Newborn Screening: Science, Policy, and People

Catharine Riley

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2012

Reading Committee:
Carolyn Watts, Chair
Wylie Burke
Barbara Burns-McGrath

Program Authorized to Offer Degree:
Public Health Genetics

University of Washington

ABSTRACT

Newborn Screening: Science, Policy, and People

Catharine Riley

Chair of the Supervisory Committee:
Professor Carolyn Watts
Health Services

Background: Newborn screening (NBS) is a large-scale, state-mandated screening program designed to screen infants for a set of rare congenital conditions, in order to identify infants that may be at risk for morbidity or mortality, with the intent to improve the health outcome for affected infants. The science behind NBS is critical to its success and data are required to make informed decisions. History has shown that advocacy efforts have also impacted the landscape of NBS policy. It is important to remember that NBS impacts infants and their families. Based on their experiences, parents can inform policy-making at the programmatic, state, and federal level.

Purpose: The purpose of this dissertation is to: 1) explore the role of advocacy in changing NBS policy, 2) describe the evidence base for understanding parents' experience with and knowledge of NBS, and 3) demonstrate policy-making at the programmatic level using a quality assurance data analysis to revise the screening algorithm for congenital hypothyroidism.

Methods: The methods used in this dissertation are: 1) a historical analysis, case studies, and the application of the Advocacy Coalition Framework, 2) a structured literature review, and 3) quantitative analysis of NBS data.

Results: The ACF provides a suitable framework for understanding NBS policy and the role of advocacy. Forty-two articles on parents and NBS revealed six major discussion topics: 1) parents' experience with NBS, 2) parents' knowledge of NBS, 3) parents' education on NBS, 4) the impact of false positive results, 5) the impact of true positive results, and 6) informed consent. The QA data analysis resulted in a revised stratification scheme that will reduce the number of false positive results by an estimated 35%.

Conclusion: It is clear that advocacy was instrumental in initiating NBS and continues to be an important factor in NBS policy. A foundation of knowledge regarding parents and NBS has been established. Future research ought to focus on how information can be translated into public health and clinical practice at the ground level. Using data to determine cutoff stratification schemes that maintain a high sensitivity and specificity is an effective way to inform policy decisions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

To the chair of my committee Carolyn Watts, PhD and committee members Wylie Burke, MD, PhD, and Barbara Burns McGrath - thank you for your mentorship.

To the entire staff of the Washington State Department of Health Office of Newborn Screening, in particular the director Mike Glass and my supervisor John Thompson - thank you for sharing your knowledge and experience. I am honored to be a part of a team committed to improving the health of babies born in Washington.

To the Institute for Public Health Genomics faculty, staff and fellow students - thank you for your continued support. I would like to acknowledge three outstanding faculty mentors outside of my committee: Kelly Edwards, PhD; Pat Kuszler, MD, JD; Anna Mastroianni, JD, MPH

To my officemates - thank you for sharing your space and providing moral support.

To my friends - thank you for keeping me sane.

To my family - you all have shaped who I am and I am very thankful for your love and support throughout this journey.

Research Funding: TL1 Training Grant - National Institutes of Health, National Center for Research Resources; Institute for Translational Health Science; UW Graduate College – Top Scholars Award; ARCS Foundation, Seattle Chapter - ARCS Fellowship
Other funding sources: NIH Center for Genomics and Healthcare Equality; UW Center for Ecogenetics and Environmental Health; Institute for Public Health Genetics

DEDICATION

I dedicate my dissertation to my parents, who have provided boundless encouragement and have let me stand on their shoulders in order to achieve my dreams.

Most importantly, it is my parents that have taught me life's most important lesson –

to love unconditionally.

PREFACE

Science, technology, public health, medicine, genomics, ethics, law, politics, and people; newborn screening encompasses all of these, creating a complex and nuanced area of research. Newborn screening, one of the first applications of testing for genetic conditions in medicine and public health, identifies rare conditions that without detection and treatment may lead to severe mental and physical harm or in some cases death. Newborn screening is a large-scale, state-mandated screening program designed to screen all babies born for a set of rare congenital conditions in order to identify infants that may be at risk for morbidity or mortality, with the intent to intervene in order to improve the outcome for affected infants. Newborn screening programs rely on health care professionals to collect and submit newborn screening specimen(s). Once the specimen is received by the state laboratory, or contracted laboratory, the appropriate tests are performed on the specimens to identify infants at risk for having one the designated conditions. Abnormal results are reported to the healthcare provider and subsequent screening or diagnostic testing is performed to determine if the infant truly has one of the conditions. **Chapter One** describes in further detail the function of newborn screening and the role of state newborn screening programs.

Historically, newborn screening has provided a clear public health benefit by preventing negative health outcomes and serious harm. Since its beginning in 1963 with screening for one disease, phenylketonuria (PKU), states have been adding conditions to the newborn screening panels, and now cover up to 53 conditions (NNSGRC 2012). Advocacy has been a significant catalyst for the expansion of newborn screening over time. For example, in 1958 Dr. Robert Guthrie was asked to develop a more effective way to detect phenylalanine levels in infants and children showing signs of mental retardation. Within a very short time, Guthrie discovered that a bacterial inhibition assay, using a very small amount of blood, could be used to determine the level of phenylalanine in the blood. His innovation did not take immediate hold in the medical community, and even encountered some resistance. However, within five years there was a transition from limited clinical use to mandating screening for all infants born in Massachusetts, a trend that would spread state to state. By the late 1970's PKU screening was mandated in all 50 states. Guthrie's story, and the stories of those that came before and after him, illustrate that innovation in science and technology does not occur in a vacuum; it transpires within the context of political and social forces. **Chapter Two** offers a historical perspective on the creation and evolution of newborn screening programs, in particular the role of advocacy and science in shaping newborn screening policy. Applying policy analysis tools to newborn screening and identifying the advocates can provide insight into how and why newborn screening policy has shifted over time.

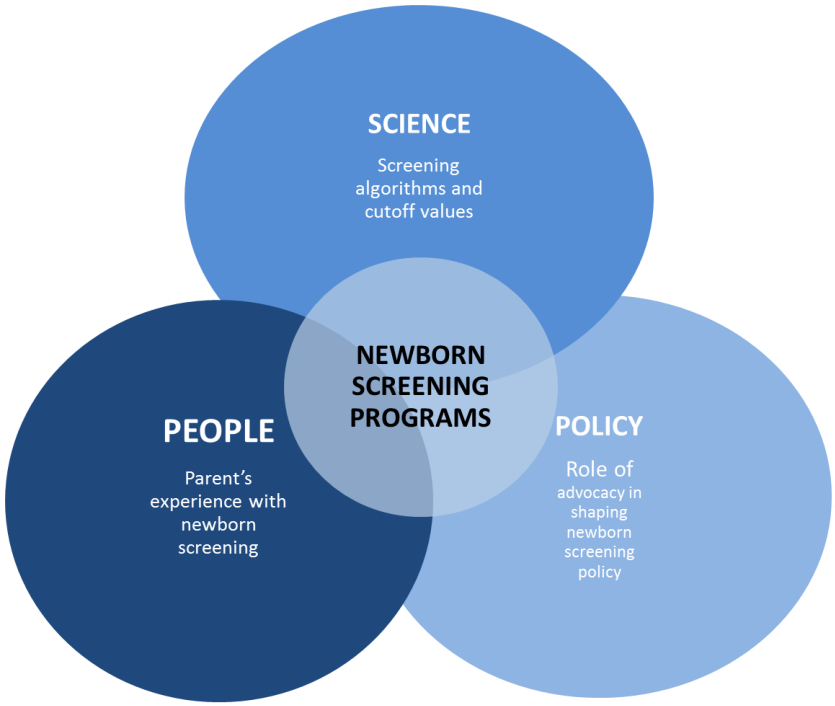
Given the public health setting of newborn screening and the goal to prevent harm, it is important to consider balancing risks and benefits. There are real and measurable benefits to infants with conditions for which effective interventions and treatments exist and are accessible. There are also potential harms, such as the burden of false positives on the public health and health care systems and the potential for negative impacts on the families of infants that receive false positive results (Botkin JR 2006). Hewelett and Waisbren (Hewlett 2006) summarized nine studies published since the early 1970's on parental stress and positive newborn screening results, both true and false positives: 76% of families reported initial anxiety after receiving news of a positive screen, 18% of a families had persistent anxiety afterwards, and one study demonstrated a 23% increase in stress levels for mothers and a 10% increase in stress levels for fathers. The number of false positive results across states increases as more conditions are added to newborn screening programs. For example, the majority of conditions added in the last decade use Tandem Mass Spectrometry (MS/MS) as the screening platform. It has been estimated that in one year in the United States the number of false positive screening results for conditions found through MS/MS would be approximately 2,575 in the best-case scenario (spec 99.995%), approximately 51,059 in the worst-case scenario (spec 99.9%), and somewhere around 25,644 infants (spec 99.95%) in an intermediate scenario, which is the most likely scenario (Tarini BA 2006).

When an infant has an abnormal screen, follow-up action is required. This may be additional screening or diagnostic testing. Either way, parents are suddenly engaged in a public health screening program that for the most part goes unnoticed or unrecognized by new parents (Davis 2006) (Kladny B 2011) (Tluczek 2005). New parents are inundated with information and instructions, especially during what is now a very short window of time that mom and baby are in the hospital. Parent education, although difficult, is very important. According to recent studies, parent education is lacking and can be improved in both the pre and post-natal environment (Tluczek 2005). To do this, we must understand parents' baseline knowledge of newborn screening and appreciate the impact newborn screening can have on parents. Due to increased exposure of newborn screening in both individual states and on the national level, there has been a notable increase in the number of articles published on newborn screening in the last decade, in comparison with previous decades. The articles published address subtopics such as: informed consent to perform screening and store specimens, use of residual specimens, addition of new conditions to the panel, the impact of screening on the medical and public health systems, and the impact of screening on parents and families. Data are limited in both type of knowledge and volume of data as it relates to parents' knowledge of, and experience with, newborn screening. **Chapter Three** offers a structured literature review of what we know, and what we don't know, about parents' experience with newborn

screening. Identifying the gaps in knowledge can direct future research in this area and offer insight into how decisions are made with limited data.

Newborn screening programs have been expanding quite rapidly in the last decade. The addition of a multitude of new conditions to newborn screening programs will unsurprisingly increase the number of infants identified with one of the conditions. In the process of identifying true cases, newborn screening programs will inevitably falsely identify infants as potentially having one of the conditions. When a screening population is large, diverse, and composed primarily of asymptomatic individuals, a balance must be made between test sensitivity and specificity. In order to effectively identify all infants with a disorder, some false positive screening results are to be expected. **Chapter Four** tackles the issue of test sensitivity and specificity and the multitude of factors that go into creating screening algorithms. Revising the screening algorithm for Congenital Hypothyroidism (CH) in the state of Washington provides a real world case study of reducing the false positive rate while maintaining a good ability to identify true cases. **Chapter 5** summarizes how these factors – science, policy, and people - are all integral to the future of newborn screening.

Figure 1: Intersection of Dissertation Components with NBS Programs



CHAPTER 1: AN INTRODUCTION TO NEWBORN SCREENING

The Basics of Newborn Screening

Newborn screening in the United States is a state-mandated public health function. The newborn screening pathway describes the journey of a screened infant (Riley 2007). It starts with a heel stick to draw blood for a newborn screening specimen. For the vast majority of infants, the process ends when the normal results are reported back to the facility that submitted the specimen or the primary care provider. For infants affected by one of the conditions, the newborn screening process ends once the infant is deemed to be a true positive case, although there is some crossover between newborn screening and the clinical domain until a diagnosis is reached.

Newborn screening is more than just processing blood spot specimens in a lab. A comprehensive program indeed has to ensure that specimens are collected after birth (and in the 15 states that utilize two screens, the program aims to receive second screens collected approximately at 7-14 days). A NBS program should monitor the timely receipt of specimens. A comprehensive screening program will continuously monitor to make sure analytically valid tests are completed in a timely and accurate manner and results are reported to the submitter of the specimen in a clear and concise manner (and to the primary care provider if applicable). A comprehensive program will also work with the healthcare community to ensure that infants with abnormal results receive the proper diagnostic testing to confirm a true diagnosis or indicate the screening result was a false positive. Additionally, newborn screening should follow all true positive cases until the patient is linked to a specialist or medical home. Long-term follow-up is also important, especially for evaluating long-term health outcomes, an area of research that requires more attention in newborn screening.

Newborn screening programs require that every infant undergo a heel prick procedure to draw blood which is then absorbed by special filter paper on the end of a newborn screening specimen card. In the hospital setting, this is done prior to the infant being discharged from the hospital. For example, in the state of Washington the recommended age at collection is between 18 – 48 hours of life. The age at collection is an important factor in screening programs because the testing platforms and cutoff schemes are designed to be sensitive and specific for a specific age range. The recommended age at collection and the number of specimens obtained (one versus two) differs across states; both of these factors influence the type of screening algorithm implemented and therefore state-specific approaches are developed to establish cutoff values and follow-up protocols. Timing of specimen collection is also influenced by the clinical status of the infant. For example, if a baby is admitted to a neonatal intensive care unit or

specialty care nursery, the specimen is more likely to be collected early (between 1-3 hours of life), prior to the administration of antibiotics, steroids, blood transfusion, or hyperalimentation. These interventions can impact the analytic validity of one or more of the tests used to screen for conditions included in the newborn screening program. Very low birth weight babies may not be stable enough between 18-48 hours of life to collect blood specimens and therefore the blood might be done at day three of life.

In the case of a hospital birth, hospital health care professionals (most commonly a nurse or phlebotomist) are responsible for collecting and submitting the newborn screening specimen card that contains the dried blood specimen to the state or regional public health laboratory, contracted private laboratory, or university-based laboratory. The public health authority to mandate screening again is state specific. One state may only have the authority to mandate screening for babies born in a hospital, thereby not covering birthing centers or home births, although the recommendation that every baby born receives screening still holds. In other states, the mandate covers all births, in which case midwives are also held responsible for submitting specimens for homebirths or babies born in a birthing center. Enforcing newborn screening for out of hospital births is a challenge for states. One approach is to be collaborative with the midwifery community and provide education for students currently studying to become a midwife. If they know why newborn screening is done and the importance of finding babies with these conditions, they will be more likely to participate.

Once the newborn screening laboratory receives the specimen, the appropriate tests are run for the previously designated set of conditions. Results are reported back to the facility that submitted the specimen. The vast majority of results are normal. For example, in the state of Washington there are approximately 87,000 births per year, and in general there are two specimens processed for each baby born, resulting in approximately 174,000 specimens screened each a year (depending on the birth rate). Approximately 5,500 of these specimens will have an abnormal result due one of the following reasons: results are outside of the normal range for a specific condition (~2,100), the infant is identified as a carrier of a hemoglobin trait (~1,200), the specimen is unsuitable due to layering or incomplete saturation or contamination (~1,750), a result for a specific condition is not valid due the infant being on substances that can interfere with the tests such as antibiotics or steroids (~300), the baby expired before the case was resolved (~75), or “other” (~75) (NBS 2012). Out of the ~2,100 abnormal screening results that are outside of the normal range for a specific condition, ~ 170-190 infants will be confirmed true positive cases, and another ~ 1,200 will be identified as having a hemoglobin trait (NBS 2012). There are over four million babies screened in the United States each year, leading to the identification of more than 6,000 babies with one of conditions being screened for across the states (NNSGRC 2012).

When an infant has an abnormal screen, or what some states refer to as "screen positive," for one of the conditions, it will end in one of three ways: 1) a true positive, when the infant is confirmed to have the disease/condition after diagnostic testing (this includes conditions designated on the state mandated panel and conditions identified in the process of screening for the state mandated conditions), 2) a false positive, when the infant screens positive but is confirmed to not have the disease/condition after follow-up screening or diagnostic testing, or 3) a false positive active, when the disease/condition for which the infant screened positive is ruled out, but equivocal laboratory results indicate the infant may or may not have something of clinical significance.

It is important to define the term false positive. The literature offers only a few definitions of the term "false positive," even though the term itself is used often. Prior to June 2006 there were 15 articles that used the term "false positive," only two of which provided a general definition of the term and five of which provided a disease specific definition (Riley 2007). Sorenson et al provide the most complete definition, "if the test results on the repeat specimen are abnormal, the infant is referred for further evaluation. If the repeat test result is normal, the initial result is considered a false positive" (Sorenson JR. 1984). But in the day to day operations of newborn screening, there are several types of false positives and not every type of false positive has the same potential to negatively impact on the parents/infants. For example, within Washington State's Newborn Screening Program (a two-screen state), three prominent types of false positives are possible:

1) *Passive Screen False Positive* – when there is an abnormal first screen with a relatively low positive predictive value, the first line of follow up is to wait for the second screen to come in. Although a second screen is not mandated, it is standard practice in Washington to collect/submit a second newborn screening specimen between 7-14 days of life (~90% compliance rate). If the second screen is normal, the first screen is considered a false positive. In this scenario the screening results, in aggregate, produced the correct outcome (normal screen), even if the first screen was borderline abnormal.
2) *Active Screen False Positive* – when there is an abnormal screen that requires follow-up staff to actively recruit a subsequent screen from the infant due to an abnormal result with a moderate to high positive predictive value. In the Washington NBS Program there are multiple reasons why a follow-up staff would recruit a subsequent screen: the abnormal result is such that an immediate second screen is required to resolve the case (based on the type of abnormal result and potential severity of the condition), the standard practice second

screen was not submitted and is required to resolve an abnormal result, or a third (or in some cases fourth) screen is required to resolve an abnormal result on a second (or third) screen.

3) *Referral False Positive* - infants with an abnormal newborn screening result that warrants an immediate referral for diagnostic testing (i.e. the screening result has a high positive predictive value and the potential negative health outcomes in the interim are severe). Diagnostic testing confirms the infant does not have the condition.

Certainly the impact of false positive results differs depending on what type of false positive result it is. One constraint of trying to understand the various types of false positive results is that state NBS programs are highly varied in both the screening process and follow-up process. For example, a single screen state will not generate a "passive screen" false positive result as described above. Any abnormal result in a single screen state requires active follow-up to request either a subsequent screen or diagnostic testing. In these states parents are more likely to be informed of an abnormal screening result because of the need to bring the infant in for follow-up screen or diagnostic testing. For two-screen states, the second screen is a standard of practice, allowing for results of the second screen to potentially resolve an abnormal result on the first screen without worrying the parents. Likewise, for the "active screen" false positive, some states recommend diagnostic follow-up for all abnormal results, while others may request subsequent screens for borderline abnormal results.

Regardless, anytime there is an abnormal result, newborn screening follow-up is required to interpret the results and notify the appropriate health care provider. Some state newborn screening programs also make recommendations to the primary care provider regarding specific follow-up actions. This may include a recommendation to consult a specialist, perform appropriate diagnostic tests, or look for clinical signs and symptoms during clinical evaluation of the patient. In most states, primary care providers communicate the information to the parents of potentially affected infants and coordinate diagnostic testing, or refer the patient to a specialist or specialty clinic. In some states, follow-up staff assists in arranging diagnostic testing and connecting the infant to a specialist or medical home. The infant is then subject to all of the benefits and drawbacks of the health care delivery system, including insurance coverage or non-coverage.

Public Health Genetics and Newborn Screening

Newborn screening is an important public health screening process, often viewed as a gold standard in demonstrating the intersection of genetics and public health; a serendipitous marriage because conditions are not chosen due to their genetic nature, rather conditions are chosen because they meet a set of

criterion for screening under a public health paradigm (Motulsky 1997). The concept of screening in public health is not unique to newborn screening. The ability to screen populations and identify individuals at an increased risk for developing disease with the intent to intervene and lessen the risk is core to public health. Newborn screening leads to diagnosis and investigation into health conditions that impact the lives of subgroups of the population. For affected infants, newborn screening mitigates morbidity and mortality associated with the rare conditions. Newborn screening actually covers nine out of ten of the essential public health services (Wang 2007).

