques for PHI data disclosure is needed (El Emam et al., 2011; Ohm, 2010).

While physiological data has the potential to unintentionally reveal otherwise unknown information about an individual, many are still willing to consent use of physiological data, including those captured from their neonates, as a resource for advancement of health research (McGregor, Heath, & Choi, 2015). The continuing issue of identifiability from PHI and biometric data motivates a need to provide participants with options, including an ability to adjust or withdraw consent for further secondary use at a later date. Consequently, the incorporation of the functionalities to allow for flexibility and change of consent longitudinally is important within a consent model.

### 2.7 The History of Informed Consent

Healthcare professionals are concerned with adherence to ethical duties to inform the patient regarding their condition, status and treatment options. Their aim is to seek, and only perform, those interventions that the patient explicitly consents to (Beauchamp, 2011). This patient-empowerment approach is in response to events during the second world war and resulting Nuremberg trials, and more recent litigations in the United States of America and elsewhere. Previously, doctors were more concerned with avoiding disclosures that could potentially harm or upset patients, giving less consideration to patient's rights and whether there was approval for treatment the doctor intended to perform (Beauchamp, 2011).

Valid informed consent consists of four elements: *competence*, *understanding*, *voluntariness* and *information*. The goal is to provide a mentally competent adult (defined as an individual who is of sufficient age and mental capacity) with sufficient information, in an appropriate language such that he or she can make a free, voluntary and adequately informed decision regarding potential participation in a research study as a research subject (Allmark & Mason, 2006; Burgess et al., 2003; Canadian Institutes of Health Research, 2005; Mason, 1997; McKechnie & Gill, 2006; Nijhawan et al., 2013; Oberle, Singhal, Huber, & Burgess, 2000). This decision should be free of coercion, undue influence or

manipulation (Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council, 2018). Research subjects can withdraw their consent at any time and in the event they do, they can also request for the withdrawal of their data or human biological materials (Canadian Institutes of Health Research et al., 2018)

Informed consent does not require perfect or complete understanding. Rather, informed consent is based on the idea that autonomous decisions reflect what an individual intends to do, which is only possible if the individual adequately comprehends the relevant information (Pedroni & Pimple, 2001). When it comes to consent to participate in research, the information provided must accurately reflect the study's purpose, methods, risks, benefits, any alternative options to the research and his or her right to choose (Nijhawan et al., 2013).

*Meta consent* (MC) is a new flexible consent model proposed by the authors such that it provides individuals the ability to choose how and when they wish to provide consent for future secondary use of biological samples or health information collected in the past or those that will be collected in the future.

### 2.8 Value of Informed Consent

While Ploug and Holm (2015) identify broad consent as its own category in Table 3, they do not list blanket consent (and refusal) as its own distinct category; rather they suggest it as an option under the MC model. The difference between broad and blanket consent should be clarified. Wendler (2013) provides a clarification of the definition for the two terms based on the collection and use of human biological samples. Broad consent refers to a process by which individuals donate their samples for a broad range of future studies, subject to specified restrictions. An example of a broad consent restriction may be that if certain types of research are known to conflict with the participant's fundamental values, this could be precluded at the point of initial consent (Wendler, 2013). Blanket consent means that an individual's collected sample would be used without further consent or restrictions. Patients provide consent once and entrust researchers and ethics committees to review and approve any future research projects (Caulfield, Upshur,

& Daar, 2003; Pentz, Billot, & Wendler, 2006; Ploug & Holm, 2015c; Wendler, 2006). Much concern and debate has been raised as to whether or not broad or blanket consent should be considered valid informed consent (Caulfield, 2002; Caulfield et al., 2003; Ploug & Holm, 2015c; Sheehan, 2011). This is especially the case with blanket consents as they are necessarily vague and thus are too general to have much legal weight. They do not enable participants to act meaningfully in their continuing interest or control their health information (Caulfield, 2002; Caulfield et al., 2003). Kaye et al. (2015) also note that the expression of individual autonomy is not static and involves making choices and decisions (e.g. the decision to participate in research) over the course of one's lifetime.

Many suggest that it is always appropriate to obtain consent, even if it is done in the most general manner (e.g. via blanket consent) because even in the event an individual makes the decision to provide their consent, they may nonetheless prefer to be asked for their permission (Kass et al., 2003). Studies focusing on adult research subjects who are capable of consenting for themselves revealed that they feel respected and valued by being asked for consent and consider the opportunity to actually give consent of secondary importance (Vermeulen et al., 2009). All of these factors once again emphasize the importance of the consent process and the need for a consent model that is flexible and allows for consent preferences to be changed in the future. In the event a patient is unable to consent, surrogate and substitute decision makers both appreciate being asked for consent and having the ability to consent on behalf of the patient.

### 2.9 Proxy Consent

The informed consent process is readily actionable when it comes to engaging with an alert and competent adult. It is generally assumed that adults have decision-making capacity until proven otherwise. However, obtaining informed consent presents significant ethical and legal difficulties where certain groups of people are concerned; this includes minors (i.e. individuals under the age of eighteen), those who are mentally disabled, those of limited capacity, and otherwise competent adults who find themselves in an emergency situations

during which they have been rendered unable to consent for themselves (Canadian Paediatric Society, 2004; Committee on Bioethics, 1995; Cooke, 2005; Mason, 1997; Nijhawan et al., 2013). With these groups, a proxy is required to make decisions on their behalf (Canadian Paediatric Society, 2004).

In most situations, a family member or a loved one acts as the proxy decision maker. There are two proxy types: *substitute* and *surrogate*. *Substitute decision makers* are individuals who know the patient well enough to have already discussed their wishes for care in such situations with them (Canadian Paediatric Society, 2004). The substitute decision maker is an adult, legally able and competent to consent on the patient's behalf. Depending on the circumstances, this adult may be designated through enduring power of attorney, or via a living will (Richards, 1993). A substitute decision maker's role is to promote the patient's expressed wishes (Canadian Paediatric Society, 2004). In other situations, the proxy consenter may not know what the patient would want done, but are tasked with the responsibility of acting (i.e. making decisions) in their best interests, in which case they become a *surrogate decision maker* (SDM) (Canadian Paediatric Society, 2004).

When the situation concerns a minor, parents generally act as SDMs, unless the child or adolescent (i.e. mature minor) demonstrates the decision-making capacity and is accepted to have the capacity to express their own wishes (Canadian Paediatric Society, 2004). Factors affecting surrogate parental decision-making in the NICU environment are discussed later in chapter 3.

This demonstrates a clear need for a consent model to enable consent to be provided by a substitute or surrogate consenter and in so doing providing the ability to record the details of the person providing consent. Within this research whether the concept of substitute and or surrogate is reflected in the use of the term surrogate though the same functionality supports the notion of a substitute consenter in the case that the study population are adults.

## 2.10 Longitudinal Consent

A longitudinal study collects information from the same set of research subjects at multiple points over time and are often conducted over an extended period of time with varying timeframes (Wood et al., 2014). Longitudinal data as they provide the opportunity for researchers to study trends throughout individual lifetimes or generations. Similarly, longitudinal consent within the context of this thesis is used to refer to requesting permission for the secondary use of collected data (or samples) over an extended period of time for a research project (or projects) with no specified end date while allowing for ongoing dynamic engagement during the process.

It must be considered whether a MC or DC model (the latter as presented by Kaye et al. (2015)) could potentially be applied to secondary uses of physiological data that can occur several years after the initial data collection. In comparison to the DC model, the MC model offers a unique feature that allows an individual to choose how they wish to provide consent for future secondary research of health data collected in the past, or of data that will be stored in the future (i.e. MC is both retrospective and prospective). Conversely, DC would only apply to future use of data and samples (Ploug & Holm, 2015a, 2015d). Hence, MC is better suited in a situation that calls for a longitudinal and flexible consent model. As the case study context for this thesis is based on the NICU, this would be the ideal opportunity to set up a MC model, gradually introducing this concept to two population groups: parents of the neonate patient, and the neonates themselves when they attain adulthood under the assumption they can demonstrate competency. As described previously, parents or legal guardians would generally be those providing surrogate consent. Thus, the consent model of choice for this thesis is the MC model. However, as there is limited literature regarding the MC model and its implementation in this manner, the following review regarding IT consent tools is drawn from papers that examine the DC model. This is due to the general similarity of these two consent models.

Initially developed in biobanking, DC refers to both a specific project as well as a general concept and approach enabling individuals to be engaged regarding

use of their personal information. It is also an interactive personalized interface that provides participants with the ability to choose their engagement levels and consent preferences in real time (Kaye et al., 2015).

The concept and approach serve as the focus of this thesis wherever the term DC appears in this thesis. DC is a participant-centered initiative (PCI) that places patients and participants in the center of decision making (Kaye et al., 2015; Oxford University Innovation, n.d.). Kaye et al. (2015) explicitly note that DC should not be seen as a replacement for: 1) existing consent models (e.g. broad consent), 2) face to face human contact at the initial point of consent, or 3) the discussion process regarding re-consent; instead, such models should be seen as "a facilitation tool to improve how that consent is obtained, understood and acted upon".

Longitudinal biobanks have recognized the limitations of one-off static consent. They have also recognized that there are often multiple researchers and multiple projects, making it challenging if not impossible to obtain informed consent for all future secondary uses of collected samples and data. Re-consenting is costly in terms of time, effort and resources and it may also be impossible to locate individuals, leading to high drop-out rates (Kaye et al., 2015). This is compounded by the expectation that research studies are traditionally hypothesis-driven, and does not take into consideration hypothesis-free research, where the research question(s), purpose, hypothesis, duration, and/or end points of the study may not yet be defined (Bates et al., 2014; Kaye et al., 2015). This is the case in ever-evolving research fields such as data mining studies, that are longitudinal in nature (Bates et al., 2014).

Similar to the acceptance of "patient-centric" approaches in health care in recent years, the adoption of "user-centric" approaches have been rapidly gaining momentum given the growth in data sharing related activities (e.g. research), leading to the development of various participant-centered initiatives (PCIs). PCIs are defined as "tools, programs and projects that empower participants to engage in the research process and, in many cases, can differentiate between a range of diverse preferences and needs" via the use of IT (Anderson, Bragg, Hartzler, &

Edwards, 2012; Kaye et al., 2012). While PCIs are each different in their own ways, all of them support the idea of placing research participants at the center of decision making via the use of an interactive IT interface to promote the idea of treating participants as active and equal partners within the research process. This approach, based on the principles of empowerment and respect of individuals, reflects the transforming mindsets towards privacy and individual involvement (Kaye et al., 2012).

### 2.11 Information Technology Consent Tools

The use of interactive IT tools for consent have great potential to support easy consent interactions over time, enabling personalized, flexible and longitudinal engagement. In comparison to traditional paper-based documentation of consent, the consent process is no longer "locked in time" and can be configured to run on a variety of IT platforms (e.g. website, tablets and mobile phones) (Kaye et al., 2015). All aspects of the interface can be customized to an individual's needs and preferences (Kaye et al., 2015). Participants can be provided with the ability to consent to new projects or change their consent choices. They can select their level of engagement, the types of information and projects they are primarily interested in receiving, how they prefer to be contacted and the frequency they wish to be contacted (Kaye et al., 2015). All these changes can be performed in real-time and individuals can have confidence that any changes in their choices will be effective immediately.

Kaye et al. (2014) provide a case study specific example of the enabling of this approach by demonstrating that specific consent provisions travel with a participant or donor's data and samples as they are shared or accessed for different purposes. These consent provisions can be electronically and cryptographically "wrapped" with the donor's samples and information (Kaye et al., 2015). All participant information (including one's consent preferences) are conveniently stored within an accessible interface.

An IT platform could also enable general research results to be returned to participants according to their preferences, either as a simple 'thank you'

acknowledgment for their contribution and involvement or by informing them how their samples and information have been used (Kaye et al., 2015). This approach encourages better communication and transparency between researchers and participants.

A flexible, multi-level consent model that ties consent data and collected study data from consenting individuals (or by proxy consenters) within a research database would allow for the aforementioned features of customizability in an initial and longitudinal manner and although outside the scope of this thesis, the possibility of informing contributors how their samples and information will be used in the future.

While the work of Kaye et al. (2015) demonstrates the potential benefits of coupling consent with the study data, they have not provided a systemic approach to support multiple studies.

### 2.12 Comparison of Consent Models

Many models are available for consenting patients in clinical practice and clinical research. Those identified during this work are listed in the furthest left column in Appendix A, which provides a brief description as well as advantages and disadvantages of the identified consent model. Consent models were found to be fundamentally based on the concept of obtaining a one-time broad consent. Addressing the research objectives of this thesis required a consent model that allows for functionalities of flexible longitudinal surrogate consent, with the ability to be coupled with data and employing the use of IT tools. Below, Table 2 provides the rationale behind the specific functionalities of interests:

| Functional Requirement | Reason |
|---|---|
| **Flexibility** | Allow SDM(s) to:<br>1) choose their consent preferences<br>2) adjust their chosen consent preferences at any time |
| **Surrogate consent** | Neonates are incompetent - parents or legal guardians often act as their neonate's SDMs |
| **Longitudinal consent** | Allow SDM(s) to:<br>1) choose their consent preferences<br>2) adjust their chosen consent preferences at any time and<br>3) ensure chosen consent preferences take effect immediately |
| **Coupled with data** | Ensure that consent data is coupled with research study data (i.e. collected PHI) |
| **Information Technology (IT)** | Enable consent model to be integrated within database infrastructures that enable secondary use of data research studies that analyze streams of data collected from sensors in relation to patient care |

*Table 2: A table depicting the various functionalities required and their reasons for including the functionalities in the proposed consent model*

To identify which existing consent models address the research objectives, Appendix A compares the consent models based on their ability to allow for the functionalities specified above in Table 2. Appendix A also provides a description as well as the advantages and disadvantages of the model, From the left to the right columns, the table is read as follows: the consent model name, the model's description, and advantages and disadvantages, with a plus sign (+) indicating the advantages while a negative sign (-) indicates a disadvantage of the consent model. An "X" in a column that identifies a functionality indicates that functionality is met by the specified consent model.

It is noted that the majority of existing consent models are geared towards primary care and may not address the needs of secondary usage of biological and genetic samples, or patient data. These consent models are not designed for longitudinal engagement and do not provide individuals, specifically proxy consenters (e.g. surrogate consenters) with the flexibility to customize their consent preferences nor do they employ IT tools.

Ploug and Holm (2015) classified informed consent models into four broad categories shown in Table 3, which includes informed, broad, dynamic and meta consent. Of particular interest to this thesis are the dynamic consent (DC) and meta consent (MC) models. *Meta consent* (MC) is a new flexible consent model proposed by the authors such that it provides individuals the ability to choose how and when they wish to provide consent for future secondary use of biological samples or health information collected in the past or those that will be collected in the future. The MC model is formally defined as any consent model that allows an individual to explicitly express their personal preference for how and when to provide consent (i.e. to design future consent requests). Ploug and Holm (2015a, 2016) note that by allowing an individual to express a preference for how and when to provide consent, one can be said to be providing consent on a meta level.

An individual can choose their preferred consent model, whether broad, dynamic, blanket consent or refusal for different types of research (Ploug & Holm, 2015b). In the event that a future research study falls under more than one category, the most restrictive consent choice would apply (Ploug & Holm, 2015a).

MC consists of the following six functional elements:

1) A limited number of categories for designing consent requests
2) A key for the prioritization of consent requests
3) A definition of the time for providing MC
4) A default setting if MC is not provided
5) A scheme for redesigning MC
6) A method and infrastructure for requesting and recording MC

Ploug & Holm (2016)

| Consent model | Focus | Use of collected samples and/or data | Issues |
|---|---|---|---|
| **Informed Consent** | Single use/particular project | Known and disclosed use at time of consent | Additional consents may be difficult or impractical |
| **Broad Consent** | All projects of particular type | Not all future uses predictable at time of consent | Future uses do not constitute valid informed consent<br><br>No regard to future uses when consenting |
| **Dynamic Consent (DC)** | Single use/particular project<br><br>Additional consents may be difficult or impractical | Known and disclosed use at time of consent | Not all potential participants will connect and opt in<br><br>Not practical for research using routine clinical data<br><br>Every request requires description of entire project; May result in consent fatigue |
| **Meta Consent (MC)** | Retrospective and prospective, participant can select broad or dynamic on a case by case basis | Retrospective = known, prospective = unknown | Original authors foresee no issues as it is their proposed model<br><br>Others have pointed out that it is possible that the costs and burdens of such a framework may be much higher than asking participants to reconsent<br><br>Care will need to be taken to avoid false positives where those who chose broad consent are contacted by mistake<br><br>Problem of what to do when participants fail to reply<br><br>Administrative systems need to be in place to track, and accurately respond to, individual fine-tuned choices |

*Table 3: A brief overview of the focus, use of collected samples and/or data and issues associated with the informed consent, broad consent, dynamic consent and meta consent models.*
Information content adapted from: (Canadian Institutes of Health Research, 2005; Kaye et al., 2015; Manson, 2019; Ploug & Holm, 2015a, 2015d)

Table 4 serves two purposes: first, it provides a comparison of the DC and MC models as described by different sources, and second, it evaluates whether the consent model meets the functional requirements of this thesis. These include whether it possesses elements for longitudinal engagement, flexibility and surrogate consent. These requirements are appropriate to the secondary use of physiological data in the NICU context: Longitudinal engagement and flexibility features allow the consenter to alter their engagement level and consent preferences as circumstances change over a long or undefined period of time. Inclusion of surrogate consent is necessary given that the target patient group in this case are neonates, who would be unable to express opinions with regards to secondary use of their own physiological data.

| Authors | Dynamic Consent (DC) | Meta Consent (MC) | Longitudinal Engagement | Flexibility | Surrogate Consent |
|---|---|---|---|---|---|
| Ploug & Holm (2015a) | • Participant centred, model: information about specific secondary use of health data or tissue and a request for consent is put to the individual via a web-based platform<br>• Protects the participant's autonomy by providing information about each new data application<br>• Makes provision or withholding of consent easy and flexible<br>• Would be asked for specific consent for every new research project in that category that uses their data | • Combines the broad and dynamic models, with additional options for blanket consent and blanket refusal<br>• One can choose how they wish to provide consent for future secondary research of data collected in the past or of data that will be stored in the future (i.e. MC is both retrospective and prospective)<br>• Individuals can choose the type of consent: DC, broad consent, blanket consent/refusal for different types of research | DC, MC | DC, MC | |
| Kaye et al. (2015) | • Personalised, communication interface to enable greater participant engagement in clinical and research activities<br>• Participant-centred initiative (PCI): places patients and research participants at the centre of decision making<br>• Provides an interactive IT interface to engage with participants<br>• Allows interactions over time; enables participants to consent to new projects/alter their consent choices in real time<br>• All aspects of the interface can be customized<br>• Individuals could provide different types of consent depending upon the kind of study | | DC | DC | |

| | | | | | |
|---|---|---|---|---|---|
| | ● Participants could choose to consent to: a broad range of uses of their samples and data, opt to be approached on a case-by-case basis, set different preferences for different types of research (i.e. such preferences can be 'opt ins' or 'opt outs'),to give a broad consent or can tailor their profiles to receive no information for specified periods of time<br>● Allows general research results to be returned to participants according to their preferences, either as a simple 'thank you' acknowledgment for their contribution and involvement or by informing them how their samples and information have been used | | | | |
| Ploug & Holm (2015d) | ● A matter of choosing how to be and stay informed and provide consent to research participation<br>● Normally offered within the context of a specific project<br>● May incorporate broad consent but not on the basis of an individual's preference for this type of consent; only if "a public engagement strategy has identified that people were content with consenting to broad types of data"<br>● Presents an individual with an overview of types of data and types of research for which different types of consent may be requested<br>● Does not introduce consent defaults<br>● Does not protect:<br>▪ 1) individual preferences<br>▪ 2) against the routinization of consent<br>▪ 3) the ability to conduct research | ● Unlike DC, MC protects an individual's preference for participation and an individual's preference for how and when to provide consent<br>● By providing an individual with an overview of different type of research and research contexts, MC allows the individuals to reflect on the consistency of their preferences concerning participation in research in general. DC does not offer such possibilities.<br>● May limit routinization by allowing an individual to limit the number of consent requests through broad or blanket consent/refusal to predefined broad types of research and predefined types of research context | DC, MC | DC, MC | |

| | | | | | |
|---|---|---|---|---|---|
| | when 1 and 2 are adequately protected | | | | |
| Oxford University Innovation (n.d.) | ● Participant-centric approach provides an interface that ensures patients can give, review and change their consent preferences while allowing two-way interactions between researchers and participants | | | DC | |

*Table 4: Comparison of definitions for dynamic consent and meta consent models by different authors (Kaye et al., 2015; Oxford University Innovation, n.d.; Ploug & Holm, 2015a, 2015d)*

Despite differing views of DC as presented by Ploug and Holm (2015) and Kaye et al. (2015), Kaye et al. (2015)'s version of DC seems to incorporate elements of both Ploug and Holm (2015)'s DC and MC models. In fact, in a later publication, Ploug and Holm (2016) openly acknowledges that the Kaye et al. (2015)'s DC model is extremely similar to their proposed MC model with regards to:

- The reliance on modern IT (Ploug & Holm, 2016)
- The underlying concept of providing individuals with a longitudinal consent model that provides them with the flexibility to change their consent and engagement preferences regarding the future use of their health data and samples (Ploug & Holm, 2016)
- Enable the aforementioned preferences to be communicated to researchers (Ploug & Holm, 2016)

There are also visible differences between the two consent models. The initial motivation for the development of the two consent models is different. DC is "normally offered within the context of a specific project" (Kaye et al., 2015; Ploug & Holm, 2015a) and was initially developed to solve consent issues in the biobanking field (Kaye et al., 2015; Ploug & Holm, 2016). Its primary focus is on the matter of choosing how to be and how to remain informed and provide consent to research participation (Ploug & Holm, 2016). In comparison, MC was developed with the goal of handling the consent preferences for a whole population for all kinds of data and biological samples for many types of research contexts in a prospective or retrospective manner (Ploug & Holm, 2015a, 2016). Given that medical research is likely to require the involvement whole population datasets in the future, it is fundamental that the MC model engage with every single citizen in their role as potential participants in big data research (Ploug & Holm, 2016). Due to the possibilities of linkage it is also imperative that researchers know if they can use a particular piece of data within a dataset that has not yet been previously used for research (Ploug & Holm, 2016). This is determined through a definitive answer provided by the individual's pre-set instructions and preferences. The MC

model enables this as it is designed to allow individuals design future consent requests on the basis of predefined types of consent, data, and contexts (Ploug & Holm, 2016). In contrast to DC, MC may introduce a default consent setting if individuals do not make consent choices themselves (Ploug & Holm, 2016). Unlike DC, MC is designed to take into consideration that research may re-use data that has been collected at various times, for different purposes and under different consent regimes, and is meant to protect: 1) individual preferences, 2) against the routinization of consent and 3) the ability to conduct research when 1 and 2 are adequately protected (Ploug & Holm, 2015b). Ploug and Holm (2016) point out that advocates of DC model could incorporate this meta aspect of the meta consent model, and thereby make their model substantially identical to the MC model. But unless and until they do so, the two consent models remain distinct (Ploug & Holm, 2016). Although it is far from obvious that 'meta choices' of this kind works in practice, Ploug and Holm (2017) later go on to report and analyze the results of a proof of concept implementation of MC as a front end application for smartphones and tablets to determine whether MC preferences can be successfully elicited in the adult Danish population via a smartphone application. Results indicated that very few people do not want their data used for any kind of research, and many are willing to allow their data be used without specific consent, especially when it comes to contributing to public research. The majority of respondents consistently across data and research types wanted some control over the use of their data. The study also revealed that people have significantly different consent preferences, thus emphasizing the requirement for a nuanced consent system such as MC (Ploug & Holm, 2017).

Ploug and Holm (2015) also suggest some valid points as well as practical implementation tips regarding their MC model that could also be applicable to the DC model. For example, while neither of the models explicitly meet the functionality for surrogate consent, they proposed that the MC model be arranged early in life which could be achieved by making it mandatory as an individual reaches the age of majority (in their jurisdiction), referring to their legal age for decision-making. Another option is that parents could provide surrogate MC for minors. They

suggested that failure to do so can be managed in different ways, such as via reminders from healthcare professionals, or perhaps a broad consent model can be applied as a default (Ploug & Holm, 2015a, 2015c). In essence, they also take into consideration the concept of proxy consent which the DC model or recommendations for its implementation does not explicitly mention. Hence this thesis addresses the gap and currently unaddressed needs by proposing a multi-level participant consent model within a research database that allows surrogate consenters the flexibility to modify their consent preferences longitudinally while ensuring that consent preference data is linked with the research data within the database.

### 2.13 Conclusions and Implications on Research

The secondary use of health data, such as real-time analysis of physiological data for research purposes requires informed consent. Despite the many available consent models that have been proposed, the majority of the traditional models in current practice are based on broad consent and are generally meant for a one-time use of the collected data or sample. The recent introduction of DC and MC are proposed as flexible models that may be appropriate for longitudinal research with secondary use data as they provide participants with the flexibility to choose and adjust their level of engagement and consent preferences via the use of interactive web-based IT tools and can be applicable in situations where proxy consent such as surrogate consent, is necessary. Although both models offer similar features, upon evaluation of both models, the concept of the MC model is used as the basis of the consent model for this thesis given its ability to provide additional flexibility by offering retrospective and prospective consent.

# 3. Literature Review – Neonatal Intensive Care Context

This chapter provides contextual information for the case study context for this research, namely that of neonatal intensive care. It begins with an overview to the physiological data continuously being generated, followed by a discussion of current approaches to, and the need for neonatal research. As parents are often the surrogate decision makers (SDMs) on behalf of their neonate, it is also important to understand their attitudes towards neonatal research and how this impacts their engagement with the consent process and research generally. The significance of delayed consent is also examined. The chapter then introduces the topics of late onset neonatal sepsis (LONS), a serious condition that can have devastating effects on the neonatal population and how heart rate variability (HRV) research could assist with providing earlier detection of LONS are provided. To emphasize and demonstrate the need for a flexible multi-level participant consent model that enables changing consent preferences longitudinally and the ability for surrogate consent, the condition will serve as a clinical case study in chapter 4.

This chapter motivates the need for consent models enabling longitudinal consent and consent performed by proxy, specifically parental surrogate consent. The chapter concludes with a summation of the functional requirements within a consent model that are necessary to incorporate these consent functional components.

## 3.1 The Neonatal Intensive Care Unit Environment

The NICU provides medical care for premature and ill term infant patients, collectively known as *neonates*. Neonates include newborn infants and infants up to twenty-eight days of age inclusive (Ligi, Boubred, Grandvuillemin, & Simeoni, 2011). This crucial transitional period is characterized by the physiological immaturity of many body organs and systems (Ligi et al., 2011). Throughout the entire human lifespan, it is also during this period that the infant is at the highest risk of dying (World Health Organization, n.d.). It is imperative to improve the infant's chances of survival and to lay the foundations for a healthy life. The NICU

is generally considered an *intensive* or *critical care* unit, that is, a hospital ward where the staff to patient ratio is maintained much lower to allow for near one-to-one care of those with severe or life threatening conditions (Lang, Hodge, Olson, Romano, & Kravitz, 2004; McGregor, 2013; Momtahan, Hetu, & Tansley, 1993; Rothschild et al., 2005; Tarnow-Mordi, Hau, Warden, & Shearer, 2000).