The genomic era has created a new and powerful gaze, one that allows us to look deep within the human body. DNA is the building blocks of life and the human genome is considered the "blueprint" of the human body (Collins 2003). The scientific and medical community has gained power with the ability to gaze into the body on a molecular level and understand and diagnose disease in a different way. All of the newborn screening conditions are either entirely or partially explained by genetics. This introduces a genomic lens under which newborn screening is scrutinized. The advances in genetic technology have generated concerns about genetic profiling or genetic discrimination and the use of personal genetic information in research. Concerns turned into action in Minnesota and Texas. Both states have had to destroy residual blood spots from newborn screening specimens and change the state laws in response to parental and community fear that the government is using baby's DNA without parental consent. The application of genomics in medicine and public health has also generated concern regarding what to do for the "unpatient," a patient deemed to be at genetic risk for disease but is not symptomatic. There is also uncertainty about what to do with patients that are not symptomatic, but may become symptomatic: "patients-in-waiting." Although this describes a very small subset of the newborn screening population, this umbrella concept emerged in the newborn screening literature, describing those under medical surveillance but who are somewhere between health and disease (Jonsen 1996) (Timmermans 2010).

Genetic or not, newborn screening is conducted under a public health paradigm, which offers its own lens in approaching programmatic issues. Screening is designed to provide information regarding the potential risk of future disease, for the purposes of prevention (Brosco 2006). With regard to conditions covered under newborn screening, if infants are not identified early, serious negative health outcomes such as severe cognitive impairment, failure to thrive, and/or death can occur(Green 2004). In the case of phenylketonuria (PKU), the first condition states mandated screening for, significant cognitive impairments occurs if the condition goes untreated (Kahler 2003). Congenital hypothyroidism, the most common preventable cause of mental retardation, if identified and treated early (through newborn screening), both physical and cognitive impairments can be avoided (LaFranchi 2011). In one of the

more recent additions to newborn screening, medium chain acyl-Coenzyme A dehydrogenase deficiency (MCADD), there is risk of seizures, coma, brain injury, or sudden death if the infant is under fasting or stressful conditions (Blois 2005) (Kahler 2003). Identification of galactosemia in the first week of life may make the difference between life and death of an affected child, the treatment for which is a switch from breastfeeding or regular formula to a soy based formula. Individuals with galactosemia cannot break down galactose, the sugar in milk, and the build-up of galactose in the body causes a metabolic crisis. Infants with untreated galactosemia most often die from overwhelming e. coli sepsis (Kaye 2006). Newborn Screening's ability to prevent morbidity and mortality makes it a ringer for satisfying one of the core functions of public health: prevention.

Newborn Screening Guidelines

As technology has advanced, new testing methodologies have been developed and new conditions have been discovered. Moving forward, new treatments and interventions for rare conditions will be developed. By the 1960's there had already been significant advances in medicine and public health with regard to the ability to use laboratory tests to assist in diagnosing conditions. Dr. Guthrie and many of his colleagues had been successful in mandating newborn screening for phenylketonuria (PKU) in most states by the end of the 1960s. At the time, there was not a systematic way to determine how, if, and when large-scale or population based screening should occur. Given the general need for screening guidelines and the use of public dollars as well as the ethical, legal and social implications of conducting population based screening, the World Health Organization produced a Public Heath Paper in 1968 entitled Principles and Practice of Screening for Disease, co-authored by Wilson and Junger (Wilson 1968). This paper identified the aim of early disease detection, different screening methods, evaluation of tests, and most importantly established the first parameters for population based screening that could be to newborn screening . These suggested criteria set an important precedent by establishing the idea that there ought to be a systematic way of evaluating whether or not large scale screening is justified for a given condition. Wilson and Junger noted that these "principles" are especially important if screening is being carried out by a public health agency. Decisions of whether or not to mandate newborn screening for a specific condition was, and remains today, an individual state responsibility; therefore the actual criteria used and subsequent breadth of screening varies across the nation. A few overarching principles have held over time, including the notion that a public health screening platform ought to accurately identify infants with the condition and there should be a reliable and available intervention that positively impacts the infant's health.

In 1975 the National Research Council of the National Academy of Sciences, critical of how newborn screening had been developed, revisited the idea of newborn screening criteria, as did the Committee for the Study of Inborn Errors of Metabolism of the National Academies of Science who published *Genetic Screening: Programs, Principles and Research* (NAS 1975) (Simopoulos 2009). In 2000 the American Academy of Pediatrics (AAP) conducted an in depth analysis of the history of newborn screening and confirmed a new set of criteria that reflected the progress and changes that had occurred since the NAS Report. Although this report came out after the introduction of MS/MS, it did not address all of the issues this new technology introduced (AAP 2000). The AAP report "raised concerns about what it saw as the potential risks of inappropriate newborn screening" (AAP 2000). The AAP was hopeful regarding the ability of newborn screening to positively impact the lives of children, but also cautious regarding how newborn screening should be accomplished.

Four main themes have maintained throughout the evolution of criteria: benefit to the newborn, adequate diagnostic capabilities, availability of treatment, and a cost-benefit and/or public health benefit. Some of the initial insight in the 1968 WHO and 1975 NAS has been deemphasized over time: understanding the history and etiology of the disease, the importance of the disorder to the public (prevalence of disease), and the inclusion of informed consent (AAP 2000) (Simopoulos 2009).

Before the initial application of MS/MS in the late 1990's, a state's ability to screen was largely limited to identifying one condition at a time, using a variety of testing mechanisms. Even after the introduction of MS/MS that allowed a high through-put way of screening for a multitude of conditions using one punch from a blood spot, there is still a need of multiple testing platforms. For example, immunoassays are used to detect abnormalities in analyte levels in screening for congenital hypothyroidism (CH), congenital adrenal hyperplasia (CAH), and cystic fibrosis (CF) (Clague A 2002). An enzyme activity test is used to detect biotinidase and galactosemia in newborns. However, the application of MS/MS to newborn screening opened the door for screening for conditions that previously required molecular testing or another methodology that was not practical for population-based screening (McCandless 2004) (Riley 2007).

The Health Resources and Services Administration (HRSA) commissioned a report from the ACMG entitled *Newborn Screening: Toward a Uniform Screening Panel and System* in which 84 conditions were evaluated; 29 conditions were recommended for mandated screening programs and an additional 25 conditions were recommended for screening within a research context. The latter are useful in assisting in differential diagnosis of conditions targeted on the mandated screening panel. These conditions are

reported to be easily identified using MS/MS technology and clinically significant (even if no treatment was available at the time of decision-making) (ACMG 2006). Some of the 29 new disorders proposed have come into question because they do not stand up to the original criteria set forth first by the World Health Organization in 1968, the National Academy of Sciences in 1975, or the American Academy of Pediatrics in 2000. Articles were published in 2006 representing different responses to the ACMG's report. Rodney Howell's article represented those on board with the ACMG recommendations who wanted to move full speed ahead with expanding newborn screening (Howell 2006). A more restrained view was put forth by Jeffrey Botkin et al presenting a "proceed with caution" approach that called for more research and heightened scrutiny over criteria (Botkin JR 2006).

The criteria have certainly evolved over time and now reach a pivotal point in their evolution. In 2008 the President's Council on Bioethics convened and neither rejected nor fully embraced the ACMG's set of criteria and suggested guidelines for newborn screening (PCOB 2008). Instead, the Council developed a hybrid model, making the following recommendations: the Wilson-Junger criteria are valid and relevant and only those disorders that clearly meet these criteria should be mandated; screening for disorders that do not meet all of the criteria can be included under a research paradigm on a voluntary basis; secondary information does not need to be passed on to the physician (i.e. detection of a poorly understood condition due to the need for a differential diagnosis); there needs to be more work toward a state consensus on which conditions should be mandated; recently added disorders should be re-evaluated to make sure they meet the original criteria; and finally the Council rejected the notion of reporting everything we are capable of finding even if the condition is poorly understood or has no known intervention (PCOB 2008). The criteria for newborn screening are somewhat of a moving target; evolving over time because of changes in technology, availability of scientific evidence, political and social pressure, and other external influences.

The onus now rests with the Health and Human Services (HHS) Secretary's Advisory Committee on Heritable Disorders in Newborns and Children (SACHDNC). With the intent of establishing consistency across states, this Committee recommends conditions that the committee has deemed appropriate for state mandated testing. The Advisory Committee makes a recommendation and the Secretary of Health and Human Services has to approve the recommendation before it can be considered an official federal recommendation. The recommended conditions are commonly referred to as the "recommended uniform screening panel," or the "universal newborn screening panel," or simple the "core panel." This uniform panel is a recommendation made by HHS, a guideline for states; currently the federal government has no authority over state newborn screening programs. States vary in adopting the Committee's

recommendation because each state has its own process for adding conditions to the state-mandated newborn screening panel. As data becomes available and advances are made in testing methodologies and medical interventions, the SACHDNC revisits the lists and considers adding conditions, based on nominations. Anyone can nominate a condition for consideration; the SACHDNC decides which nominated conditions will move on to an external evidence review. The external review looks at three components: systematic evidence review (including unpublished evidence), an estimation of benefit and harm, and an assessment of laboratory readiness, and the feasibility of implementing comprehensive NBS from the state perspective (HRSA 2011). The number of recommended uniform newborn screening conditions is up to 31 now, with the recent addition of Severe Combined Immunodeficiency and Critical Congenital Heart Disease (HRSA 2011) (SACHDNC 2010).

How, why, and when states decide to add conditions to newborn screening are tough questions to answer. Who is making the decisions is not clear cut either. Like other aspects of newborn screening, the authority to make changes to newborn screening services depends on the state and the type of decision being made. In general, the final decision is the responsibility of an advisory committee, a board of health, the legislature, or agency leadership. A state may require legislative approval to add new conditions to the newborn screening program whereas programmatic decisions to change testing platforms or adjust cutoffs may fall to the director of the newborn screening program. Regardless of who is responsible for making the decision, or at what programmatic level the decision is made, one thing is clear: the individual or entity making the decision needs up-to-date information on scientific discoveries, technological advances, new screening modalities, newly discovered conditions, new or improved treatments for existing conditions, and data from other states on screening platforms and establishing normal/abnormal thresholds. Chapter 4 discusses one aspect of how science informs the practice of newborn screening - through utilizing data - subsequently impacting the population being served by newborn screening. Another factor in the practice of newborn screening over time is advocacy, discussed in further detail in Chapter 2.

CHAPTER 2: ADVOCACY AND NEWBORN SCREENING POLICY

Newborn screening policy had been flying somewhat under the radar screen; it is not a "hot topic" that gleans a lot of media attention and there has been limited involvement of the general public with regard to newborn screening research and practice. However, in the last five years, there has been a surge in public interest in newborn screening, undoubtedly a result of advocacy. There is a web of advocates, some have been involved from the beginning and some of weaved into the web over time, that are active and vocal in setting the newborn screening agenda. Specific subsets of the population acutely concerned with newborn screening policy include disease specific advocacy organizations actively such as: the Cystic Fibrosis Foundation, Hunter's Hope Foundation, the Sickle Cell Disease Foundation, March of Dimes, and Genetic Alliance. There is also a subset of the research community that specializes in rare metabolic and genetic disorders that has an interest in promoting newborn screening for all infants born in the U.S. Likewise, a subset of the research community is interested in ensuring that newborn screening policies are balancing benefit and harm and acknowledging the ethical, legal, and social implications. A historical look at the role of advocacy sets the stage for evaluating shifts in newborn screening over time.

PURPOSE

Explore the role of advocacy in changing newborn screening policy over time.

SPECIFIC AIMS

1. Describe the historical context of advocacy in newborn screening
2. Analyze case studies of parent driven advocacy in newborn screening
3. Apply the Advocacy Coalition Framework to facilitate our understanding of how advocates inform newborn screening policies

THE HISTORY OF NEWBORN SCREENING, AN ADVOCACY PERSPECTIVE

The year was 1934 and a Norwegian physician, Dr. Asbjorn Folling, was determined to find the reason for the severe mental retardation in two siblings. The mother of these siblings, Mrs. Borgnyn Egeland, was "relentless" in her pursuit of understanding her children's condition and was somewhat of a thorn in the side of Dr. Folling (Koch 1997). This mother could very well be considered the first advocate for NBS. Dr. Folling, as a result, went on to discover that some children with mental retardation had high levels of phenylpyruvic acid in their urine and was able to make the connection between this and mental retardation as a phenotype. He had discovered phenylketonuria (PKU). He would later be recognized in

1962 by President John F. Kennedy as the first recipient of the Joseph P. Kennedy International Award in Mental Retardation (Centerwall 2000).

Following this discovery, Garrod recognized that "variation in Mendelian heredity could explain an 'inborn error of metabolism'" and concluded that changes in metabolic pathways had the potential to create different health outcomes in individuals (Scriver 2001). In 1951 another mother entered the picture. Laura Jones had a young toddler with severe mental retardation who was diagnosed with PKU by Professor Horst Bickel at the Children's Hospital in Birmingham, England (Bickel 1996). According to Professor Bickel's accounts, this mother continuously hounded him about treatment options. Bickel hypothesized that there was a causal relationship between excess levels of phenylalanine and the brain damage seen in PKU patients. He worked with several colleagues to develop the first formula in which phenylalanine was removed. This young toddler was the first to receive this formula and the research team noted significant improvement within a couple of months (Bickel 1996). These two mothers were perhaps the first "advocates" in newborn screening because they pushed the researchers and physicians to investigate further and determine the cause and ultimately a prevention strategy for mental retardation in infants with PKU (Riley 2011).

Now that PKU had been discovered and a treatment has been identified Dr. Willard Centerwall, a fellow at the Children's Hospital of Los Angeles, and a father of a child with mental retardation, discovered that if a solution of ferric chloride was applied to the wet diaper of a baby with PKU it would produce a green color (Koch 1997). This "diaper test" provided a way to screen infants, however this did not translate into a mass screening effort. Many cases of PKU were missed at birth and identified much too late, after severe mental retardation had occurred. A year later in 1958 Dr. Robert Warner, who was caring for patients with mental retardation, knew a physician and researcher by the name of Dr. Robert Guthrie and approached him about developing a method for detecting this disorder earlier and with greater effectiveness than the urine test. Dr. Guthrie, considered by most to be the father of newborn screening, accepted the challenge and ended up being very successful. He went on to discover that the bacterial test he had been using in cancer patients to identify various substrates in the blood could actually be modified to identify the presence of phenylalanine in the blood (Guthrie 1996). This bacterial inhibition assay (BIA) is now commonly known as the "Guthrie Spot."

It is interesting to pause and take note of the involvement of physicians and researchers that either have a close personal relationship with someone affected by mental retardation. Personal passion is often a catalyst and motivating factor for researchers and clinicians alike. "Dr. Guthrie was the father of a child

with mental retardation (unrelated to PKU) and was quite active in the local Buffalo Chapter of the New York State Association for Retarded Children (Guthrie 1996)" (Riley 2011). Within a few years of his discovery, Guthrie would become personally affected by PKU when one of his own family members, a niece, was diagnosed with PKU. At this point many clinicians were still using the old diaper test (Koch 1997), and Guthrie's niece ended up being diagnosed at 15 months of age when she was already suffering from severe cognitive impairment, prompting Guthrie to take action. He realized it could be possible to use this test to screen newborns before developmental problems occurred, so in 1961 he pitched the idea to Dr. Alfred Yankhauer, the Director of Maternal and Child Health for New York State (Riley 2011). Dr. Yankhauer encouraged Guthrie to do a pilot project and by1963 the project had expanded to 29 states (Guthrie 1996). Guthrie and colleagues had screened 400,000 infants and identified 39 cases of PKU without missing any cases among those screened (Guthrie 1996). The personal involvement in developing the technique, his own experience with caring for a family member with mental retardation, and his personal relationship to someone affected by PKU created a strong advocate in Dr. Guthrie. He would end up dedicating much of his life's work to making sure children had access to screening.

In her 2000 article recalling the history of genetic disease, M.S. Lindee stated that "disease is often a place where a culture's moral narratives, social organization and economic pressures are made manifest. Just as a mutant fly provides a window into the genome, so disease provides a window into culture" (Lindee 2000). In the late 1950's, federal dollars were appropriated for the purposes of treating and caring for the mentally retarded. It was a topic that was heavy on the minds of many in society. Leaders were looking for a way to alleviate this "burden" on society. In the early 1960's when Dr. Guthrie and colleagues were ramping up to advocate state-by-state for mandated newborn screening there was significant interest in understanding mental retardation and finding cures or preventing future incidence of mental retardation. John F. Kennedy was President of the United States and he had a sister who had been institutionalized for mental retardation (Paul 2000); certainly he was sympathetic to families dealing with this issue. Going back to the beginning of newborn screening and highlighting advocates along the way can teach us about the important role advocacy has played in shaping newborn screening policy over time and perhaps provide insight into how it will continue to impact newborn screening policy.

PARENT DRIVEN ADVOCACY ORGANIZATIONS: THREE CASE STUDIES

Indeed parents have contributed to and influenced newborn screening research and practice through advocacy. Before newborn screening even developed into a population-based screening program, it was the mothers of infants with phenylketonuria (PKU) that made their presence known and spurred on researchers to find out why their children were cognitively impaired and what could be done to prevent it.

As time went on advocacy efforts became increasingly coordinated and centralized, yet the majority of these advocacy organizations were started by a parent, or a small group of parents. Small parent-driven advocacy efforts grew into well-oiled advocacy machines. In order to discuss the broader question of whether parents *should have an impact* on the direction of advocacy, it is important to determine whether or not parents *have had an impact* on advocacy. Three advocacy efforts that were started by parents of children with a genetic-based disorder stand out as informative case studies in demonstrating that parents can indeed be effective in influencing the direction of advocacy: PXE International/Genetic Alliance (pseudoxanthoma elasticum), Cystic Fibrosis Foundation (CFF), Hunter's Hope Foundation (krabbe leukodystrophy). These genetic-specific disease advocacy stories provide models for others to follow.

Case Study 1: PXE International/The Genetic Alliance

Inspired by the work of Nancy Wexler and the Hereditary Disease Foundation in the 1980's which led to the identification of the genetic marker for Huntington Disease, Sharon Terry founded PXE International in 1985, an advocacy organization that has been able to accelerate research on pseudoxanthoma elasticum by encouraging researchers to move beyond the genetic basis of disease and to research on developing prognostic, diagnostic, and therapeutic strategies. Pseudoxanthoma elasticum is an autosomal recessive disease that primarily affects the connective tissue, dermis, retina and cardiovascular system (Naouri 2009). PXE has an estimated prevalence ranging from 1 in 25,000 to 1 in 70,000 births (Naouri 2009). PXE International was successful in part because of a "hybrid" approach to advocacy that pulled ideas from academia, industry, and other advocacy organizations (Terry 2007). PXE International realized researchers could use some incentive to focus on rare conditions. Thanks to the founders of PXE International, a community of affected individuals had been established, which due to the rarity of the disease did not previously exist. The organization realized they had a "commodity" that this community could offer researchers: access to a rare disease population. The idea of a biobank was born (Terry 2007). The biobank would be populated with samples from individuals affected by this rare disease, and access to the samples would be controlled by the advocacy group. This proved to be a useful endeavor. In 2000 the genetic mutations that cause PXE were identified using samples from this biobank. The founders and researchers later went on to patent the PXE gene (ABCC6) and license the test for this gene to outside laboratories, making genetic testing for PXE a possibility (Terry 2007). PXE International successfully used its status as an organized advocacy coalition to develop and maintain a biobank, which in turn was a catalyst for those interested in conducting research on rare genetic disorders. The biobank turned a hard to reach sample population into a population accessible for research.

However, advocacy efforts backing one rare disease lacks power because it lacks large numbers of constituents whom the organization represents. The leadership of the smaller disease groups realized their organizations' missions were similar, as were the barriers they were facing related to researching rare conditions. Sharon Terry and her colleagues who were fighting for rare conditions such as psoriasis, alopecia areata and ataxia telangiectasia decided to join forces (Terry 2007). PXE International aligned with the Alliance for Genetic Support Groups. This support group was founded in 1986 by executive director Joan Weiss; she had a son with a congenital heart defect. Mary Davidson (a parent of a child with a rare genetic condition) later became the executive director and grew the organization to a new level with more involvement in policy issues. This larger organization is now known as the Genetic Alliance (GA). PXE International still focuses on improving the health outcome for people living with PXE, but now the organization has joined forces with hundreds of other small advocacy and support groups to make up The Genetic Alliance, currently being led by Chief Executive Officer Sharon Terry (Lindee 2000). This consortium of smaller disease-specific advocacy groups have banded together to increase their numbers, increase their power, and make sure their voices are heard. The mission of the Genetic Alliance is to be:

> "an international genetics coalition comprised of millions of individuals with genetic conditions and more than 600 advocacy, research and health care organizations representing them. As a broad-based coalition of key stakeholders, the Alliance identifies solutions to emerging problems and endeavors to reduce obstacles to rapid and effective applications of research into accessible technologies" (GA 2012).

Their largest advocacy role has been to support public policies that "speed research applications"(GA 2012). The main purpose behind this organization is to support those with genetic diseases and fight for progress in science that will allow for a better understanding of genetic disorders and perhaps find ways to prevent or treat them.

This case provides an example of how savvy advocates can leverage social, political, and financial will to address health issues, even if the issue is rare. Pooling resources and playing to the strengths of both advocates and scientists gave PXE International and the Genetic Alliance leverage over the direction of research The Genetic Alliance has been able to bring together diverse stakeholders and create partnerships that have facilitated access to information to enable translation of research into services and individualized decision-making. This has primarily been done through the Genetic Alliance BioBank and providing researchers with the rare sample populations (GA 2010). In 2003 the biobank expanded and is now the Genetic Alliance Biobank & Registry with over 10,000 samples and thousands of clinical records. "The BioBank collects, stores and distributes samples in accordance with the procedures and specifications determined by the Genetic Alliance BioBank´s Advisory Committee and approved by the

Genetic Alliance BioBank´s Institutional Review Board (GA 2010). The Biobank has been successful discovering the gene for PXE and CFC and continues to strives to ultimately generate the research needed to develop therapies for genetic conditions.

Case Study 2: The Cystic Fibrosis Foundation (CFF)

Founded in 1955 by a small group of families affected by cystic fibrosis (CF), the Cystic Fibrosis Foundation encourages and supports the development of a cure and discovery of new treatments to improve the quality of life for those living with CF (Marshall BC 2009) (CFF 2012) The families that started the CFF were "determined to find help for their children, who were suffering from a rare disease that no one understood very well, these parents turned fear and frustration into hope and action by creating the Cystic Fibrosis Foundation" (CFF 2009). The foundation has grown from a few families back in 1955 to an organization with 600 employees, 250,000 volunteers, and a budget of $250 million a year (Marshall BC 2009). One of the ways the CFF was so successful is its nationally coordinated effort that is backed by over 80 chapters and branches. The CFF now accredits more than 115 CF care centers, 95 adult care programs and 50 affiliate programs (CFF 2012). These statistics are quite impressive for any advocacy organization. The CFF has really focused on three major functions: 1) acting as the accrediting body for CF Care Centers and Laboratories, 2) facilitating the advancement of research, in many instances by directly funding research, and 3) facilitating and promoting advocacy on the state and federal level. The CFF has also provided a venue for parents of infants and children with CF to express their concern, tell their stories, and connect with other affected families.

Cystic Fibrosis is a chronic, progressive, genetic disease that primarily affects the respiratory and digestive systems. The median age of survival in 1955 when the CFF was formed was less than five years. Now the median age of survival is 37 years (CFF 2009). The CFF has played an active role in moving the research forward that is responsible for the increased life span. The foundation has put significant resources into finding new treatments and improved ways to care for individuals living with CF. As the number of adults living with CF increases due to improved therapies, there are higher demands on the CF Care Centers and clinicians treating adults with CF. The CFF has therefore also focused on improving clinical care, primarily though the Quality Improvement Initiative, which includes Learning and Leadership Collaborative, disseminating best practices, developing evidence-based clinical practice guidelines, initiating a monitoring program, and supporting the development of patient/family advisory boards at the CF Care Centers (Marshall BC 2009). Some of the milestones of the foundation include: establishing a patient registry, formation of a Research Development Program, identification of the CF gene, establishment of the Therapeutics Development Program and Network, and the recent

establishment of the Patient-Assistance Foundation (Marshall BC 2009). The Therapeutics Development Program and Network is a nationwide clinical trials network which has been able to conduct over 50 clinical trials for potential therapeutics, all under the supervision of an independent data safety and monitoring board (Marshall BC 2009).

The CFF is well known for its effective advocacy efforts and large scale national campaigns. They have built trusting relationship with those living with CF and the affected families as well as the scientists and clinicians working on CF; "trust and communication among consortium members are important elements in moving this research paradigm forward" (Marshall BC 2009). The foundation has been able to influence the direction of research not only by directly funding research, but by bringing more attention to the disease and engaging researchers from multiple disciplines. The foundation is dedicated to bringing the best scientist together to solve the tough issues CF presents; "pooling the talent and energy of multiple scientists will hopefully expedite the pace of scientific discovery" (Marshall BC 2009). The National Institutes of Health, *Forbes*, *The New York Times*, *The Wall Street Journal* and *BusinessWeek* have all recognized the CFF for its sound business model (CFF 2012). The CFF is dedicated to improving patient care and the overall quality of life for those living with CF.

The case of the Cystic Fibrosis Foundation is a testament to the power of partnership, in this case the partnership between an advocacy organization and the scientists conducting the research. And much like the PXE case in which the organization realized they need a "commodity" or a resource to offer researchers, the CFF took the lead in organizing affected individuals for the purposes of conducting clinical trials. The CFF is also a great example of the importance of a strategic plan and a good business model in the world of advocacy and non-profits.

Case Study 3: Hunter's Hope Foundation (krabbe leukodystrophy)

Even though advocacy efforts have become centralized with national organizations such as the CFF, there have still been individuals that have had influence over policy; one person can still make a difference. One parent with a powerful story (and a lot of political and financial support) can influence policy, as was the experience in New York. Jim Kelly, a professional football player and hall of famer with the Buffalo Bills, proved this point. Jim Kelly has three children, two girls and a son named Hunter who was born with a disease called krabbe leukodystrophy. Krabbe is a rare disease that affects about one in 100,000 live births in the United States. Krabbe disease was not screened for as part of the New York Newborn Screening program at the time of Hunter's birth. The Kelly family wanted to do something about this disease, as they did not want other parents to have to go through what they went through. The family

started Hunter's Hope Foundation (HHF) in 1997 with a "commitment to increase public awareness of krabbe disease and other leukodystrophies, as well as to increase the likelihood of early detection and treatment" (HHF 2009). HHF is focuses on funding research efforts to identify new treatments, therapies and hopefully someday a cure for krabbe disease (HHF 2009). The mission is four-fold: 1) to broaden public awareness of krabbe disease and other leukodystrophies thus increasing the probability of early detection and treatment, 2) to gather and provide current, functional information and service linkages to families of children with leukodystrophies, 3) to fund research efforts that will identify new treatments, therapies and ultimately, a cure for krabbe disease and other leukodystrophies, and 4) to establish an alliance of hope that will nourish, affirm and confront the urgent need for medical, financial and emotional support of family members and those afflicted with leukodystrophies (HHF 2009). In 2004 the Foundation partnered with the University of Buffalo to establish the Hunter James Kelly Research Institute (HJKRI), which now coordinates all research funded by HHF, both internally at the Institute and externally with researchers across the country (HHF 2009). Research conducted at the Institute will focus on remyelination techniques, while external research focuses on the biology and pathophysiology of krabbe's disease(HHF 2009). The two research foci will hopefully lead to ways to preemptively treat the disease as well as find ways to repair brain damage that has already occurred in children with krabbe's disease. To date HHF as raised over $14 million which researchers working on leukodystrophy and other neurological disease (HHF 2009).

In an effort to encourage early detection and treatment, the Kelly family advocated that the state of New York add krabbe leukodystrophy to the newborn screening panel. In large part due to Jim Kelly's celebrity status and social capitol as well as endless hours of advocacy group by the Kelly family and others, including the March of Dimes, the goal of screening all newborns for krabbe disease in New York was achieved in August 2006. Between August 2006 and June 2008 the state of New York screened 550,000 newborns and picked up four high-risk cases, six moderate-risk cases, and fifteen low-risk children. As of August 2008 none of the moderate or low-risk children have presented with clinical evidence of disease. Two of the four high-risk infants were treated with umbilical cord blood transplantation (Duffner 2009).

New York state is currently the only state screening for krabbe disease (NNSGRC 2012), however it is under consideration in several states. Hunter's Hope Foundation has also been a big supporter of expanded universal newborn screening. In fact, when you go to HHF's website, you are redirected to a page that states "Every child, Every time, Everywhere" (HHF 2009). Jim Kelly continues to be an advocate at the state and national level. This case demonstrates the how effective a high profile

"champion" can be in leveraging political and social capitol. They can bring attention to a rare disease that often is difficult to build momentum around.

Summary of Case Studies

A person who experiences an event is uniquely positioned to explain the event. This first-hand perspective can be beneficial in that the account of the experience is authentic. It can also be misleading in that it represents one account of an experience that may not translate to what others have experienced in a similar situation. To date, parents who have been participating in the newborn screening policy process and sharing their experiences have been parents of affected children. This is certainly the scenario in the three case studies presented. Moving forward with advocacy in newborn screening, advocates ought to be thinking about how to include parents of infants that receive a false positive screening result, a cohort of parents that are generally missing from the conversation. I posit one reason for the omission of this group of parents is that the "cause" is not enough to engage parents in advocacy efforts. There is not a strong enough tie between parents who experience false positives, not in the same way there is a tie between parents of infants with a rare condition. These parents have a stake in issues and make quite passionate advocates. How then can parents of infants with false positive results participate in the advocacy process? What would they advocate for and how would they come together and be effective in influencing the direction of advocacy and newborn screening policy? Additionally, newborn screening currently screens for rare genetic and metabolic disorders. Lessons learned from these advocacy organizations could be applied to future efforts, should newborn screening expand to include other types of conditions. Using a policy framework can help to identify the advocates, their positions, and how different advocates come together around common goals, while maintaining opposing views on finer points of the broader newborn screening conversation.

APPLYING THE ADVOCACY COALITION FRAMEWORK

Background

When approaching policy analysis, models and frameworks are not necessarily interchangeable. Sabatier's work differentiates the two based on the purpose or goal of each one. Models can be used to test, modify, and/or explain theories, but models do not have to be linked to a specific theory (Sabatier 2007). Models, and theories alike, can inform a framework. Frameworks are more often used as tools in determining how theories and models play out in real situations. There are numerous models of policy making, some that were developed with certain areas of policy in mind such as energy, education, or banking and others that can be applied to a variety of issues. According to Sabatier there really are only two papers that have used frameworks for analyzing the policy process, the Advocacy Coalition

Framework (ACF) and the institutional analysis and development (IAD) framework (Sabatier 2007). A model that has a range of applications is likely to be more effective in framing the issues associated with newborn screening (i.e. state mandates; state decision-making versus national guidelines/potential oversight; benefits and harms; opportunity costs; influence of scientific information; influence of advocacy). A framework that has been successfully applied to an issue could also prove to be an effective tool for analyzing newborn screening policy.

Modeling and building frameworks is an important endeavor because the policy process is not solely driven by problems and issues. The political environment can foster new and old issues, bringing them to the forefront; it can also burry issues, making it impossible to get traction even if the issue is worthy of recognition. Kingdon provides a great metaphor of seeds and soil to approach agenda setting and issue identification, "seeds come from many places. Why they germinate, and flourish is much more interesting than their origins" (Kingdon 1995). The origins of newborn screening and the discoveries, developments, and policy issues over time certainly make it an interesting seed to flower story.

Newborn screening is a narrow policy process compared to energy or environmental policy, but many of the components and issues that arise are similar, perhaps just smaller in scale. For example, there are issues of control and power of regulation with regard to environmental policy between the states and the federal government. Likewise, this has the potential to become an issue with newborn screening, given recent developments with national guidelines and new federal funding to support the expansion of newborn screening programs through the "Newborn Screening Saves Lives Act," or SHINE Act. Using a policy framework, such as the Advocacy Coalition Framework, can offer insight into newborn screening policies and stimulate further thinking on questions such as: how has NBS policy shifted overtime, who has influence over the issues, who has been missing from the discussion, what events have influenced change, and what direction is newborn screening policy headed? There are no definitive answers to these questions. However, application of a policy framework will shed much needed light on newborn screening policies, providing a basis for discussion on these important questions.

After considering several different tools/frameworks for analyzing policy, including two models of policy processes - disjointed incrementatlism and the 2-stage model of policy making – the Advocacy Coalition Framework (ACF) emerged as useful way to approach newborn screening policy and provide insight into the changes in newborn screening over time. Sabatier indicates that the ACF addresses "belief change and policy change over long periods," and provides a frame for presenting a multi-faceted approach to understanding policy changes (Sabatier 2007). Newborn screening has been around since the late 1960's

and has been evolving ever since. The most recent evolution is also one of the most significant, involving a shift in venue from state to federal involvement, as well as a shift in the type of scientific evidence and technological advances applied to newborn screening. One benefit to using the ACF is its emphasis on the role of scientific and technical information, both of which are important in newborn screening policy. Some researchers argue that the expansion of newborn screening has been driven by technology; in particular MS/MS. Chalcraft and Britz-McKibbin posit that "the introduction of tandem MS/MS technology has resulted in expanded newborn screening programs since it offers a selective, sensitive, and high throughput platform for the quantitative analysis of multiple biomarkers simultaneously" (Chalcraft 2009). If technology is the primary driving force, how does that impact the other factors playing a role in newborn screening decision-making? Or, is the impact of technology perceived to be more influential than it is? What other factors should be considered?

Another important factor that the ACF emphasizes is that political context matters (Sabatier, 2007). Decisions to expand newborn screening are not being made in a vacuum. There are other public health and non-public health interests at play. The ACF allows for the identification of advocates beyond the usual suspects: legislators, lobbyists, and government agencies (i.e. "the iron triangle") (Sabatier, 2007). The ACF not only includes legislators, agency officials and interest groups, the framework expands to include researchers, journalists, judicial officials, university scientists, and analysts (Sabatier, 2007). This becomes an important issue because physicians, researchers, and university scientists play an especially important role in the most recent policy changes surrounding newborn screening. It is also important to note that when dealing with informal networks, advocates and advocacy coalitions can be difficult to define and there is limited evidence that actors with similar beliefs actually come together in a coordinated fashion (Sabatier, 2007). This will prove to be a challenge with newborn screening policy.

The framework assists in identifying the "actors," or interested parties, and provides a way to group advocates - individuals, agencies, organizations - into advocacy coalitions. Sabatier and Jenkins-Smith define advocacy coalitions as a group comprised of "actors from a variety of governmental and private organizations at different levels of government who share a set of policy beliefs and seek to realize them by influencing the behavior of multiple governmental institutions over time" (Sabatier 1993). The framework encourages approaching a policy issue by identifying who is at the table and assessing how individual "actors" align with other interested parties.

Methods

This data collection and analysis is a pilot test to determine if the ACF can be applied to newborn screening; it does not represent a complete analysis. Data were gathered through December 2008. There are several sources of data for this type of analysis. In the majority of applications of the ACF the data sources are Congressional hearing transcripts, legislation, mainstream print media, and academic journal articles. Lexis Nexis (Academic and Congressional) search engine was used to find all congressional hearings that deal with newborn screening by using the search terms "newborn," "newborn screening," "metabolic screening," and "Phenylketonuria (PKU)," and filtering out other issues and topics regarding newborns that do not apply to public health newborn screening. The Lexis Nexis search is capable of finding other resources as well, such as U.S. and world publications, regulations, and records.

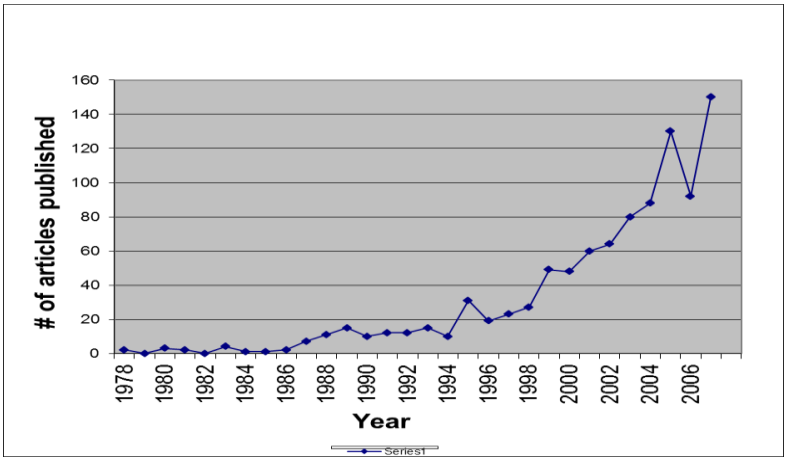
An initial scan of congressional hearings found only five hearing transcripts that directly relate to newborn screening. Because data from federal hearings are very limited, an adjusted approach of utilizing organizations' mission statements or position statements on newborn screening were used to identify where the players are situated and which advocacy coalition best categorizes them. This search was done using Google™ Search (www.google.com), search terms included: newborn screening AND legislation, expanded newborn screening, and newborn screening policy. These searches made it possible to find governmental and non-governmental organizations linked to the issue of expanded newborn screening. Additionally, organizations that were mentioned by name in a congressional testimony were specifically targeted using the Google™ search engine. This type of search is especially important for capturing non-governmental organizations involved in newborn screening advocacy. An organization's mission statement and/or position statement on newborn screening was considered data and assigned to a advocacy coalition accordingly.

TABLE 2.1: Lexis Nexis Search Results

Lexis Nexis Search – December 2008				
Type of Search	**Search Terms**	**Advanced Search Options**	**Number of Hits**	**Years Covered**
major U.S. and world publications	newborn screening	previous 50 years	997	1978 – 2008
Congressional	newborn screening	searching within "hearings" within the previous 50 years	5	2002

The number of academic and mainstream articles about newborn screening has dramatically increased between the 1970's and the 2000's (see Figure 2 below). Speculatively, this is due to the expansion of state newborn screening programs, leading to exponentially more conditions being screened for in states across the country. This generated more scientific and societal attention to newborn screening.

Figure 2: Increase in Newborn Screening Articles over Time



Recognizing that scientific evidence and expert opinion are used in the issue of newborn screening expansion, it is important to make sure the academic/research voice is captured as well. Given the large number of academic articles published on this topic, inclusion of every article is not possible for the scale of a pilot test of the ACF and NBS. A subset of articles related to policy, advocacy, community, risk/benefit, ethical/legal/social issues, guidelines, and standards of practice were considered.

Federal Congressional hearing transcripts, relevant organization's webpage content, relevant U.S. and world publications, and select academic published articles were reviewed to: determine who the advocates are, identify the advocacy coalitions that have formed, and determine which advocacy coalition the advocate(s) belong to. Preliminary data were used to summarize the newborn screening policy subsystem, which provides baseline information about the underlying context of newborn screening as a policy issue.