Within the NICU, a variety of medical equipment is used to support care of the neonate, monitoring vital signs, performing or assisting with breathing, maintaining body temperature, or delivering necessary drugs and nutrients to the child (Donchin & Seagull, 2002; Make et al., 1998; McGregor, 2013; Momtahan, Hetu, & Tansley, 1993). Many of these devices continuously collect physiological or other clinical data streams that fluctuate on a second-by-second basis that caregivers are ultimately responsible for translating into actionable information (Donchin & Seagull, 2002; Make et al., 1998; McGregor, 2013; Momtahan et al., 1993). These devices produce audible and visual alerts when a measurement produced by the neonate breaches that of the devices' pre-set standardized population thresholds, potentially suggesting an abnormality in one's health condition (McGregor, 2013; Momtahan et al., 1993). Considering that many of these physiological and other data streams produce thousands of readings per minute per day for a single patient (McGregor, 2013), the data generated in critical care units is considered a Big Data problem. As noted in the introduction chapter, there is great potential for these continuous real-time physiological and other data streams to provide earlier recognition of the deteriorating patient. However, research is required to explore this which requires consent.

### 3.2 Surrogate Consent

Research in minors is permitted only if the child is exposed to no more than minimal risk and has the opportunity to benefit from the study (Mason, 1997). Consent is required regardless of whether a study is assessed as minimal risk, in order to preserve patient autonomy. Even data sharing for secondary research has potential for privacy and security breaches or incidental findings. Advising the patient or consenter of risk is a necessary component to securing informed

consent. Neonates, lacking competency, represent a unique group of research subjects (Ballard, Shook, Desai, & Anand, 2004). As they are obviously unable to provide consent on behalf of themselves, a surrogate decision maker must decide by proxy.

Since parents/legal guardians are traditionally and legally allowed to provide consent for their neonate's medical care under the assumption that they have their neonate's best interests at heart and act accordingly, it is only appropriate that they are typically the ones who researchers and health care professionals turn to when they seek permission to enrol the neonate into research (Ballard et al., 2004; Burgess et al., 2003; Cooke, 2005; McKechnie & Gill, 2006). Parents are obliged to undertake the responsibility of making decisions on behalf of their infant; a concept described in the legal domain as *surrogate consent* (Cohen, Trentalange, & Fried, 2015; Flaherty, 2017). However, the proxy consent at best represents parental discretion, preferences and family values (Cooke, 2005; Mason, 1997). Although pediatric researchers generally agree that parents should act in the role of SDM, opinion is divided with some feeling that when emotional investiture and other issues are factored in, parents may not actually be the most suitable decision maker (Canadian Paediatric Society, 2004; Golec et al., 2004; Oberle et al., 2000). Other issues can include: 1) the parental lack of decision-making capacity, 2) irresolvable differences between the parents with respect to the minor's care, 3) when parents have evidently relinquished responsibility for the minor, or 4) another legal guardian has been appointed (Canadian Paediatric Society, 2004). As a result, a consent model needs to enable surrogate consent, even if a parent is not the individual consenting.

### 3.3 Parental Attitudes Towards Neonatal Research

It has been demonstrated that many parents are very willing to enrol their neonate in research studies (Ballard et al., 2004; Stenson et al., 2004) even if it is known that there are significant gaps in knowledge about the study (Ballard et al., 2004). Many would also be willing to enrol their neonate again, if presented with

the opportunity (Ballard et al., 2004). This is further supported by a study conducted by Morley et al. (2005).

Parental opinions regarding the enrolment of their premature neonate(s) into several research studies in the days following birth were examined via the use of a questionnaire. Parents of preterm infants in the NICU who were invited to participate in two or more research studies were approached with this survey. Amongst the 92 invited participants who completed the questionnaire, 10% declined to allow their infants to join any studies. The majority of parents were willing to have their infant(s) be enrolled in multiple studies. 78% of these parents were willing to give their permission to enrol their infant(s) in two or more than studies, 58% willing to consent for three or more studies and 20% were willing to have their infant(s) participate in more than ten studies (Morley, Lau, Davis, & Morse, 2005). Many parents choose to consent for their neonate's participation in research studies because they are hopeful that it would somehow benefit their infant (Ballard et al., 2004) while contributing to the advancement of health research. According to findings by Morley et al. (2005), 94% of the parents thought that if their baby joined a research study, the care of infants in the future would either be "better" or "very much better". Parental altruism was further demonstrated when parents were asked, "Who will benefit from these studies?" in which 91% responded that "future babies", 67% said "researchers", 25% mentioned "my baby" and 2% said "no one".

It has been suggested that the above findings may be related to the limited parental understanding of risk associated with participation in a study (Harth, Johnstone, & Thong, 1992), parents who consent for their newborn's participation may be emotionally vulnerable and socially disadvantaged in comparison to those who refuse to consent (Harth et al., 1992), or it could simply be a reflection of parents' trust in physicians and the respect they have for physicians' opinions (Singhal, Oberle, Burgess, & Huber-Okrainec, 2002). A MC model would provide parents with the flexibility to review the study information and re-evaluate their consent preferences at their convenience and perhaps in a setting away from the hospital, such as in the comfort and privacy of their own home.

In terms of who should be the ones consenting, most parents felt that they should be the ones consenting (Singhal et al., 2002). In Singhal et al. (2002)'s study, it was found that 90.1% of parents of NICU newborns and 91.1% of parents of normal newborns respectively believed that ‘‘all forms of research with babies, no matter how minor, should be carried out only after parents have given informed consent’’. A mere 7.0% of NICU parents and 3.9% of parents of healthy newborns felt that, ‘‘doctors should make the decisions about which babies should be in research; I do not think the parents should have to make a decision’’. This was further supported by Burgess et al. (2003) as they found that 93% of parents in a retrospective study and 91% in a prospective study were against the option that a doctor rather than the parent should make the decision if a newborn should be enrolled into a study. Stenson et al. (2004) found that 83% of parents would be ‘‘unhappy to forgo the consent process for trials passed by the institutional ethics committee’’. Singhal et al. (2002) also found that parents did not want their newborn to be enrolled in studies without their consent, even if it was considered to be low risk. Such data indicates that parents expect to be asked for consent to allow for their infant's participation in research and do not feel it should be left to the discretion of physicians.

Korotchikova et al. (2010) provides a further perspective supporting the need for flexible longitudinal consent when they found that it was most common for mothers to be approached for consent alone (76%). However, in situations where mothers originally agreed to provide consent, half later chose to withdraw after discussion with their husbands/partners (Korotchikova, Boylan, Dempsey, & Ryan, 2010). The investigation revealed that presence of both parents during the consent process in the early postnatal period increases a positive response and makes obtaining consent more likely. In a recent European Delphi survey of parent representatives and clinicians that sought to establish a consensus between parent representatives of neonatal associations and healthcare professionals concerning the study information deemed essential by both parties to improve the recruitment of neonates into clinical trials, it was found that information should be provided to both parents at the same time (Neyro, Elie, Thiele, & Jacqz-Aigrain,

2018). However, clinicians agreed, in both Delphi rounds that information could be provided to only one of the parent if the other was not available but emphasized that the other parent should be informed as soon as possible. Parent representatives disagreed with the suggestion that only one parent could be informed (Neyro et al., 2018). Consequently, a flexible consent model that can be accessed at a convenient location and time while providing the ability to modify one's consent preferences after making a decision after further consultation with trusted individuals.

Even though many studies found that the majority of parents felt that they should be the ones making the final decision with regards to providing consent (Singhal et al., 2002; Zupancic, Gillie, Streiner, Watts, & Schmidt, 1997), studies have also shown that parents want their infant's physicians to advise them on study enrolment. The influence physicians can have on parents raises the notion of shared decision making where physicians generally provide the facts and parents supply the value (Golec et al., 2004). Parent representatives of neonatal associations and healthcare professionals were also in favour of having a third person present during the informed consent process, such as a family member or their family doctor (Neyro et al., 2018) but healthcare professionals did not share this view (Neyro et al., 2018). Other parents have suggested that that a consultant be present during all meetings related to neonatal clinical trial participation between parents and clinicians with the amount of consultant support provided during the consenting process varying in respect to the type of research to which parents are being asked to consent their newborn to (O'Shea, Doran, Ryan, & Dempsey, 2018). While shared decision-making appears to be a favourable solution in helping to reduce the parental burden while helping to maintain parental autonomy and family individuality during a stressful time, in reality it can result in more ethical problems (Golec et al., 2004). Over time, parents' opinions may become less dependent on the physician though and results in the need to enable changing consent over time.

A MC model would allow parents to consult their neonate's healthcare professionals, and then be provided with an opportunity to rethink about their

discussions and decisions outside of the hospital even if they had previously decided regarding enrollment in research. It also allows parents to choose to make a decision in the privacy of home without on-the-spot influence from anyone or if they prefer, the chance to reach out to other people for support if they so wish, such as clergy, neighbors, relatives, (formal) support groups or even contact with other families who had had a high-risk infant or a multiple birth (Pinch & Spielman, 1990).

### 3.4 Delayed Consent

Excluding situations of emergency research, in the event that a parent is to be approached for consent for more than one non-urgent study, Golec et al. (2004) recommended that such requests should be made 48-hour apart although McKechnie & Gill (2006) specifically points out that there is an absence of evidence to support their recommendation (McKechnie & Gill, 2006). It is possible that Golec et al. (2004)'s suggestion of delayed consent may have been based on the results of studies by Loue et al. (1996), Bosk (2002) and Kuczewski and Marshall (2002). The idea of delayed consent arose from a workshop that took place in Uganda to examine the system of bioethical principles governing biomedical research and clinical trials conducted in the country. Of importance between Western and Ugandan cultural differences was the need to involve family members in the decision making and consent process for Ugandan citizens. Participants at the workshop suggested a waiting period of 48 hours between the time individuals were approached about participation in research and the time they make the decision to sign the consent form (Loue, Okello, & Kawuma, 1996). Although Loue et al. (1996)'s study is not based on the neonatal context, Golec et al. (2004) suggested that the findings regarding the importance of respecting cultural influences can be applicable to the multicultural settings of many NICU environments. Other researchers such as Bosk (2002) and Kuczewski and Marshall (2002) likewise supported the concept of delayed consent. They pointed out that some sort of "cooling off" period should be implemented such that one can have time to evaluate their options (Bosk, 2002) and that delayed consent may

help accommodate cultural norms with regards to family involvement in the decision-making process all the while without reducing the importance of respect for the individual (Kuczewski & Marshall, 2002). It is noted that the significance of the waiting period is even more prominent when language issues are involved (Golec et al., 2004).

The proposed *Consent of Infants for Neonatal Secondary-use research* (CoINS) model in this thesis (based on a MC model) can accommodate these suggestions by providing parents with the ability and opportunity to review study details and modify their consent preferences at anytime and anywhere, such as without the immediate influence of the health care setting if they prefer. By accommodating for the surrogate consenter's needs and preferences may strengthen one's understanding of what they are consenting to thus ensuring better informed consent and by providing a sense of control may help increase their comfort in permitting the secondary use of their neonate's data. This in turn may lead to an increase in consent for neonatal data to be used for immediate and future longitudinal research studies thus benefiting society as a whole.

### 3.5 Late Onset Neonatal Sepsis

An estimated 15 million – more than one in ten – neonates are born preterm annually and the number continues to increase (March of Dimes, The Partnership for Maternal Health, Newborn and Child Health, Save the Children, & World Health Organization, 2012; World Health Organization, 2018). Preterm birth is a significant cause of long-term loss of human potential amongst survivors (March of Dimes et al., 2012). Birth complications resulting from a premature birth is the leading cause of neonatal deaths and is also the second cause of mortality after pneumonia in children under five years of age, accounting for over one million deaths in 2012 (March of Dimes et al., 2012) In addition to preterm birth, infections is another major cause of neonatal deaths (World Health Organization, 2019). Advancements in neonatal intensive care has contributed to the increased survival of premature neonates at lower gestational ages. However, even with access to neonatal intensive care, the premature and/or critically ill neonates are at an

increased risk of facing complications, such as the onset of one or more comorbidities of prematurity and the development of multiple conditions concurrently or over time following their preterm birth (March of Dimes et al., 2012). Premature neonates are more likely to develop severe infections and are at an increased risk of dying if they contract an infection.

Neonatal sepsis is a significant global health concern. It is a major cause of morbidity and mortality in the high-risk newborn population admitted to the NICU (Griffin et al., 2003; Stoll et al., 1996). The majority of neonates who die from neonatal sepsis are preterm (March of Dimes et al., 2012). Considering that many of the preterm neonates who survive often face a lifetime of disability such as the development of neurodevelopmental functioning impairments, learning disabilities, and visual and hearing problems (March of Dimes et al., 2012; World Health Organization, 2018), a clinical decision support tool that can also improve knowledge could dramatically reduce these implications or at the very least improve their quality of life.

Approximately one in four infants born with a very low birth weight (VLBW) (< 1500 g) who survive their first three days of life experience at a minimum of one episode of LONS while 25% of infected infants experience multiple episodes, ultimately resulting in a greater rate of mortality and longer hospital stay (Stoll et al., 1996). A timely and accurate diagnosis and treatment of LONS is crucial for improved survival rates and quality of life by decreasing complications and other adverse outcomes. In order to improve knowledge and understanding about LONS requires the contributions arising from health care research, which in turn requires the collection of data for such analysis. Analysis of physiological and clinical data has the potential to be used as an early detection tool for patients at risk of developing various medical conditions. However, the collection of physiological data, such as those stored within the Artemis Cloud could potentially be used as identifying information. It would only be respectful and appropriate to ask for surrogate consent before collecting, identifying eligible patients and use such data from a NICU patient.

LONS, also commonly referred to as simply late-onset sepsis is distinguished from early onset (neonatal) sepsis according to the onset of age (Dong & Speer, 2015). The onset of LONS develops three days (i.e. 72 hours) after birth (Dong & Speer, 2015; Griffin et al., 2003) at which point it is considered appropriate to differentiate LONS from EONS based on the causative pathogens (Dong & Speer, 2015). It has been noted that the incidence of LONS is inversely associated with birth weight (Dong & Speer, 2015) and gestational age (GA) with the incidence of infection increasing with decreasing birth weight and GA (Stoll et al., 1996). This signifies that immaturity is a great risk factor. Other risk factors for LONS include the long-term use of invasive interventions (e.g. mechanical intervention and intravascular catheterization), failure of early enteral with breast milk, a prolonged duration of parenteral nutrition, hospitalization, surgery and underlying cardiovascular and respiratory diseases (Dong & Speer, 2015; Stoll et al., 1996), all of which are likely possibilities for preterm and/or VLBW neonates admitted to the NICU.

Whilst the advancement of medicine, medical knowledge and technology have led to improved survival of preterm infants (i.e. <37 weeks gestation), in particular those who were born with a VLBW, the incidence of LONS has also increased in parallel, indicating that the postnatal nosocomial or community environment (i.e. hospitalization) and life-sustaining medical devices play a role in the pathogenesis of the condition (Cohen-Wolkowiez et al., 2009; Dong & Speer, 2015).

Early detection of LONS is difficult due to nonspecific and inconspicuous clinical signs (Dong & Speer, 2015). Neonates with sepsis are often detected when they are seriously ill. This not only increases morbidity and mortality rates, but also reduces the opportunity for a timely and complete recovery with antibiotic therapy (Griffin et al., 2003). To date, a blood culture remains as the “gold standard” diagnostic test for neonatal sepsis (Meem et al., 2011). However this method is not ideal as it is time-consuming and unreliable as it is prone to producing false positive and negative results due to the difficulty in determining between a true infection from a sample contamination (Dong & Speer, 2015). A timely and

accurate diagnosis and treatment of LONS is crucial for improved survival rates and quality of life by decreasing complications and other adverse outcomes.

### 3.6 Background on Heart Rate Variability Research

Monitoring of physiological data have shown promise as a potential diagnostic tool. Heart rate variability (HRV) is defined as the oscillation in the interval between consecutive heart beats (Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, 1996). It has been observed that abnormal HRV and transient HR decelerations were present in some neonates 12 to 24 hours before a clinical diagnosis of sepsis is confirmed (Griffin & Moorman, 2001). This phenomenon is attributed to the likely consequence of a systematic inflammatory response syndrome, partially mediated by inflammatory cytokines (Fairchild, 2013; Griffin et al., 2003). However, that work does not take into account that other conditions and situations can result in reduced HRV (McGregor et al., 2012). Consequently, many early detections of LONS approaches are prone to reporting false positives.

For example, study results have revealed that the majority of raised HRV actually occurred in patients without bloodstream infection (BSI), whereas patients with BSI seldom experienced elevations implying that heart rate characteristic (HRC) score is neither very sensitive nor very specific in its ability to for BSI (Coggins et al., 2016). Thus at this point in time the predictive value of HRV monitoring in clinical practice remains undetermined. The results also suggest that more than just HRV behaviours should be analyzed when attempting to rely on physiological indications rather than physical signs for an earlier diagnosis of LONS.

de Beer et al. (2004) have confirmed that the administration of atropine, a muscarinic receptor caused large variations in HRV (de Beer et al., 2004). Loforte et al. (2006) examined HRV and the relationship between the heart's RR-wave intervals and the spontaneous respiration in a selected population of ill premature infants. It was discovered that lower relationship values were strongly correlated with a diagnosis of sepsis. Thus they suggested that further analysis of HRV and

respiration relationships could be a potential indicator of infection in premature newborns (Loforte, Carrault, Mainardi, & Beuche, 2006). McGregor et al. (2012) discovered that the dependence on low HRV alone was limited in helping to differentiate between patients who developed LONS from patients who had low HRV due to confounding factors such as surgery and/or the use of narcotics or other medications. As a result, further clinical retrospective research studies are required.

### 3.7 Summary and Implications on Research

This chapter has presented an appraisal of various consent issues and factors that affect the parental consent process with respect to neonatal research performed within the NICU settings. Given that parents often act as their neonate's surrogate decision makers during this challenging time, there is a need to support changing proxy consent over time. This motivated the first and second research questions: Specifically, a new approach to informed consent is required. It is proposed that a consent model that enables flexible, longitudinal and surrogate consent would be beneficial to parents. To enable such functionalities requires a consent model that can couple consent data with study data within an IT platform and be integrated within a systematic platform that can support multiple research studies. The consent framework is presented in chapter 5.

# 4. Artemis: A Big Data based Health Analytics for Clinical Decision Support and Clinical Research

This chapter introduces Artemis, a Big Data based Health Analytics platform that supports the analysis of high-speed physiological data in real-time to support clinical decision support and enable clinical research. The Artemis platform is relevant to this thesis as the infant consent model proposed would guide the use of data via the Artemis system. At present the platform does not have a consent mechanism.

Artemis is a high frequency, multisource, real time, online health analytics platform developed through a collaboration between the University of Ontario Institute of Technology (Ontario Tech University) and the IBM T.J Watson Research Center (Blount et al., 2010; McGregor, 2017). It is a Big Data platform that allows for concurrent multi-patient, multi-diagnosis and multi-stream temporal analysis in real-time for purposes of clinical management and research (Blount et al., 2010; McGregor, 2013).

Artemis was named after the Greek goddess of childbearing (Blount et al., 2010) (Blount et al., 2010; McGregor, 2017). The platform obtains high frequency physiological data from neonatal monitors that are located at the infant's bedside. Along with the necessary clinical information, physiological data collected include electrocardiogram (ECG), derived signals from the ECG including the heart rate (HR), respiratory rate (RR) and respiratory impedance for purposes of breath detection. Other signals captured provide information such as blood pressure (i.e. diastolic, systolic and mean blood pressure) together with pulse wave plethysmography, blood oxygen saturation ($SPO_2$), transcutaneous oxygen and carbon dioxide ($CO_2$) measurements when such data is available. The information can be simultaneously processed and returned in real-time or stored for later analysis and research (McGregor, 2013, 2017; Pugh & McGregor, n.d.).

## 4.1 Technical Architecture

The Artemis platform is composed of five different components. This consists of the: 1) Data Acquisition, 2) Online Analysis, 3) Data Persistency, 4) Knowledge

Extraction and 5) redeployment components as shown in Figure 6. Along with pertinent clinical information, physiological data streams are continuously collected from medical devices, which are then inputted into the data acquisition component of the platform. From the Data Acquisition component, the data is then sent to the Online Analysis component where IBM's InfoSphere middleware system is used to process the data in real-time. In conjunction with the newly generated analytics, the raw data is stored in the data persistency component of the platform. For clinical research purposes, the Knowledge Extraction component performs the task of data mining (DM) on the data based on the medical condition of interest.

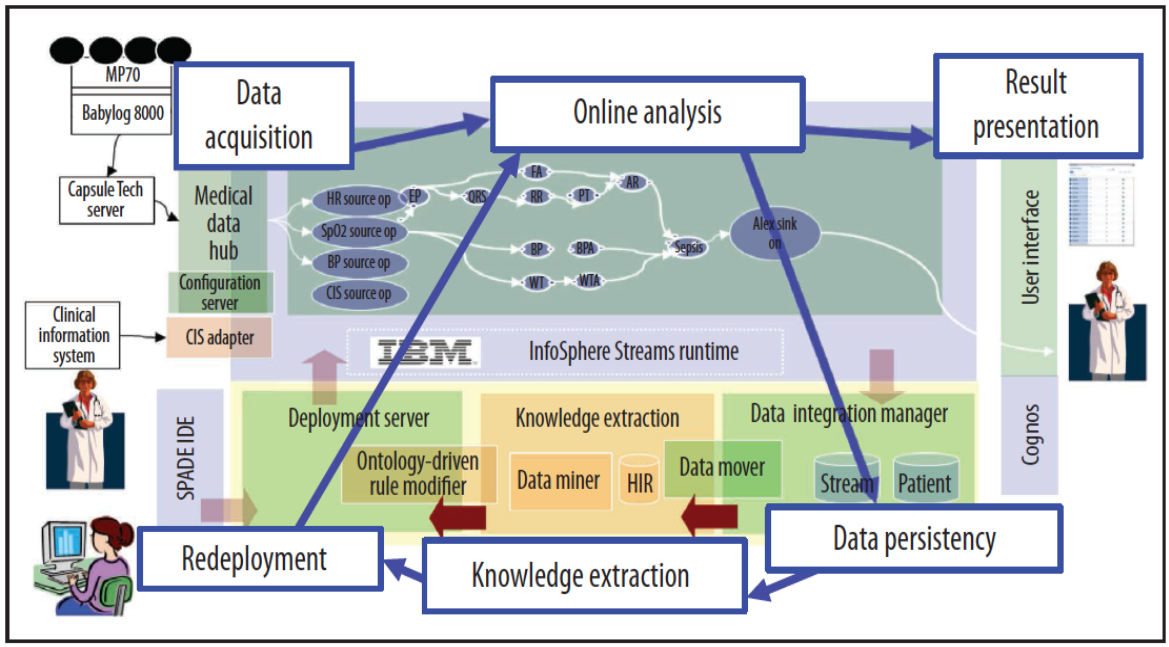


*Figure 6: The Artemis platform*
*(McGregor, 2013)*

### 4.2 Retrospective Clinical Research Using Artemis

Examples of neonatal research using Artemis include late onset neonatal sepsis (LONS) (McGregor, Catley, Padbury, & James, 2013), apnoea of prematurity (Thommandram et al., 2013; Thommandram et al., 2014), anemia of prematurity (Pugh, Ng, McGregor, Belik, & James, 2013), premature infant pain (Naik et al., 2013) and sleep wake cycling detection (Eklund et al., 2014).

Although Artemis implementations currently have no influence in the clinical setting, by comparing the analytical results that are gathered in the platform with current treatment practices, new patterns in real-time physiological data can be identified ultimately contributing to the goal of enabling earlier detection and prevention of various health conditions before clinical symptoms are visible (McGregor, 2013).

### 4.3 Implementation Hospitals

The Artemis platform has been implemented to support clinical research studies within NICUs at various hospitals. Inclusion of the consent model proposed in this thesis as a key component of the Artemis platform implementation would strengthen the existing approach through inclusion of surrogate consent. Artemis was initially deployed at The Hospital for Sick Children (SickKids) in 2009 for a pilot case study regarding late onset neonatal sepsis (LONS) (McGregor, 2017; Pugh & McGregor, n.d.). In that deployment, a waiver of consent was granted by REB at SickKids. Following the successful Artemis implementation at SickKids, Artemis Cloud, a newer cloud computing-based version of Artemis was deployed within NICUs at the Women and Infants Hospital of Rhode Island (WIHRI) and the Children's Hospital of Fudan University. Within that deployment, Artemis was a sub-study of another study where consent for the primary and sub-study were obtained. However, the consent data and study data were not located electronically together. More recently was the implementation of the expanded Artemis Cloud at the NICUs at McMaster Children's Hospital (MCH) which will allow for high frequency data to be used in real time for analysis at much higher speeds than its predecessors (McGregor, 2017; Pugh & McGregor, n.d.). The expansion of Artemis Cloud platform provides the ability to service multiple healthcare facilities (Khazaei, Mench-Bressan, McGregor, & Pugh, 2015).

### 4.4 Service Based Multidimensional Temporal Data Mining

The Data Persistency and Knowledge Extraction components of Artemis are an instantiation of the Service Based Multidimensional Temporal Data Mining

($STDM^n_0$) Framework that represent a system, method and computer program proposed by McGregor to support multi-dimensional temporal analysis (TA) and data mining (DM) (McGregor, 2013; McGregor et al., 2011, 2012). The main focus of the $STDM^n_0$ framework is to bridge the gap between clinical management and clinical research by facilitating the secondary use of data collected from various medical monitoring devices. Through the use of DM in physiological data streams, $STDM^n_0$ enables the identification of previously unknown pathologies by supporting the discovery of patterns within physiological data streams that exist prior to a clinical condition under investigation (McGregor et al., 2011; McGregor, Smith, et al., 2013).

The original and current $STDM^n_0$ architecture is presented in Figure 7.