RESULTS

The ACF provides a suitable framework for working through the venue shifts and policy changes over time, which is useful to the process of understanding newborn screening policy and the role of advocacy. Newborn screening is also useful to the ACF, providing a non-traditional application of the ACF and thus new case study for the policy framework. An initial scan of the information reveals that newborn

screening is posited as a valence issue because the debate is over "how," and to what extent should newborn screening services expand, including whether the services should expand at all. The issue is not whether newborn screening is an important public health service. The policy issues are: what are the most effective and ethical ways to implement newborn screening; which conditions ought to be included in state newborn screening programs; and at what level should the federal government be involved?

Table 2.2 below summarizes the newborn screening policy subsystem, which is essentially the landscape. What are the belief systems at play, the strongly rooted core beliefs and the more pliable secondary beliefs. The table is adapted from Weible and Sabatier's "Table Two. Summary of Application of the ACF applied to the Lake Tahoe Basin" from the book chapter, "A Guide to the Advocacy Coalition Framework" (Weible 2007).

Table 2.2: Newborn Screening Expansion Policy Subsystem (Weible 2007)

ACF COMPONENT	NEWBORN SCREENING EXPANSION APPLICATION
Relatively Stable Parameters	
a. Basic Attribute of the Problem Area	Variation across states of newborn screening services
b. Basic distribution of the natural resources	not applicable
c. Fundamental cultural values and social structure	Autonomy, informed consent, beneficence, "do no harm"
d. Basic constitutional structure	Newborn screening is a state mandated public health service
Policy Subsystem	
a. Territorial Scope	All 50 states
b. Substantive Scope	Newborn Screening Expansion Policy
c. Policy Participants	**State Health Departments** (50 different state programs), **Federal Public Health/Healthcare Agencies/Commissions** (i.e. Health Resources and Services Administration, Association of Maternal and Child Health Programs, and the U.S. Secretary's Advisory Committee on Heritable Disorders and Genetic Diseases in Newborns and Children), **Professional Organizations** (i.e. American College of Medical Genetics and American Academy of Family Physicians), **Advocacy Driven Organizations** (i.e. The Association for Retarded Citizens, Hunter's Hope Foundation, The FOD (Fatty Oxidation Disorders) Family Support Group, and March of Dimes), **Academic/Research Community** (i.e. AACC-an international scientific/medical society of clinical laboratory professionals, physicians, research scientists and other individuals involved with clinical chemistry and the Screening Technology and Research in Genetics group).

Table 2.2: Newborn Screening Expansion Policy Subsystem (continued) (Weible 2007)

ACF COMPONENT	NEWBORN SCREENING EXPANSION APPLICATION
Belief Systems	
a. Deep Core Beliefs	Promote, Preserve, Protect Population Health (Should Newborn Screening be a state-mandate public health program?)
b. Policy Core Beliefs	Evidence-based decisions (What conditions should be included in newborn screening and how is this decided?)
c. Secondary Beliefs	Standardization (Should there be federal guidelines or oversight or should newborn screening remain a state public health function?)
Advocacy Coalitions	Full speed ahead v. Proceed with caution
Policy Broker	American College of Medical Genetics & The Secretary's Advisory Committee on Heritable Disorders in Newborns and Children.
Resources	Clinical research, basic science and translational research, state newborn screening programmatic research
Venues	Currently, the primary venue is state legislation. Other venues include federal legislation and clinical standards of care.
Mechanisms of Policy Change	
a. External Shocks	Application of Tandem Mass Spectrometry (MS/MS) to newborn screening.
b. Policy-oriented learning	Scientific studies and new evidence, clinical experience, state programmatic responses, patient advocacy organizations that fund and endorse specific research
Strategies Employed	
a. Information	Build support for and against expansion, using new and emerging information. Information used to convince decision-makers, influence the public, and hopefully influence other advocacy coalition members.
b. Financial Resources	Newborn screening is primarily a state funded program (fees or state general funds). Federal funds could potentially be distributed to state newborn screening programs.

This analysis focused on an advocacy coalition at each tier of the believe system (see Table 2. 3 on the next page). Tier 1 focuses on deep core beliefs - whether or not newborn screening should be a public s - health program. Tier 2 focuses on policy core beliefs – when and how should newborn screening services expand? Tier 3 focuses on secondary beliefs – state versus federal oversight.

Table 2.3: Advocacy Coalitions

Tier 1: Deep Core Beliefs	
Support NBS	**Do Not Support NBS***
Senator Dodd	*no advocates indicated that they do not support NBS
March of Dimes	
Health Resources and Services Administration - Maternal & Child Health Administration	
National Conference of State Legislatures	
State Newborn Screening Programs	
American Academy of Pediatrics	
National Newborn Screening and Genetics Resource Center	
Rodney Howell (Howell 2006)	
Botkin et al (Botkin 2006)	
Tier 2: Policy Core Beliefs	
Support Full Expansion	**Use Caution re: Expansion**
Senator Dodd	American Academy of Family Physicians
March of Dimes	Screening, Technology And Research in Genetics Project
Hunter's Hope Foundation	State Newborn Screening Programs (split)
AACC - Improving Healthcare through Laboratory Medicine	Association of Public Health Laboratories American
Jill Wood, Parent Advocate	Botkin et al (Botkin 2006)
Rodney Howell (Howell 2006)	
Scott A. Rivkees, MD	
Tier 3: Secondary Beliefs	
Federal Oversight	**Maintain State Authority**
Health Resources and Services Administration - Maternal and Child Health Administration & National Newborn Screening and Genetic Resource Center	National Conference of State Legislatures
Centers for Disease Control and Prevention	State Newborn Screening Programs
Rodney Howell (Howell 2006)	Association of Public Health Laboratories (split)
Botkin et al (Botkin 2006)	
American Academy of Pediatrics	

There were two main advocacy coalitions identified, a "full expansion" vs. "cautious expansion" advocacy coalitions. Alliance with one of these groups was not predictive of supporting federal versus states with regard to newborn screening oversight. It is likely that advocacy coalitions are changing venues from state to federal legislation to increase the swiftness of newborn screening expansion and impact of NBS policy change on a larger scale.

DISCUSSION

The initial scan of available data suggests that the ACMG Report was indeed a focusing event and the members of the advocacy coalitions have rallied around the publication of the report. The second proposition is that the push for a venue shift from state to federal oversight in newborn screening is in response to varied uptake of a policy change to expand newborn screening services at the state level. Data gathered were less clear on this issue. Most members of both advocacy coalitions actually reference the disparity across states regarding how newborn screening programs are developed, implemented, and run. One could infer from this data that there is a strong likelihood that advocacy coalitions are changing venues from state to federal legislation in order to move forward at a faster pace, as well as provide consistency across states.

Newborn screening is a public health service that has historically been designated as a state function. Recent events have shifted newborn screening policy into the realm of federal recommendations and federal legislation. The ACF can be used to address whether there was a focusing event or a change in social or cultural influences; whether this is an issue of technology or scientific evidence driving change; and what role advocacy coalitions play in the venue shift and movement toward policy change. Along with several factors, how policy shifts transpire and who is involved determine the inclusion or exclusion of issues. When decisions are being made, is the issue of an increase in the number of false positives even part of the discussion?

Due to the nature of the policy issues surrounding newborn screening and the involvement of advocacy organizations, government, and academia, the Advocacy Coalition Framework provides a useful framework in working through the venue shifts and policy change over time. What makes this an important issue and what this pilot study can contribute to the ACF discourse is that the framework can be applied to non-traditional valence issue; an issue in which there is not two dichotomous groups arguing completely opposing viewpoints, rather the disagreement comes into play in the details of the issue. Newborn screening is a classic example of a valence issue; the interested parties involved agree that there should be screening for newborns, but do not agree to what extent services should expand and be funded

with public dollars and infrastructure. More specifically, the two advocacy coalitions differ in regards to what the screening programs should entail, how the programs should be implemented, and who should be responsible for coordinating and funding the expansion of newborn screening services. Although it should be noted that being a member of the "full speed ahead" or "proceed with caution" advocacy coalition does not determine whether the actor believes regulation and oversight should remain at the state level or involve the federal government. The majority of actors in the newborn screening policy subsystem believe there is utility in federal involvement, whether that be through funding, legislation, or an oversight committees is dependent on the individual actor's belief system.

CONCLUSION

Looking back to the pioneers of PKU, it is clear that advocacy was instrumental in initiating newborn screening and continued to be so every step of the way. Dr. Guthrie was a prominent researcher; however he was met with opposition when it came time to implement screening for PKU. He was persistent and garnered support from several other organizations – National Association for Retarded Citizens and the March of Dimes – to continue what some have referred to as his crusade. Dr. Guthrie dedicated his life's work to PKU and NBS. There have certainly been many advocates to follow. Certainly parents continue to be advocates, as evidenced by the case studies presented. Individual parents, like Jim Kelly in New York, have been key factors in changing newborn screening policy by tapping into their social and political capitol. Other parents have used their talents to bring organizations together in order to garner a bigger voice by representing and organizing difficult to reach populations, like Sharon Terry and her collaborators did with Genetic Alliance. Each time a new condition is considered for newborn screening, each time funding is raised for research on a condition that is already being screen for, or a condition that a group of advocates would like for states to include in NBS.

CHAPTER 3: PARENTS AND NEWBORN SCREENING

Every state is now screening for at least 27 conditions, ranging in severity as well as effectiveness, cost, and accessibility of corresponding treatments or interventions. Given newborn screening continues to expand within a public health setting, there is a need to think about balancing benefit and risks. There is the obvious benefit of decreased morbidity and mortality when infants with a treatable condition are identified early and started on treatment before negative sequelæ present. There are also potential risks involved with newborn screening; the literature mainly points to the psychosocial implications of false positive results. Given the advances in technology, research, and medicine, NBS will continue to evolve. As this evolution happens, it is important to identify how NBS programs affect parents (or primary care givers) and infants. Parents are, by proxy, the population being served by this public health program. There has been research done on different aspects of parental knowledge and experience with newborn screening. What is unclear at this point is what is known in aggregate and what gaps in knowledge exist.

PURPOSE

The purpose of this structured literature review is to describe the evidence base for understanding parents' experience with, and knowledge of, newborn screening - both parents of infants with normal results and parents of infants with abnormal results. This review provides information that will help researchers, newborn screening practitioners, clinicians, and affected families better understand the family's perspective as they move through the newborn screening pathway. This includes how parents are informed about newborn screening, when education about newborn screening is provided and in what modality, if parents remember the initial specimen collection (heel stick), was there informed consent, when and how are results communicated to parents, and in the event of an abnormal result – when, how, and by whom are abnormal results being reported to the family, and how families experience both true and false positive abnormal newborn screening results.

Synthesizing the knowledge generated in the last decade about parents and newborn screening paints a more complete picture of parents' experiences with, and knowledge of, newborn screening. The compiled information ought to inform decision-makers about newborn screening in a way that provides a more complete depiction of the overall impact of expanding newborn screening services.

SPECIFIC AIMS

1. Review the recent literature on parents' knowledge and experiences with NBS (2000-2012)
2. Characterize recent literature in terms of methodology, disease of interest, and conclusions.
3. Identify gaps in knowledge about parents and newborn screening

BACKGROUND/SIGNIFICANCE

To date, there has not been a comprehensive literature review of what is known about the relationship between newborn screening and parents. This structured review will fill this gap in the newborn screening literature and identify areas that require more research. There have been a few reviews written that focus on certain aspects of newborn screening (i.e. identifying carrier status or false positives) or specific conditions (i.e. cystic fibrosis or hearing screening). For example, Waisbren and Hewlett conducted a review of newborn screening studies focused on the impact of false positive results (Hewlett 2006). As of June 2006, there had been nine studies published since the early 1970's on parental stress and positive newborn screening results, both true and false positives, that report "false-positive screening results have been associated with increased anxiety and stress in parents of infants who require follow-up testing" (Hewlett 2006). Two of their observations were that parental stress and psychosocial effects of following-up positive newborn screening results ranged from 76% of families reporting initial anxiety, to 18% of a families reporting persistent anxiety afterwards and they found that studies have documented a 23% increase in stress levels for mothers and a 10% increase in stress level for fathers of infants receiving false positive results (Hewlett 2006). The structured review presented here identifies 11 more articles that have been published since 2006 that focus on false positive results. In the last ten years, the prominence of the topic of false positive results in the newborn screening literature indicates it is a salient issue in the larger newborn screening conversation.

There have also been review articles that concentrate on one condition, for example cystic fibrosis (CF). In 2003, Parsons and Bradley reviewed the evidence on the psychosocial issues raised by NBS for CF, in particular the issues of parental attitudes to screening. The article addressed four major issues with regard to psychosocial implications of CF NBS. The first issue was parental attitudes toward NBS and the authors concluded the literature provided evidence of widespread support of NBS for CF, even though people's reasons for supporting it may be different (Parsons 2003). Secondly, the authors addressed families with an affected infant and wanted to know if NBS causes more distress or disruption than picking CF clinically? The authors reported that "it has been seen from the published research that diagnostic distress, although considerable, is no more extreme following a newborn diagnosis than a later clinical one, nor is the mother-baby relationship more negatively affected. It has also been shown that

newborn screening offers the potential to avoid the stresses of diagnostic delay and the psychological benefit of involvement in prophylactic treatment" (Parsons 2003). Thirdly, they addressed an additional issue that can arise when screening for CF - families may be confronted with having an infant identified as a carrier for a CF mutation. The authors suggest this could result in possible stigmatization, potentially burden the family with information that could impact other family members, or uncover misattributed paternity (Parsons 2003). The recommendation to address this was aimed at policy-makers, leaving it to the policies in place to ensure sensitivity in service delivery in order to minimize distress and make sure families are given "sufficient, appropriate information," and ensure health professionals are aware that some families may continue to worry and require ongoing support (Parsons 2003). Lastly, the authors looked at lessons learned regarding service delivery and concluded that a few changes in practice could alleviate some of the anxiety and stress associated with NBS for CF. First, minimizing the time it takes to determine if a positive screen is a true or false positive. The author indicated more research is needed on the type, timing, and amount of information given prior to the sweat test in order to ensure appropriate support is given to parents following the sweat test (Parsons 2003).

The reviews published to date have been helpful in synthesizing the information available, however, there has not been a comprehensive, structured review of the existing literature as it relates more broadly to parents and newborn screening. Because parents are a crucial target audience for newborn screening, meaning parents are ultimately the ones dealing with newborn screening results, it is important to systematically articulate the knowledge base - what we know, and don't know, about parents' experience with, knowledge of, and understanding of newborn screening. Clayton posits that NBS is a public health activity that actually discourages parental involvement, in part because "the practical and ethical factors that shape typical clinician-parent interactions simply did not apply in the same ways to newborn screening" (Clayton 2005) when the state programs were developed. The consolidation of information will be a resource for state NBS programs because the state programs do want to engage parents and maximize the benefit of screening. State NBS programs are continuously faced with policy decisions to add conditions to the mandated newborn screening panels or programmatic policy decisions about where to set cutoff values or how to conduct short and long-term follow-up. This structured review will be a resource for state newborn screening programs and will be useful for future policy decisions. This review will provide a guide for researchers interested in the impact of newborn screening on parents; identifying the knowledge that has already been generated and the gaps in knowledge will inform future research efforts.

METHODS

A structured literature review was conducted in order to determine what has been reported in the literature about parents' experience with, and knowledge of, newborn screening as well as the educational needs in order to identify gaps in knowledge and discuss how decisions are made with limited data. All searches were conducted using OVID's PubMed database, accessed through the University of Washington Libraries web interface, and limited to articles written in English. Prior to conducting PubMed searches, inclusion/exclusion criteria were established. Much of the focus and attention on newborn screening has come after state newborn screening programs started to substantially expand the number of conditions on screening panels. The rapid expansion of newborn screening began in the early 2000s, prompted by technological advances in newborn screening techniques that made it possible to conduct high throughput screening for multiple conditions with one punch of the blood spot.

To define the timeframe of interest, articles published prior to the year 2000 were excluded. Articles published during or after 2000 were included. The second criterion for inclusion was that the article had to be related to state newborn screening (blood spot screening) or newborn hearing screening. For example, some articles were related to screening newborns for other types of conditions or risk factors. Articles were excluded that were not directly related to state newborn blood spot or hearing screening. The next criterion for inclusion was that the research had to be conducted in the United States and articles had to be published in a journal published in the United States. The reasoning here is that newborn screening, both blood spot and hearing, are structured and administered differently in other countries. Due to the variation across countries, research and information gathered in non-US populations is not generalizable to the US population, therefore articles not based in the US were excluded. Finally, articles had to present a parent component; the article had to provide new data, information, analysis, review, or discussion about parents in relation to newborn screening. Articles that did not include a parent component were excluded. The search terms used are listed in Table 3.1 below.

RESULTS

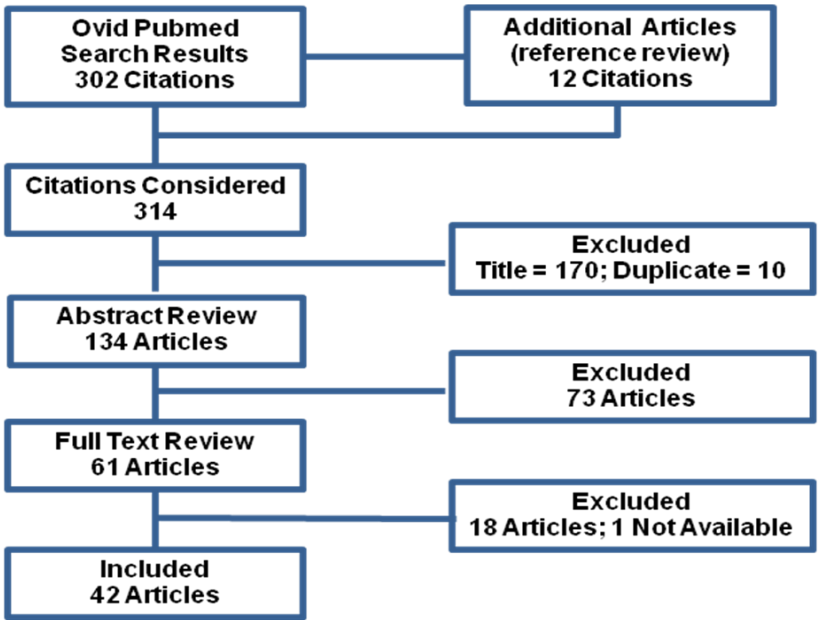
The initial search yielded 302 citations. The references in each article included were reviewed to determine if there were additional articles that were appropriate for inclusion in the literature review. An additional 12 articles were identified from the references used in one of the articles identified in the PubMed search. A total of 314 articles were considered for inclusion. Based on an initial review of the title and short description, 170 articles were excluded from the literature review because the article was either not related to state blood spot or hearing newborn screening, or was published before 2000. Duplicate articles were assigned to the first search that generated them and removed from subsequent

searches. Ten articles were removed from one of the searches due to duplication. The next step in the process involved full abstract review for each article. Upon abstract review, 78 articles were excluded: 11 were not related to state blood spot or hearing newborn screening, 50 were not conducted in the United States, and 12 did not have a parent component to the article. Sixty articles qualified for a full review. After a full review of the articles, another 17 articles were excluded: two were not related to state blood spot or hearing newborn screening, three were not in the US, and 12 did not have a parent component. One of the sixty articles considered for full review was not attainable, leaving 42 articles that were given a full review and included in the literature review. Thirty one of the 42 articles present original data, seven are review articles, one is a workshop summary, two are ethical or legal analyses, and one is a policy analysis. The flowchart below (Figure 3) documents the search results, current through October 2011. See Appendix A for a detailed account of the literature review.