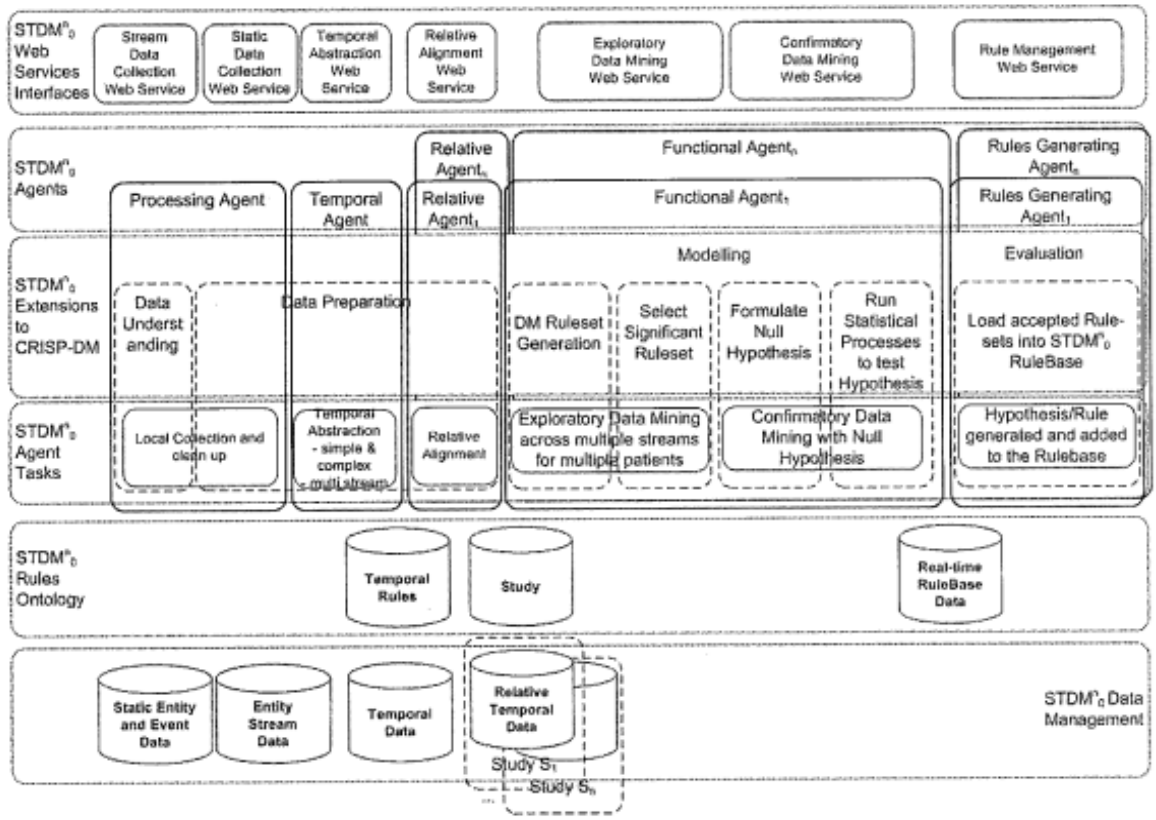


*Figure 7: Original and current $STDM^n_0$ database (McGregor, Smith, et al., 2013)*

The $STDM^n_0$ framework is comprised of six layers. As per Figure 7, from the top to the bottom this includes the $STDM^n_0$ *Web Services Interfaces*, $STDM^n_0$

*Agents*, STDM$^n_0$ *Extensions to CRISP-DM*, STDM$^n_0$ *Agent Tasks*, STDM$^n_0$ *Rules Ontology*, and STDM$^n_0$ *Data Management* layers (McGregor, Smith, et al., 2013). The multi-agent system within the STDM$^n_0$ *Agents* layer drives the overall framework, the extended CRISP-DM model layer defines the data mining tasks and the STDM$^n_0$ *Agent Tasks* layer are the three main components or layers of the STDM$^n_0$ framework. The n and 0 in the STDM$^n_0$ framework acronym represent the extensions enabling support for multiple research studies and null hypothesis testing respectively (Dhanoa, 2011). STDM$^n_0$ incorporated the null hypothesis approach introduced in Heath (2006) and (Heath, 2006; Heath & McGregor, 2010)

STDM$^n_0$ is built on the foundation of CRISP-DM and enables knowledge discovery. The CRoss Industry Standard Process for Data Mining (CRISP-DM) as an approach to use for the process of knowledge discovery that is utilized to perform clinical research studies is detailed next. The infant consent model considered in this thesis would enable the utilization of the STDM$^n_0$ with world best practice approach to consent.

### 4.5 Cross Industry Standard Process for Data Mining

*CRoss Industry Standard Process for Data Mining* (CRISP-DM) was developed in 1996 as a methodology for data mining (DM) processes within the knowledge discovery function (Figure 8). CRISP-DM has been established and accepted as a cross-industry standard model for DM and knowledge discovery (Catley, Smith, McGregor, & Tracy, 2009; Huang, McGregor, & James, 2014; McGregor et al., 2011; Moorman et al., 2011; Shearer, 2000). CRISP-DM breaks down the life cycle of a data mining project into six high-level phases, being: 1) business understanding, 2) data understanding, 3) data preparation, 4) modeling, 5) evaluation and 6) deployment (Chapman et al., 2000). The sequence of phases is neither a strict nor linear process. Rather, the outcome of each phase determines which phase or specific task within a phase should be implemented next (Wirth & Hipp, 2000), meaning; one may have to move back and forth between the phases depending on the specific project being undertaken.

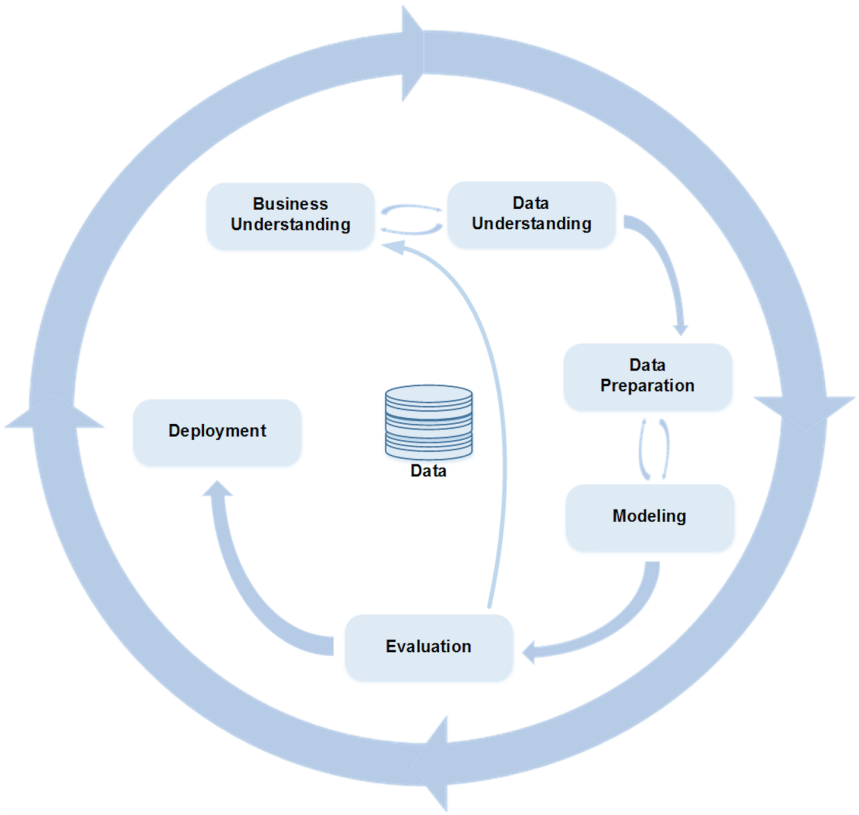


*Figure 8: CRISP-DM diagram*
*Adapted from (Jensen, n.d.)*

Since the conceptualization and development of the initial CRISP-DM methodology, “the needs of data mining users, technologies available, types of data harvested and types of deployment have all evolved” (Hildebrandt & Gutwirth, 2008). In response to the many environmental changes, a Special Interest Group (SIG) was formed to in an attempt to enhance CRISP-DM 1.0 to CRISP-DM 2.0 (Catley et al., 2009; Hildebrandt & Gutwirth, 2008; Mariscal, Marban, & Fernandez, 2010) in the mid-2000s although to date it appears that initiatives are no longer active (Vorhies, 2016).

Catley et al. (2009) proposed that when the CRISP-DM methodology is applied to the healthcare industry, specifically in the context of the (neonatal) intensive care environment, its limitations become evident especially when it is used to model Intelligent Data Analysis (IDA) systems that perform Temporal Data Mining (TDM) of time series data such as streams of physiological data. The use

of CRISP-DM methodology is inadequate in meeting the integrated needs of TDM making it extremely challenging, if not unfeasible to compare and evaluate systems from a clinical or IDA viewpoint (Catley et al., 2009). They note that despite the ongoing efforts to advance CRISP-DM 1.0 to CRISP-DM 2.0, there is no provision for multi-dimensional time series data, which is increasingly forming the data source for intricate DM systems that incorporates the use of Temporal Abstraction (TA) as a pre or post-processing step (Catley et al., 2009).

Temporal abstractions (TA) are features or patterns of time-orientated raw data, a task that typically employs the use of temporal reasoning techniques (Fisher, Gabbay, & Vila, 2005). When the TA task is perceived as a process, it involves being given a set of time-stamped data, external events, and abstraction goals for the purpose of producing abstractions of the data that interpret past and present states and trends, that are pertinent for the given set of goals (Shahar, 1997). TA is often utilized as a pre-processing step prior to DM (Catley et al., 2009).

In the medical domain, temporal data mining (TDM) techniques are commonly designed either for exploration or for prediction purposes. Exploratory methods involve processing a database to detect groups of time series with similar patterns of frequent intervals and temporal relationships. To determine if they represent useful or previously unidentified relationships between data types may require or rely on the application of clinical domain knowledge to these groups or clusters. The target of predictive techniques may be a particular diagnosis, therapeutic response, or other clinical or patient care process. Such techniques scrutinize for combinations of intervals that often transpire with some temporal relationship to the target (Post & Harrison, 2008). The consent model proposed in this thesis enables the utilisation of NICU data for TDM activities through capture of surrogate consenter preferences.

To date, the majority of NICU medical records are recorded in a manual manner (i.e. by hand) on paper records by health care professionals. This means that physiological data streams values are typically summarized once at time intervals (e.g. every 30 or 60 minutes). However, significantly atypical variations in measured physiologic parameters are prevalent in the critically ill population. This

means that constantly changing data, at a second-by-second basis often result in missed events and are not recorded. Research has shown that these missed events can be crucial in predicting survival and quality of survival free of significant disability (Lister et al., 2000; McGregor, 2013). In addition, neonatal research has advanced to the stage where the presence of specific physiological measurements or indicators can help forecast future physiological episodes. Consequently, opportunities exist to use TA-based DM to detect medical conditions prior to their onset via the identification of these physiological patterns and then incorporating them as rules within database models to allow for alert based reporting and actionable preventive treatments to be undertaken (Catley et al., 2009; McGregor, Bryan, Curry, & Tracy, 2002).

Examination of the CRISP-DM methodology has revealed inadequacies particularly in a NICU context, in nearly all the phases. These include limitations in their ability to describe the following including:

1. Clinically relevant and population-based information in phase 1 (business understanding)
2. Temporal aspects of the multidimensional data along with the clinical study in phase 2 (data understanding)
3. TA of relevant details and knowledge management in phase 4 (data modeling)
4. System integration in phase 4 (data modeling); CRISP 1.0 concentrates on applying several DM techniques to arrive at one which offers the best results, rather than providing support for integration of techniques, such as DM and TA
5. Assessment of process mining results based on temporally abstracted data in phase 5 (evaluation)
6. Storage issues, knowledge sharing and representation issues in phase 6 (deployment), including mechanisms used for knowledge representation, such as adherence to health care standards

(Catley et al., 2009; McGregor, Catley, & James, 2011)

As the number of integrated TDM systems continues to multiply, there is a need to extend the CRISP-DM methodology to support temporal data mining (Catley et al., 2009) (Catley et al., 2009). Within that extension, additional activities were proposed as detailed in Figure 9. Extended tasks are distinguished with an asterisk (*) and extended attributes are marked with double asterisks (**) (Catley et al., 2009).

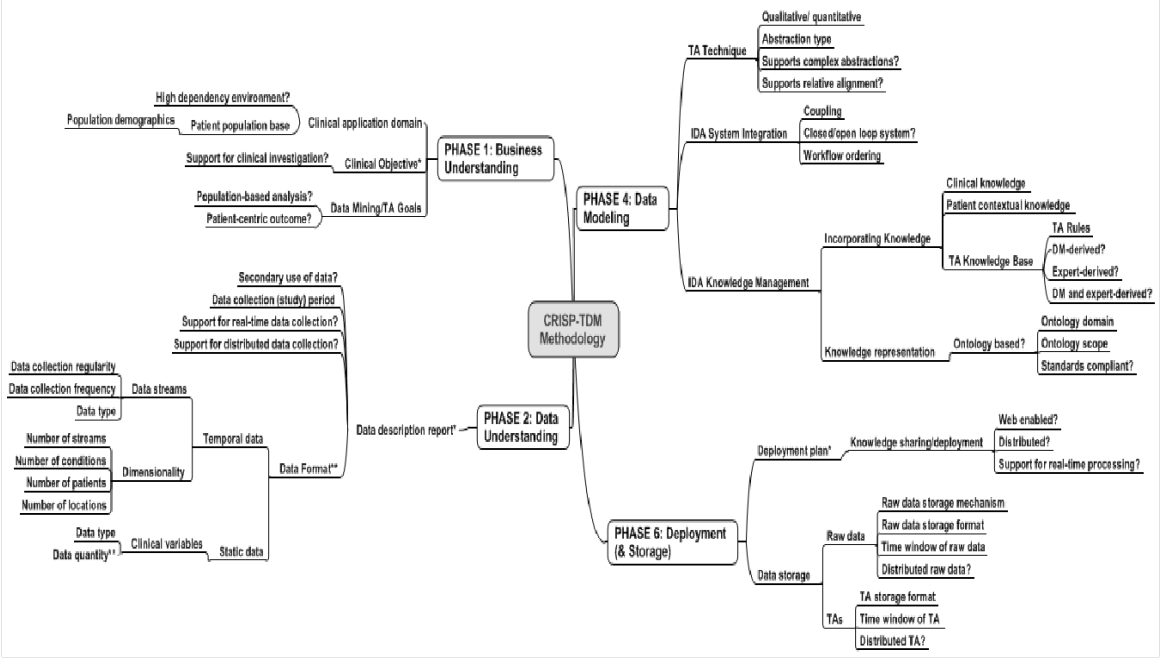


*Figure 9: Extended CRISP-TDM methodology for CRISP-TDM phases (Catley et al. (2009))*

McGregor et al. (2011) introduced the CRISP-DM into the intensive care environment. To support temporal DM specifically that of multidimensional time series DM, they extended it to create the CRISP-TDM model (Catley et al., 2009; Huang et al., 2014). This was accomplished via the addition of elements pertinent to the mining of clinical data, consisting of: 1) business understanding, 2) data understanding, 4) modeling, and 6) deployment phases via the utilization of the $STDM^{n}_{0}$ framework that enables knowledge discovery of new condition onset pathologies from physiological data streams (McGregor et al., 2012).

The enhanced CRISP-DM model has been labeled as CRISP-TDM to emphasize its capability to support temporal (T) and multidimensional (n) clinical

data (McGregor et al., 2011). Both the CRISP-DM and CRISP-TDM models can be implemented in processes and sectors outside of healthcare.

While the CRISP-TDM builds upon the standard, existing CRISP-DM model, neither model explicitly considers the secondary use of health data consent process.

# 5. Database Model Design to Support Flexible, Longitudinal and Surrogate Consent

This chapter presents a database design supporting research studies involving the analysis of patient physiological data that supports a flexible, longitudinal, individual and surrogate consent model. This is proposed via the inclusion of additional tables as an extension to the current Service Based Multidimensional Temporal Data Mining Framework ($STDM^{n}_{0}$). This supports three of the four objectives namely enabling the *Consent of Infants for Neonatal Secondary-use research* (CoINS) model to be integrated within database infrastructures to allow for secondary use of data research studies that analyzes streams of data collected from sensors in relation to patient care, instantiation of the consent model within a research database and demonstration of the CoINS consent model within an ethically approved secondary use of data research study where surrogate consent is required. The latter is demonstrated via a Late Onset Neonatal Sepsis (LONS) clinical case study in chapter 6.

## 5.1 Data Elements Required for Capturing Consent

The CoINS consent model offers the primary functionalities as concluded from the literature review and neonatal context reviews in chapters 2 and 3 respectively. The resultant functional requirements are repeated here in Table 5.

| Functional Requirement | Reason | Consent model data requirements |
|---|---|---|
| **Flexibility** | Allow SDM(s) to:<br>1) choose their consent preferences<br>2) adjust their chosen consent preferences at any time | • Consent levels |
| **Surrogate consent** | Neonates are incompetent - parents or legal guardians often act as their neonate's SDMs | • Proxy consenter's contact information<br>• (Neonatal) patient information |

| | | |
|---|---|---|
| **Longitudinal consent** | Allow SDM(s) to:<br>1) choose their consent preferences<br>2) adjust their chosen consent preferences at any time and<br>3) ensure chosen consent preferences take effect immediately | ● Ability to filter out patients who may be eligible as a participant for a particular research study |
| **Coupled with data** | Ensure that consent data is coupled with research study data (i.e. collected PHI) | |
| **Information Technology** | Enable consent model to be integrated within database infrastructures that enable secondary use of data research studies that analyze streams of data collected from sensors in relation to patient care | ● Original $STDM^n_0$ platform |

*Table 5: Required functionalities, the purpose(s) of the required functionality and the consent model data requirements for the required functionality*

To allow for the functionalities of flexible, surrogate consent and longitudinal consent and to ensure that consent data is coupled with research study data (i.e. collected PHI) required the CoINS consent model to be able to identify and link the patient with their (surrogate) consenter, the chosen consent level of engagement and the research study/studies for which the patient is enrolled in. This required the addition of new tables to the existing $STDM^n_0$ architecture as seen in figure 11.

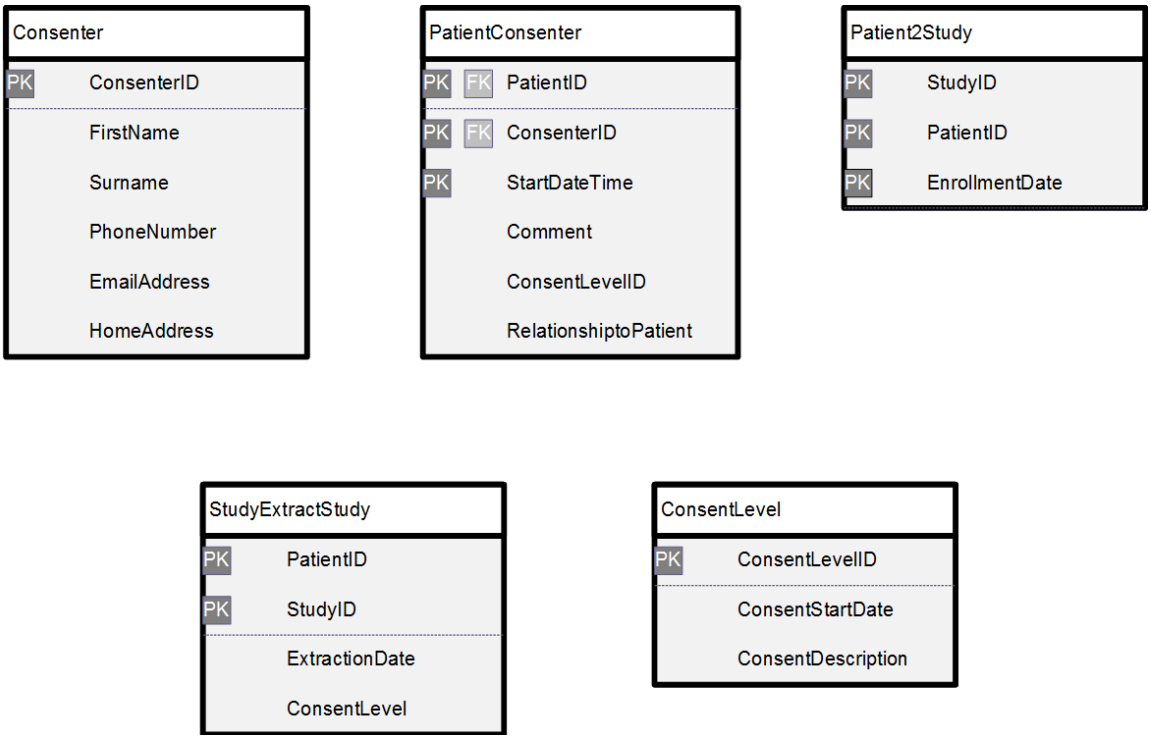


*Figure 10: New tables to be added to the existing STDM$^{n}_{0}$ framework*

The purpose of adding each table is defined below in Table 6:

| Table | Purpose |
|---|---|
| Consenter | • Captures the identity and contact information of the (surrogate) consenter |
| PatientConsenter | • Connects the patient with the (surrogate) consenter, identifies the (surrogate) consenter's relationship with the patient and displays the consenter's consent preference via a consent level |
| Patient2Study | • Connects the patient to (a) research study/studies that they are consented by the (surrogate) consenter to be enrolled in and records the date enrollment begins |
| StudyExtractStudy | • Identifies the patient profile extracted for enrollment in research study/studies |
| ConsentLevel | • Defines the various consent level of engagement based on the selected available consent preference and the start date of the selected consent option<br>• In the event that the consent level option is modified at a later date, the start date will be modified to reflect the date that the change was implemented |

*Table 6: New tables and the purpose of adding each specific table to be added to the original and current STDM$^{n}_{0}$ database model*

Table 7 clearly identifies the primary key(s), foreign key(s) and attributes associated with each of the new tables.

| Table | Primary key(s) | Foreign key(s) | Attributes |
|---|---|---|---|
| Consenter | ConsenterID | | FirstName, Surname, PhoneNumber, EmailAddress, HomeAddress |
| PatientConsenter | PatientID, ConsenterID, StartDateTime | PatientID, ConsenterID | Comment, ConsentLevelID, RelationshiptoPatient |
| Patient2Study | StudyID, PatientID, EnrollmentDate | | |
| StudyExtractStudy | PatientID, StudyID | | ExtractionDate, ConsentLevel |
| ConsentLevel | | | ConsentStartDate, ConsentDescription |

*Table 7: Primary key(s), foreign key(s) and attributes for each new table to be added to the original and current STDM$^n_0$ database model*

### 5.2 Amended STDM$^n_0$ Architecture and Database

This research extends the database design in STDM$^n_0$ to enable the flexible, longitudinal surrogate consent process as well as the ability for a patient to be enrolled in multiple studies. New tables were added to the STDM$^n_0$ database for purposes of supporting the definition of the consenter, their relationship to the patient and their level of consent (i.e. engagement level) at any given point in time noted in Figure 12 as Consent Data. Study Consent Data were added to support the instantiation of consent on a study-by-study basis. Consented Relative Temporal Data contains the patient temporal data that have been extracted from the clinical system for purposes of a given research study which satisfy the rules for consent.

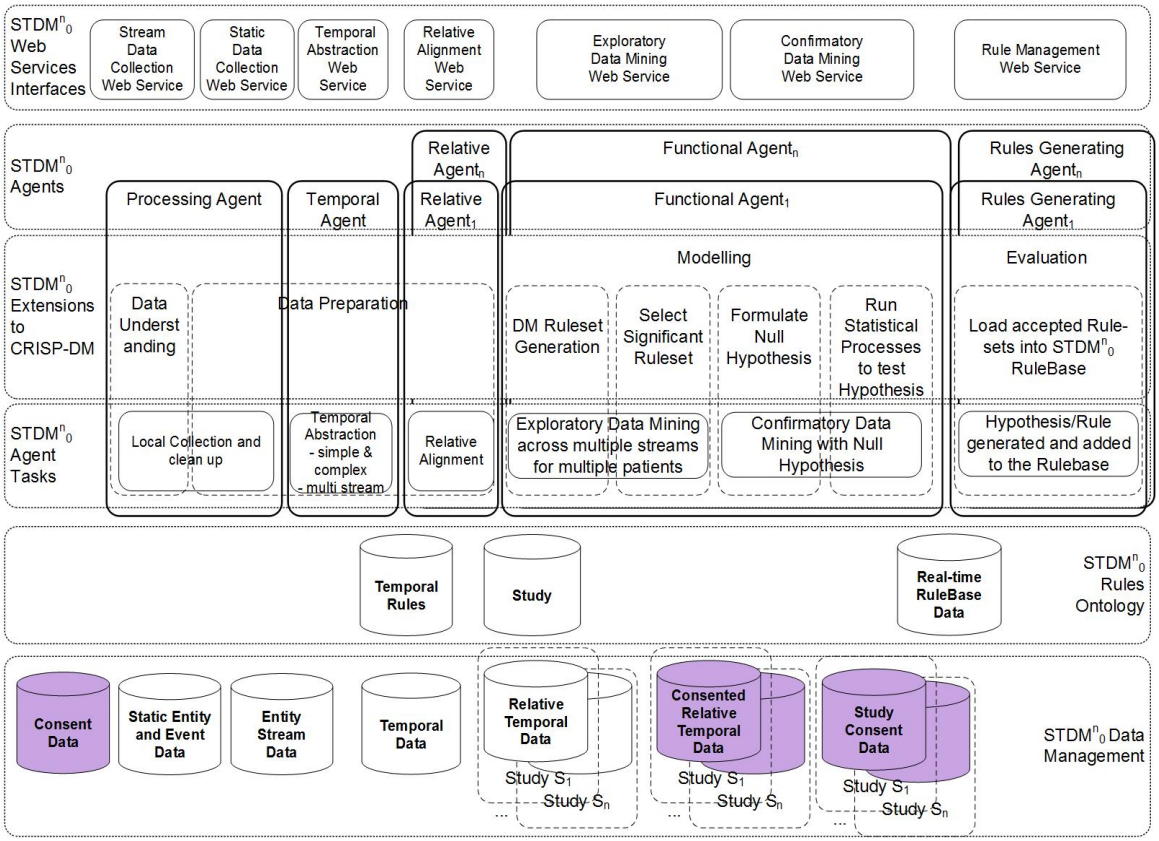


*Figure 11: ERD of the current STDM$^n_0$ framework along with the addition of new tables (outlined in bold)*

The original, existing tables within the STDM$^n_0$ database are depicted on the top while the additional tables are outlined in bold at the bottom third of the diagram.

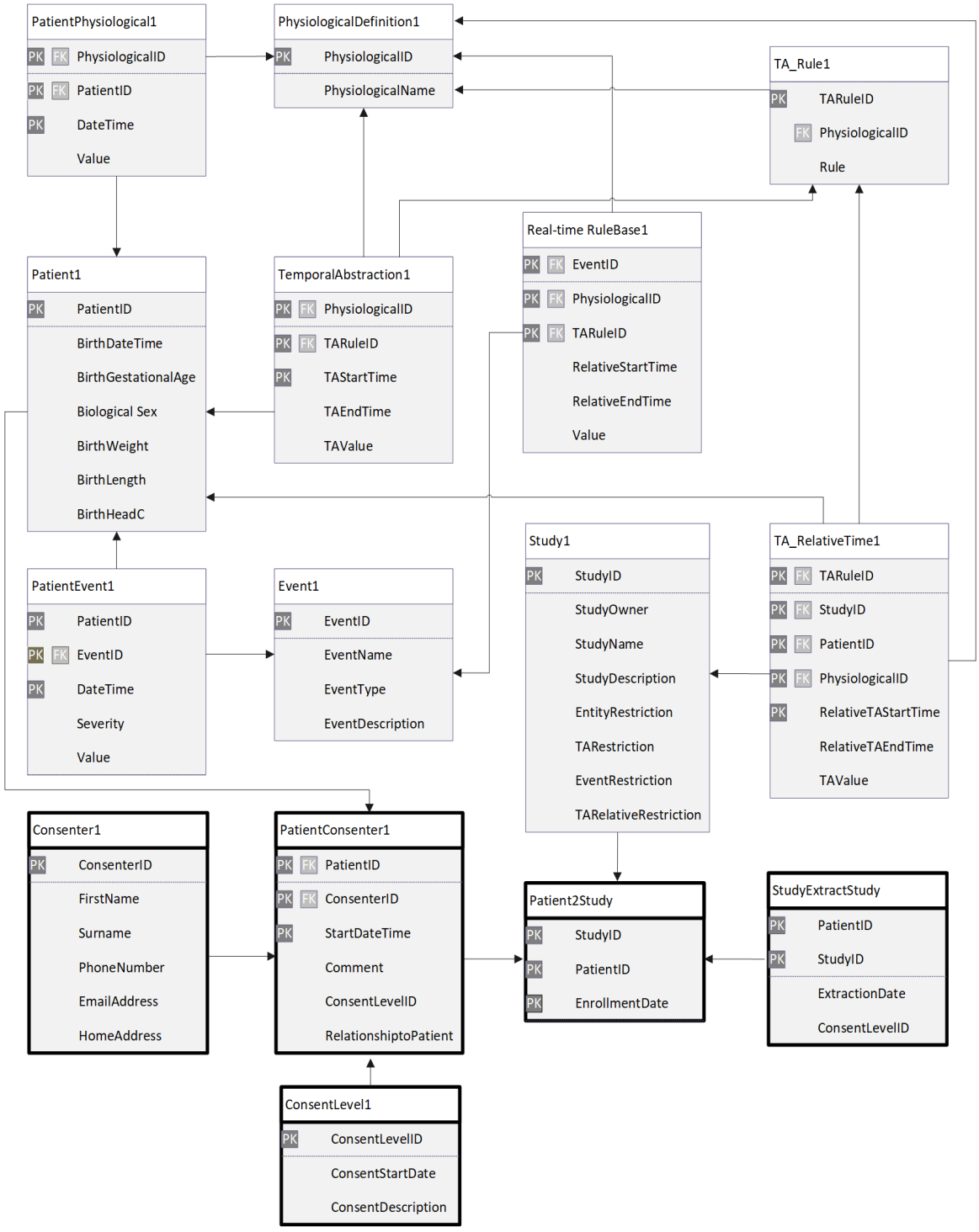


*Figure 12: ERD of the current STDM$^{n}{}_{0}$ framework along with the addition of new tables (outlined in bold)*

### 5.3 Amendments to the CRISP-TDM Method

To support the population of the data within the additional $STDM^n_0$ database tables, additional activities are added to the CRISP-TDM's phase 4 Modelling step.

Implementation of a flexible consent model will significantly impact the consent process, affecting the data collection process and enriching the volume and quality of data flowing to other processes within CRISP-$TDM^n_0$. The consent process begins before application of the CRISP-$TDM^n_0$ approach to knowledge discovery.

Flexible consent models such as the one proposed can split the general consent process into three phases as per the following:

- Consent to Collect (C2C): Permission to collect patient data
- Consent to Identify whether patient fits into research Cohort (C2IC): Based on the research study's inclusion and exclusion criteria, this is permission to be able to review the profile of a patient to determine if they are (in)eligible for a particular study
- Consent to Use (C2U): Permission to use the patient data for analysis

In Catley et al. (2009), additional activities were proposed within CRISP-TDM as detailed in Figure 9 in section 4.5 and repeated below in Figure 13. However, as noted in section 2.15, activities relating to consent were not included.

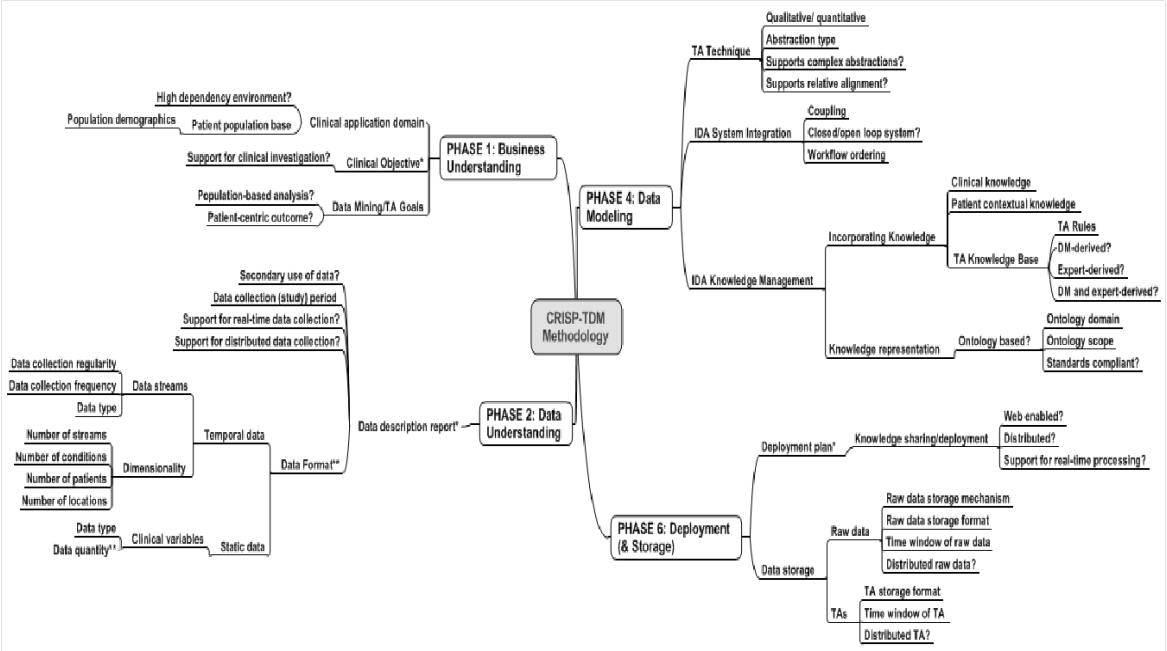


*Figure 13: Extended CRISP-TDM methodology for CRISP-TDM phases (Catley et al. (2009))*

Through the method proposed in this thesis, a fourth high element is added to the Data Modeling extensions proposed by Catley et al. (2009) to cater for Consent. That Consent element contains details of the consent level appropriate for the proposed research study together with details of the consent data for the data that will be used for the secondary use of data study.

The Data Preparation step is extended with an additional process step to enable the consent process for data collection for the secondary use of data. In this case, we consider the LONS study portion only and consider the secondary use of the data collected previously by the Artemis Cloud Database study. The process begins with identifying the eligible patient population. Selection of the eligible patient population is based on the predetermined inclusion and exclusion criteria of the REB approved study. The next step is to check the SDM's chosen consent level to determine if the neonatal patient's collected health data can be used for secondary research. Data for consented patients is then populated within the study consented dataset as the study $S_{LONS}$ instance of the Study Consent Data.

## 5.4 Summary

STDM$^n_0$ is built on the foundation of CRISP-DM and allows for knowledge discovery. However, its original design the database was not designed to capture consent. This chapter demonstrated that STDM$^n_0$ can be extended such that it can allow for the incorporation of a flexible, longitudinal surrogate consent process as well as the ability for a patient to be enrolled in multiple studies via the addition of tables to the existing database.

This chapter presents the application of the extended STDM$^n_0$ and the CoINS consent model into a retrospective late onset neonatal sepsis (LONS) case study.

# 6. Case study: Late Onset Neonatal Sepsis

This chapter presents an instantiation of the proposed flexible consent model that was introduced in the previous chapter. The instantiation within the Artemis research database study is detailed to demonstrate the process of capturing the consent for that study. The results of that consenting process are then presented. A demonstration of its application in a LONS study at McMaster Children's Hospital (MCH) is then presented...

Artemis Cloud has been developed to provide a robust real-time cloud to provide health analytics as a service. The primary purpose of undertaking the LONS clinical study is to evaluate the deployment of Artemis Cloud and to demonstrate its reliability, functionality and efficacy. McMaster Children's Hospital (MCH) is the first health care site providing data to the new Artemis Cloud and will simultaneously serve as the first real-time test of this new infrastructure that has the capability of storing and processing multiple channels of data provided at up to 1000 Hz. MCH is a pediatric academic health science centre with a tertiary NICU located in Hamilton, Ontario, Canada (McMaster Children's Hospital, 2018a, 2018b; Pugh et al., 2018). The secondary use of the data collected for the LONS clinical study was required for the purpose of performing final validation and verification of the implementation of Artemis Cloud (Pugh et al., 2018) and the heart rate variability and respiration variability calculations used as an early onset detection mechanism for LONS.

## 6.1 Implementation in Artemis

The implementation of the flexible consent model in Artemis Cloud required the creation of new tables within the Persistent storage component of Artemis that has been instantiated in IBMs DB2 database management software. This component also instantiates the extended Service Based Multidimensional Temporal Data Mining Framework ($STDM^{n}_{0}$) framework data management component. The data from the consenting process for this case stay was loaded into these tables in DB2.

### 6.2 McMaster Children's Hospital Artemis Cloud Database

The McMaster Children's Hospital Artemis Cloud Database (henceforth referred to as “Artemis Cloud Database”) is the research database created in collaboration between McMaster Children's Hospital and the University of Ontario Institute of Technology (Ontario Tech University) for the data collection part of the Artemis Cloud Database study. The purpose of this is a research database is to collect and save all the information that is produced by the neonatal patient that would otherwise be lost after it disappears from the monitors. This includes the collection of physiological data from the bedside monitors and electronic equipment along with clinical information from their electronic medical record (EMR) during the NICU stay. The information may be used for future health research studies in which analyzing information about the patient and other similar infants may help improve the future care of term and preterm infants. McMaster Children's Hospital (MCH) is a pediatric academic health science centre with a tertiary NICU located in Hamilton, Ontario, Canada (McMaster Children’s Hospital, 2018a, 2018b; Pugh et al., 2018).

### 6.3 Process Implemented for Artemis Cloud Database Consent

The McMaster Children's Hospital Artemis Cloud Database study (HiREB 3859-D) was approved on January 4, 2018. Data from the bedside monitors began streaming on March 9, 2018 and the date of the first SDM consent was obtained on March 12, 2018. At that point in time, the LONS protocol had not yet been approved, but the data collection for LONS study is being collected through the Artemis database approval. The secondary use of data study entitled *Artemis Cloud Clinical Decision Support Framework – An algorithm to accurately detect late onset neonatal sepsis* (HiREB 4833-C) was approved on April 26, 2019.

The subject inclusion criteria for the Artemis Cloud Database study consisted of the following two requirements:

1. Infants who are admitted to the NICU at McMaster Children’s Hospital

2. Parents/guardians who have provided signed informed consent for the enrolment of their infant(s) into the McMaster Children's Hospital Artemis Cloud Database (HiREB 3859-D)

The sole subject exclusion criteria were for infants whose parental/guardians' consent for the Artemis Cloud Database is not obtained or is withdrawn.

The sole RC responsible for obtaining surrogate consent would first review the neonatal admission note for each infant prior to approaching any potential surrogate consenter. The note provided the RC with information regarding the infant's general condition and characteristics including the physical examination, reason for admission, the assessment and treatment plan as well as information regarding maternal history, labour and delivery. Given each family's unique situation, the RC would then make a subjective decision as to determine whether or not it would be appropriate to approach the surrogate consenters regarding participation in the Artemis Cloud Database study. If deemed appropriate, the RC approached available surrogate consenter(s) to verbally explained the purpose of the McMaster Children's Hospital Artemis Cloud Database study and was available to answer any questions they may have regarding their potential infant's/infants' study enrollment before presenting them with a written Participant Information Sheet (Appendix B) for each of their admitted infants. In the event that (a) surrogate consenter(s) chose to consent their infant for the participation, they would select one of the two consent options as presented in the research study consent form (Appendix B). Figure 14 summarizes the research subject selection and consent process in a unified modelling language (UML) model.

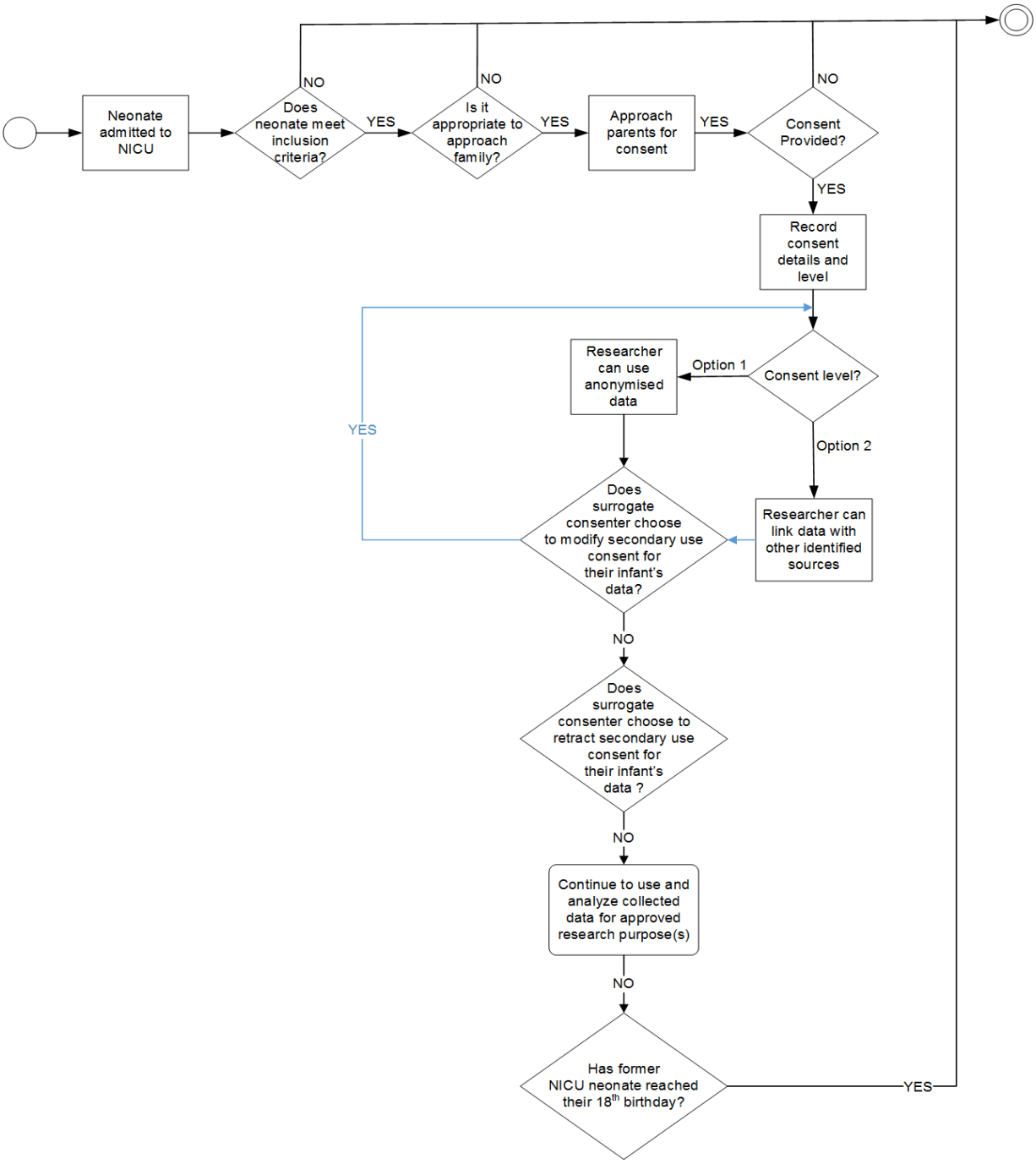


*Figure 14: The consent process for the Artemis Cloud Database displayed in an UML diagram*

Although Artemis collects data from the moment the infant is connected to the bedside monitor, it did not enter the research database tables until consent was obtained. This meant that the surrogate consenter(s) was approached for consent after the admission of their neonate into the NICU and after the collection of their infant's data has already commenced. Only one parent's written consent was required for a neonate's enrollment into the Artemis Cloud Database study.

No standardized script was utilized by the RC when approaching the surrogate consenters. While the consent approach process remained consistent throughout the consent time range, there was a change with regards to the RC's presentation and explanation of the available consent options during the initial study discussion. Initially it was explained to potential surrogate consenters why two consent options were available. The discussion placed emphasis on the privacy and confidentiality aspects of health data in a research ethics context with consent option 1) being presented as McMaster data and consent option 2) presented as data that can be linked. It was noted that there was a general disconnect between the surrogate consenters understanding between the privacy and confidentiality in a research ethics context versus how their infants' data contribution would be beneficial for future research. Consequently, the later change involved providing parents with less technical detail about consent option 2. Rather than focusing the discussion on explaining why there were two "yes" options, option 2 was presented as a practical method to facilitate longitudinal research by demonstrating the importance of being able to conduct future research on what happens to McMaster Children's Hospital's NICU infants. Unfortunately, the date of when this change in consent process occurred was not tracked. For this reason, consent option selection has been longitudinally analyzed to see if changes in distribution of consent option occurred over time that could be attributed potentially to this change in method for recruitment.

### 6.4 Results of Artemis Cloud Database Consent

In the span of less than one month, beginning in early March to April of 2018, fifty-one NICU bed spaces at McMaster Children's Hospital were set up to acquire

data collection for the Artemis Cloud Database. Of the 51 beds, there were 48 beds in the Level 3 NICU and 3 beds in the delivery suite. Table 8 shows the exact dates, specific bed numbers and the total number of beds that were set up to collect patient data from consented patients.

| Date NICU Beds Were Set Up | Specific Bed Numbers | Total Number of Beds |
|---|---|---|
| March 9, 2018 | Bed 1 | 1 |
| March 14, 2018 | Beds 2 – 10 | 10 |
| March 20, 2018 | Beds 11– 20 | 20 |
| March 22, 2018 | Beds 21–30 | 30 |
| April 3, 2018 | Beds 31-50<br>● 31 – 35 at 11 am<br>● 36 – 40 at 12 pm<br>● 41 – 45 at 1 pm<br>● 46 – 50 at 2 pm<br>● 51 | 50 |

*Table 8: When NICU bed spaces at McMaster Children's Hospital were set up to acquire patient data collection for the Artemis Cloud Database Study*

The consent duration that is examined and analyzed ranged from March 12, 2018 to December 20, 2018. During this time, there were a total of 766 unique admissions to the MCH's NICU. The monthly average was 76.60 ± 6.35 admissions. From those enrolments, A total of 239 NICU patient consents were received during this timeframe. Individual patients were identified by their unique VINES identifier (ID). Several infants who had a short NICU stay were not consented due to logistics of securing consent in a very short timeframe. Figure 13 depicts the total number of unique neonatal infants who were connected to Artemis per month.

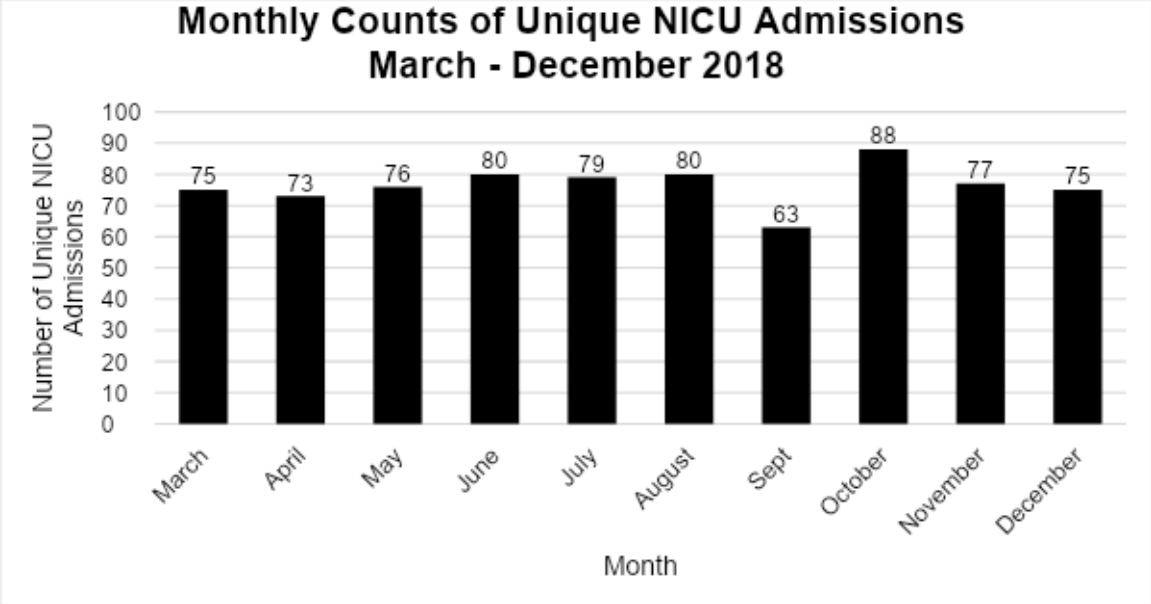


*Figure 15: Monthly counts of unique admissions to MCH's NICU*

Available patient characteristics and an explanation available for analysis are shown in Table 9 along with possible options or examples as applicable:

| Characteristic | Explanation | Options/Examples |
|---|---|---|
| **Vines ID** | Individual infant identifier | |
| **Multiple identifier** | Whether infant was part of a multiple birth | Singletons, twins, triplets |
| **Infant identifier** | Order of siblings in a multiple birth situation | A or B for twins<br>A, B, or C for triplets |
| **Biological sex** | Biological character or quality that is assigned/determined at birth based on an individual's gonadal, morphologic (internal and external), chromosomal and hormonal characteristics (Bockting, 2019; Planned Parenthood, 2020; Stedman's Medical Dictionary, 2004) | Female, Male |
| **Date of birth (DOB)** | Day/Month/Year in DD-Month-YYYY format | 26-Apr-2018 |
| **Gestational age (GA)** | Calculated based on the addition of infant's GA weeks and days | 35 weeks and 2 days |

| | | |
|---|---|---|
| **Date of consent** | Date surrogate consent was obtained for data collection; this was based on the infant's admission to the NICU<br><br>Day-Month-Year in DD-Month-YYYY format | 28-Jun-2018 |
| **Consent option** | Surrogate consenters can select one of the three consent options | **Option 1:** Anonymous data can be used in combination with other similar infant's data without further consent for all future studies that are approved by the institutional research ethics boards.<br><br>**Option 2:** Infant's information stored in this database may be linked using their name, date of birth or health card number with other research data sets, without seeking further consent from surrogate consenter. Future studies that are approved by the institutional research ethics board will be able to use this information. This linked data will be made anonymous before it is used in combination with other similar infants.<br><br>**Option 3:** Consent is not given |

*Table 9: Characteristics available for analysis regarding consented infants for the Artemis Cloud database study*

### 6.5 NICU Infants Characteristics

Of the 239 infants who were consented into the Artemis Cloud Database study, 105 are females and 134 are males (Figure 16 and Table 10). Both graphics show that a higher percentage of NICU patients were consented via Option 2 regardless of the infant's sex.

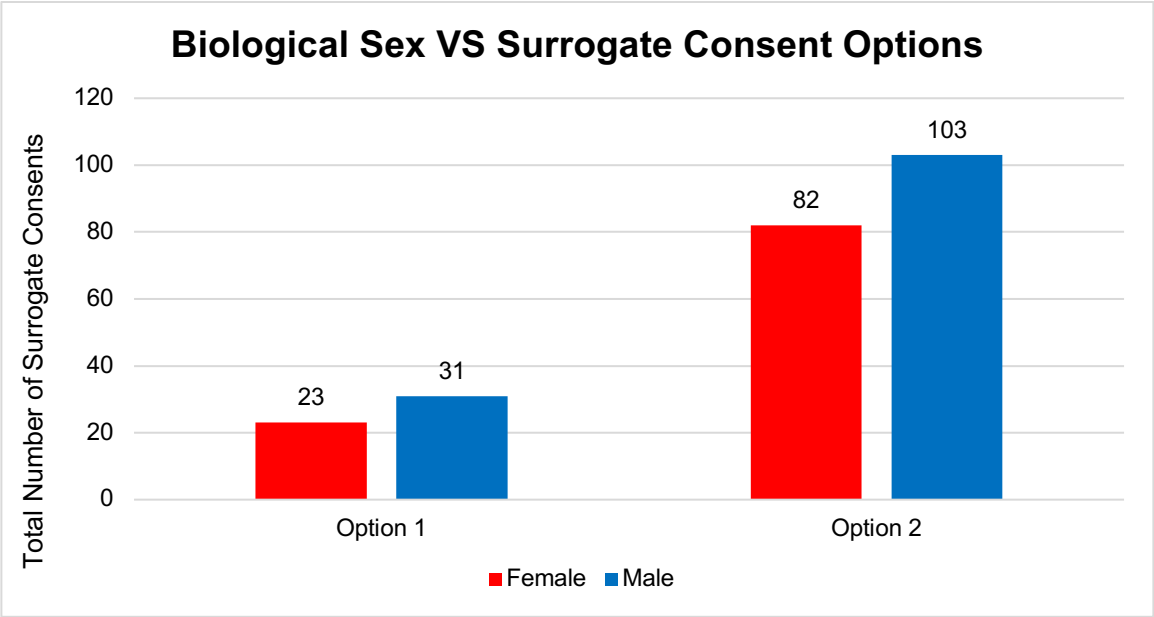


*Figure 16: Breakdown of surrogate consent choices based on infant's biological sex*

| Biological Sex | Option 1 (# and %) | Option 2 (# and %) |
|---|---|---|
| Female | 23 (42.59%) | 82 (44.32%) |
| Male | 31 (57.41%) | 103 (55.68%) |
| | 54 (100%) | 185 (100%) |

*Table 10: Number and percentage of SDM's preferred consent option choice based on biological sex of neonate*

Types of births included singletons, twins and triplets (Figure 16 and Table 11).