Table 3.1: Structured Literature Review Search Terms

PUBMED SEARCH TERMS	# of Citations
newborn screening, parents, experience	108
newborn screening, parents, views	32
newborn screening, parents, perspective	25
newborn screening, false positives, parents	19
false positive, newborn screening, impact	48
psychosocial, false positive, newborn screening	14
newborn screening, parents, decision-making	56
articles added in based on Reference Review	12
TOTAL	**314**

Figure 3: NBS and Parents Structured Literature Review – Search Results



Each article was carefully reviewed and results were abstracted for the following: author(s), title, citation, purpose, condition (if specific), study type, sample size, methods, findings, and conclusions. In addition, discussion topics were tracked to determine what areas of newborn screening have been presented in the literature. Six major discussion topics were identified: 1) parents' experience with NBS, 2) parents' knowledge of NBS, 3) parents' education on NBS, 4) the impact of false positive results on parents, 5) the impact of true positive results on parents, and 6) informed consent.

FINDINGS

The majority of articles reviewed included more than one of the six discussion topics identified. Of the 42 articles reviewed, 25 dealt with parents experience with NBS, 26 discussed parental knowledge, 27 addressed parent education, 19 delved into the impact of false positive results, 10 investigated the impact of true positive results, and 9 discussed informed consent. Thirty-two of the articles were condition specific; thirteen of the articles were specific to Cystic Fibrosis (CF), seven addressed newborn hearing screening, six were related to metabolic/biochemical conditions, two focused on endocrine (one specifically on congenital hypothyroidism), one article focused on hemoglobinopathies, and one on conditions associated with identifying carrier status. Conditions not currently included in newborn screening are also represented in the literature: three articles covered Fragile X and one was related to

Pompe Disease. These conditions have been or will be considered for newborn screening. The remaining nine articles did not focus on specific conditions, rather discussed NBS in general (see Table 3.2 below).

Table 3.2: Articles Included in the Literature Review by Condition

Condition	# of Citations
Cystic Fibrosis	13*
All NBS - General	9
Hearing	7
Metabolic Conditions	6
Fragile X	3
Endocrine	2*
Hemoglobinopathies	1
Conditions associated with identification of carrier status	1
Pompe Disease	1
*one article covered CF and CH	

Both quantitative and qualitative methodologies were represented. Fifteen of the studies employed qualitative approaches, six utilized quantitative approaches, and eight studies used a mixed methods approach using a combination of qualitative and quantitative methodologies. Of the remaining articles, seven are reviews, one is a workshop summary, one is a case study, one is an evaluation, and three are ethical, legal, or policy analyses. See summary Table 3.3 below and Table 3.4 for details on the study type, sample size, and methods used.

Table 3.3: Articles Included in the Literature Review by Methodology

Methodology	# of Citations
Qualitative	15
Quantitative	6
Mixed Methods (Qualitative & Quantitative)	9
Review	5
Analysis (Ethical, Legal, Policy)	4
Workshop Summary	1
Case Study	1
Evaluation	1

With regard to the distribution of articles over time, there was a peak between – 2003 and 2006 – and a few quieter years – 2000, 2001, and 2002. MS/MS entered the newborn screening scene in the late 1990s/early 2000s, prompting more research on how this testing platform could maximize newborn screening efforts. States began to consider adding a plethora of conditions to newborn screening panels, generating even more research interest in these rare conditions, but also conjuring more attention to the ethical, legal, social, and policy issues that arise when expanding newborn screening services. This may, in part, explain the increase in articles published in 2003. Then the ACMG came out with their first set of universal newborn screening recommendations in 2005, publishing the report in 2006, which is another prolific year for newborn screening articles. This review focuses on parents and newborn screening, but the increase in the number of articles published in the last six years parallels the increase in the number of articles published on all aspects of newborn screening.

The 42 articles in this literature review represent the last 12 years of research and thinking on parents and newborn screening: four articles were published in 2011, five in 2010, four in 2009, four in 2008, two in 2007, seven in 2006, five in 2005, one in 2004, seven in 2003, one in 2002, one in 2001, and one in 2000 See Table 3.4 below for a complete list of all 42 articles.

TABLE 3.4: Summary of Literature Review of Newborn Screening and Parents

Author/Year	Condition	Sample Size	Study Type (Methods)	Exper-ience with NBS	Know-ledge of NBS	Educ-ation on NBS	False Pos	True Pos	Infor-med Con-sent
(Kladny B 2011)	Hemoglobin-opathies	114	Qualitative (Survey)	X	X				
(Morrison 2011)	Metabolic and Endocrine	60	Qualitative & Quantitative (Interviews & PSI)	X	X	X	X		
(Tluczek 2011)	Cystic Fibrosis	87	Qualitative (Interviews)	X	X		X		
(Skinner 2011)	Fragile X (FMR1)	1930 mothers	Qualitative & Quantitative (Survey)		X	X			X
(Timmermans 2010)	Metabolic Conditions	55	Qualitative (Observational & Interviews)	X	X		X		
(Russ 2010)	Hearing	8 states	Case Study (Learning Collaboratives)	X		X			
(Tluczek A 2010)	Cystic Fibrosis	10 parents	Qualitative (Interviews)	X	X	X	X		
(Lipstein 2010)	All NBS	40 (6 FGs; 4 Interv.)	Qualitative (Semi-structured / Vignettes)		X	X	X		X
(Dillard 2010)	Cystic Fibrosis	38 parents	Qualitative (Questionnaire/ Observ./Tele-Survey)	X	X	X	X		
(Lipstein 2009)	Biochemical Conditions	337 (200 FP;137con)	Quantitative (Phone Survey & PSI)					X	
(Miller 2009)	Conditions with Carrier Status	n/a	Policy Analysis		X			X	
(Bailey Jr 2009)	All NBS	n/a	Policy Analysis						
(Tluczek 2009)	Cystic Fibrosis & Congenital Hypothyroidism	193	Qualitative (semi-structured interviews)	X	X	X	X	X	X
(Grob 2008)	Cystic Fibrosis	35	Qualitative (Interviews)	X				X	
(Tarini 2008)	Pompe Disease	n/a	Ethical/ Legal Analysis						X
(Bailey Jr 2008)	Fragile X Syndrome	n/a	Ethical Analysis	X		X	X	X	X
(Vohr 2008)	Hearing	145 (33 T;42 FP;70 con)	Quantitative (PSI)	X			X		
(Comeau 2007)	Cystic Fibrosis	n/a	Workshop Summary			X			
(Feuchtbaum 2007)	MS/MS Conditions		Qualitative & Quantitative (Focus groups, Interviews, Hospital Stats)	X	X	X			
(Davis 2006)	General	22 Focus Groups; 3 Interv.	Qualitative (Interviews & Focus Groups)	X	X	X			
(Hewlett 2006)	All Conditions	9 studies reviewed	Review	X	X	X	X		
(Arnold 2006)	Hearing	29 Focus Groups; 23 Interv.	Qualitative (Focus Groups & Interviews)	X	X	X		X	
(Burton 2006)	Hearing	44 (5 Focus Groups)	Qualitative (Focus Groups)		X				X
(Gurian 2006)	Biochemical Conditions	240 (173 FP; 67 cont)	Qualitative & Quantitative (Interview & PSI)	X	X	X	X		
(Tluczek 2006)	Cystic Fibrosis	33 families	Qualitative (Interviews)	X	X	X	X	X	

TABLE 3.4: Summary of Literature Review of Newborn Screening and Parents (continued)

Author/Year	Condition	Sample Size	Study Type (Methods)	Exper-ience with NBS	Know-ledge of NBS	Educ-ation on NBS	False Pos	True Pos	Infor-med Con-sent
(Arnold 2006)	General	49 NBS Prog.	Quantitative/Evaluation of NBS brochures using the Flesch reading ease (FRE) formula and assessment of user-friendliness using 22 criteria			X			
(Tluczek 2005)	Cystic Fibrosis	28 int; 51 CESD; 35 cont	Qualitative & Quantitative (Interviews & CES-D)	X	X	X	X		
(Fant 2005)	All NBS	51 NBS Prog.	Qualitative (Telephone Survey)		X	X			
(Lagoe 2005)	Cystic Fibrosis	61 parents	Quantitative (Questionnaire)	X	X	X	X		
(Comeau 2005)	Cystic Fibrosis	3 groups/ models	Group/Model Comparison	X		X	X		
(Clayton 2005)	All NBS	n/a	Review		X	X			X
(Dillard 2004)	Cystic Fibrosis	17 families	Qualitative/ Observational (video recorded sessions w/ medical staff, GC, and parents/family)	X		X			
(Waisbren 2003)	Biochemical conditions	258	Qualitative & Quantitative (Interviews & PSI)	X	X		X	X	
(Kurtzer-White 2003)	Hearing	n/a	Review	X			X	X	
(Gracey 2003)	Hearing	n/a	Review			X		X	
(Campbell 2003)	General	102 (in 12 Focus Groups)	Qualitative (Focus Groups using a semi-structured interview schedule)		X	X			X
(Kim 2003)	All NBS	51 state NBS Programs	Quantitative Survey (27 closed ended questions)			X			
(Parsons 2003)	Cystic Fibrosis	n/a	Review						
(Skinner 2003)	Fragile X Syndrome	442 (279 moms; 163 dads)	Qualitative & Quantitative (survey)		X				
(Mandl 2002)	All NBS	51 NBS Prog.	Quantitative (Survey - 34 questions)						X
(Ciske 2001)	Cystic Fibrosis	138	Primarily Quantitative (survey) & Qualitative (tele-interview)	X	X	X	X		
(Clemens 2000)	Hearing	49 int; 5010 record rev	Qualitative & Quantitative (Interviews & Record Review)	X	X	X	X		

The Newborn Screening Experience

Twenty-five articles in this literature review address topics related to parents' experience with: reporting results, the follow-up process, genetic counseling after NBS, dealing with both true and false positive results, understanding the concept of "patients in waiting," contradictory or ambiguous findings, identifying and understanding carrier status, parents' experience with educational efforts, parents' information seeking behavior after an abnormal result, and the overall interaction parents have had with the newborn screening process. Eleven of these articles are based on qualitative studies, two are based on quantitative studies, and seven articles were based on mixed methods research. One of the articles is based on a case study, one is an ethical analysis, two are review articles, and one is a model comparison.

Parents are inherently supportive of newborn screening (Feuchtbaum 2007). Parents want the opportunity to understand why some babies have false positive results, especially if their child has a false positive result. Parents surveyed indicated they are willing to deal with false positive results if it means making sure another baby with the condition gets the care he/she needs (Tluczek 2005). That being said, throughout the literature, there are contradictory findings about what the newborn screening experience is like for parents. This makes it difficult to come to a single conclusion about the overall findings. The subset of articles addressing parents' experience with newborns screening is no different. Research has been published that indicates parents of children that fail the first hearing screening or have a false positive result on the first blood spot screening have higher anxiety and stress. At the same time, other research findings indicate there is no difference in the stress level of parents between groups (i.e. parents of infants with normal results versus false positive results). Morrison and Clayton reported that parents of infants with false positive results "expressed concern about having more children and identified numerous problems with how they were told about newborn screening" and 10% reported "significant stress as well as worry about their child's health and future" (Morrison 2011). Gurian et al reported that "false positive results may lead to increased parental stress," especially for parents who have not been provided sufficient information about newborn screening (Gurian 2006). Contrary to these two findings, Lipstein et al reported that "despite the reported negative psychosocial effects of false-positive results, our study found no impact on early health care utilization" (Lipstein 2009). And Clemens et al reported that "regardless of anxiety, 94% of all respondents were glad that their children had a hearing test and thought that UNHS was a good idea" (Clemens 2000). Suggesting that the potential negative impact of false positives is not enough to detour parents from screening. Vohr et al implemented the Parenting Stress Index (PSI) and found no statistically significant difference in stress levels between mothers of infants with hearing loss versus mothers of infants with false positives versus mothers of infants with normal screens (Vohr 2008).

Timmermans and Buchbinder discuss the concept of "patients-in-waiting" as individuals under medical surveillance because of an abnormal newborn screening result, specifically an abnormal result for one of the metabolic conditions, but do not necessarily have a known condition. These individuals find themselves between health and disease and for some it can be "prolonged liminality between a state of normal health and pathology" (Timmermans S 2010). This is similar to the "unpatient" concept put forth by Jonsen et al in 1996 in regards to individuals that carry a susceptibility gene: they are neither patients under treatment nor free of a medically relevant condition (Jonsen 1996). This concept was further explored by Bailey et al when they worked through a case study on whether or not to add Fragile X to newborn screening (Bailey Jr 2008). The most likely results that lead to this state of limbo are abnormal metabolic screening results. In some cases, the diagnostic testing rules out the condition for which the infant was initially identified, yet the metabolic profile is not entirely normal and something else is going on that may or may not be clinically significant for the infant. Although the infants are the unpatient or patient-in-waiting, the parents are the individuals navigating this unknown territory. The public health and medical community need to know about parents' experiences with this type of uncertainty. Armed with this information, newborn screening programs and clinicians can implement processes that help the parents navigate the journey.

The difference in findings about the parents' experience and understanding of the nuances of newborn screening could be attributed to different study populations, sample size, or methodology utilized. The difference could also be attributed to the fact that parents in different sample populations faced different scenarios due to the type of condition that the false positive calls into question (i.e. hearing loss versus a metabolic condition versus cystic fibrosis). Although the literature may not come to a consensus on the overall experience of parents, each study provides information that state newborn screening programs and policy makers ought to consider when developing and revising daily programmatic functions as well as broader newborn screening policies.

In a number of ways, the 25 articles that address some aspect of parents' experiences with newborn screening inform the dialogue moving forward. Several studies illustrate that false positive results and true positive results can cause stress, anxiety, worry, over protective behavior, or parent-child attachment issues. This knowledge is important to keep in mind for revisiting existing newborn screening policy and developing new policies moving forward – both at the programmatic and the larger systems level. It is also important to recognize that several studies shown no major psychosocial impact of newborn screening, even when parents receive false positive NBS results for their infants. Collectively, the literature suggests it is important not to assume parents will be impacted one way or the other. Every

parent will experience NBS in a unique way. The different experiences can be attributed to a number of factors: type of result, follow-up protocol, type of diagnostic testing, interaction with health care professionals (primary care and specialists), or the parent's age, socioeconomic status, education level, primary language, and social network. Even with this variation, there are many instances where the experience is similar or overlaps. It is in the common experiences that policies and procedures can make the most impact on improving the overall experience for parents. The literature suggests this boils down to consistent and timely education and improved communication during the newborn screening process.

Summary Points:

Parents' Experience	**Information** - based on their experiences, parents have indicated they want to be informed about NBS and in the event of an abnormal result they want accurate resources to turn to
	Uncertainty - surrounding what abnormal screening results mean
	Parental Response - reaction to the newborn screening process, in particular receiving abnormal results, varies depending on: their background (age, gender, education, SES), the type of result, type of condition, how the information is delivered, whom delivers the information, and resources available.
	Informing Decision-making - the variety of parents' perspective ought to be considered when making decisions about NBS

Knowledge and Education

The parents experience is clearly dependent on the timing and type of education received and the parents' knowledge about NBS. But interestingly, "although parents were instrumental in the initiation of the earliest NBS programs, evidence has repeatedly shown that most parents are uninformed or misinformed about NBS" (Tluczek 2009). Thirty-four articles in the literature review discuss NBS knowledge or education and often both. Twelve articles were based on qualitative studies, five were quantitative studies, and seven were mixed methods. Three of the articles were review articles, one was a case study, one was a policy analysis, one an ethical analysis, one was a workshop summary, and one reported a model comparison.

In 2005 Fant et al reviewed the educational materials for all 51 NBS Programs, the authors noted that: none of the educational materials (provided by 47 states) contained all of the elements recommended by the AAP; median readability grade level was 10; 46 programs had materials that contained statements about benefit of screening; 9 programs discussed the direct risks associated with screening (not including false positives); 6 programs explicitly mentioned that false-positive results are a possibility and only 3

states mention any risk associated with this; 41 programs communicate how parents will be notified and only 15 mention that parents will only be notified if there is a problem; 33 address the potential for the need for retesting or the associated uncertainty about results; 16 programs provide specific reasons regarding the importance of follow-up after an abnormal result; and 5 programs mention storage and residual use of samples (Fant 2005). Many of the states have updated education materials since 2005, in particular with regard to the last point mentioned. States had to respond to the increased awareness and concern regarding the storage and use of residual blood spots. The article was critical in pointing out where state newborn screening programs could improve educational efforts and communication.

In 2005 Tluczek set out to gain a better understanding of parents' phenomenologic experience of the CF NBS process. In doing so they documented that "shocked response of most parents seemed to stem from their general lack of knowledge" about NBS and "although most mothers remembered receiving a NBS brochure at the time of their infants' births, few had read it" (Tluczek 2005). Parents indicated they would have liked more information about NBS during pregnancy, after giving birth, and in particular at the time the blood sample was taken (Tluczek 2005) (Arnold 2006). Parents want healthcare providers to take an active role in helping them to understand the process (Dillard 2004). This poses a challenge in many states. Often results are first reported to the healthcare provider and the provider then passes the information on to the parents. In the case of an abnormal result, the provider may not be familiar with the condition, since some of the conditions included in newborn screening are quite rare (> 1:100,000). Newborn screening programs have years of experience and embody a much more nuanced understanding of abnormal results than does the primary care provider. For example, when reporting cystic fibrosis results, a follow-up specialist has an idea of the positive predictive value of the level of elevation in the immunoreactive trypsinogen (IRT). Understanding an abnormal newborn screen has a 10% likelihood of being a real case might make the information easier to hear than a parent thinking an abnormal newborn screening result means their child has the disease. However, this understanding of abnormal results can get lost in translation in the newborn screening process since often abnormal results are reported by phone to a nurse or medical assistant who relays the information to the physician, who then relays the information to the family.

According to the literature, parents want to know that newborn screening is being done and that their baby might have to be retested (Davis 2006). State newborn screening programs grapple with how to better inform and educate parents. In an attempt to distill parents' targeted desire for information, Davis et al found that there is some basic information about newborn screening that is important to parents. Parents want to know that their infant will be screened, how screening benefits their infant, that sometimes

retesting is necessary and why retesting does not necessarily mean the infant has one of the conditions (Davis 2006). Parents want to know how they will be notified about results and if retesting is necessary who will contact them and how will they know how quickly they need to act (Davis 2006). In the event of an abnormal result, parents want accessible and understandable information, especially when further diagnostic testing is required (Tluczek 2006).

Newborn screening has the potential to identify carrier status of infants, individuals with one copy of a recessive genetic mutation that will not result altered health outcome. Lagoe et al reported that for every affected infant detected by NBS, there are 11 infants identified as carriers (Lagoe 2005). Given this reality, the authors identified several issues that need to be considered regarding CF carrier identification: cascade testing (testing to the first-degree relatives of a known carrier), "teachable moments" (the point at which a patient is most receptive to new information on a given subject, and being respectful that people may have reasons for not wanting to be tested. Genetic counseling has been reported to be a positive and helpful experience for parents (Kladny B 2011) in understanding what it means to carry a hemoglobin trait.