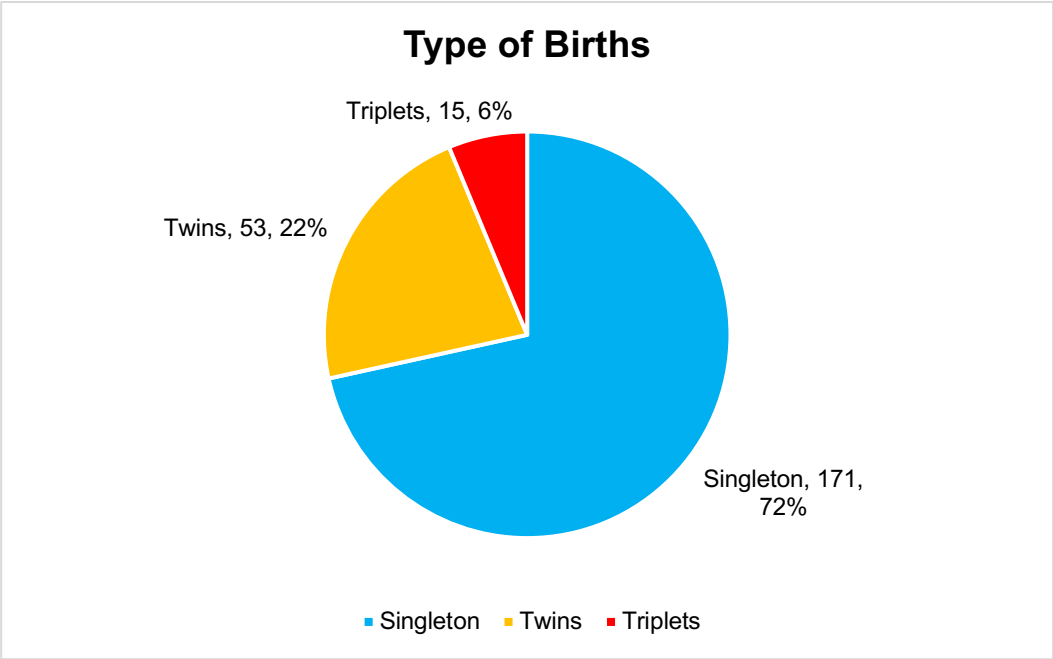


*Figure 16: Breakdown of consented infants by types of births*

| Multiple Identifier | Total Number of Infants |
|---|---|
| Singletons | 171 |
| Twins | 53 |
| Triplets | 15 |
| | 235 |

*Table 11: Breakdown of consented infants by types of births*

The gestational age (GA) of the neonates at birth ranged from 22 weeks and 6 days to 41 weeks and 3 days of age. The mean GA was 228.07 ± 35.24 days. The number of consented infants in each GA classification are summarized as follows in Table 12.

| Gestational Age Classification | Gestational Age Range (Week and Days) | Number consented |
|---|---|---|
| **Extremely preterm** | < 27 weeks and 6 days | 59 |
| **Very preterm** | 28 weeks and 0 days - 31 weeks and 6 days | 42 |
| **Moderate preterm** | 32 weeks and 0 days - 33 weeks and 6 days | 26 |
| **Late preterm** | 34 weeks and 0 days - 36 weeks and 6 days | 57 |
| **Early term** | 37 weeks and 0 days - 38 weeks and 6 days | 27 |
| **Full term** | 39 weeks and 0 days - 40 weeks and 6 days | 23 |
| **Late term** | 41 weeks and 0 days - 41 weeks and 6 days | 5 |
| **Post term** | > 42 weeks | 0 |

*Table 12: GA classification*
*GA ranges confirmed based on (American College of Obstetricians and Gynecologists, 2013 [Reaffirmed 2017]); March of Dimes, n.d.; Raju, 2017; Raju, Higgins, Stark, & Leveno, 2006; Spong, 2013; World Health Organization, 2018)*

As shown in Figure 17, a quarter of the infants consented were extremely preterm (25%). There were no post term infants within the consented population.

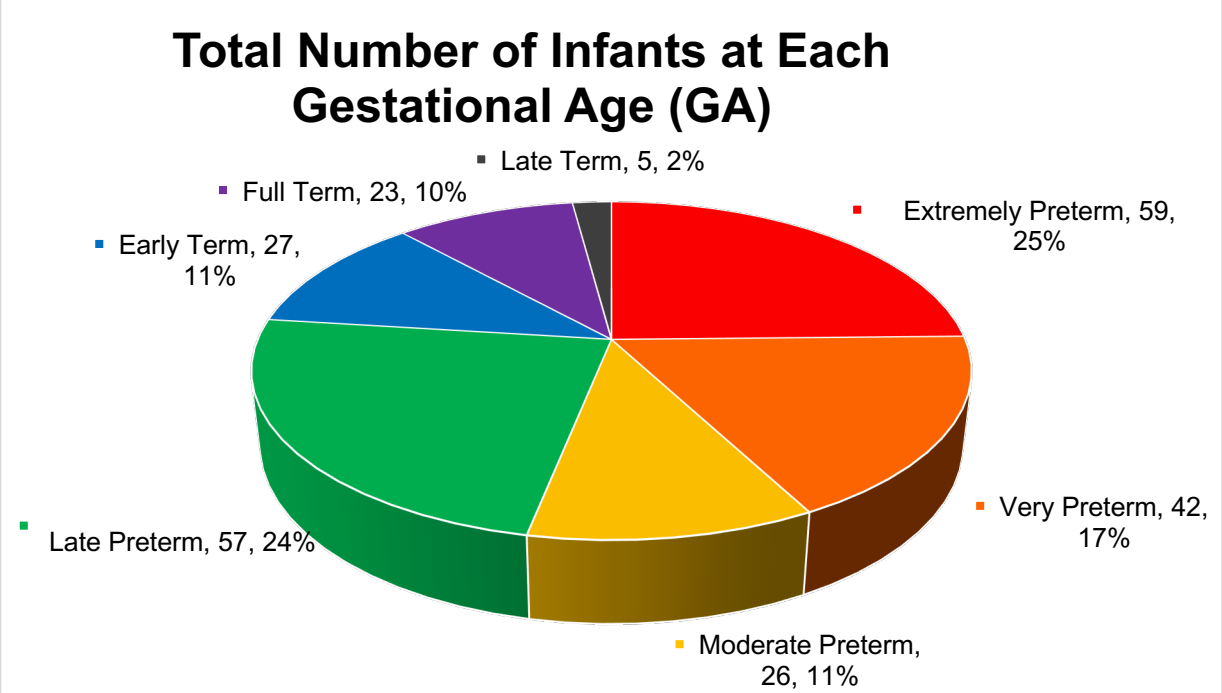


*Figure 17: Breakdown of consented infants by gestational age (GA)*

### 6.6 Surrogate Consenters and Consent Choices

Of the 134 parents or guardians who consented, 40 parents (32.26%) chose consent option 1 and 84 parents (67.74%) chose option 2 (Table 13). Reasons for choosing to decline participation were not captured although the RC noted that in two of these cases, the parents who declined to provide consent were fathers of extremely preterm infants (i.e. between 25-27 weeks of GA).

The surrogate consenters (i.e. parents or guardians) refers to those who provided consent levels 1 or 2 for the 239 study subjects. The majority of parents approached were mothers who were admitted to McMaster University Medical Centre (MUMC) as an inpatient.

| Mother's Status | Total Approached | Total Consented | Total Declined |
|---|---|---|---|
| **MUMC InPatient** | 96 | 89 | 7 |
| **Non-Inpatient** | 49 | 45 | 4 |
| **Total** | **145** | **134** | **11** |

*Table 13: Total number of SDMs who were approached and whether they consented or declined for their infant's enrollment in the Artemis Cloud Database. SDMs (i.e. mothers) were also classified based on their hospital admission status*

Not all surrogate consenters were approached. However, the number of parents not approached and their reasons for declining consent was not tracked. Reasons for not approaching parents included but are not limited to:

- Lack of adequate language skills to give informed consent on the part of the surrogate decision maker
- Consideration of family stress due to post-partum maternal health issues, and/or neonatal instability (including planned withdrawal of neonatal care)
- Family instability requiring Social Work or Children's Aid Society intervention
- Inability to contact parents/guardians
- Rapid transfer, discharge or death of neonate prior to obtaining consent

In the case of multiples, only one consent approach to a parent was made but resulted in the enrolment of multiple subjects. There were no cases where some neonates of the same multiple birth were enrolled and others were not. In this study, the surrogate consenter for a multiple birth made the same consent level preference for all of their infants.

No analysis has been conducted to determine if there were any noticeable characteristics that are predictive of willingness to consent as no information on the consenter was collected.

During the first month of consent in March of 2018, more infants were enrolled under consent Option 1. However, from April onwards, Option 2 became the preferred consent option (Figure 18).

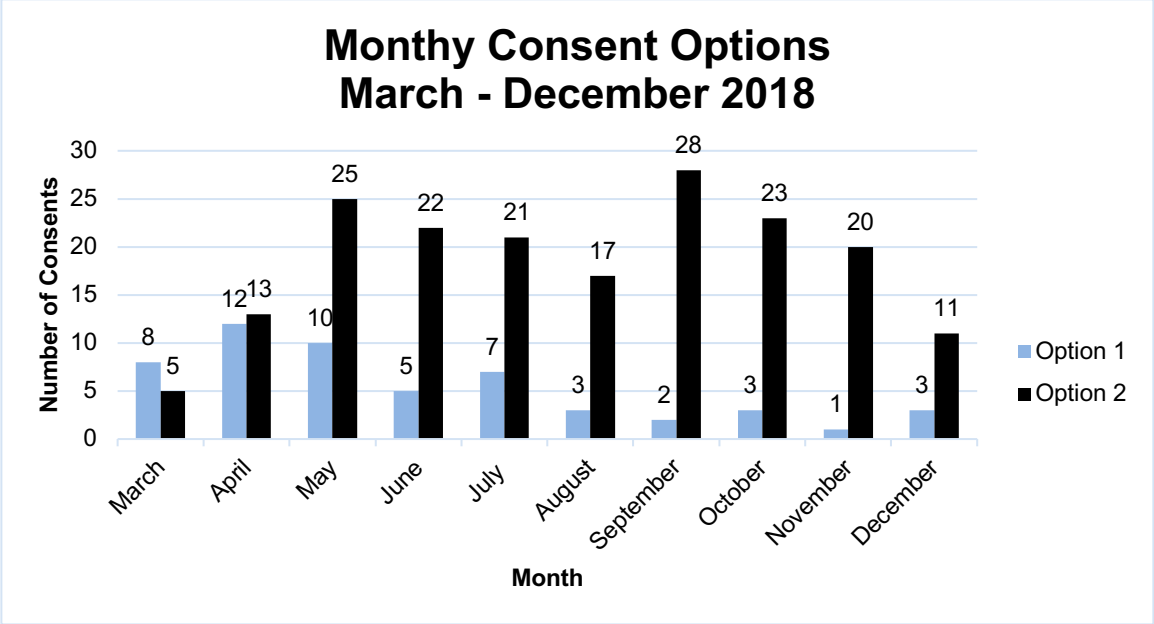


*Figure 18: Total monthly consent options preferences per infant between March to December of 2018*

Due to small number of consents for some categories, it was not appropriate to stratify this monthly data into singleton, twin and triplet monthly analysis. Figure 19 shows the total number of surrogate consents based on the type of birth.

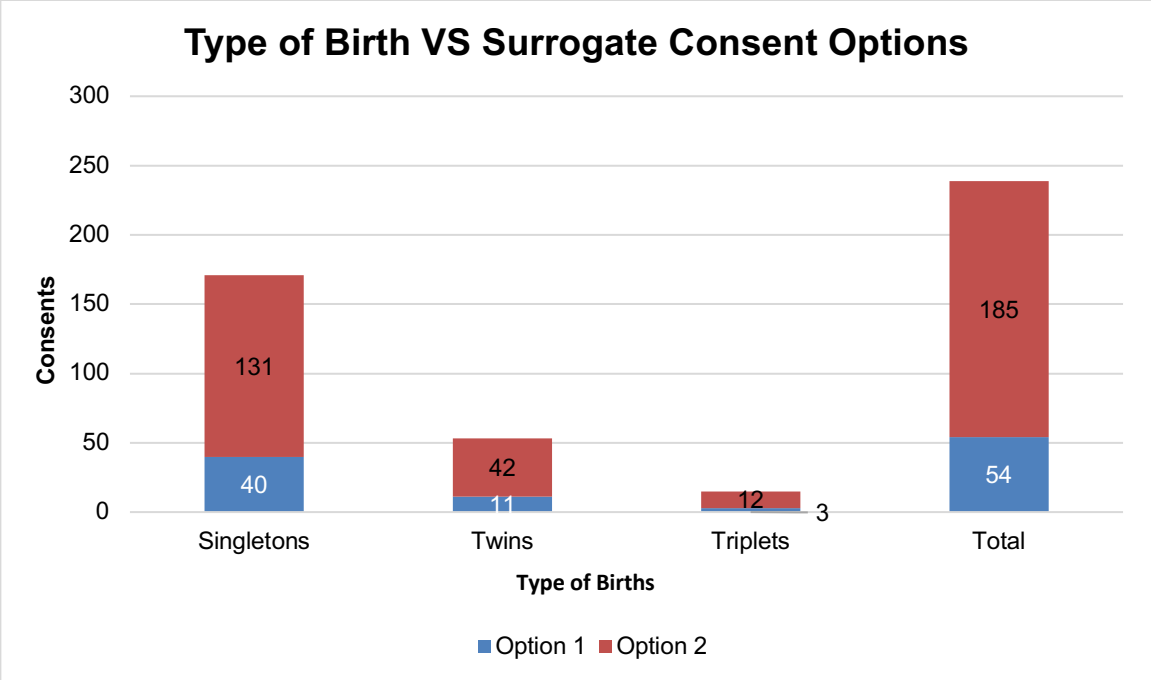


*Figure 19: Total number of surrogate consents based on type of birth*

In all GA classification groups, of those infants who were enrolled in the study, the surrogate consenter's preference for consent Option 2 surpasses that of Option 1 (Table 14). Preference for consent option 1 ranged between 11.11% to 29.63% with an average of 22.59% (54/239) surrogate consenters choosing this option. Preferences for consent option 2 ranged between 70.37% to 88.89% for all patient groups with an average of 77.41% surrogate (185/239) consenters choosing this option (Table 14).

| Gestational Age Classification | Consent Option 1 | Consent Option 2 |
|---|---|---|
| **Extremely preterm** | 22.03% | 77.97% |
| **Very preterm** | 23.81% | 76.19% |
| **Moderate preterm** | 19.23% | 80.77% |
| **Late preterm** | 29.63% | 70.37% |
| **Early term** | 11.11% | 88.89% |
| **Full term** | 18.18% | 81.82% |
| **Late term** | 20.00% | 80.00% |

| | | |
|---|---|---|
| **Post term** | N/A | N/A |

*Table 14: Percentage breakdown of each infant's surrogate consent options based on infant's gestational age (GA) classification*

Although an infant's date of birth and consent dates were captured in the data collection, it was not always clear how long each surrogate consenter required to make a decision about their infant's participation in the Artemis Cloud Database study as some of the neonates enrolled in the study had been in the NICU prior to their enrollment. Figure 20 provides a visual displaying how many weeks it took each surrogate consenter to consent to option 1 or 2. It is clear that the bulk of surrogate consent is achieved during the first week of the infant's life.

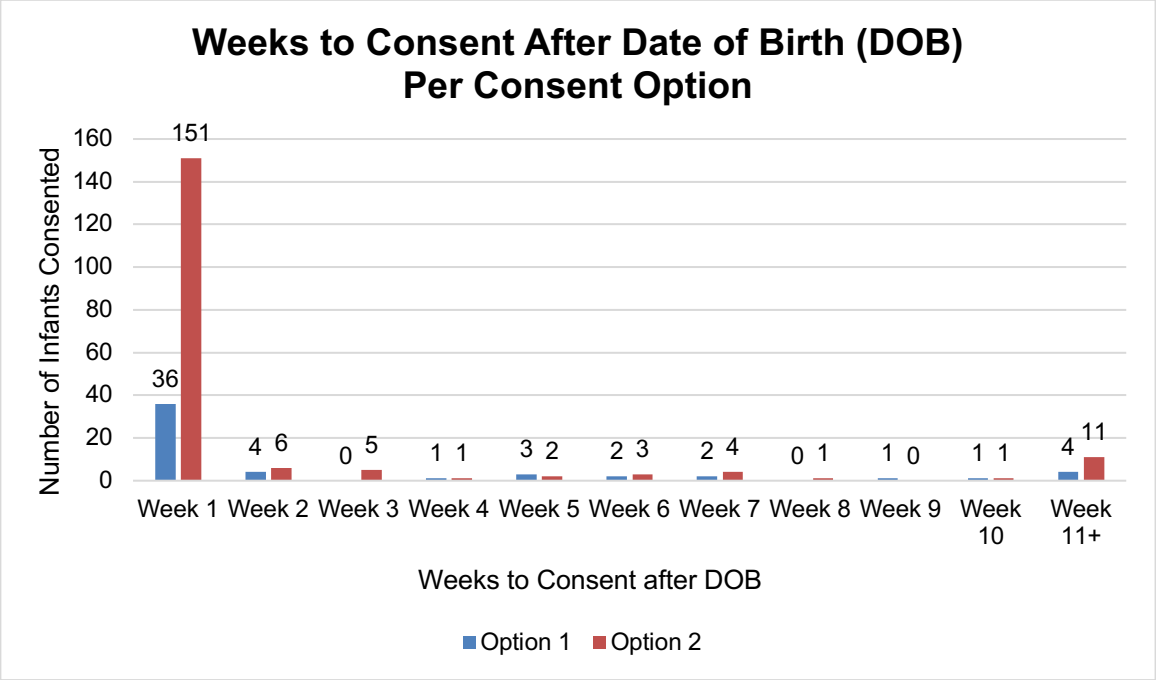


*Figure 20: Number of infants versus which week consent was received after DOB per consent options*

No SDMs retracted consent and withdrew from the study at any point (e.g. during the NICU stay, after discharge or death of their infant).

## 6.7 Application of the CRISP-TDM$^n_0$ Framework in a LONS Case Study

The LONS study will be presented via the use of the CRISP-TDM$^n_0$ framework as detailed:

### 1) Business Understanding

The first phase of CRISP-TDM$^n_0$ is to perform business understanding. This phase involves the identification of the *Clinical Application Domain*, *Clinical Objective* and *Data Mining TA Goals*.

- **Clinical application domain**: The Level 3 NICU located at the McMaster Children's Hospital (MCH), Hamilton, Ontario, Canada
- **Condition:** Late onset neonatal sepsis (LONS)
- **Availability of a high dependency environment:** All infants admitted to the Level 3 NICU at MCH
- **Patient population base:** All infants admitted to the NICU at MCH will be eligible to be enrolled in this research study. There are no exclusion criteria.
- **Clinical objectives:**
    - 1) To install and test an advanced minimal presence, time aligned high frequency physiological data platform in a busy tertiary NICU
    - 2) To gather a time aligned database of high frequency physiological data with time aligned low frequency medical record data for future studies
    - 3) To perform verification and validation of the Artemis LONS algorithms
- **Data mining temporal abstraction (TA) goals**: Perform patient centric heart rate variability (HRV) and respiration rate variability (RRV) temporal abstractions and associate these with the suspicion and diagnosis of LONS

Application for ethics approval was sought for the research at this stage. Approval was received for the data collection study from the Hamilton Health

Sciences REB (HiREB 2859-D) and Ontario Tech University (#14736). Approval was received for the LONS study from the Hamilton Health Sciences REB (HiREB 4833-C) and Ontario Tech University (#15536).

**2) Data Understanding**

The second phase of CRISP-TDM$^{n}_{0}$ is to perform Data Understanding and Data Preparation. Physiological and other clinical data for the LONS study was captured utilizing the Artemis platform.

- **Primary or secondary use of data:** Secondary use of data
- **Data collection period:** Artemis Cloud Database data collection at McMaster Children's Hospital began on March 9, 2018 and the first consent was obtained on March 12, 2018. For purposes of this thesis, data collection until December 31, 2018 was analyzed.
- **Support for real-time data collection:** Artemis allows for simultaneous real-time data collection and data analysis from bedside monitors. Physiological streams are captured by the Knowledge Extraction component of the Artemis platform.
- **Support for distributed time data collection:** The Artemis Cloud platform is capable of supporting multi-centre studies. This research is based on a single site.
- **Data collection regularity:** Physiological data was collected for the duration of the infant's admission within the NICU provided that they occupied a bed space that was connected to the Artemis platform.
- **Data collection frequency:** All data produced by bedside medical equipment at frequencies up to 1000 samples a second are captured in real-time
- **Streams:** De-identified physiological data and clinical data that has been captured in REB approved study (HiREB 2859-D)
- **Conditions:** Clinical info was obtained from electronic medical records (EMR)

- **Consent Data:** Parental/guardian demographic data from consent forms (full name, consenter ID, phone number, email address and full home address).
- **Consent Level:** Inclusion consent levels for this study are: Level 1, Level 2 and Level 3 (see Table 15)
- **Clinical Data:** patient identification (ID), date and time of birth, gestational age (GA) at birth, gender, birth weight, birth length and head circumference at birth. In addition, physiological data such as Heart Rate (HR) captured from the patient's electrocardiogram (ECG), respiratory rate (RR), oxygen saturation ($SPO_2$) values will be collected along with the associated values as well as patient events (i.e. change of bed, transfusion) and will include the patient event, severity value of event and date and time of event.
- **Patients (Subjects):** All NICU patients who are admitted to the NICU at MCH
- **Location:** MCH's NICU
- **Primary inclusion criteria:** All NICU patients who are admitted to MCH
- **Secondary inclusion criteria:** NICU patients whose surrogate decision maker(s) (SDM(s)) have consented to the LONS study

Surrogate consenters (i.e. parents or legal guardians) of all infants who were admitted to an Artemis bed space within the NICU at McMaster Children's Hospital were individually approached by a research coordinator (RC) to allow for the usage of their infant's collected data for future research purposes in the Artemis Cloud Database within fourteen days of their infant's hospital admission. In the event that consent was not obtained within fourteen days of the infant admission to the NICU, then all of the infant's data is not included within the Artemis research database tables

Surrogate consent was concurrently sought for the Artemis Cloud data collection and for the secondary use of data for the LONS clinical study respectively. The physiological and other clinical data that is gathered had no effect

on the clinical management during the study and all infants received the current standard treatment in the NICU.

To facilitate the surrogate consent process for the LONS case study, SDMs (i.e. parents or guardians) were presented with three consent levels (i.e. options) as follows in Table 15:

| Level | Consent Preference |
|---|---|
| 1 | Anonymous data can be used in combination with other similar infant's data without further consent for all future studies that are approved by the institutional research ethics boards. This means that no one will know that the information that is collected came from your infant. |
| 2 | Your infant's information stored in this database may be linked using their name, date of birth or health card number with other research data sets, without seeking further consent from you.<br><br>Future studies that are approved by the institutional research ethics board will be able to use this information. This linked data will be made anonymous before it is used in combination with other similar infants. This means that no one outside of the research team will know that the information that is collected came from your infant. |
| 3 | Consent not given - all buffered data will be deleted from the research system. |

*Table 14: Description of consent preference associated with selection of consent preference level as presented to the infant's SDM*
*(Pugh et al., 2018)*

Consent was manually collected on paper for the LONS study and stored in a locked office at the MCH site (Pugh et al., 2018). This was due to the fixed end date of the project and availability of resources (i.e. the decision not to translate the consent form into different languages). A research coordinator (RC) then manually entered the SDM's chosen consent level preference into an Excel spreadsheet. This resulted in a Comma Separated Value (CSV) file of consent data to load into Artemis.

All patient clinical data were stored under the unique patient's research identification (VINES ID) within the research database. The translation of the patient's hospital number to the VINES ID was held in a separate secure data-table secured away from the medical record (Pugh et al., 2018).

**3) Data Preparation**

As noted in the prior chapter, the Data Preparation step is extended with an additional process step to enable the consent process for data collection for the secondary use of data. In this case, we consider the LONS study portion only and consider the secondary use of the data collected previously by the Artemis Cloud study.

The process begins with identifying the eligible patient population. Selection of the eligible patient population is based on the predetermined inclusion and exclusion criteria of the REB approved study. This requires the use of tables that focus on the patient and patient events (e.g. confirmation of a diagnosis). For this specific LONS study, there were no exclusion criteria and all neonates were eligible for study enrollment.

This study data is then populated within the related study consented data as the study $S_{LONS}$ instance of the Study Consent Data. Data for patients with consent level 1 would be automatically loaded into the consented datamart instance as study $S_{LONS}$ instance of a Consented Relative Temporal Data subset. Data for patients with consent level 2 who have indicated 'YES' would be automatically loaded into the consented datamart instance as study $S_{LONS}$ instance of a Consented Relative Temporal Data subset. This requires the use of tables that contain their temporal data as abstracted to create HRV and respiration rate variability (RRV) temporal abstractions and for those that were suspected of developing LONS, the relative alignment of that data as a relative distance from the LONS suspicion event. The time of suspicion is chosen in this case rather than the time of confirmation of LONS as confirmation can sometimes be as much as 24 hours after the time of suspicion when the blood was drawn for testing for LONS.

**4) Data Modeling**

The Data Evaluation phase of CRISP-TDM$^n_0$ model step is beyond the scope of this thesis.

**5) Evaluation**

The second last phase of CRISP-TDM$^n_0$ model is the evaluation phase. The broad aim is to evaluate the clinical algorithm, software and consent model. The clinical and technical evaluation is beyond the scope of this thesis.

**6) Deployment**

The final phase of CRISP-TDM$^n_0$ is deployment. This stage is outside the scope of this thesis.

### 6.8 Application of Extended STDM$^n_0$ to a Retrospective LONS Case Study

In the previous chapter, five new tables were proposed to be added to the current STDM$^n_0$ framework to capture the necessary data required to support the flexible consent process with additional functionality for that to be provided by a surrogate. These tables were related to the identity of the consenter, the linking of the consenter to the patient, the selected consent level preference, the extraction of the patient for a study based on consent level and the linking of the patient to specific research studies based on the selected consent level associated with the patient.

The following steps demonstrate how the extended STDM$^n_0$ can be used to select potential neonatal patients who are eligible for a retrospective study from a researcher's perspective.

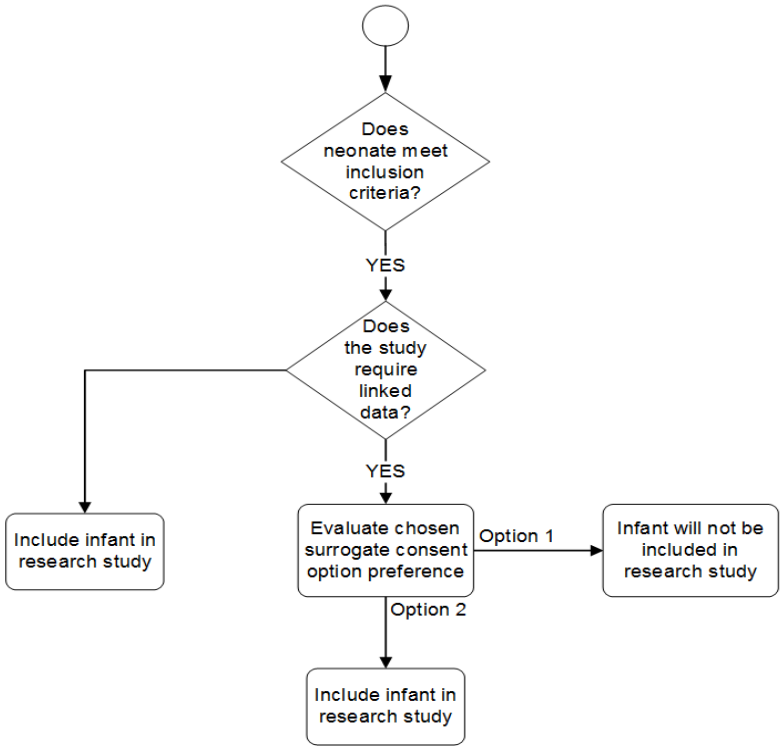


*Figure 21: How to filter for potential neonatal patients who are eligible for a LONS retrospective study from a researcher's perspective.*

## 6.9 Summary

This research demonstrates the use of the consent model proposed within the context of a retrospective research study to investigate whether there are common patterns in physiological data before the clinical suspicion and diagnosis of late LONS.