Summary Points:

Parent Knowledge	**Lacking** -Parents own lack of knowledge, and if the provider is lacking in knowledge about NBS or the condition in question, contributes to parental stress and anxiety **Reliability** -Parents have a hard time retaining knowledge about NBS, so may need multiple forms and points of communication to reinforce their knowledge base **Results -** parents have indicated they want to know the results of a test, even if results are not definitive; however, other parents have described the time period of not knowing as stressful and difficult

Parent Education	**More -** education is needed for parents, both before delivery and at the time the newborn screening specimen is collected or the hearing screening test is done **Targeted -** education is needed for providers (pediatricians/primary care providers/nurses) to ensure appropriate communication throughout the newborn screening process **Useful** -state NBS programs need to ensure parent educational materials are: understandable (readability should be at least < grade 10, preferably < grade 8), assessable, balanced (discussing both benefit and potential risks), accurate, current, and available in multiple formats (verbal, brochures, videos, and online resources). **Genetic Counselors -** can educate parents about NBS results and help parents interpret results, in terms of the baby's current health status and future disease risk

False Positive Results

Nineteen articles in the literature review discussed the impact of false positive results on parents. Of these, twelve articles were reporting primary data. Of the 19 articles: nine are specific to CF, six are generalizable to metabolic screening or all NBS conditions, three are specific to hearing, and one is about Fragile X -a condition not currently included in NBS but may be considered by states in the near future. Seven articles presented research that used a qualitative approach, six of the articles approached the research from both a qualitative and quantitative perspective, two were solely quantitative, one article is an ethical analysis, two articles are reviews, and one article is based on a group/model comparison.

State programs grapple with the issue of false positives daily. Many state programs have implemented protocols that help to mitigate the impact of false positives by lowering the number of abnormal results generated and/or tailoring follow-up based on severity and immediate threat of the condition. Even still, given false positive results are expected with a large-scale screening program, it is crucial to have a better understanding of what parents know about newborn screening and if, and how, parents of normal screened infants differ from parents of infants with false positive results.

Findings from the research presented in the nineteen articles included in this subset of articles range in results. There is research that finds no negative impact of false positives and research indicating a potential negative impact of false positives, but a willingness among parents to endure this to make sure true positive cases are identified. Some of the research indicates there is a significant potential for false positive NBS results to cause negative psychosocial impacts and other research that reports the negative consequences as experienced by parents. The New England Consortium of Metabolic Programs concluded that "false-positive screening results may place families at risk for increased stress and parent-child dysfunction" (Waisbren 2003). As of June 2006, there had been nine studies published on parental stress and positive newborn screening results, both true and false positive, that report "false-positive screening results have been associated with increased anxiety and stress in parents of infants who require follow-up testing" (Hewlett 2006). Findings of parental stress and psychosocial effects of following-up positive newborn screening results ranged from 76% of families reporting initial anxiety, to 18% of a families reporting persistent anxiety afterwards, to a 23% increase in stress levels for mothers and a 10% increase in stress level for fathers of infants receiving false positive results (Hewlett 2006). Gurian et al (2006) collected data from parents of infants with a false positive newborn screening result for one of the metabolic disorders using parents of infants that screened normal as a comparison group. Short structured interviews were conducted and analyzed using content analysis and the Parental Stress Index (PSI) was administered. The data concluded that "false-positive screening results may affect parental stress and the

parent-child relationship" and that "improved communication with parents regarding the need for repeat screening tests may reduce the negative impact of false-positive results" (Gurian 2006).

Contrary to Waisbren's findings, in 2008 Prosser et al found that "parents have a high tolerance for false-positive newborn screening results" (Prosser 2008). Using "willingness-to-pay" and "time trade-off" models the researchers found that parents of children with false-positive screening results were actually willing to trade less time to avoid a false-positive than were parents of children with normal screening results (Prosser 2008). The parents were asked how much of their own life they would be willing to trade to "prevent" a disease. The researchers reported that fifty-seven percent of parents of infants that had received a false-positive newborn screen indicated they were not willing to trade any time to avoid a false-positive, whereas 33% of parents of children with normal screening results were not willing to trade any time (Prosser 2008). The researchers go on to comment that their research suggests that the false-positive may not be as bad of an experience as once thought, given that parents who have actually experienced a false-positive were less willing to trade time (of their own life) and money in order to avoid a false-positive (Prosser 2008). Specific to hearing screening, Clemens et al found that "regardless of anxiety, 94% of all respondents were glad that their children had a hearing test and thought that UNHS was a good idea" (Clemens 2000).

However, Morrison and Clayton concluded that false positive NBS results cause some parents to experience stress and long-term worry and recommend that this be taken into consideration when deciding to add new conditions to state NBS programs (Morrison 2011). Tluczek et al (2005) researched the impact of false positive CF NBS results and concluded the following about parents: "despite their distress, most (90%) were supportive of NBS for CF, with the belief that early detection and intervention would improve the health and welfare of affected children." However, the authors went on to conclude that the important overarching question yet to be rigorously explored is, "what are the long-term psychosocial effects of true-positive and false-positive results for a potentially life-shortening disease? (Tluczek 2005)." Beyond long-term consequences, additional issues were raised in the literature. Timmermans and Buchbinder (2010) identified another potential consequence of a false positive, and that is that not all results fall neatly into one category over the other (i.e. false positive v. true positive).
It is not clear how research results specific to the impact of false positive results are being utilized. Specific to hearing screening, Kurtzer-White et al (2003) concluded that although parental stress is noted in the literature, there is little evidence that it is being programmatically addressed.

It is difficult to discern how findings presented in these 19 articles, and research conducted prior to 2000, are utilized by newborn screening programs, policy/decision-makers, and advocacy groups. The intent and hope is that this information is making its way into the hands of decision-makers, in the hopes of painting a more complete picture of the newborn screening process. Besides the Kurtzer-White et al (2003) article, there have not been any publications regarding if and how the information generated about false positive results in newborn screening is being used.

Summary Points:

Impact of False Positives	**Real Risks** - there is a risk of psychosocial issues when abnormal newborn screening results are reported (note: research has shown some parents experience stress/anxiety/worry and research has also shown that these parents do not experience any more stress than normally associated with having a new baby).
	Mitigating the Impact - by 1) reducing the number of false positives and 2) by improving communication with parents
	Decision-making - the potential risk associated with false positives should be taken into consideration when regulatory and programmatic decisions are being made

Impact of True Positive Newborn Screening Results

Ten of the articles in the review broached the subject of the impact of true positives on parents. Of the ten articles, three focus on hearing impairment, three concentrate on CF (one of these articles co-focused on CH), two articles focus on biochemical conditions, one on conditions that result in identification of carrier status, and one of the articles centers on Fragile X (a condition which may be proposed for NBS in the future). Of the ten articles, four are based on a qualitative approach, two implement mixed methods (quantitative and qualitative), one utilizes a strictly quantitative approach, two are reviews, and two are presented as an ethical or policy analysis.

Grob focuses her research on the experience of identifying cystic fibrosis through newborn screening compared to clinically (Grob 2008). She makes the argument, based on qualitative interviews with parents of children diagnosed with CF, that “early, unsought diagnosis deeply affects parents' feeling of competence to care for their newborn and their sense of who the child is, and places the disease - rather than the process of "falling in love with” the new baby - at center stage during the child's early weeks and months,” and results in health professionals and the medical complex to have larger than necessary presence in the family's life at this formative time of newborn/parental bonding.

Another article also focused on CF, provided a more broadly applicable description of parents' experience with newborn screening. Tluczek et al (2005) found that the "shocked response of most parents seemed to stem from their general lack of knowledge about the Wisconsin NBS program." The article also points out an important point about education and general awareness of newborn screening and how that can impact the overall experience: "Although most mothers remembered receiving a NBS brochure at the time of their infants' births, few had read it. Several parents said they wished they had received more information about NBS before their child was born or at the time the blood sample was taken" (Tluczek 2005). This is consistent with what Davis et al extrapolated from parent focus groups, that "almost no parents were familiar with the term "newborn screening," although some were familiar with terms such as "heel-stick" or "PKU," however, were often unaware that a phenylketonuria test involved > 1 test" (Davis 2006).

Summary Points:

Impact of True Positives	**Resources** - parents want reliable resources to turn to in the event their infant has one of the newborn screening conditions
	Communication - how the information was initially communicated, and how well providers communicated with the parents throughout the diagnosis process impacts the parents' overall experience

Informed Consent

Whether or not informed consent should be part of the newborn screening paradigm is an ongoing debate. Given newborn screening has been a state-mandated function for fifty years now, the argument can be made that when screening is of great public health import, as NBS has been determined to be, traditional informed consent models are not appropriate. However, given the recent climate regarding the concern over privacy, a few states have implemented informed consent screening models. The majority of states conduct NBS under a state-mandated model ; 38 of 51 states/Washington DC notify parents but do not get consent (Mandl 2002). Ten states do not notify parents of screening, nor ask for consent, and three states actually do require parental consent (Mandl 2002). Most states have an "opt out" capability; however, there are two states that do not permit refusal under any circumstances (Mandl 2002). An interesting point to come out in the literature is that although some parents are concerned about being informed about screening prior to the event, parents are also concerned about what is being done with the residual blood spots and the lack of consent surrounding the use of residual blood spots.

The other issue to emerge from the literature was how newborn screening may change in the future, to include expanded genomic-based tests and what that will mean in terms of reporting results and protecting privacy and the paradigm may shift, bringing the parent's right to know to the forefront (Tluczek 2009). "As testing programs expand to include results of a less critical nature, the patient should be in greater control of the release of personal genomic data. Approaches to informed consent necessarily must evolve as newborn screening programs broaden from public health service providers to custodians of a population-based genomic databank that is used for research" (Mandl 2002).

It is important to understand where parents stand on the issue of consent. This becomes even more crucial as NBS services expand to include new types of screening, which may cause more concern over the protection of privacy. However, in the environment of an expanding public health program, it is also imperative to think about maximizing the benefits of screening (i.e. making sure all babies are screened). So there is balance between potential benefits and potential harms. Including parents in developing the best methods to balance the benefits and harms, especially as NBS programs continue to expansion, is an important step (Lipstein 2010). In the context of hearing screening, when parents were asked about how they felt about molecular testing (genetic testing) for hearing loss, under a screening paradigm, more than half of the parents indicated that consent should be obtained before testing done (Burton 2006). More parents may indicate they want to know about this type of screening because it involves genetic testing and they have concerns about stigma, discrimination, or insurance coverage.

There is also the standing question of whether informed consent is helpful or harmful. The intent is for informed consent to be helpful, to provide information before conducting testing so that the individual (or their proxy, who in the case of NBS is the parent, making decisions on behalf of the infant) can make a decision as to whether or not he/she wants to participate in the testing. Trying to make NBS conform to a research model that utilizes informed consent could also potentially be detrimental, making it more difficult to ensure all babies are screened and negative health outcomes are averted (Tarini 2008).

Summary Points:

Informed Consent	**Variation** - a few states have implemented informed consent NBS models. The majority of states operate under an "opt out" model or a full mandated model (with no opt out). **Research** - informed consent has become a more prominent issue in NBS due to 1) the expansion of the types of conditions being screened for and the type of testing being done (i.e. molecular-based testing), 2) concerns regarding privacy of health information, especially genetic information, and 3) the potential use of residual blood spots for future research.

DISCUSSION

The main gap in knowledge is the translational component, determining how we use the information we already have to improve newborn screening programs at the ground level. Current research has established that some parents experience anxiety, worry, or prolonged stress after receiving false positive results. Of particular concern are parents of infants that are picked up through newborn screening and the primary condition is ruled out, but the infant's metabolic profile has abnormal results of uncertain clinical significance. This is previously uncharted territory in newborn screening that is now a reality given the type of conditions included in newborn screening now. One thing parents have indicated would be useful is more education and communication before newborn screening is done so they are prepared to receive the results. Parents have also indicated that better communication with the providers and being provided with resources up front would help them with the newborn screening process. How to improve communication and what type of resources are the most accessible to parents is something the newborn screening and clinical research communities have yet to identify.

Another missing piece to this puzzle is long-term outcomes. Research to date has primarily assessed parent's experience (stress, anxiety, worry, etc) shortly after the parents have been through newborn screening. Some research tracks parent's up to one year post receiving the results. For both true and false positive results, the long-term impact on parents and the infant are largely unknown. This is challenging research to conduct because the population is disperse and difficult to access. Again, as newborn screening continues to expand, long-term outcomes are important. The long-term health outcomes of affected infants are vitally important to measure, as are long-term impacts of the newborn screening process on parents.

One of the challenges with the newborn screening program is that one program is responsible for screening for a group of conditions that greatly vary with regard to clinical presentation, type and effectiveness of treatment, need for specialty care, and long-term health outcomes. Because the type and severity of the condition can influence the parental experience, research needs to be done on how the variation in conditions (including the different follow-up required for abnormal results with each condition) influences the parents' experience. How newborn screening programs and clinicians handle communication under these different scenarios will lead to alleviating some of the stress, anxiety, and worry following abnormal newborn screening results that has been documented.

Another gap is actually a disparity in representation of different types of conditions in the newborn screening literature. Since 2000, there has been a heavy emphasis on cystic fibrosis, followed by hearing

and metabolic conditions. This is in part due to the funding available for research on cystic fibrosis, in comparison to hemoglobinopathies for example (Smith 2006). Another theory is that it took several years for cystic fibrosis to be added to the list of conditions endorsed by the Secretary's Advisory Committee on Heritable Disorders in Newborns and Children, prompting more research to be done to inform the process of adding CF to newborn screening programs across the country. Endocrine, hemoglobinopathies (e.g. sickle cell disease) are underrepresented in the newborn screening literature focused on parents. Conditions such as biotinidase, and galactosemia are not represented in this literature, except when encompassed in articles that reference newborn screening in general.

Newborn screening services will continue to expand. When states go through the process of adding conditions to the state mandated newborn screening program, it is important to include the parent perspective. As evidenced by the literature, there is variation in parents' knowledge, experience, and reaction to newborn screening; because these are dependent on many factors, it is not feasible to anticipate every type of scenario. It is also not realistic to think that advisory committees would consist of twenty different parents, in hopes of capturing all of the different types of experiences. Traditionally this has been done by enlisting one or two parents of children with one of the conditions. Although this parent perspective is extremely important, it represents one viewpoint. One way to balance the representation of the parent experience is to include a parent of an infant that had a false positive result. The assumption is that the experience is going to be negative or the parents will not support newborn screening. The research to date has shown that an overwhelming majority of parents support NBS, even if their baby has had a false positive result. An example of this is actor Scott Baio and his wife's experience with receiving a false positive result for a potentially deadly metabolic disorder known as glutaric acidemia type 1 (GA-1). They had to wait 10 weeks of the definitive all clear. Scott and wife, Renee, are now avid spokespeople, speaking out in support of expanded newborn screening and making sure all babies are screened. They have teamed up with www.savebabies.org to help spread the word of the importance o newborn screening. On the flip side of this is a parent that will become dissatisfied with the public health or healthcare system, or become over-protective of their child, even after diagnostic testing rules out the condition. Again, every parent will have a unique experience. Perhaps another way to inform the decision-making process about the variation in parent knowledge and experience is to supplement in-person testimony with this summary of the research findings to date.

CONCLUSION

This structured literature review demonstrates that there has been an interest in research on parents and newborn screening, covering a variety of topics within newborn screening. The research is not exhaustive; more research would be beneficial for all aspects of newborn screening. Focusing in on specific aspects, the lack of parental knowledge about newborn screening has been documented. Future research ought to address how parent education can be improved and more importantly, how that information can be translated into public health and clinical practice at the ground level. The research has also shown there is a lack of parental knowledge about NBS. Moving forward, research ought to focus on different types of feasible educational interventions. State newborn screening programs can then implement parent education models that have been proven to be successful.

Likewise, the issues surrounding both true and false positives have been elucidated. However, there are aspects of the true positive and false positive experience that could be further ferreted out using different qualitative approaches. In particular, identifying specific points in the process that newborn screening and clinical medicine can improve. One recommendation is the use of a narrative approach that can tell the story of the parents' experience. Every parent is going to have a different experience. Nevertheless, commonalities will emerge and parents can identify ways to improve the process. It is important to recognize that there will always be false positive results in a large-scale screening program like newborn screening. State newborn screening programs have to grapple with how to handle false positives on a daily basis and strive to minimize the impact of false positives. Therefore, newborn screening research ought to shift focus from identifying the possible negative ramifications of false positives, which has been nicely documented, to how developing strategies to minimize the impact of false positives. These data would greatly benefit state newborn screening programs. Additionally, determining how public health and clinical medicine can better serve the families of infants that have one of the conditions identified through newborn screening may lead to improved outcomes.

CHAPTER 4: ESTABLISHING CUTOFF VALUES FOR NEWBORN SCREENING

Fifty years of experience with newborn screening has generated a wealth of knowledge and has shaped the culture of state newborn screening programs. Newborn screening knowledge and data provide the backbone on which policy decisions are made. This chapter provides a real world application of utilizing data to guide decision-making in newborn screening, amongst other factors and influences. In the case of newborn screening, there is a balance between being able to correctly identify true cases while at the same time correctly identifying non-cases. Using a thyroid stimulating hormone (TSH) as the primary marker for screening for congenital hypothyroidism (CH), requires understanding that there is a known physiologic surge of TSH within a few hours after birth, followed by a decline in TSH over time. Due to the natural physiology of TSH, establishing the most effective algorithm for CH requires results to be stratified by the infant's age at the time of specimen collection and varying the cutoff level across these stratifications. Optimal age stratifications and cutoff values can be identified through careful analysis of screening data. When a new screen is introduced, an initial cutoff strategy is implemented and over time (anywhere from one week to one month to several years) a newborn screening program can gather enough data to adjust the algorithm in order to improve the balance between sensitivity, specificity, and positive predictive value. However, data are not the only factor influencing the decision-making process. The newborn screening culture, programmatic functions, administrative issues, laboratory logistics, and politics also weigh into the decision-making process.

The Washington State Newborn Screening Program, as part of ongoing quality assurance efforts, opted to revisit the screening algorithm for congenital hypothyroidism (CH). The University of Washington Human Subjects division determined the quality assurance data analysis is not human subjects (see Appendix B). The Washington State Newborn Screening (NBS) Program has been using thyroid stimulating hormone (TSH) as the primary marker for CH screening since July 2004. There are no known severe CH cases missed by NBS in the interim. In the last six years there has been a ten-fold increase in the number of specimens collected early (0-3 hours of life), 40% of which had an elevated TSH that resolved on a subsequent screen. Thus, the first screen was a false positive. Additionally, pediatric endocrinology collaborators have expressed concern that some cases of mild CH may be missed due to TSH cutoff values that were being used for older infants, in particular infants greater than two weeks of age. These factors prompted a review of the CH screening algorithm in order to 1) reduce the number of false positive screens while maintaining the sensitivity of the test, and 2) improve the screen's capability of identifying mildly elevated TSH values that may correspond with a mild form of CH.

PURPOSE

Demonstrate policy-making at the programmatic level using the process of revising the Washington State Newborn Screening Program CH screening algorithm as a case study.

SPECIFIC AIM

1. Analyze six years of CH screening data using primary TSH
2. Compare different stratification strategies by TSH value and age at collection
3. Identify an age stratification strategy, and cutoff values within the stratifications, that balance sensitivity, specificity, and positive predictive value
4. Propose a revised CH Screening Algorithm that improves the positive predictive value of the CH screening algorithm, while maintaining or enhancing the sensitivity of the first newborn screen
5. Describe factors beyond data analysis in the decision-making process

BACKGROUND/SIGNIFICANCE

The Washington State Newborn Screening (WA NBS) Program has been screening for CH since 1977**.** Approximately 15-20% of cases can be attributed to genetic defects; the remaining cases are of unknown etiology or are due to de novo mutations that have not been identified yet (LaFranchi 2011). Congenital Hypothyroidism is the most prevalent condition currently screened for through newborn screening programs. The prevalence in the US population has been estimated to be 1:3300 (LaFranchi 2011); however the prevalence of CH in the US population has been increasing (Harris 2007). The observed prevalence in the state of Washington has been trending upward over the last decade and is higher than national estimates. Important to note is that Washington identifies mild forms of CH due to using TSH as the primary marker and because Washington is a two-screen state. Approximately a third of the CH cases in Washington are identified on the second newborn screen. Table 4.1 below illustrates Washington's increasing prevalence that is higher than previous national estimates (NBS 2012).