# 7. Discussion

This chapter discusses issues related to the results presented in chapter 6, within the context of the additional functionalities provided by the *Consent of Infants for Neonatal Secondary-use research* (CoINS) model. Here we further discuss how consent, and research data made accessible by that consent, are coupled within the Service Based Multidimensional Temporal Data Mining Framework ($STDM^n_0$) platform and how these are implemented by Artemis in a systemic platform that supports multiple concurrent research studies. This chapter also highlights changes made to the study protocol and any identified limitations.

## 7.1 Flexible Consent

CoINS offers multi-level participant consent allowing surrogate decision-makers (SDMs) to select consent preferences from a potentially wide range of options. Within the Artemis Cloud Database study, while consent option 2 was the most popular option selected by SDMs, there are still those who selected option 1, or who chose not to consent at all. This emphasizes the importance of providing parents with the ability to choose among a variety of available consent options as a 'one size fits all' approach to the consent process does not exist. Additionally, the high consent rates support results previously seen in adult studies in that many SDMs also accept that secondary use of their personal EHR data in research contributes positively to the community (Hill et al, 2013), and generally see health research as valuable and unproblematic (Vermeulen et al, 2009). Also, and perhaps most notably, that they appreciate an opportunity to decide on research participation on behalf of their infant, thus emphasizing the importance of the flexibility functionality of the CoINS model. This also reinforces prior research findings that most parents were displeased when the consent requirement is supplanted by blanket approvals for distribution of health data made by an ethics committee, leading to enrollment of children in studies absent disclosure or consent of the responsible parent (Singhal et al., 2002; Stenson et al., 2004) and

that they did not want their newborn to be enrolled in studies without their consent, even if it was considered to be low risk, such as that of a data collection study.

### 7.2 Surrogate Consent

In order to capture surrogate consent, the research system is required to capture the consenter's information and their chosen consent preference, which must both be linked to the record of the individual patient.

### 7.2 Longitudinal Consent

CoINS provides an ability to capture longitudinal consent, that is, consent which can be altered by the participant over time as their circumstances and preferences change. Requesting additional consent for future studies is costly in terms of time, effort and resources. There can also be a significant loss of participants over time as people's contact details change (Kaye et al., 2015). With CoINS, SDMs are offered the option to provide broad consent to allow anonymous data linking for all future studies that are approved by institutional research ethics boards (REBs), meaning that should they so choose, re-consent with each new research study would not be necessary. Some SDMs may choose to update their consent preferences throughout the life of the Artemis Cloud Database study, as the data is consented for use until the child attains 18 years of age. CoINS enables recording of these updated consent preferences and ensures the most recent level of consent is applied when assessing whether an infant's data may be included in each new component of the wider study.

### 7.3 Coupling Consent and Research Data

Within the Artemis database, consent and research data is linked, together with the SDM's contact information, with the patient record during the consent process. In that way, the SDM's consent preferences are indelibly linked to the patient. Whenever an approved researcher filters for potential research participants, the individual neonatal patient is either included as part of the candidate pool or excluded based on the recorded consent.

### 7.4 Coupling Consent and Research Data within an IT Solution

There are numerous advantages of implementing an IT-based consent solution. All participant information (including consent preferences) are conveniently stored within an accessible interface for researchers and SDMs. Specific consent provisions travel with a participant or donor's data and samples as they are shared or accessed for different purposes – in the case of this thesis, for different studies. These consent provisions can be electronically and cryptographically "wrapped" with the donor's samples and information (Kaye et al., 2015). Should the SDM wish to retract consent, their infant's data can immediately be excluded from the research system whilst still remaining in its original location in the clinical system, thus not affecting ongoing treatment.

Furthermore, as a systematic IT platform already exists, future improvements to the proposed consent model can allow for development and incorporation of features that extend the solution with broader consent options. For example, an ability to filter for and choose one's preferred categories of research, for example: clinical trials, genetic research, public health research, observational studies, research related to particular medical conditions or illnesses. Other examples might include allowing the choice to participate in research conducted by commercial organizations, or to request general research results either as a simple 'thank you' acknowledgment for their contribution and involvement, or demonstrating how their samples and information have been useful (Kaye et al., 2015) thus encouraging future participation and engendering transparency between researchers and participants. This would improve the public's understanding and trust for secondary use research using collected health data and samples and could potentially lead to higher consent rates.

### 7.5 Changes to and Limitations of the Study

It was initially proposed that four consent levels be offered to the neonate's SDM. This would have included all of the options as shown under the "consent option" in Table 12 along with the choice that "the parent/guardian of the infant wishes to be contacted and consented each time their infant's data is to be used

in a new research study". However, McMaster Children's Hospital (MCH)'s neonatal research committee deemed that offering such a broad degree of individual consent was far too labor intensive to implement for the entire study. They proposed to conduct a separate study assessing the percentage of uptake should this option be available. Removal of this option meant it was not possible to implement and evaluate the original flexible consent model. However, results still showed parents were willing to contribute to this research.

As discussed in Section 6.3, there was no standardized script and subtle changes were made by the research coordinator (RC) during the consent approach explanation to surrogate decision makers (SDMs). Initially when presenting the study to parents, the RC focussed on explaining why two 'yes to consent' options were available, and emphasis was placed on the privacy and confidentiality of data collection in a research ethics context as part of that explanation. The changes involved explaining that the second option was a way to better facilitated longitudinal research (e.g. for purposes of doing future research on what happens to MCH's NICU infants). The changes resulted from the RC noticing SDMs' slight confusion between the two available consent options as they were not as aware that data usage in option two was still anonymous despite linkage in the database to ensure consent (i.e. it is emphasized that results dissemination is still anonymous). Consequently, by providing the information in more relatable and easily understood terms, it was felt this change would result in increased participation via the second consent option. This hypothesis has proven difficult to substantiate as, by default, option 2 was already the popular option (Figures 18 and 20). It is unfortunate that the date of when this change in the process was applied was not tracked. For this reason, consent option selection has been longitudinally analyzed to monitor for changes in distribution of consent option as they occurred, and which might be attributed to this change in recruitment approach. Apart from choosing between consent options, no parents have made requests to exclude specific types of data or asked for specific data to be removed after it had been collected.

### 7.6 Limitations Regarding Data Collection

Most results from the data collected during this study reflect the demographics of the infants who were consented for the study and do not actually reflect whether SDMs would use the CoINS model or demonstrate the nuances of how the proposed consent model would function in reality. Additionally, having many multiple births in a month can also skew the data as SDMs tended to choose the same consent level for all infants over which they exercise responsibility. Many neonates are admitted to the NICU after premature births, which puts them at greater risks for congenital or emergent complications. To better understand users of the CoINS model, it would be useful to examine the range of medical conditions these neonates were diagnosed with to consider whether their condition affects their SDM's decision for participation in research.

Either parent's written consent was all that was necessary to enrol a neonate in the Artemis Cloud Database study. It is unclear if only birth mothers were approached or if fathers (or a second partner) consented. The rationale for this was because it was not uncommon for only one parent (usually the mother) to have sole care and custody of the child, which meant consent from a second parent was believed to not be necessary. In situations where neonates are in custody of the Children's Aid Society (CAS) or Catholic Children's Aid Society (CCAS), it was technically possible to seek consent for their participation in research studies. However, this is often discouraged by the research ethics board for two reasons relating to: (a) potential legal implications should the custody status of the child change in the future, and (b) practicality issues, as it is usually very difficult to arrange to obtain consent from case or social workers. In the event that an infant had two parents it is not known whether there were situations where one parent consented for one infant while the second parent consent for the sibling. Certain data was not tracked, including the number of parents not approached and the reasons of those parents who declined consent to participate. It is also unclear from the data if there were SDMs who provided and later revoked consent, and if so, how soon revocation occurred. If any reasons for revoking consent are provided that relate to the consent model or process, these could have been

significant to processes for further improvement of the consent model and process. Similarly, it is also unclear if any SDMs declined consent when approached initially, but later went on to consent for the usage of their infant's data. Should SDMs have chosen to change their consent preference, it would have been demonstrative of the practicality of CoINS to accommodate such flexibility. These data should be collected in any future research of consent models.

# 8. Conclusion

This thesis presented the *Consent of Infants for Neonatal Secondary-use research* (CoINS) model, which we believe is the first proposed flexible and longitudinal consent model targeted to the secondary use of physiological data for medical research studies and incorporating the ability for surrogate consent. CoINS offers a range of consumer engagement features as drawn from the PF4SUMD model (Heath, 2012), and incorporating the flexibility of MC. The CoINS consent model can be integrated and instantiated within a research database of the Artemis platform. CoINS is suitable for use in any healthcare environment supporting secondary use of patient data in research studies including those seeking to analyse streaming data collected from physiometric sensors. Use of CoINS was demonstrated as part of ethically approved secondary use research where surrogate consent was collected using the application for the retrospective *Late Onset Neonatal Sepsis* (LONS) study at McMaster Children's Hospital

## 8.1. Thesis Contributions

This thesis presents the following contributions:

1) An up-to-date review of existing consent models;
2) The CoINS consent model;
3) Demonstration and evaluation of CoINS applied to a retrospective study in a neonatal intensive care environment.

The following research objectives were addressed in the chapters:

1) Creation of CoINS, a flexible multi-level consent model enabling the research participant to adapt consent preferences as their wishes change, along with the ability to do the same with regard to surrogate consent for minor children in their care (chapter 5);
2) Development of an approach for integrating CoINS within databases supporting secondary use of data in research studies, including analysis of streams of data collected from physiometric sensors attached to the patient (chapter 5)

3) Instantiate that model within a research database (chapter 5)
4) Demonstrate the use of that model within an ethically approved secondary use of data research study where surrogate consent is required (chapter 6)

### 8.2 Future work

Additional functionality to allow consenters to pick "how" and "when" consent is given in a more holistic way may be developed to further develop CoINS. Given the relative recency of the Artemis Cloud Database study, the consent options offered presently only apply to prospective, not retrospective, consent, it will take time to see how many SDM's will take advantage of the ability to alter their consent retrospectively. Additionally, the existing CoINS model is intended for use in a single site setting and not tested in a multi-site environment. However, successful implementation of any novel or non-traditional consent model will require change management and the cooperation and feedback of the responsible research ethics board, IT department and researchers who will use the collected data. End users (e.g. patients and substitute decision makers) contributing their neonate's/neonates' data would also need to be engaged in this process. We propose that Patient and Public Involvement (PPI) sessions would be required to ensure that developers of these solutions understand and anticipate how non-clinical users engage with these systems.

Many neonates have two parent SDMs, and it is possible that one parent may agree while the other disapproves providing consent to allow the use of their child's clinical data in research. A multidisciplinary team including bioethicists, lawyers, PPI representatives and clinical staff would be required to develop protocols to address this and other exceptional situations as and when they arise.

Some related topics and issues did not fall within the scope of this thesis and were not examined. These include approaches for informed consent and the ongoing ethical issues of 'ownership' of the clinical record and 'monetising' of what may be valuable consented or *donated* data. It may also be difficult to know whether SDMs who consented inclusion of their infant's data actually gave

*informed* consent and their degree of comprehension regarding the purpose for undertaking the study and uses to which their child's data may be applied.

CoINS is capable of expansion beyond serving as a consent model that identifies and links patients and SDMs, potential research participants, and researchers who may require their data. After successful implementation of a modernised consent model, the model should be expanded to become a platform that promotes communication and transparency of research and results to SDMs (and children as they age and achieve Gillick competence).

There is little doubt that when a person or their child requires hospitalization, it can be an extremely distressing time for loved ones. Ethical and emotional questions regarding *appropriateness* must also be considered: *When is the most appropriate time to introduce parents to the CoINS concept; when should be approached regarding consent to access to their or their child's health data*; and *whether parents of stillborn or neonates who have died during the NICU stay should be approached at all*.

Informed consent is a dynamic process that requires the engagement and partnership of multiple stakeholders including the donor or consenter, collectors, and users of the data. While the donor and/or consenter may be the same individual in the case of a competent adult, this is not the case in the NICU where a surrogate consenter is required. A flexible, longitudinal consent model that encourages parental autonomy that caters to proxy consent such as surrogate consent and leverages the use of information technology (IT) tools that are already collecting streams of data collected from patient care sensors would potentially encourage more surrogate decisions makers to enroll their infants in research studies, thus enriching current evidence-based medical knowledge.

# REFERENCES


Adler-Milstein, J., & Jha, A. K. (2013). Healthcare's "Big Data" Challenge. *The American Journal of Managed Care, 9*(7), 537-538.

Albersheim, S. G., Lavoie, P. M., & Keidar, Y. D. (2010). Do neonatologists limit parental decision-making authority? A Canadian perspective. *Early Hum Dev., 86*(12), 801-805. doi:10.1016/j.earlhumdev.2010.09.007

Allmark, P., & Mason, S. (2006). Improving the quality of consent to randomised controlled trials by using continuous consent and clinician training in the consent process. *Journal of medical ethics, 32*(8), 439-443. doi:10.1136/jme.2005.013722

Amarasingham, R., Patzer, R., Huesch, M., Nguyen, N. Q., & Xie, B. (2014). Implementing electronic health care predictive analytics: considerations and challenges. *Health Affairs, 33*(7), 1148-1154. doi:10.1377/hlthaff.2014.0352

American Cancer Society. (2019). What Is Informed Consent? Retrieved from https://www.cancer.org/treatment/finding-and-paying-for-treatment/understanding-financial-and-legal-matters/informed-consent/what-is-informed-consent.html

Anderson, N., Bragg, C., Hartzler, A., & Edwards, K. (2012). Participant-centric initiatives: Tools to facilitate engagement in research. *Applied & Translational Genomics, 1*(1), 25-29. doi:10.1016/j.atg.2012.07.001

Angell, M. (2005). *The truth about the drug companies: How they deceive us and what to do about it*. New York, United States: Random House Trade Paperbacks.

Article 29 Data Protection Working Party. (2015). ANNEX - health data in apps and devices. Retrieved from https://ec.europa.eu/justice/article-29/documentation/other-document/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf

Ballard, H. O., Shook, L. A., Desai, N. S., & Anand, K. J. S. (2004). Neonatal research and the validity of informed consent obtained in the perinatal period. *Journal of Perinatology, 24*, 409-415. doi:10.1038/sj.jp.7211142

Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs, 33*(7), 1123-1131. doi:10.1377/hlthaff.2014.0041

Beauchamp, T. L. (2011). Informed consent: Its history, meaning, and present challenges. *Cambridge Quarterly of Healthcare Ethics, 20*(4), 515-523.

Black, N. (2003). Secondary use of personal data for health and health services research: Why identifiable data are essential. *Journal of Health Services Research & Policy, 8*(suppl 1), 36-40. doi:10.1258/135581903766468873

Blount, M., Ebling, M. R., Eklund, J. M., James, A. G., McGregor, C., Percival, N., . . . Sow, D. (2010). Real-time analysis for intensive care: development and deployment of the artemis analytic system. *IEEE Engineering in Medicine and Biology Magazine, 29*(2), 110-118. doi:10.1109/MEMB.2010.936454

Bockting, W. O. (2019). Nelson Textbook of Pediatrics E-Book. In R. M. Kliegman & J. St. Geme (Eds.), *Nelson Textbook of Pediatrics* (pp. 1021). Canada: Elsevier Canada.

Bosk, C. L. (2002). Obtaining Voluntary Consent for Research in Desperately Ill Patients. *Medical Care, 40*(9 Suppl), V-64-V68.

Burgess, E., Singhal, N., Amin, H., McMillan, D. D., & Devrome, H. (2003). Consent for clinical research in the neonatal intensive care unit: a retrospective survey and a prospective study. *Archives of Disease in Childhood. Fetal and Neonatal Edition., 88*(4), F280–286.

Canadian Institutes of Health Research. (2005). *CIHR Best Practices for Protecting Privacy in Health Research*.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council. (2018). *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2* Ottawa, Canada

Canadian Paediatric Society. (2004). Treatment decisions regarding infants, children and adolescents. *Paediatrics & Child Health, 9*(2), 99-103.

Cardona-Morrell, M., Prgomet, M., Lake, R., Nicholson, M., Harrison, R., Long, J., . . . Hillman, K. (2016). Vital signs monitoring and nurse-patient interaction: A qualitative observational study of hospital practice. *International Journal of Nursing Studies, 56*, 9-16. doi:10.1016/j.ijnurstu.2015.12.007

Catley, C., Smith, K., McGregor, C., & Tracy, M. (2009). *Extending CRISP-DM to incorporate temporal data mining of multi-dimensional medical data streams: A neonatal intensive care unit case study.* Paper presented at the 22nd IEEE International Symposium on Computer-Based Medical Systems (CBMS 2009).

Caulfield, T. (2002). Gene banks and blanket consent. *Nature Reviews Genetics, 3*(8), 577.

Caulfield, T., Upshur, R. E. G., & Daar, A. (2003). DNA databanks and consent: A suggested policy option involving an authorization model. *BMC medical ethics, 4*(1). doi:10.1186/1472-6939-4-1

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, W. (2000). CRISP-DM 1.0.

Cios, K. J., & Moore, G. W. (2002). Uniqueness of medical data mining. *Artificial Intelligence in Medicine, 26*(1-2), 1-24.

Coggins, S. A., Weitkamp, J. H., Grunwald, L., Stark, A. R., Reese, J., Walsh, W., & Wynn, J. L. (2016). Heart rate characteristic index monitoring for bloodstream infection in an NICU: a 3-year experience. *Archives of Disease in Childhood. Fetal and Neonatal Edition., 101*(4), F329-332. doi:10.1136/archdischild-2015-309210

Cohen, A. B., Trentalange, M., & Fried, T. (2015). Patients with next-of-kin relationships outside the nuclear family. *Journal of the American Medical Association, 313*(13), 1369-1370.

Cohen-Wolkowiez, M., Moran, C., Benjamin, D. K., Cotten, C. M., Clark, R. H., Benjamin, D. K., Jr., & Smith, P. B. (2009). Early and late onset sepsis in late preterm infants. *The Pediatric Infectious Disease Journal, 28*(12), 1052-1056.

Committee on Bioethics. (1995). Informed consent, parental permission, and assent in pediatric practice. *Pediatrics, 95*(2), 314-317.

Considine, J., Trotter, C., & Currey, J. (2016). Nurses' documentation of physiological observations in three acute care settings. *J Clin Nurs, 25*(1-2), 134-143. doi:10.1111/jocn.13010

Cooke, R. W. I. (2005). Good practice in consent. *Seminars in Fetal and Neonatal Medicine, 10*(1), 63–71. doi:10.1016/j.siny.2004.09.009

Covington & Burling LLP. (2015). Article 29 Working Party Clarifies Scope of Health Data in Apps and Devices. *Inside Privacy.* Retrieved from https://www.insideprivacy.com/international/article-29-working-party-clarifies-scope-of-health-data-in-apps-and-devices/

Crnkovic, G. D. (2010). Constructive Research and Info-computational Knowledge Generation. *Model-Based Reasoning in Science and Technology. Studies in Computational Intelligence., 314*, 359-380.

Cunningham, S., O'Doherty, K. C., Senecal, K., Secko, D., & Avard, D. (2015). Public concerns regarding the storage and secondary uses of residual newborn bloodspots: an analysis of print media, legal cases, and public engagement activities. *J Community Genet, 6*(2), 117-128. doi:10.1007/s12687-014-0206-0

Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social science & medicine, 64*(1), 223-235. doi:10.1016/j.socscimed.2006.08.045

de Beer, N. A. M., Andriessen, P., Berendsen, R. C., Oei, S. G., Wijn, P. F., & Oetomo, S. B. (2004). Customized spectral band analysis compared with conventional Fourier analysis of heart rate variability in neonates. *Physiological Measurement, 25*(6), 1385-1395. doi:10.1088/0967-3334/25/6/004

Deeny, S. R., & Steventon, A. (2015). Making sense of the shadows: priorities for creating a learninghealthcaresystem based onroutinely collected data. *BMJ Quality & Safety 2015, 24*(8), 505-515. doi:10.1136/bmjqs-2015-004278

Dhanoa, A. (2011). *SDTDM$^{n}_{0}$: A multidimensional distributed data mining framework supporting time series data analysis for critical care research* University of Ontario Institute of Technology, Oshawa, Ontario.

Donchin, Y., & Seagull, F. J. (2002). The hostile environment of the intensive care unit. *Current opinion in critical care, 8*(4), 316-320.

Dong, Y., & Speer, C. P. (2015). Late-onset neonatal sepsis: recent developments. *Archives of Disease in Childhood. Fetal and Neonatal Edition, 100*(3), F257-263. doi:10.1136/archdischild-2014-306213

Eklund, J. M., Fontana, N., Pugh, E., McGregor, C., Yielder, P., James, A., . . . McNamara, P. (2014). *Automated sleep-wake detection in neonates from cerebral function monitor signals.* Paper presented at the IEEE 27th International Symposium on Computer-Based Medical Systems, New York, NY, USA.

El Emam, K. (2010). Risk-based de-identification of health data. *IEEE Security & Privacy, 8*(3), 64-67.

El Emam, K., Buckeridge, D., Tamblyn, R., Neisa, A., Jonker, E., & Verma, A. (2011). The re-identification risk of Canadians from longitudinal demographics. *BMC Medical Informatics and Decision Making, 11*(46). doi:10.1186/1472-6947-11-46

Fabres, J., Carlo, W., Phillips, V., Howard, G., & Ambalavanan, N. (2007). Both extremes of arterial carbon dioxide pressure and the magnitude of fluctuations in arterial carbon dioxide pressure are associated with severe intraventricular hemorrhage in preterm infants. *Pediatrics, 119*(2), 299-305. doi:10.1542/peds.2006-2434

Fairchild, K. D. (2013). Predictive monitoring for early detection of sepsis in neonatal ICU patients. *Current Opinion in Pediatrics, 25*(3), 172-179.

Fairclough, S. (2014). Physiological data must remain confidential. *Nature, 505*(7483), 263.

Fischbach, F. T. (1991). *Documenting Care Communication, the Nursing Process and Documentation Standards*. Philadelphia, PA, USA: F.A. Davis.

Fisher, M., Gabbay, D., & Vila, L. (2005). *Handbook of Temporal Reasoning in Artificial Intelligence*: Elsevier.

Flaherty, E. (2017). AGS Sets Sights on Better Care, More Responsive Policies for “Unbefriended” Older Adults. *Journal of Gerontological Nursing, 43*(3), 51-52.

Foley, T., & Fairmichael, F. (2015). Site visit to Geisinger Health System. Retrieved from http://www.learninghealthcareproject.org/publication/6/63/site-visit-togeisinger-health-system

Goldacre, B. (2014). *Bad pharma: how drug companies mislead doctors and harm patients*: Macmillan.

Golec, L., Gibbins, S., Dunn, M. S., & Hebert, P. (2004). Informed consent in the NICU setting: an ethically optimal model for research solicitation. *Journal of Perinatology, 24*(12), 783–791. doi:10.1038/sj.jp.7211198

Graham, I. D. (2012). Knowledge synthesis and the Canadian Institutes of Health Research. *Systematic Reviews, 1*(6). doi:10.1186/2046-4053-1-6

Grande, D., Mitra, N., Shah, A., Wan, F., & Asch, D. A. (2014). The importance of purpose: Moving beyond consent in the societal use of personal health information. *Annals of Internal Medicine, 161*(12), 855-862.

Griffin, M. P., & Moorman, J. R. (2001). Toward the early diagnosis of neonatal sepsis and sepsis-like illness using novel heart rate analysis. *Pediatrics, 107*(1), 97-104.

Griffin, M. P., O’Shea, T. M., Bissonette, E. A., Harrell, F. E., Lake, D. E., & Moorman, J. R. (2003). Abnormal heart rate characteristics preceding neonatal sepsis and sepsis-Like Illness. *Pediatric Research, 53*(6), 920-926. doi:10.1203/01.PDR.0000064904.05313.D2

Harth, S. C., Johnstone, R. R., & Thong, Y. H. (1992). The psychological profile of parents who volunteer their children for clinical research: a controlled study. *Journal of medical ethics, 18*, 86-93.

Heath, J. (2006). *A framework for an intelligent decision support system (IDSS), including a data mining methodology, for fetal-maternal clinical practice and research.* Western Sydney University,

Heath, J., & McGregor, C. (2010). *CRISP-DM0: A method to extend CRISP-DM to support null hypothesis driven confirmatory data mining*. Paper presented at the Advances in Health Informatics Conference.
Heath, J. A. (2012). *A privacy framework for secondary use of medical data.* (Doctor of Philosophy). University of Wollongong, Retrieved from http://ro.uow.edu.au/theses/3667/
Herland, M., Khoshgoftaar, T. M., & Wald, R. (2014). A review of data mining using big data in health informatics. *Journal of Big Data, 1*(2). doi:10.1186/2196111512
Hildebrandt, M., & Gutwirth, S. (2008). *Profiling the European Citizen: Cross-Disciplinary Perspectives*: Springer.
Hill, E. M., Turner, E. L., Martin, R. M., & Donovan, J. L. (2013). "Let's get the best quality research we can": public awareness and acceptance of consent to use existing data in health research: a systematic
review and qualitative study. *BMC Medical Research Methodology, 13*(72).
Huang, W., McGregor, C., & James, A. (2014). *A Comprehensive Framework Design for Continuous Quality Improvement within the Neonatal Intensive Care Unit: Integration of the SPOE, CRISP-DM and PaJMa Models*. Paper presented at the IEEE Biomedical and Health Informatics Conference, Valencia, Spain.
Humer, C., & Finkle, J. (2014). Your medical record is worth more to hackers than your credit card. Retrieved from https://www.reuters.com/article/us-cybersecurity-hospitals/your-medical-record-is-worth-more-to-hackers-than-your-credit-card-idUSKCN0HJ21I20140924
Institute for Healthcare Improvement. (n.d.). Triple aim for populations. Retrieved from http://www.ihi.org/Topics/TripleAim/Pages/Overview.aspx
International Biometric Group. (2010, 2010). BioPrivacy FAQ's and definitions. Retrieved from http://www.bioprivacy.org/
Israel, S. A., Irvine, J. M., Cheng, A., Wiederhold, M. D., & Wiederhold, B. K. (2005). ECG to identify individuals. *Pattern Recognition, 38*(1), 133–142. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0031320304002419
Iyer, P., & Camp, N. M. (1995). *Nursing Documentation: A Nursing Process Approach* (2nd edn. ed.). St. Louis, MO, USA: Mosby-Year Book.
Jones, D. A., Shipman, J. P., Plaut, D. A., & Selden, C. R. (2010). Characteristics of personal health records: Findings of the medical library association/national library of medicine joint electronic personal health record task force. *Journal of the Medical Library Association (JMLA), 98*(3), 243-249.
Kasanen, E., Lukka, K., & Siitonen, A. (1993). The Constructive Approach in Management Accounting Research. *Journal of Management Accounting Research, 5*, 243-264.
Kass, N. E., Natowicz, M. R., Hull, S. C., Faden, R. R., Plantinga, L., Gostin, L. O., & Slutsman, J. (2003). The use of medical records in research: What do patients want? *The Journal of Law, Medicine & Ethics, 31*(3), 429–433. doi:10.1111/j.1748-720X.2003.tb00105.x
Kaye, J., Curren, L., Anderson, N., Edwards, K., Fullerton, S. M., Kanellopoulou, N., . . . Winter, S. F. (2012). From patients to partners: participant-centric initiatives in biomedical research. *Nature Reviews Genetics, 13*(5), 371-376. doi:10.1038/nrg3218
Kaye, J., Whitley, E. A., Lund, D., Morrison, M., Teare, H., & Melham, K. (2015). Dynamic consent: a patient interface for twenty-first century research networks. *European Journal of Human Genetics, 23*, 141-146. doi:10.1038/ejhg.2014.71
Khazaei, H., Mench-Bressan, N., McGregor, C., & Pugh, J. E. (2015). Health informatics for neonatal intensive care units: An analytical modeling perspective. *IEEE Journal of Translational Engineering in Health and Medicine, 3*, 1-9. doi:10.1109/JTEHM.2015.2485268
Kim, K. K., Joseph, J. G., & Ohno-Machado, L. (2015). Comparison of consumers' views on electronic data sharing for healthcare and research. *Journal of the American Medical Informatics Association, Advance online publication*. doi:10.1093/jamia/ocv014
Kopelman, L. M., & Kopelman, A. E. (2007). Using a new analysis of the best interests standard to address cultural disputes: whose data, which values? *Theoretical medicine and bioethics, 28*(5), 373-391.