TABLE 4.1: Prevalence of Congenital Hypothyroidism in Washington

Year	# Screened	# True	Prevalence
2005	77,299	55	1: 1,405
2006	82,609	45	1: 1,835
2007	84,925	49	1: 1,733
2008	86,058	83	1: 1,037
2009	84,780	73	1: 1,161
2010	82,930	74	1: 1,120
2011*	87,376	106	1:824

*Preliminary data for 2011 (note: Primary TSH was used for screening for all years listed)

WA NBS has been using thyroid stimulating hormone (TSH) as the primary marker for congenital hypothyroidism (CH) screening since July 2004. Prior to 2004, thyroxine (T4) was the primary marker used to screen for CH. Cutoff levels were adjusted in the first few months of screening using TSH as the primary marker and have been performing well. There are no known severe CH cases missed by newborn screening. Washington State NBS Program categorizes CH cases based on the serum TSH level at the time of diagnosis - severe CH (serum TSH $\geq$ 30 μIU/mL) and mild CH (serum TSH < 30 μIU/mL).
In the last six years there has been a ten-fold increase in the number of specimens collected early (0-3 hours of life), 40% of which had a false positive TSH that resolved on the second screen. Analysis of six years of screening data demonstrates that implementing different stratification protocols impacts the sensitivity, specificity, and positive predictive value of using primary TSH to screen for CH. Comparing various stratification options provides the data needed to make an informed decision about the most effective algorithm for interpreting TSH values in order to determine appropriate follow-up actions. The goal in revising the TSH cutoff algorithm is to improve the positive predictive value while maintaining the sensitivity, thereby reducing the number of false positive screens - mitigating the impact of false positive results and reducing the workload for the NBS program and health care system.

Sensitivity is how well a test produces a positive result when conducted on an individual who has the condition. Sensitivity can be calculated by dividing the number of true positives by the sum of true positives and false negatives. Specificity is how accurate a test is in producing a negative result when the

individual does not have the condition. Specificity is calculated by dividing the number of true negatives by the sum of true negatives and false positives. The sensitivity and specificity of a test are crucial in determining the test's applicability in a clinical setting or a screening scenario. With newborn screening in particular, states want to maximize the detection of affected infants. To do this, the threshold or cutoff levels for the screening mechanism are set at a level that will pick up as close to 100% of the true positive cases without giving up too many false positives.

The positive predictive value (PPV) of a test is the probability that an individual has the condition when a positive test result is observed. The calculation for PPV is the number of true positives divided by the sum of true and false positives. The negative predictive value (NPV) of a test is the probability that an individual does not have the condition when the test results observed indicate the absence of that condition. The calculation for NPV is the number of true negatives divided by the sum of true and false negatives. PPV and NPV are dependent on disease prevalence; the higher the known prevalence, the higher the PPV will be for any given sensitivity.

Due to the standard of practice in Washington to get a second newborn screen between 7-14 days of life, the algorithm on the first newborn screen will be different than other states which only get one screen. Table 4.2 below indicates the age stratifications and cut off values used by the state of Washington from October 28, 2004 through February 14, 2012.

Table 4.2: Categorization of Thyroid Screening Results Using Primary TSH (10/28/04 – 2/14/12)

TSH µIU/mL	1 to 12 hrs	13 to 24 hrs	25 to 36 hrs	37 to 48 hrs	49 to 504 hrs	> 504 hours
0 – 14.99	Normal	Normal	Normal	Normal	Normal	Normal
15.00 – 19.99	Normal	Normal	Normal	Normal	Normal	Borderline[a]
20.00 – 24.99	Normal	Normal	Normal	Normal	Borderline[a]	Borderline[c]
25.00 – 29.99	Normal	Normal	Normal	Borderline[b]	Borderline[a]	Borderline[c]
30.00 – 44.99	Normal	Normal	Borderline[b]	Borderline[a]	Borderline[a]	Borderline[c]
45.00 – 54.99	Normal	Borderline[b]	Borderline[a]	Borderline[a]	Borderline[a]	Borderline[c]
55.00 – 59.99	Borderline[b]	Borderline[a]	Borderline[a]	Borderline[a]	Borderline[a]	Borderline[c]
60.00 – 99.99	Borderline[b]	Borderline[c]	Presumptive[c]	Presumptive[c]	Presumptive[c]	Presumptive[c]
≥100.00	Presumptive[d]	Presumptive[d]	Presumptive[d]	Presumptive[d]	Presumptive[d]	Presumptive[d]

Note: Superscript letters ([a,b,c,d]) denote general guidelines for the different follow-up responses to be initiated by Follow-up staff.

[a] - if first test, wait for routine second specimen; if second test, call health care provider immediately to request a follow-up specimen
[b] - wait for second specimen
[c] - call health care provider immediately to request follow-up specimen
[d] - call health care provider immediately to recommend immediate diagnostic work-up

Sensitivity and specificity in newborn screening

Screening tests are done on asymptomatic individuals to provide information about future disease risk, whereas diagnostic tests are usually done on symptomatic patients to identify the cause of his/her illness. If screening tests are not sensitive enough, then an individual with a disease may get missed, and he/she will not get timely, appropriate follow-up care. If a screening test is not specific enough, then individuals who do not have the disease will actually be picked up by the screen as being at an increased risk of disease. These infants have to undergo diagnostic testing to rule out disease, and may even have unnecessary treatment in the interim. Once diagnostic testing rules out the condition, the screen is considered a false positive. The harm associated with a false positive is dependent on the: severity of the condition in question, risks involved with the diagnostic follow-up procedures, parents knowledge base and interpretation of results, effectiveness of the provider's communication, educational information provided, and potential negative side effects of interim treatment. Interim treatment is not typical with

newborn screening because newborn screening programs recommend follow-up diagnostic testing to confirm whether the infant has the condition or not, prior to initiation of treatment. In rare cases, such as a child with low enzyme activity for galactosemia, the provider may put the infant on soy formula until confirmatory testing is complete. The risk of not putting a baby on treatment is high because if the infant does have galactosemia, he/she may suffer a metabolic crisis – causing seizures, sepsis, and even death. The risk of putting a baby on soy formula for a few days until confirmatory testing is complete is low in comparison.

Because newborn screening tests have a much broader target population than do traditional clinical diagnostic tests, the sensitivity and specificity of the tests are different. Newborn screening programs use testing methods, cut-off values, and follow-up protocols that result in high sensitivity because one of the goals is to produce the least number of false positive results while ensuring legitimate cases are not missed. To maximize detection of affected infants, NBS programs generally set cutoff levels for each of the conditions at a level that will inevitably pick up some infants that do not have the disease, thus increasing the sensitivity and lowering the specificity.

Kwon and Farrell conducted a study using data collected by the Council of Regional Networks for Genetic Services in 1993 and 1994 on five newborn screening disorders and found that the sensitivity reported by states was ~100% and the specificity was above 99% for all five conditions (Kwon 2000). An interesting finding was that the total number of false positive test results reported was over 183,000 in a 2-year period, a false positive to true positive ratio of 61:1 in 1993 and 56: 1 in 1994 (Kwon 2000). "Despite the high sensitivity and specificity of screening with MS/MS, expanded screening, such as that recommended by the American College of Medical Genetics Newborn Screening Expert Group, will lead to an increased number of infants with false positive results due to the low positive predictive values inherent when screening for multiple independent diseases of low prevalence" (Lipstein 2009). Interestingly, newborn screening tests have higher sensitivities and specificities than most diagnostic tests, yet newborn screening will produce far more false positives than any given diagnostic test because of the sheer volume: between all fifty states, over four millions infants are screened each year.

DATA ANALYSIS

TSH values, age at collection, diagnosis code, and final case disposition (True, False, and Normal) data were analyzed. These variables were extracted from a larger dataset maintained by the Washington State Department of Health, based on newborn screening results processed between November 2004 and December 2010. Data were analyzed using Microsoft Excel 2007 and SAS® software version 9.2.

Descriptive statistics for TSH values (mean, median, 5th centile, 25th centile, 75th centile, 95th centile, 99th centile, min, max, variance, standard deviation, CLM (two-sided confidence limit for the mean, 95% CI)) were generated for the normal population and confirmed CH population for 140 age at collection (AAC) categories, with a focus on the first 72 hours of life. The initial AAC categories analyzed were defined as follows: every hour of life for the first 72 hours (73 age ranges), every 6 hours of life for 73 - 168 hours (3-7 days) of life (16 age ranges), every 12 hours of life for 169 – 336 hours (8 – 14 days) of life (14 age ranges), every 24 hours from 337 – 672 hours (15 - 28 days) of life (14 age ranges), every seven days > 28 days (4 weeks) of life (23 age ranges).

Descriptive statistics, with an emphasis on the 99th centile and median, for each of potential age categories were evaluated to determine which age categories could be grouped together to form age stratifications. Ten different age stratification options were considered (see Table 4.3 below). Receiver operator characteristic (ROC) curves (see Appendix C for an example) were utilized to plot the probability of a true positive result against the probability of a false positive result for each age group within the different stratification schemes. Evaluation of the 99th centiles, in combination with the ROC curves, generated a potential cutoff value within each of the age ranges. Using this cutoff value as a starting point, an iterative process was then employed to evaluate cutoff values within age categories to improve sensitivity, specificity, and positive predictive values (PPV). In total, 15 different stratification schemes were considered. These stratification schemes were tested against the existing TSH, AAC, and case disposition data in order to check what the sensitivity, specificity, and positive predictive value of the 15 different stratification schemes would have been, had they been used originally. This gives an idea of what to expect for future cohorts.

TABLE 4.3: Age At Collection (AAC) Stratification Options

Option 1	Option 2	Option 3	Option 4	Option 5
0 - 3 hrs 4 - 16 hrs 17 - 24 hrs 25 - 39 hrs 40 - 48 hrs 49 - 96 hrs 97 - 168 hrs ≥ 169 hrs	1 hr 2 - 4 hrs 5 - 11 hrs 12 - 16 hrs 17 - 21 hrs 22 - 35 hrs 36 - 48 hrs 49 - 96 hrs 97 - 384 hrs ≥ 385 hrs	1 hr 2 - 5 hrs 6 - 16 hrs 17 - 41 hrs 42 - 96 hrs 97 - 168 hrs 169 - 384 hrs ≥ 385 hrs	1 hr 2 - 7 hrs 8 - 17 hrs 18 - 23 hrs 24 - 47 hrs 48 - 144 hrs ≥ 145 hrs	1 hr 2 - 3 hrs 4 - 7 hrs 8 - 11 hrs 12 - 17 hrs 18 - 23 hrs 24 - 35 hrs 36 - 47 hrs 48 - 144 hrs 145 - 336 hrs ≥ 337 hrs
Option 6	**Option 7**	**Option 8**	**Option 9**	**Option 10/Final**
1 hr 2 hr 3 hr 4 hr 5 - 7 hrs 8 - 12 hrs 14 - 18 hrs 19 - 22 hrs 23 - 43 hrs 44 - 384 hrs ≥ 385 hrs	1 hr 2 - 7 hrs 8 - 17 hrs 18 - 23 hrs 24 - 35 hrs 36 - 47 hrs 48 - 144 hrs 145 - 336 hrs ≥ 337 hrs	1 hr 2 - 7 hr 8 - 17 hr 18 - 22 hrs 23 - 25 hrs 26 - 35 hrs 36 - 47 hrs 48 - 144 hrs 145 - 336 hrs 337 - 504 hrs ≥ 505 hrs	1 hr 2 - 7 hrs 8 - 17 hrs 18 - 22 hrs 23 - 25 hrs 26 - 35 hrs 36 - 47 hrs 48 - 336 hrs 337 - 504 hrs ≥ 504 hrs	1 hr 2 - 7 hrs 8 - 17 hrs 18 - 22 hrs 23 - 25 hrs 26 - 35 hrs 36 - 47 hrs 48 - 72 hrs 73 - 144 hrs 145 - 504 hrs ≥ 505 hrs

The previous follow-up protocol designated two types of abnormal results for congenital hypothyroidism: borderline and presumptive. Presumptive abnormal results initiated a much more rigorous follow-up than borderline abnormal results. For most borderline abnormal results, the protocol was to wait for the second screen. If the second screen was not received by 3-4 weeks of life, a follow-up specialist called to request a second screen. A smaller proportion of borderline results had a higher positive predictive value, requiring an immediate second screen. Presumptive results require immediate follow-up diagnostic testing (serum thyroid studies) and a consult with a pediatric endocrinologist. The WA NBS Program wanted to maintain discernment between borderline abnormal and presumptive abnormal. So, once the age stratifications and cutoff values between normal and abnormal results were identified, analyses were done to determine the cutoff between borderline abnormal and presumptive abnormal. The same iterative approach of generating the sensitivity, specificity, and PPV for different borderline/presumptive cutoff

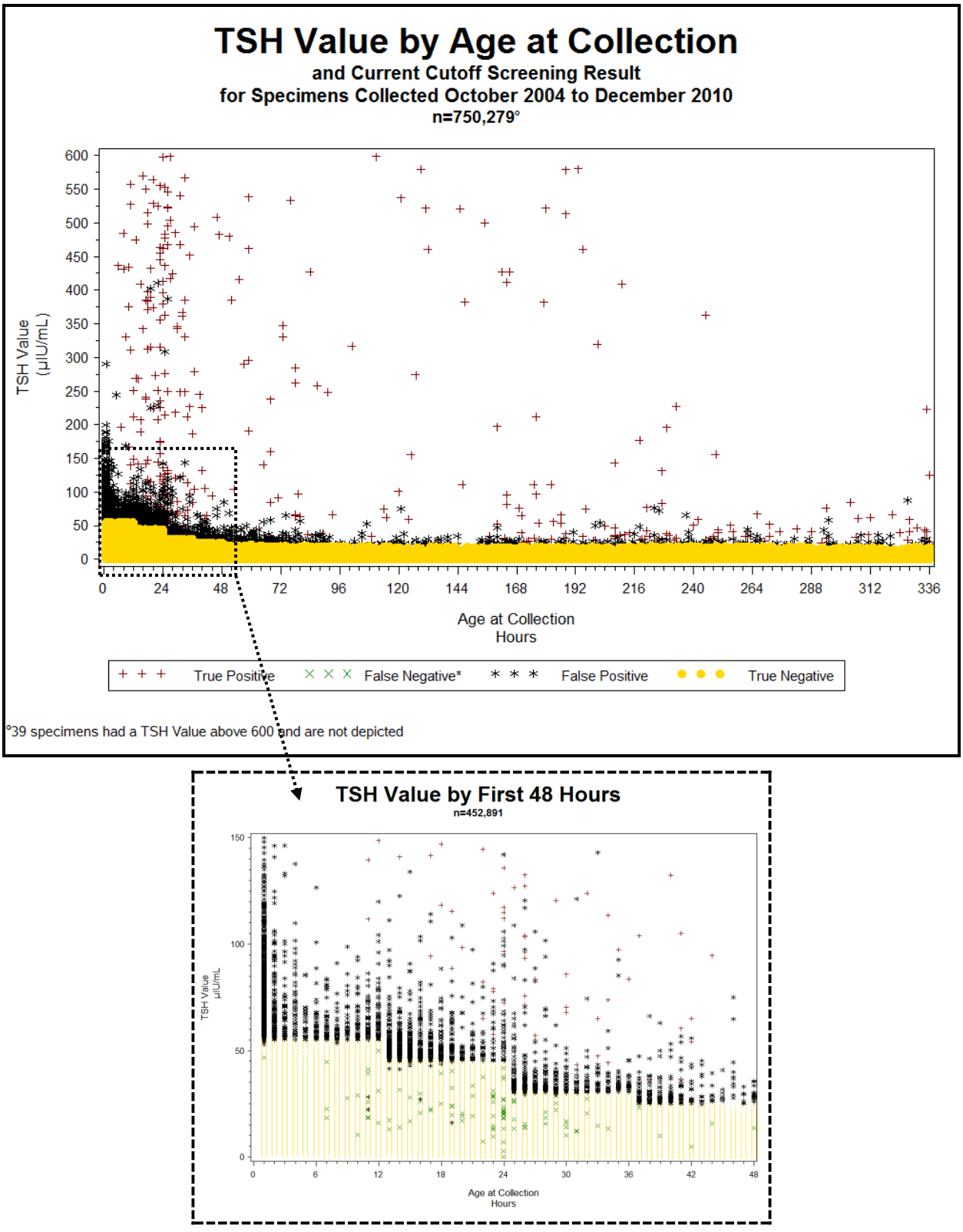
values was employed to determine the appropriate TSH value to distinguish between borderline and presumptive abnormal results.

RESULTS

Below is a scatterplot that charts 750,279 TSH values by age at collection (see Figure 4 on the next page); screening results are based on cutoff scheme used by the WA NBS Program from November 2004 – February 2012. As expected, the vast majorities of TSH values are true normal (yellow circles) and cluster at the bottom of the plot, exhibiting higher TSH values in the first few days of life. The black asterisks indicate a false positive result, which also exhibit higher TSH values in the first few days of life. There is a known physiologic surge of TSH immediately after birth, followed by a decline in TSH over time. Due to the natural physiology of TSH and variation in the rate of decline, the first 72 hours of life were stratified into seven age categories. These smaller age ranges are expected to both reduce the number of false positive screens and increase the number of true cases (red +'s) identified on the first screen. To reflect the large bulk of specimens collected around 24 hours of life and the difference of the median and 99th centile TSH value at this age of collection from earlier and later collections, a narrow age stratification of 23-25 hours was created. The false negative specimens (green x's) in the scatterplot are normal TSH values that are linked to true CH cases identified by a subsequent abnormal screen. The WA NBS Program routinely receives two screens for each baby (three screens for NICU babies), the first newborn screen preferably collected between 18-48 hours of life and a second screen preferably collected between 7-14 days of life. A benefit of being a two-screen state is that higher cutoffs for early ages can be implemented to reduce the number of false positive screens on the initial specimen. On the other end, conservative cutoffs after 72 hours of life can improve our ability to correctly identify babies with CH, in particular mild forms of CH. A third screen at 4-6 weeks is also recommended for low birth weight or sick babies to maximize detection of CH cases that present with a delayed rise in TSH due to an immature hypothalamic-pituitary-thyroid axis (HPTA).

Given the de-identified dataset, the false negative TSH values on the first screens cannot be linked to the subsequent abnormal screen; however, according to the WA NBS Program, in the same six-year timeframe (2004 – 2010) only three false negative cases were brought to their attention. These three CH cases, identified clinically, were mild CH cases (defined by the WA NBS Program as having a baseline serum TSH value < 30 μIU/mL). Two of the three cases involved very low birth weight babies, which may have been due to an underdeveloped HPTA. In general, with low birth weight babies, the concern is missing a severe CH case.

Figure 4: Scatterplot of TSH Value by Age at Collection



The revised age stratifications and cutoff values improve the sensitivity on the first screen, and the overall specificity and PPV, facilitating faster diagnoses for some babies and greatly reducing the number of false positives. Application of the new cutoff scheme to the existing dataset predicts the overall screening sensitivity will be maintained and identification of all severe CH cases will continue. Both the previous and revised cutoff schemes are presented (see Table 4.4 on the next page) for comparison. The new cutoff scheme is stratified into 11 age groups providing a balance between sensitivity and specificity. The new cutoff scheme reduces the number of false positives by 35% while maintaining the ability to detect all true cases of CH.