Korotchikova, I., Boylan, G. B., Dempsey, E. M., & Ryan, C. A. (2010). Presence of both parents during consent process in non-therapeutic neonatal research increases positive response. *Acta Paediatrica, 99*(10), 1484-1488.

Kuczewski, M. G., & Marshall, P. (2002). The decision dynamics of clinical research: The context and process of informed consent. *Medical Care, 40*(9 Suppl), 45–54.

Lang, T. A., Hodge, M., Olson, V., Romano, P. S., & Kravitz, R. L. (2004). Nurse-patient ratios: a systematic review on the effects of nurse staffing on patient, nurse employee, and hospital outcomes. *Journal of Nursing Administration, 34*(7-8), 326-337.

Lehtiranta, L., Junnonen, J., Kärnä, S., & Pekuri, L. (2015). The Constructive Research Approach: Problem Solving for Complex Projects. In B. Pasian (Ed.), *Designs, Methods and Practices for Research of Project Management*. Surrey, England: Gower Publishing.

Ligi, I., Boubred, F., Grandvuillemin, I., & Simeoni, U. (2011). Clinical research in newborn infants: difficulties and specificity. *European journal of clinical pharmacology, 67*(1), 29-32.

Lister, R., Bryan, G., & Tracy, M. (2000). *The e-Babies project: Integrated data monitoring and decision making in neo-natal Integrated data monitoring and decision making in neo-natal intensive care.* Paper presented at the 8th European Conference on Information Systems, Trends in Information and Communication Systems for the 21st Century (ECIS 2000), Vienna, Austria.

Loforte, R., Carrault, G., Mainardi, L., & Beuche, A. (2006). Heart rate and respiration relationships as a diagnostic tool for late onset sepsis in sick preterm infants. *Computers in Cardiology, 33*, 737−740.

Loue, S., Okello, D., & Kawuma, M. (1996). Research bioethics in the Ugandan context: A program summary. *The Journal of Law, Medicine & Ethics, 24*(1), 47-53.

Lunshof, J. E., Chadwick, R., Vorhaus, D. B., & Church, G. M. (2008). From genetic privacy to open consent. *Nature Reviews Genetics, 9*(5), 406-411. doi:10.1038/nrg2360

Make, B. J., Hill, N. S., Goldberg, A. I., Bach, J. R., Criner, G. J., Dunne, P. E., . . . Robert, D. (1998). Mechanical ventilation beyond the intensive care unit: Report of a consensus conference of the American College of Chest Physicians. *Chest, 113*(5), 289S-344S.

Manson, N. C. (2019). The biobank consent debate: Why 'meta-consent' is not the solution? *Journal of medical ethics, 45*, 291–294. doi:10.1136/medethics-2018-105007

March of Dimes, The Partnership for Maternal Health, Newborn and Child Health, Save the Children, & World Health Organization. (2012). *Born Too Soon: The Global action report on preterm birth*. Retrieved from Geneva:

Mariscal, G., Marban, O., & Fernandez, C. (2010). A survey of data mining and knowledge discovery process models and methodologies. *The Knowledge Engineering Review, 25*(2), 137-166.

Mason, S. A. (1997). Obtaining informed consent for neonatal randomised controlled trials - an "elaborate ritual"? *Archives of Disease in Childhood. Fetal and Neonatal Edition, 76*(3), F143–145. doi:10.1136/fn.76.3.F143

McGregor, C. (2013). Big data in neonatal intensive care. *Computer,* 54-59.

McGregor, C. (2017). *SRHC minimal risk protocol/research plan: Late onset neonatal sepsis study with Artemis Cloud*. Southlake Regional Health Centre.

McGregor, C. (2018). *Using constructive research to structure the path to transdisciplinary Innovation and its application for precision public health with big data analytics*.

McGregor, C., Bryan, G., Curry, J., & Tracy, M. (2002). *The e-baby data warehouse: A case study*. Paper presented at the 35th Hawaii International Conference on System Sciences, Big Island, HI, USA.

McGregor, C., Catley, C., & James, A. (2011). *A process mining driven framework for clinical guideline improvement in critical care.* Paper presented at the Learning from Medical Data Streams 13th Conference on Artificial Intelligence in Medicine (LEMEDS), Bled, Slovenia.

McGregor, C., Catley, C., & James, A. (2012). *Variability analysis with analytics applied to physiological data streams from the neonatal intensive care unit.* Paper presented at the 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS 2012), Rome, Italy.

McGregor, C., Catley, C., Padbury, J., & James, A. (2013). Late onset neonatal sepsis detection in newborn infants via multiple physiological streams. *Journal of Critical Care, 28*(1), e11-12. doi:10.1016/j.jcrc.2012.10.037

McGregor, C., Heath, J., & Choi, Y. (2015). *Streaming Physiological Data: General Public Perceptions of Secondary Use and Application to Research in Neonatal Intensive Care*. Paper presented at the MEDINFO 2015: eHealth-enabled Health, São Paulo, Brazil.

McGregor, C., James, A., Eklund, M., Sowe, D., Ebling, M., & Blount, M. (2013). *Real-time multidimensional temporal analysis of complex high volume physiological data streams in the neonatal intensive care unit* Paper presented at the The 14th World Congress on Medical and Health Informatics (MedInfo, 2013), Copenhagen.

McGregor, C. P., Smith, K. P., & Dhanoa, A. (2013). *System, method and computer program for multi-dimensional temporal and relative data mining framework, analysis & sub-grouping*. University Of Ontario Institute Of Technology.

McHaffie, H. E., Laing, I. A., Parker, M., & McMillan, J. (2001). Deciding for imperilled newborns: medical authority or parental autonomy? *J Med Ethics, 7*(2), 104-109.

McIntosh, N. (2000). Clinical diagnosis of pneumothorax is late: Use of trend data and decision support might allow preclinical detection. *Pediatric Research, 48*(3), 408-415. doi:10.1203/00006450200009000000025

McKechnie, L., & Gill, A. B. (2006). Consent for neonatal research. *Archives of Disease in Childhood. Fetal and Neonatal Edition, 91*, F374–F376.

McLachlan, S., Potts, H. W. W., Dube, K., Buchanan, D., Lean, S., Gallagher, T., . . . Fenton, N. (2018). The Heimdall framework for supporting characterisation of learning health systems. *Journal of Innovation in Health Informatics, 25*(2), 77-87.

McMaster Children's Hospital. (2018a). McMaster Children's Hospital. Retrieved from http://www.mcmasterchildrenshospital.ca/default.cfm?id=1

McMaster Children's Hospital. (2018b). Neonatal Nurseries. Retrieved from http://www.mcmasterchildrenshospital.ca/body.cfm?id=85&fr=true

Meem, M., Modak, J. K., Mortuza, R., Morshed, M., Islam, M. S., & Saha, S. K. (2011). Biomarkers for diagnosis of neonatal infections: A systematic analysis of their potential as a point-of-care diagnostics. *Journal of Global Health, 1*(2), 201-209.

Miller, F. G. (2008). Research on medical records without informed consent. *The Journal of Law, Medicine & Ethics, 36*(3), 560-566.

Mirowski, P., & Van Horn, R. (2005). The contract research organization and the commercialization of scientific research. *Social Studies of Science, 35*(4), 503-548.

Momtahan, K., Hetu, R., & Tansley, B. (1993). Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergonomics, 36*(10), 1159-1176.

Moorman, J. R., Carlo, W. A., Kattwinkel, J., Schelonka, R. L., Porcelli, P. J., Navarrete, C. T., . . . O'Shea, T. M. (2011). Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial. *Journal of Pediatrics, 159*(6), 900-906.e901. doi:10.1016/j.jpeds.2011.06.044

Morley, C., Lau, R., Davis, P. G., & Morse, C. (2005). What do parents think about enrolling their premature babies in several research studies? *Archives of Disease in Childhood. Fetal and Neonatal Edition, 90*(3), F225–228.

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *Journal of the American Medical Association, 309*(13), 1351-1352. doi:10.1001/jama.2013.393

Naik, T., Bressan, N., James, A., & McGregor, C. (2013). Design of temporal analysis for a novel premature infant pain profile using Artemis. *Journal of Critical Care, 28*(1), e4. doi:10.1016/j.jcrc.2012.10.024

Nair, K., Willison, D., Holbrook, A., & Keshavjee, K. (2004). Patients' consent preferences regarding the use of their health information for research purposes: A qualitative study. *Journal of Health Services Research & Policy, 9*(1), 22-27.

Neamatullah, I., Douglass, M. M., Li-wei, H. L., Reisner, A., Villarroel, M., Long, W. J., . . . Clifford, G. D. (2008). Automated de-identification of free-text medical records. *BMC Medical Informatics and Decision Making, 8*(32).

Neyro, V., Elie, V., Thiele, N., & Jacqz-Aigrain, E. (2018). Clinical trials in neonates: How to optimise informed consent and decision making? A European Delphi survey of parent representatives and clinicians. *PLOS ONE, 13*( 6), e0198097. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5999079/pdf/pone.0198097.pdf

Nijhawan, L. P., Janodia, M. D., Muddukrishna, B. S., Bhat, K. M., Bairy, K. L., Udupa, N., & Musmade, P. B. (2013). Informed consent: Issues and challenges. *Journal of Advanced Pharmaceutical Technology & Research, 4*(3), 134-140.

O'Shea, N., Doran, K., Ryan, C. A., & Dempsey, E. M. (2018). Parental And Clinician Views Of Consent In Neonatal Research. *Irish Medical Journal, 111*(3), 706. Retrieved from http://imj.ie/parentaland-clinician-views-of-consent-in-neonatal-research/

Oberle, K., Singhal, N., Huber, J., & Burgess, E. (2000). Development of an instrument to investigate parents' perceptions of research with newborn babies. *Nursing Ethics, 7*(4), 327–338.

Ohm, P. (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymisation. *UCLA Law Review, 57*, 1701-1777.

Oxford University Innovation. (n.d.). Dynamic Consent. Retrieved from https://innovation.ox.ac.uk/licence-details/dynamic-consent/

Pedroni, J. A., & Pimple, K. D. (2001). A brief introduction to informed consent in research with human subjects.

Pentz, R. D., Billot, L., & Wendler, D. (2006). Research on stored biological samples: Views of African American and white American cancer patients. *American Journal of Medical Genetics, 140*(7), 733–739. doi:10.1002/ajmg.a.31154

Pinch, W. J., & Spielman, M. L. (1990). The parents' perspective: ethical decision-making in neonatal intensive care. *Journal of Advanced Nursing, 15*(6), 712-719.

Pinsky, M. R., & Dubrawski, A. (2014). Gleaning knowledge from data in the intensive care unit. *American Journal of Respiratory and Critical Care Medicine, 190*(6), 606-610. doi:10.1164/rccm.201404-0716CP

Planned Parenthood. (2020). Sex, Gender, and Gender Identity. Retrieved from https://www.plannedparenthood.org/learn/sexual-orientation-gender/gender-gender-identity

Ploug, T., & Holm, S. (2015a). Meta consent: a flexible and autonomous way of obtaining informed consent for secondary research. *BMJ, 350*(h2146), h2146. doi:10.1136/bmj.h2146

Ploug, T., & Holm, S. (2015b). Rapid response to: Meta consent: a flexible and autonomous way of obtaining informed consent for secondary research. *BMJ*, 350:h2146. Retrieved from https://www.bmj.com/content/350/bmj.h2146/rr-1

Ploug, T., & Holm, S. (2015c). Routinisation of Informed Consent in Online Health Care Systems. *International Journal of Medical Informatics, 84*(4), 229-236. doi:http://dx.doi.org/10.1016/j.ijmedinf.2015.01.003

Ploug, T., & Holm, S. (2015d). Why meta consent is necessary and appropriate in the context of secondary use of health data for research. *BMJ, 350*.

Ploug, T., & Holm, S. (2016). Meta consent - A flexible solution to the problem of secondary use of health data. *Bioethics, 30*(9), 721-732.

Ploug, T., & Holm, S. (2017). Eliciting meta consent for future secondary research use of health data using a smartphone application - a proof of concept study in the Danish population. *BMC medical ethics, 18*(1), 51.

Post, A. R., & Harrison, J., J.H. (2008). Temporal data mining. *Clinics in Laboratory Medicine, 28*, 83-100.

Pugh, E. T., A. , Ng, E., McGregor, C. E., J.M. Narang, I., Belik, J., & James, A. (2013). Classifying Neonatal Spells Using Real-Time Temporal Analysis of Physiological Data Streams – Algorithm Development. *Journal of Critical Care, 28* (1), e9.

Pugh, J. E., Keir, A. K., McGregor, C., & James, A. G. (2013). *The impact of routine blood transfusion on heart rate variability in premature infants.* Paper presented at the American Medical Informatics Association.

Pugh, J. E., & McGregor, C. (n.d.). *Artemis Cloud clinical decision support framework – An algorithm to accurately detect late onset neonatal sepsis*. University of Ontario Institute of Technology and McMaster Children's Hospital.

Pugh, J. E., Twiss, J., El-Helou, S., Doyle, I., Glass, J., Inibhunu, C., . . . Gates, A. (2018). Implementation of Artemis Cloud at McMaster Children's Hospital: An innovative, high fidelity minimal physical presence physiological data platform. In.

Ramshaw, E. (2010). DNA Deception. Retrieved from https://www.texastribune.org/2010/02/22/dshs-turned-over-hundreds-of-dna-samples-to-feds/

Richards, E. P. (1993). Proxy consent Retrieved from http://biotech.law.lsu.edu/Books/lbb/x302.htm

Robling, M. R., Hood, K., Houston, H., Pill, R., Fay, J., & Evans, H. M. (2004). Public attitudes towards the use of primary care patient record data in medical research without consent: a qualitative study. *Journal of medical ethics, 30*(1), 104-109.

Rothschild, J. M., Landrigan, C. P., Cronin, J. W., Kaushal, R., Lockley, S. W., Burdick, E., . . . Bates, D. W. (2005). The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. *Critical care medicine, 33*(8), 1694-1700.

Safran, C., Bloomrosen, M., Hammond, W. E., Labkoff, S., Markel-Fox, S., Tang, P. C., & Detmer, D. (2007). Toward a national framework for secondary use of health data: An American medical informatics society white paper. *Journal of the American Medical Informatics Association, 14*(1), 1-9. doi:10.1197/jamia.M2273

Sahoo, S. S., Jayapandian, C., Garg, G., Kaffashi, F., Chung, S., & Bozorgi, A. (2014). Heart beats in the cloud: Distributed analysis of electrophysiological "Big Data" using cloud computing for epilepsy clinical research. *Journal of the American Medical Informatics Association, 21*(2), 263-271. doi:10.1136/amiajnl-2013-00215

Sanderson, S. C., Brothers, K. B., Mercaldo, N. D., Clayton, E. W., Antommaria, A. H. M., Aufox, S. A., . . . Holm, I. A. (2017). Public Attitudes toward Consent and Data Sharing in Biobank Research: A Large Multi-site Experimental Survey in the US. *American Journal of Human Genetics, 100*(3), 414-427.

Sandhu, E., Weinstein, S., McKethan, A., & Jain, S. H. (2012). Secondary uses of electronic health record data: benefits and barriers. *Joint Commission Journal on Quality and Patient Safety, 38*(1), 34-40.

Saria, S., Rajani, A. K., Gould, J., Koller, D. L., & Penn, A. A. (2013). Integration of Early Physiological Responses Predicts Later Illness Severity in Preterm Infants. *Science Translational Medicine, 2*(48), 48ra65.

Shahar, Y. (1997). A framework for knowledge-based temporal abstraction. *Artificial Intelligence, 90*(1-2), 79-133.

Shankaran, S., Langer, J., Kazzi, S. N., Laptook, A. R., & Walsh, M. (2006). Cumulative index of exposure to hypocarbia and hyperoxia as risk factors for periventricular leukomalacia in low birth weight infants. *Pediatrics, 118*(4), 1654-1659. doi:10.1542/ peds.2005-2463

Shearer, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing, 5*, 13–22.

Sheehan, M. (2011). Can broad consent be informed consent? *Public Health Ethics, 4*(3), 226-235.

Singhal, N., Oberle, K., Burgess, E., & Huber-Okrainec, J. (2002). Parents' perceptions of research with newborns. *Journal of Perinatology, 22*(1), 57–63. doi:10.1203/00006450-199904020-00214

Sriwatanakul, K., Weis, O. F., Alloza, J. L., Kelvie, W., Weintraub, M., & Lasagna, L. (1983). Analysis of narcotic analgesic usage in the treatment of postoperative pain. *Journal of the American Medical Association, 250*(7), 926-929.

Stedman's Medical Dictionary. (2004). *Stedman's Medical Dictionary for the Health Professions and Nursing, Illustrated* (5 ed.): Lippincott Williams & Wilkins.

Stenson, B. J., Becher, J. C., & McIntosh, N. (2004). Neonatal research: The parental perspective. *Archives of Disease in Childhood. Fetal and Neonatal Edition, 89*(4), F321-324. doi:10.1136/adc.2002.021931

Stoll, B. J., Gordon, T., Korones, S. B., Shankaran, S., Tyson, J. E., Bauer, C. R., . . . Wright, L. L. (1996). Late-onset sepsis in very low birth weight neonates: A report from the National Institute of Child Health and Human Development Neonatal Research Network. *The Journal of pediatrics, 129*(1), 63-71.

Sweeney, L. (2000). Simple Demographics Often Identify People Uniquely. *Data Privacy Working Paper 3*. Retrieved from http://dataprivacylab.org/projects/identifiability/paper1.pdf

Tanner, A. (2017). Strengthening Protection of Patient Medical Data. Retrieved from https://tcf.org/content/report/strengthening-protection-patient-medical-data/

Tarnow-Mordi, W. O., Hau, C., Warden, A., & Shearer, A. J. (2000). Hospital mortality in relation to staff workload: a 4-year study in an adult intensive-care unit. *The Lancet, 356*(9225), 185-189.

Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. (1996). Heart rate variability standards of measurement, physiological interpretation and clinical use. *Circulation, 93*, 1043-1065. doi:10.1161/01.CIR.93.5.1043

The National Health and Medical Research Council, The Australian Research Council and Universities Australia, & Commonwealth of Australia. (2007 (Updated 2018)). *The National Statement on Ethical Conduct in Human Research*: National Health and Medical Research Council.

Thommandram, A., Eklund, J. M., & McGregor, C. (2013). *Detection of apnoea from respiratory time series data using clinically recognizable features and kNN classification.* Paper presented at the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

Thommandram, A., Eklund, J. M., McGregor, C., Pugh, J. E., & James, A. G. (2014). *A rule-based temporal analysis method for online health analytics and its application for real-time detection of neonatal spells*. Paper presented at the 2014 IEEE International Congress on Big Data, Anchorage, AK, USA.

Tuzcu, V., Nas, S., Ulusar, U., Ugur, A., & Kaiser, J. R. (2009). Altered heart rhythm dynamics in very low birth weight infants with impending intraventricular hemorrhage. *Pediatrics, 123*(3), 810-815. doi:10.1542/peds.2008-0253

van Veen, E. B. (2008). Obstacles to European research projects with data and tissue: solutions and further challenges. *European Journal of Cancer, 44*(10), 1438-1450. doi:10.1016/j.ejca.2008.03.011

Vermeulen, E., Schmidt, M. K., Aaronson, N. K., Kuenen, M., & van Leeuwen, F. E. (2009). Obtaining 'fresh' consent for genetic research with biological samples archived 10 years ago. *European Journal of Cancer, 45*(7), 1168-1174. doi:10.1016/j.ejca.2008.11.022

Vorhies, W. (2016). CRISP-DM – A Standard Methodology to Ensure a Good Outcome. Retrieved from https://www.datasciencecentral.com/profiles/blogs/crisp-dm-a-standard-methodology-to-ensure-a-good-outcome

Weiskopf, N. G., Hripcsak, G., Swaminathan, S., & Weng, C. (2013). Defining and measuring completeness of electronic health records for secondary use. *Journal of biomedical informatics, 46*(5), 830-836.

Wendler, D. (2006). One-time general consent for research on biological samples. *BMJ, 332*(7540), 544–547. doi:10.1136/bmj.332.7540.544

Wendler, D. (2013). Broad versus blanket consent for research with human biological samples. *Hastings Center Report, 43*(5), 3–4.

Willison, D. J., Schwartz, L., Abelson, J., Charles, C., Swinton, M., Northrup, D., & Thabane, L. (2007). Alternatives to project-specific consent for access to personal information for health research: what is the opinion of the Canadian public? *Journal of the American Medical Informatics Association, 14*(6), 706-712. doi:10.1186/1472-6939-9-18

Willison, D. J., Steeves, V., Charles, C., Schwartz, L., Ranford, J., Agarwal, G., . . . Thabane, L. (2009). Consent for use of personal information for health research: do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC medical ethics, 10*, 10. doi:10.1186/1472-6939-10-10

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In *In Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (pp. 29-39).

Wood, A., O'Brien, D., Altman, M., Karr, A., Gasser, U., Bar-Sinai, M., . . . Wojcik, M. J. (2014). Integrating Approaches to Privacy Across the Research Lifecycle: Long-Term Longitudinal Studies. *Berkman Center Research Publication, 2014-12*. doi:10.2139/ssrn.2469848

World Health Organization. (2018). Preterm birth. Retrieved from http://www.who.int/news-room/fact-sheets/detail/preterm

World Health Organization. (2019). Newborns: reducing mortality. Retrieved from https://www.who.int/news-room/fact-sheets/detail/newborns-reducing-mortality

Yao, W., Chu, C., & Li, Z. (2010). *The use of RFID in healthcare: Benefits and barriers*. Paper presented at the 2010 IEEE International Conference on RFID-Technology and Applications (RFID-TA), Guangzhou, China.

Zupancic, J. A., Gillie, P., Streiner, D. L., Watts, J. L., & Schmidt, B. (1997). Determinants of parental authorization for involvement of newborn infants in clinical trials. *Pediatrics, 99*(1), E6.

## Appendix A – Description of Consent Models

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| **Standard/ general/ generic/ traditional consent** | Involves approaching parents for a study at the time that the infant becomes eligible. SDMs are provided with verbal and written information and encouraged to ask questions before they make a decision to participate. If they consent to participate in the study, they must sign a consent form. [7] | (+) Most common and widely used consent model; familiar process to many researchers [1][7]<br><br>(-) Participants'/ SDMs' decisions may change after initial consent<br><br>(-) A variety of decisions are made at a single point in time [1]<br><br>(-) Explaining study information can add on to a potentially lengthy and complex consent process [1]<br><br>(-) Difficult to apply when time is limited (e.g. in studies involving emergency scenarios or procedures) [7]<br><br>(-) May lead to decreased enrollment due to invasiveness of research protocols, lack of benefit for the infant enrolled, and/or the infant's illness severity [7]<br><br>(-) Negative impact of formal wording utilized in written information and consent forms [7]<br><br>(-) Participants/SDMs may experience information overload regarding a study and potential incidental findings [1] | | | | **X** | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| **Advanced/ antenatal consent** | Involves approaching parents in advance of their infant's eligibility for a study. Parents are solicited in anticipation that their infant will meet the study inclusion criteria at a future date. [7]<br><br>Often used prior to birth (e.g. during the antenatal period) seeking postnatal enrollment [7]<br><br>Most suitable for studies that involve the immediate postnatal period [7] | (-) Possibility of overburdening parents with unnecessary information in the situation that their infant does not meet the study's inclusion criteria especially in situations where parents are given information about neonatal studies in the prenatal period and maternal studies are solicited simultaneously [7] [12]<br><br>(-) Not possible to know with certainty when an infant will be eligible for many studies; possibility that parents may not be able to recall the specific details of a study or perhaps ever being approached for consent at all, following their infant's birth [7]<br><br>(-) Not recommended for majority of studies in the NICU environment [7]<br><br>(-) Validity of consent obtained during labour is questionable [12]<br><br>(-) Obtaining consent from high risk pregnancies may put undue stress on parents [12] | | | | **X** | |
| **Blanket consent** | An individual's collected sample (or data) would be used without further consent or restrictions. Patients provide consent once and entrust researchers and ethics committees to review and approve | (+) One-time consent simplifies the research process [3]<br><br>(-) Legal challenges associated with blanket consent models have been underplayed [3]<br><br>(-) As blanket consents are necessarily vague, they are too general to have much legal weight [3] | | | | **X** | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | any future research projects [19] | (-) Do not allow patients to act meaningfully on their continuing right to control their health information [3]<br><br>(-) Samples or data may be used in future studies that conflict with an individual's fundamental values [3]<br><br>(-) Does not provide participants with the ability to control the specific projects as to which their samples/data are to be used in; an independent body (e.g. REB) makes these decisions on behalf of the individual who provides consent | | | | | |
| **Bona fide consent** | A patient makes informed consent after a collaborative discussion with the clinician [5][6] | | | | | | |
| **Broad consent** | A process by which an individual donates their samples (or data) for a broad range of future studies, subject to specified restrictions [19]<br><br>An individual is provided the option of consenting to future research of a particular type with regards to both content and context rather than to just a specific research project [19] | (+) A broad consent restriction may be that if certain types of research are known to conflict with the participant's fundamental values, this could be precluded at the point of initial consent [19]<br><br>(-) Does not provide participants with the ability to control the specific projects as to which their samples/data are to be used in; an independent body (e.g. REB) makes these decisions on behalf of the individual who provides consent<br><br>(-) No regard to future uses when consenting; Future uses do not constitute valid informed consent | | | | **X** | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| **Cascading consent** | A process that offers donors successive choices, starting with blanket consent as the default (e.g. those who are uncomfortable with blanket consent could be offered "broad" consent. If the donor objects, then they may be offered the chance to opt out of certain forms of research, or, offered the chance to opt in to certain types of research) by using computing technology [9] | (+) As part of a cascade of choices, "broad" consent may be an efficient way to cluster likely objections without being too cognitively burdensome on patients. By retaining the full range of potential consent scope, from blanket to opt-in, the regime can better respect autonomy and reflect donor preferences. [9]<br><br>(+) Nudges subjects toward blanket consent while retaining options for donors who would prefer a more limited form of consent. This may help biobanks minimize costs and maximize use of their samples. [9]<br><br>(-) Some donors may want to only allow work on a particular disease affecting their family. Some biobanks may find that it is not cost-effective to accept such restricted samples, and consequently may decline to accept them. [9]<br><br>(+) Yet there are likely instances in which a specimen is so important to research that narrow consent is better than none at all. [9] | **X** | | **X** | | |
| **Committee consent** | Involves the Research Ethics Board (REB)/ Local Research Ethics Committee (LREC)/ Institutional Review Board (IRB)/ Ethics Review Board (ERB) examining trials on | (-) Ethical and practicality concerns:<br>• Questionable if REBs should make decisions for individuals considering how preoccupied they already are [7]<br>• REBs cannot know the specific detail regarding each and every case or participant situation [7] | | | | **X** (not by family/ legal guardians) | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | behalf of potential patient subjects [7]<br><br>The REB makes a decision on behalf of individual patients that are eligible for the study (i.e. on a case by case basis) [7] | • Likely that REBs are unable to make decisions within some given timeframes [7]<br>• In non-emergency situations, health care professionals are restrained from making decisions on behalf of competent patients [7] | | | | | |
| **Contrived consent** | An individual is presented with a "menu" of choices and a response is elicited. This menu may be accompanied by a large quantity of information, most of which is not specific to patient or on the other extreme, little or no information may be provided. [6] | (-) Emphasis of contrived consent is not on ensuring the patient's understanding of information but on his/her indication that clinician may proceed with the proposed procedure or treatment. [6] | | | | | |
| **Deferred consent** | Legislation in certain countries permit research without prior consent when the following conditions are met [20]:<br>1) treatment is required urgently;<br>2) urgent action is required for the purposes of the trial<br>3) it is not reasonably practicable to | (+) Seeking informed consent requires time which is limited in emergencies when even minimal treatment delays can be harmful to the patient [20]<br><br>(+) Parent(s)/guardian(s) are not always present when a child requires emergency treatment or a mother of a critically ill neonate may be sedated<br><br>(-) Clinicians with no experience of research without prior consent may be concerned that this model would be detrimental to the parent–practitioner relationship. In nations such as the | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | obtain consent prospectively; and<br>4) an ethics committee has approved the consent procedure<br><br>A child research subject will have already received an intervention as part of a trial before any information is provided or consent is sought. Fundamentally permission is sought post-intervention to use data that have already been collected and consent for the child research subject to continue to take part in the trial. [20] | United States and those in the European Union, informed consent can only occur prior to enrollment. "Deferred" consent is not an acceptable term. Consequently, the waiver of consent process is the only available alternative. [18]<br><br>(-) Removes parents' (and children's) autonomy [10]<br><br>(+) In contrast, practitioners with experience of this consent model described how families were receptive to the method as long as discussions were appropriately timed and demonstrated sensitively. [20] | | | | | |
| **Mandatory Return** | Advise participants at time of consent of mandatory issues or items that may be returned regardless of consent to return findings. [1] | (+) Researchers' obligations to return incidental findings are clearly defined from the beginning of consent process [1]<br><br>(-) Considerable information about possible findings and how they would be dealt with will need to be provided as part of the initial consent process. [1]<br><br>(-) Individuals' choices about receiving incidental findings are restricted | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | | to those findings they desire to receive and those they would rather not know about, leading to disclosures that could be both under- and overinclusive, depending on their preferences. [1]<br>• May discourage study enrollment [1]<br><br>(-) Re-contacting participants is costly in terms of finances, time and effort [1] | | | | | |
| **Open consent** | Volunteer consents to unrestricted re-disclosure of data originating from a confidential relationship (i.e. their health records) and to unrestricted disclosure of information that emerges from any future research on their data set, the information content of which cannot be predicted. [10] | (+) The leading moral principle is veracity which should precede autonomy [10]<br><br>(-) No promises of anonymity, privacy or confidentiality are made [10]<br><br>(-) Donor's data could be included in an open-access public database [10]<br><br>(-) Participation involves a certain risk of harm to themselves and their relatives [10]<br><br>(-) Participation does not benefit the participants in any tangible way. [10]<br><br>(-) While withdrawal from the study is possible at any time, complete removal of data that have been available in the public domain may not be possible [10] | | | | **X** | |
| **Opting out** | Parental consent is presumed. An eligible infant is automatically enrolled unless the | (+) Attempts to lessen parental distress by removing some of the decision-making on behalf of their infant, as they are not solicited for their consent [7] | | | | **X** | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | parents explicitly object to the enrollment. [7] | (+) Results in the generation of valuable knowledge earlier due to increased enrolment [7]<br><br>(-) Does not respect parental autonomy [7]<br><br>(-) May negatively affect parents' perception of the study as they may be under the impression that the experimental intervention is equivalent to treatment that is provided to their infant is the health care unit/facility (in the case that the study is a clinical trial) and/or the possibility of thinking that asking for their permission to enroll their infant is almost unnecessary [7] | | | | | |
| **Outsourced consent** | Give participants their own raw data and allow them to take to another specialist who can assist them to decide what information they want returned from the study. [1] | (+) Participants are spared from the immediate task of deciding which findings to receive and allows them to choose their own services to interpret their own raw data if they wish [1]<br><br>(+) No need to address issues such as return of incidental findings thus saving time and costs [1]<br><br>(+) Researchers have less obligations [1]<br><br>(-) Participants who do not use an interpretive service may not learn of medically significant data – however interpretive services are not widely available and accessible to everyone [1] | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | | (-) Without readily accessible interpretive genetic services, the outsourced model cannot function at all [1] | | | | | |
| **Presumed consent** | Parental consent is not sought prior to enrolling their infant in a study. [7] | (+) Removes the decision-making burden from the parents [7]<br><br>(-) Parental autonomy is not considered. This issue can be extremely sensitive and controversial in certain situations, such as when a study involves experimental options for different treatments that have different risks and benefits compared to other treatment options that may be available. [7]<br><br>(-) Unlike a clinical trial, the secondary use of physiometric data collected from infants may not have an immediate effect on their treatment or result in a life or death situation. However, it fails to comply with laws in many jurisdictions that relate to consent, parental autonomy and surrogate decision-making. [7]<br><br>(-) May increase stress if parents are concerned that their child has been enrolled in a study without their knowledge [7]; made worse if they find out years later as seen with unconsented Guthrie test use in several countries. [4]<br><br>(-) Parents may make the assumption that the entire NICU supports the | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | | research that is being done and will place a certain level of responsibility regarding the research on all health care professionals who work in that setting. [1] | | | | | |
| **Randomization without consent** | Infants are randomized for the purpose of comparing two standard methods of treatment without parental knowledge or consent. [7]<br><br>The purpose of such a trial is to contrast two treatment approaches, each of which is deemed to be acceptable clinical practice which can be individually implemented without involvement of the parents. [7] [13]<br><br>Similar to the presumed consent model. [7] | (+) Alleviate the stressful burden associated with decision-making from the parents from what would otherwise have been routine practice [7]<br><br>(-) Does not take into account parental autonomy [7] | | | | | |
| **Waiver of consent** | Consent is sometimes waived by the REB. [7]<br><br>Typically, this is done for innocuous studies that are assessed as presenting minimal | (+) Removal of some of the decision-making burden on the parents of having to make a decision on behalf of their infant [7] | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | or no risk to the subject (e.g. retrospective chart reviews). [7]<br><br>Although rarely accepted in NICU research practice, this model may be used in studies that require immediate randomization (e.g. new emergency procedures for patient resuscitation) and/or in situations where the research could not be practicably accomplished without the waiver. In such situations, the subjects (or their surrogate decision maker) should be provided with the additional relevant information after participation in a study. [7] | | | | | | |
| **Dynamic Consent (DC)** | Information regarding secondary use of health data or sample along with a request for consent is provided to the individual via a web-based platform. [8] | (+) DC is currently a biobanking project but has the ability to be expanded to other applications and fields [8]<br><br>(+) In the case where there are multiple researchers and research projects, it is difficult to obtain informed consent for all future research uses at the time of | **X** | **X** | **X** | **X** | **X** |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | Every request for consent would be presented with a description of the entire project. [8] | recruitment into the biobank or before such research commences. DC recognizes that the expression of individual autonomy is also not static: it involves making choices and decisions, such as the decision to consent to participate in research, over the course of one's lifetime [8]<br><br>(+) Individuals could provide different types of consent depending upon the kind of study. These consent preferences travel securely with their samples or data so that third parties know the scope of the consent that applies. [8]<br><br>(+) A secure consent interface allows participants to [8]:<br>• Modify their consent preferences reliably [8]<br>• Alter their contact details [8<br>• Receive information on the use of their samples and data [8]<br>• Enroll in new studies [8]<br>• Complete online surveys [8]<br>• Engage with the research study in their own time, as often or as little as they choose [8]<br><br>(+) Available preferences can be adapted to suit the capabilities and needs of institutions, researchers and participants [8]<br><br>(+) Could enable researchers to gain a better sense of participants' views on incidental findings and could ultimately | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | | provide the opportunity for participants to be better informed and to set specific preferences relating to what feedback they would like, and when and how it is received [8]<br><br>(-) Requires cultural change for both health-care professionals and individuals. It requires partnerships for health that are open, transparent and engaging, and which understand and value the central role that patients have in research as the providers of information and biological material. [8]<br><br>(-) Requires an investment of resources such as time, money, expertise and a commitment to such a vision by clinicians and researchers, health-care services, research institutions and governments is necessary for [8]:<br>• Development of new policies, standards and ways of working that can accompany this approach [14]<br>• System must have the technical capacity to interface with the systems of the various research organizations so it can provide information and feedback [8]<br><br>(-) Online services run the risk of excluding individuals and communities with limited or no internet access [8]<br><br>(-) New processes and technology for testing and monitoring the integrity of DC technologies must be developed, to | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | | provide operational control for managing risk throughout the lifetime of the data holding [8]<br><br>(-) Not all potential participants will connect and opt in<br><br>(-) Not practical for research using routine clinical data<br><br>(-) Every request requires description of entire project; May result in consent fatigue [8] | | | | | |
| **Meta Consent (MC)** | MC denotes the idea that individuals should be asked how and when they would like to provide consent [16]<br><br>Allows an individual to choose their preferred consent model (i.e. broad consent/refusal, dynamic consent, blanket consent or refusal) for different types of research for future secondary research for previously collected data or data that will be collected in the future via a web-based platform (i.e. retrospective and prospective). [14] | (+) Combines the broad and dynamic models, with additional options for blanket consent and blanket refusal [14]<br><br>(+) One can choose how and when they wish to provide consent for future secondary research of data collected in the past or of data that will be stored in the future (i.e. MC is both retrospective and prospective) [14] [15]<br><br>(+) Individuals can choose the type of consent: DC, broad consent, blanket consent/refusal for different types of research [14]<br><br>(-) Costs and burdens of such a framework; may be much higher than asking participants to reconsent [11]<br><br>(-) Need to take care to avoid false positives where those who chose broad consent are contacted by mistake [11] | **X** | **X** | **X** | **X** | **X** |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | MC denotes the idea that individuals should be asked how and when they would like to provide consent [19]<br><br>Future consent requests can be generated and communicated via information and communication technology based on individual's preferences. | (-) Problem of what to do when participants fail to reply [11]<br>• (+) Failure to provide MC could be handled in several ways (e.g. MC could be arranged early in life, which could be achieved by making it mandatory as an individual comes of age, the individual could receive reminders from healthcare professionals when seeking treatment, reminders could be linked to the use of various official web services, or a default position of broad consent could be applied). [14]<br><br>(-) Administrative systems need to be in place to track, and accurately respond to, individual fine-tuned choices [11]<br><br>(-) Online services run the risk of excluding individuals and communities with limited or no internet access. [8][14] | | | | | |
| **Staged Consent** | Obtain consent in stages, with brief mention of incidental findings at the time of initial consent but with more detailed consent obtained if and when reportable results are found [1] | (+) Staged approaches allow participants to retain control [1]<br><br>(-) However, when a decision is deferred until findings exist, the door is open to communication of unwanted information about potential risks. The likelihood of such findings will vary across studies, depending on the scope of reportable results defined in each protocol. [1] | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | | (+) Various alternatives have been proposed for this model, including:<br>• Obtaining consent to return of incidental findings at the time of enrollment would be to defer the process of decision-making about returning them until later in the process<br>• Postpone the process regarding decisions about receipt of incidental findings until it was clear whether there would be such findings for a given participant.<br>• When consent to participation was obtained, participants would be told that incidental findings might be detected in the relevant categories and that they would be given an opportunity, if such results were found, to learn more about them and decide whether to receive them.<br><br>(-) Participants/SDMs decide whether to enroll in a study without receiving full information regarding incidental findings [1]<br><br>(-) Effectiveness partly depends on funding to create the infrastructure on for a system for communicating with participants, soliciting preferences, and returning information over time [1]<br><br>(-) Re-contacting and following up with participants/SDMs could be challenging and costly [1] | | | | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | | (-) In certain types of studies, the act of re-contacting the participant/ SDM can reveal unwanted information [1] | | | | | |
| **Tiered Consent** | Provide patient with Generic consent as Tier 1, then as Tier 2 provide them with specific content to meet the needs of that particular participant/ patient. [2] | (+) Allows for some tailoring to participant's consent preferences | | | **X** | | |
| **Binned Consent** | Where a large amount of information is relevant to consent, this information is divided into smaller 'bins' of relevant information. [2] | (+) Approach was proposed as a framework to organize genes to consider the return of incidental findings in genomic studies [2]<br><br>(+) Binning genes could also be a useful approach for consent and patient education [2]<br><br>(+) Rather than providing specific risks for each individual gene, genes can be organized into bins, defined by their risk, clinical utility, evidence supporting medical management, or other relevant features [2] | | | | | |
| **Tiered-Binned Consent** | The different information areas are divided into bins. A generic Tier 1 is developed from each bin and the combination of all tier 1's are presented to the participant/ patient and they are asked if they prefer | (+) Combines tiered and binned consent for those who wish to have more information [2]<br><br>(+) Minimize information overload and support informed decision making [2]<br><br>(-) Consent model initially designed to evaluate the outcomes of receipt of multiplex testing for cancer | | | **X** | | |

| Consent Model | Description | Advantages and Disadvantages | Flexibility | Longitudinal engagement | Multi-level | Proxy consent | Coupled with data |
|---|---|---|---|---|---|---|---|
| | more information on each item during the tier 1 presentation. [2] | susceptibility; however, the outcomes of the tiered-binned approach are currently unknown. [2]<br><br>(-) This model is proposed for genetic susceptibility testing for common cancers and may not apply in other contexts, such as genetic testing for rare cancers, for which panels are commonly utilized and do not have the same wide variability in the disease spectrum, risks, or management options. [2]<br><br>(-) Designers of model debated the merits of providing specific information for particular genes when patients have expressed varying testing preferences.[32] The value of utilizing specific examples remains unknown. [2]<br><br>There is also value in evaluating various visual formats and content of risk information, particularly among populations of varying genomic literacy [2] | | | | | |

Definitions and/or descriptions adapted from:

[1] Appelbaum, P. S., Parens, E., Waldman, C. R., Klitzman, R., Fyer, A., Martinez, J., . . . Chung, W. K. (2014). Models of consent to return of incidental findings in genomic research. *Hastings Center Report, 44*(4), 22-32. doi:10.1002/hast.328

[2] Bradbury, A. R., Patrick-Miller, L., Long, J., Powers, J., Stopfer, J., Forman, A., . . . Domchek, S. M. (2015). Development of a tiered and binned genetic counseling model for informed consent in the era of multiplex testing for cancer susceptibility. *Genet Med., 17*(6), 485-492. doi:10.1038/gim.2014.134

[3] Caulfield, T. (2002). Gene banks and blanket consent. *Nature Reviews Genetics, 3*(8), 577.

[4] Cunningham, S., O'Doherty, K. C., Senecal, K., Secko, D., & Avard, D. (2015). Public concerns regarding the storage and secondary uses of residual newborn bloodspots: an analysis of print media, legal cases, and public engagement activities. *J Community Genet, 6*(2), 117-128. doi:10.1007/s12687-014-0206-0

[5] Edozien, L. C. (2016). Montgomery and informed choice: Not yet uhuru. *Clinical Risk, 22*(1-2), 1-3. doi:10.1177/1356262216671472

[6] Edozien, L. C. (2015). UK law on consent finally embraces the prudent patient standard. *BMJ, 350*(h2877), 1-2. doi:10.1136/bmj.h2877

[7] Golec, L., Gibbins, S., Dunn, M. S., & Hebert, P. (2004). Informed consent in the NICU setting: an ethically optimal model for research solicitation. *Journal of Perinatology, 24*(12), 783–791. doi:10.1038/sj.jp.7211198

[8] Kaye, J., Whitley, E. A., Lund, D., Morrison, M., Teare, H., & Melham, K. (2015). Dynamic consent: a patient interface for twenty-first century research networks. *European Journal of Human Genetics, 23*, 141-146. doi:10.1038/ejhg.2014.71

[9] Loe, J., Robertson, C. T., & Winkelman, D. A. (2015). Cascading consent for research on biobank specimens. *The American Journal of Bioethics, 15*(9), 68-70. doi:10.1080/15265161.2015.1062174

[10] Lunshof, J. E., Chadwick, R., Vorhaus, D. B., & Church, G. M. (2008). From genetic privacy to open consent. *Nature Reviews Genetics, 9*(5), 406-411. doi:10.1038/nrg2360

[11] Manson, N. C. (2019). The biobank consent debate: Why 'meta-consent' is not the solution? *Journal of medical ethics, 45*, 291–294. doi:10.1136/medethics-2018-105007

[12] McKechnie, L., & Gill, A. B. (2006). Consent for neonatal research. *Archives of Disease in Childhood. Fetal and Neonatal Edition, 91*, F374–F376.

[13] Modi, N. (1994). Clinical trials and neonatal intensive care. *Arch Dis Child, 70*, F231–232.

[14] Ploug, T., & Holm, S. (2015a). Meta consent: a flexible and autonomous way of obtaining informed consent for secondary research. *BMJ, 350*(h2146), h2146. doi:10.1136/bmj.h2146

[15] Ploug, T., & Holm, S. (2015b). Why meta consent is necessary and appropriate in the context of secondary use of health data for research. *BMJ, 350*.

[16] Ploug, T., & Holm, S. (2016). Meta consent - A flexible solution to the problem of secondary use of health data. *Bioethics, 30*(9), 721-732.

[17] Ploug, T., & Holm, S. (2017). Eliciting meta consent for future secondary research use of health data using a smartphone application - a proof of concept study in the Danish population. *BMC medical ethics, 18*(1), 51.

[18] Rich, W. D., & Katheria, A. C. (2019). Waived Consent in Perinatal/Neonatal Research—When Is It Appropriate? *Front. Pediatr., 7*, 493. doi:10.3389/fped.2019.00493

[19] Wendler, D. (2013). Broad versus blanket consent for research with human biological samples. *Hastings Center Report, 43*(5), 3–4.

[20] Woolfall, K., Frith, L., Dawson, A., Gamble, C., Lyttle, M. D., the CONNECT advisory group, & Young, B. (2016). Fifteen-minute consultation: an evidence-based approach to research without prior consent (deferred consent) in neonatal and paediatric critical care trials. *Arch Dis Child Educ Pract Ed., 101*(1), 49–53.

# Appendix B – Participant Information Sheet

**PARTICIPANT INFORMATION SHEET**

**Title of Database:** McMaster Children's Hospital Artemis Cloud Database

**Local Investigator, Department/Hospital/Institution:**
Dr. Edward Pugh, Division of Neonatology, McMaster Children's Hospital

**Co-Investigators, Department/Hospital/Institution**

Jennifer Twiss MD[1,2], Salhab El-Helou[1,2], Ian Doyle[3,4], Jonah Glass[3,4], Catherine Inibhunu[3,4], John Madill[3,4], Carolyn McGregor PhD[3,4], Aaron Gates [3]

[1]Division of Neonatology, McMaster Children's, Hamilton. [2]Department of Paediatrics, Faculty of Medicine, McMaster University; [3]Faculty of Business & Information Technology, and the [4]Health Informatics Research Laboratory, University of Ontario Institute of Technology, Oshawa, Ontario, Canada.

**Sponsor:** University of Ontario Institute for Technology/Federal Economic Development Agency for Southern Ontario

---

You are being invited to participate in a research database, which is being created by Dr. Edward Pugh and the division of Neonatology in collaboration with University of Ontario Institute for Technology.

The purpose of this database is to collect your infant's data from the bedside monitors and electronic equipment (for example heart rate, blood pressure) in the neonatal intensive care unit (NICU) along with clinical information from their electronic medical record (EMR). This information may be used for future health research studies in which analyzing information about your infant and other similar infants may help improve the future care of term and preterm babies.

The monitors and equipment around your infant's bedside display information that is used to help provide your infant with clinical care. Most of this information is not saved and is lost after it disappears from the monitors. This database will collect and save all the information that is produced, so that it can be used for future research projects that will try to improve the care of NICU infants. As this information is already gathered and displayed at your infant's bedside, the research database will not present any additional harms, risks or discomfort. You will not be asked to do anything additional or different.

All infants admitted to the NICU at McMaster Children's Hospital will be eligible to contribute their health information to this database for future research purposes.

If you are happy to hear about this research, a research assistant will meet with you to describe the data collection process and answer any specific questions you may have

about data storage and use in future research. It is important for you to know that all studies that make use of your infant's information will do so as part of a combined data set. This means your baby will not be identifiable in any way. Your infant's information will not be published or disclosed unless you give permission by signing a separate consent form. If you would like your infant's data to be part of the database we offer two options for consent:

Option 1) Anonymous data can be used in combination with other similar infant's data without further consent for all future studies that are approved by the Hamilton Integrated Research Ethics Board. This means that no one will know that the information that is collected came from your infant.

Option 2) Your infant's information stored in this database may be linked using their name, date of birth or health card number with other research data sets, without seeking further consent from you. Future studies that are approved by Hamilton Integrated Research Ethics Board will be able to use this information. This linked data will be made anonymous before it is used in combination with other similar infants. This means that no one outside of the research team will know that the information that is collected came from your infant.

If at any time you wish to stop having your infant's data used for research purposes, you can do so by calling Dr. Edward Pugh at McMaster Children's Hospital at 905-521-2100 ext. 76342. Once this notification is received, your infant's data will be removed from the research database and will not be available for any future studies. Your child's data will continue to be used in studies where analysis of your infant's data has been completed up until the time you removed your consent.

There are no medical benefits to you or your infant from taking part in this database. However, by participating, you may assist us in continuously learning how to improve the care of future infants admitted to the NICU.

**It is important for you to know that you can choose not to have your infant's data collected for research purposes. You may choose not to provide your consent in which case all of your infant's data will be deleted from the research system and only held within the clinical system.**

Choosing not to participate in this database will in no way affect your infant's care or treatment.

As a parent/guardian, you will not have access to the collected bedside data. The collected data will be maintained until your child's 18th birthday, at which time we will endeavour to re-consent your child for continued database participation. If this is not possible, your child's data will be irreversibly de-identified at this time.

Your infant's data will not be shared with anyone without your consent or as required by law. By having your infant's personal health information added to a research database

there is the potential for a breach of confidentiality. Every effort will be taken to ensure that your information is kept private. The only people who will have access to the information in the database are Dr. Edward Pugh, Dr. Salhab El Helou, Dr. Jennifer Twiss and Dr. Carolyn McGregor (PhD). These individuals understand the laws regarding privacy and have signed a confidentiality agreement. Identifiable data from your infant's medical chart will be stored within the hospital computing server rooms and maintained with the same level of security as their medical chart. Since your infant's bedside data that is collected from the medical monitors and equipment is too large to be stored on-site, it will be given a unique identifying code that will be stored with your infant's electronic medical records in the hospital. Your infant's personal identifiers will then be removed from the bedside machine data. Finally, the bedside machine data will be sent to a secure data centre over secure data cables after encryption and maintained at this site. When your infant's data is used for research, it will be re-identified on computers within the hospital network. Data will be collected, de-identified of names and dates and then analyzed in accordance with the rules of the Hamilton Integrated Research Ethics Board.

All research studies that will be conducted using the database information will be approved by the Hamilton Integrated Research Ethics Board.

For the purposes of ensuring the proper monitoring of the research database, it is possible that a member of the Hamilton Integrated Research Ethics Board may consult your research data and medical records. By signing this consent form, you or your legally acceptable representative authorize such access.

If you volunteer to be part of this database, you may withdraw at any time and this will in no way affect the quality of care you receive at this institution.

If you have any questions about the research database now or later, please contact Dr. Edward Pugh, Division of Neonatology, Office 4F5, Department of Paediatrics, 1200 Main Street West, Hamilton, Ontario, L8N3Z5 or through switchboard 905 521 2100.
If you have any questions regarding your rights as a research participant, you may contact the Office of the Chair of the Hamilton Integrated Research Ethics Board at 905-521-2100, ext. 42013.

**CONSENT STATEMENT**

**I have read the preceding information thoroughly. I have had the opportunity to ask questions, and all of my questions have been answered to my satisfaction. I agree to have my personal health information included in this database. I understand that I will receive a signed copy of this form.**

**I have selected the following option for consent:**

☐ **Option 1) Anonymous data can be used in combination with other similar infant's data without further consent for all future studies that are approved by the** Hamilton Integrated Research Ethics Board**. This means that no one will know that the information that is collected came from your infant.**

☐ **Option 2) Your infant's information stored in this database may be linked using their name, date of birth or health card number with other research data sets, without seeking further consent from you. Future studies that are approved by the institutional research ethics board will be able to use this information. This linked data will be made anonymous before it is used in combination with other similar infants. This means that no one outside of the research team will know that the information that is collected came from your infant.**

_____
Name of Participant

_____
Name of Legally Authorized Representative (if applicable)

_____ _____
Signature of Participant or Legally Authorized Representative Date

**Consent form administered and explained in person by:**

_____
Name and title

_____ _____
Signature Date