Table 4.4: Comparison between the Previous and New TSH Cutoff Scheme

Previous TSH (µIU/mL) Cutoff Scheme: October 28, 2004 – February 14, 2012

Number of TSH Values	Previous Age Stratification	Previous TSH Cutoff	Sensitivity	Specificity	PPV	True Pos	False Pos	True Neg	False Neg*
42223	1-12 hr	55	60.98	95.99	1.46	25	1692	40490	16
216184	13-24 hrs	45	60.00	99.57	8.60	87	925	215114	58
148170	25-36 hrs	30	75.73	99.62	12.23	78	560	147507	25
46314	37-48 hrs	25		99.44	7.47	21	260	46028	5
382638	49-504 hrs	20	80.07	99.84	28.99	245	600	381732	61
32237	505 hrs -6 mo	15	82.26	99.83	48.57	51	54	32121	11
867766	**OVERALL**		**74.23**	**99.53**	**11.03**	**507**	**4091**	**862992**	**176**

Note: 61 TSH values were collected at zero hours of life and are not included

New TSH (µIU/mL) Cutoff Scheme: February 15, 2012 - current

Number of TSH Values	Proposed Revised Age Stratification	Revised TSH Cutoff	Sensitivity	Specificity	PPV	True Pos	False Pos	True Neg	False Neg*
4835	1 hr	115	50.00	98.68	3.03	2	64	4767	2
7388	2-7 hr	100	28.57	99.66	7.41	2	25	7356	5
91370	8-17 hr	60	58.06	99.77	14.52	36	212	91096	26
85000	18-22 hrs	40	68.89	99.48	6.51	31	445	84510	14
95280	23-25 hrs	35	59.04	99.73	16.07	49	256	94941	34
116083	26-35 hrs	30	76.47	99.66	14.01	65	399	115599	20
49435	36-47 hrs	26	85.19	99.52	8.81	23	238	49170	4
39044	48-72 hrs	20	59.09	99.17	7.41	26	325	38675	18
20870	73-144 hrs	18	80.85	99.19	18.45	38	168	20655	9
326224	145-504hrs	16	90.32	99.87	31.72	196	422	325585	21
32237	505 hrs -6 mo	13	87.10	99.61	29.83	54	127	32048	8
867766	**OVERALL**		**76.43**	**99.69**	**16.30**	**522**	**2681**	**864402**	**161**

Note: 61 TSH values were collected at zero hours of life and are not included

**** CH cases with "False Negative" results were identified based on elevated TSH levels on subsequent screens.***

OUTCOME

The outcome of the programmatic decision-making process was to implement a revised screening algorithm, based on the data analysis, but also taking into account laboratory logistics, newborn screening cultural norms, programmatic feasibility, and the politics of screening, inherent to every state newborn screening program. These factors are discussed in more detail below. The Washington State Department of Health Office of Newborn Screening implemented the revised screening algorithm on February 15, 2012. Table 4.5 below shows the new classification categories of TSH results in μIU/mL by age at collection that the WA NBS Program now uses

Table 4.5: Categorization of Thyroid Screening Results Using Primary TSH

Age at Collection	*Borderline Passive TSH ≥*	*Borderline Active TSH ≥*	*Presumptive Referral TSH ≥*	*Urgent Presumptive TSH ≥*
1 hr	*115*	*175*	*190*	*300*
2-7 hr	*100*	*150*	*180*	*300*
8-17 hr	*60*	*100*	*125*	*300*
18-22 hrs	*40*	*75*	*80*	*300*
23-25 hrs	*35*	*75*	*80*	*300*
26-35 hrs	*30*	*50*	*80*	*300*
36-47 hrs	*26*	*50*	*60*	*100*
48-72 hrs	*20*	*50*	*60*	*100*
73-144 hrs	*18*	*40*	*50*	*100*
145-504hrs	*n/a*	*16*	*35*	*100*
> 505 hrs	*n/a*	*13*	*30*	*100*

Follow-up notes:

Borderline Passive - wait for 2nd screen

Borderline Active - request subsequent screen

Presumptive Positive - refer for thyroid studies and recommend endocrinology consult

Urgent Presumptive Positive - refer for thyroid studies, recommend endocrinology consult, and recommend immediate treatment after labs are drawn

DISCUSSION

Certainly programmatic logistics and feasibility of implementing stratified screening algorithms comes into play. Given the variation in the natural surge in TSH just after birth and how fast the TSH drops down, one approach would be to plot every value on a continuous curve to determine if the result is abnormal. From a software and database perspective, the laboratory is not set up for this type of cutoff

system. Another option would be to make every hour for the first 72 hours of life its own age stratification. Adding on smaller age stratifications after that would result in at least 100 age stratifications. Even if this makes sense statistically, logistically it is not feasible. Laboratory personal review all potentially abnormal results; meaning any result within a designated percentage of the threshold is repeated in duplicate to determine the final result. Final results that are abnormal are assigned the appropriate mnemonic in the system to trigger an abnormal report. There are two issues with having too many age stratifications. First, there is more room for error in assigning the appropriate mnemonic because of the overwhelming number of cutoff options. Second, in order to be able to stratify by age, the age of the infant at the time of collection has to be known. The newborn screening specimen cards are designed to obtain this information by requiring the date and time of birth and the date and time of collection. The database calculates the age at collection based on these four data points. If one of these data points is missing, the database will not calculate the age at time of collection. Newborn screening staff members spend valuable time calling hospitals, labs, and clinics to obtain the missing information. The greater number of age stratifications, the more specific the age at the time of collection has to be to interpret results, thus increasing the number of times a call has to be made to obtain the missing data. There was a noticeable increase in the number of calls to get missing data even with going from six to eleven age stratifications in the new cutoff scheme. Availability of staff and appropriate use of staff time is also therefore a factor in the decision-making process.

Decision-making: it is based on more than data. Policymakers may balance benefits and risks differently depending on: 1) what they consider benefits and risks, 2) how they interpret the significance of the benefits and risks, and 3) commensurability (Wilfond 2005). In their analysis of the decision to add cystic fibrosis to the NBS panel, Wilfond et al point out that policy decisions that are based on balancing benefits and risks are not just an empirical calculus, it is value-based judgment about what constitutes a benefit and a risk (Wilfond 2005). This value-based judgment component is inherent in most policy decisions, even in a decision that is primarily based on data, as was the case in revising the TSH screening algorithm. There are different value-based judgments involved in a decision about establishing cutoffs versus a decision to add a new condition to a newborn screening program, but present none-the-less.

In the context of this decision the concept of value-based judgment came into play before the analysis even began. Over time, a culture of newborn screening has developed. Within this culture, is the firmly held notion that every baby ought to be screened and every true case ought to be identified through screening. Intrinsic to the decision to invoke public health power to mandate a screen, is that the benefit to identifying individuals in the population at risk for these conditions is so great that the public health

system needs to step in and ensure that every baby be screened. The use of public health authority is similar to requiring car seats for infants and children. This overarching decision certainly influences programmatic level decisions such as where to set thresholds for screening tests. If it is imperative to identify the condition, as suggested by a state mandate to do so, then you want to make sure that the screening test implemented identifies as close to 100% of the true positive cases as possible. This informs the decisions about where to set cutoff values and thresholds in newborn screening. There needs to be a balance between sensitivity and specificity. The sensitivity needs to be high in order to ensure true positives are picked up, but there is a limit to how much specificity can be compromised, especially in a large-scale screening program, because the nature of screening a large population is that the vast majority of individuals being screened do not have the condition. So the test needs to be highly specific as well, or too many false positives will be identified. What constitutes too many false positives? Therein lies the value judgment, which is interpreted differently by state newborn screening programs. States recognize that some false positive results come with the ability to identify true cases.

On the programmatic level, there are many things that can be done to mitigate the impact of false positives; the first being to reduce the number of them. The TSH analysis demonstrates that large datasets are useful in revising and fine-tuning screening algorithms to accomplish this task. What comes into play next is the follow-up process for abnormal results, including false positives, and what the public health and healthcare systems can do to mitigate the impact of false positives on the parents and infants. How aggressive follow-up has to be, how results are communicated to health care provider who in turn communicate the results to the parents. One of the benefits of reporting results through providers is that they have an established relationship with the patient and the patient's parents. One of the potential drawbacks is that the type of result or accompanying information that newborn screening would like to report to the parents gets lost in translation. Chapter three presents a literature review that sheds light on what we can learn from the parents about the process and how to make it better.

CONCLUSION

The revised CH cutoff scheme will dramatically decrease the number of false positives, while maintaining the ability to identify true cases of CH. Reducing the number of false positive screens mitigates the impact of false positive results on infants and their families and reduces the workload for the NBS program and the burden on the health care system. The revised CH cutoff scheme also improves the NBS Program's ability to identify mildly elevated TSH values that may correspond with mild forms of CH. A two-screen methodology (three screens for NICU babies) allows for higher cutoff values for early ages and also improves detection of CH cases associated with a delayed rise in TSH and milder forms of CH.

In this case, the large dataset has proven beneficial in identifying effective age stratifications and cutoff values. Using data to determine cutoff stratification schemes that maintain a high sensitivity and specificity is an effective way to inform policy decisions. The ability to analyze a large dataset is crucial to the policy process, as is an understanding of the logistical, programmatic, and political factors that impact decision-making.

CHAPTER 5: SUMMARY AND WHAT IS NEXT FOR NEWBORN SCREENING

When it comes to newborn screening, it is all important. The science behind newborn screening is critical to its success and certainly data are required to make informed decisions about newborn screening. Science is not the only factor in decision-making. History has shown that individual people, and people coming together in organized groups, can advocate for change and impact the landscape of newborn screening. From its roots, newborn screening has been developed, propagated, and sustained by individuals dedicated to the cause. And in the midst of the science and the policy-making, it can be easy to forget that this public health program is about real babies, real parents, and real families. Newborn screening can mean the difference between a healthier life and long-term disability or death. For parents of affected children, they know this, which is one of the reasons parents are such powerful advocates. The parents who experience false positive results are also have experiences to share; these experiences are an important piece of the knowledge base in NBS. Parents can inform policy-making at the programmatic, state, and federal level because parents of infants with normal, false, and true results have knowledge to share. We need to discern this knowledge and translate it into more effective programs.

Moving forward with newborn screening services it is essential to remember that advocacy matters. It is important to know who the advocates are and where they stand on the issues within NBS. It is also crucial to acknowledge that the parents' experiences matter. This is one of the populations this public health program serves and parents can contribute information that will enable NBS programs across the country improve services. Public health has the responsibility to make the best effort to lessen the impact, of particular import is the impact of false positives. The quality assurance data analysis demonstrated that data matters. One way to mitigate the impact of false positives is to reduce the number of false positive results being produced. This can be accomplished by using data to make programmatic decisions about adjusting cutoffs. This is a lesson that could be applied to other conditions currently being screened for in states, or conditions under consideration for expanded NBS.

REFERENCES

AAP, Newborn Screening Task Force. (2000). "Serving the family from birth to the medical home. Newborn screening: a blueprint for the future - a call for a national agenda on state newborn screening programs." Pediatrics **Aug 106**(2 Pt 2): 389-422.

ACMG, American College of Medical Genetics. (2006). "Newborn screening: toward a uniform screening panel and system." Genet Med **8**(Suppl 1): 1S-252S.

Arnold C, Davis TC, Frempong JO, Humiston SG, Bocchini A, Kennen EM, Lloyd-Puryear M (2006). "Assessment of newborn screening parent education materials." Pediatrics **May;117**(5 Pt 2): S320-5.

Arnold C, Davis TC, Humiston SG, Bocchini Jr JA, Bass PF, Bocchini A, Kennen EM, White K, Forsman I (2006). "Infant Hearing Screening: Stakeholder Recommendations for Parent-Centered Communication." Pediatrics **May;117**(5 Pt 2): S341-54.

Bailey Jr, D (2009). "The blurred distinction between treatable and untreatable conditions in newborn screening." Health Matrix Clevel **Winter;19**(1): 141-53.

Bailey Jr, D, Skinner D, Davis AM, Whitmarsh I, Powell C (2008). "Ethical, legal, and social concerns about expanded newborn screening: fragile X syndrome as a prototype for emerging issues." Pediatrics **Mar;121**(3): e693-704.

Bickel H (1996). "The first treatment of phenylketonuria." Eur J Pediatrics **155** (Suppl 1): S2-S3.

Blois B, Riddell C, Dooley K, Dyack S (2005). "Newborns with C8-acylcarnitine level over the 90th centile have an increased frequency of the common MCAD 985A>G mutation." Journal of inherited metabolic disease **28**(4): 551-556.

Botkin J, Clayton EW, Fost NC, et al (2006). "Newborn screening technology: proceed with caution." Pediatrics **117**(5): 1793-1799.

Botkin JR, Clayton EW, Fost NC, et al (2006). "Newborn screening technology: proceed with caution." Pediatrics **117**(5): 1793-1799.

Brosco JP, Seider MI., Dunn AC (2006). "Universal Newborn Screening and Adverse Medical Outcomes: A Historical Note. ." Mental Retardation and Developmental Disabilities Research Reviews(12): 262-269.

Burton ., Withrow K, Arnos KS, Kalfoglou AL, Pandya A (2006). "A focus group study of consumer attitudes toward genetic testing and newborn screening for deafness." Genet Med **Dec;8**(12): 779-83.

Campbell E, Ross LF (2003). "Parental attitudes regarding newborn screening of PKU and DMD." Am J Med Genet A **Jul 15;120A**(2): 209-14.

Centerwall S, Centerwall WR (2000). "The discovery of phenylketonuria: the story of a young couple, two retarded children, and a scientist." Pediatrics **Jan;105**(1 Pt 1): 89-103.

CFF, Cystic Fibrosis Foudnation (2009). 2009 Strategic Planning Report. P. Robert J. Beall, President and CEO. Bethesda, MD.

CFF, Cystic Fibrosis Foundation (2012). "Cystic Fibrosis Foundation Website." from http://www.cff.org/.

Chalcraft K, Britz-McKibbin P (2009). "Newborn Screening of Inborn Errors of Metabolism by Capillary Electrophoresis - Electrospray Ionization-Mass Spectrometry: A Second-Tier Method with Improved Specificity and Sensitivity." Anal. Chem **81**: 307 – 314.

Ciske D, Haavisto A, Laxova A, Rock LZ, Farrell PM (2001). "Genetic counseling and neonatal screening for cystic fibrosis: an assessment of the communication process." Pediatrics **Apr;107**(4): 699-705.

Clague A, Thomas A (2002). "Neonatal biochemical screening for disease." Clin Chim Acta. **315**(1-2): 99-110.

Clayton EW (2005). "Talking with parents before newborn screening." J Pediatr **Sep;147**(3 Suppl): S26-9.

Clayton EW (2005). "Informed consent and biobanks." J Law Med Ethics **33**(1): 15-21.

Clemens C, Davis SA, Bailey AR (2000). "The false-positive in universal newborn hearing screening." Pediatrics **Jul;106**(1): E7.

Collins F, Green ED, Guttmacher AE, Guyer MS (2003). "A Vision for the Future of Genomics Research: A Blueprint for the Genomic Era." Nature Reviews **422**(835).

Comeau A., Accurso FJ, White TB, Campbell 3rd PW, Hoffman G, Parad RB, Wilfond BS, Rosenfeld M, Sontag MK, Massie J, Farrell PM, O'Sullivan BP (2007). "Guidelines for Implementation of Cystic Fibrosis Newborn Screening Programs: Cystic Fibrosis Foundation Workshop Report." Pediatrics **Feb;119**(2): e495-518.

Comeau A., Parad R, Gerstle R, O'Sullivan BP, Dorkin HL, Dovey M, Haver K, Martin T, Eaton RB (2005). "Communications systems and their models: Massachusetts parent compliance with recommended specialty care after positive cystic fibrosis newborn screening result." J Pediatr **Sep;147**(3 Suppl): S98-100.

Davis T, Humiston SG, Arnold CL, Bocchini Jr JA, Bass PF, Kennen EM, Bocchini A, Kyler P, Lloyd-Puryear M (2006). "Recommendations for effective newborn screening communication: results of focus groups with parents, providers, and experts." Pediatrics **May;117**(5 Pt 2): S326-40.

Dillard J, Carson CL, Bernard CJ, Laxova A, Farrell PM (2004). "An analysis of communication following newborn screening for cystic fibrosis. ." Health Commun **16**(2): 197-205.

Dillard J, Shel L, Robison JD, Farrell PM (2010). "Parental information seeking following a positive newborn screening for cystic fibrosis." J Health Commun **Dec; 15**(8): 880-94.

Duffner P, Caggana M, Orsini JJ, Wenger DA, Patterson MC, Crosley CJ, Kurtzberg J, Arnold GL, Escolar ML, Adams DJ, Andriola MR, Aron AM, Ciafaloni E, Djukic A, Erbe RW, Galvin-Parton, P, Helton, LE, Kolodny EH, Kosofsky BE, Kronn DF, Kwon JM, Levy PA, Miller-Horn J, Naidich TP, Pellegrino JE, Provenzale JM, Rothman SJ, Wasserstein MP (2009). "Newborn screening for Krabbe disease: the New York State model." Pediatr Neurol **Apr;40**(4): 245-52.

Fant K, Clark SJ, Kemper AR (2005). "Completeness and complexity of information available to parents from newborn-screening programs." Pediatrics **May;115**(5): 1268-72.

Feuchtbaum L., Cunningham G, Sciortino S. (2007). "Questioning the Need for Informed Consent: A Case Study of California's Experience with a Pilot Newborn Screening Research Project." J Empir Res Hum Res Ethics **Sep;2**(3): 3-14.

GA (2010). Genetic Alliance Annual Report 2010. Washignton, DC.

GA. (2010). "Genetic Alliance Registry & Biobank - 2010 Executive Summary." from http://www.biobank.org/english/view.asp?x=1364&id=76.

GA. (2012). "Genetic Alliance Website." from www.geneticalliance.org.

Gracey K. (2003). "Current concepts in universal newborn hearing screening and early hearing detection and intervention programs." Adv Neonatal Care **Dec;3**(6): 308-17.

Green J., Hewison J, Bekker HL, Bryant LD, Cuckle HS (2004). "Psychosocial aspects of genetic screening of pregnant women and newborns: a systematic review (executive summary)." Health and Technology Assessment **8**(33).

Grob R. (2008). "Is my sick child healthy? Is my healthy child sick?: changing parental experiences of cystic fibrosis in the age of expanded newborn screening." Soc Sci Med **Oct;67**(7): 1056-64.

Gurian E, Kinnamon D, Henry J, Waisbren S (2006). "Expanded Newborn Screening for Biochemical Disorders: The Effect of a False-Positive Result." Pediatrics **17**(6): 1915–1921.

Guthrie R. (1996). "The introduction of newborn screening for phenylketonuria: A personal history." Eur J Pediatrics **155**(Suppl 1): S4-S5.

Harris K, Pass KA (2007). "Increase in congenital hypothyroidism in New York and in the United States." Molecular Genetics and Metabolism **91** 268-277.

Hewlett J, Waisbren S (2006). "A review of the psychosocial effects of false-positive results on parents and current communication practices in newborn screening." J Inherit Metab Dis **29**(677-682).

HHF. (2009). "Hunter's Hope Foundation." Retrieved July 10, 2009, from http://www.huntershope.org.

Howell R. (2006). "We need expanded newborn screening." Pediatrics **117**: 1800-1805.

HRSA (2011). Recommended Uniform Screening Panel* of the Secretary's Advisory Committee on Heritable Disorders in Newborns and Children. Rockville, MD Health Resources and Services Administration.

Jonsen A, Durfy SJ, Burke W, Motulsky AG (1996). "The advent of the "unpatients'." Nat Med **2**(6): 622-4.

Kahler S, Fahey MC (2003). "Metabolic disorders and mental retardation." Am J Med Genet C Semin Med Genet **Feb 15;117C**(1): 31-41.

Kaye C, Accurso F, La Franchi S, Lane PA, Hope N, Sonya P, G Bradley S, Michele LP (2006). "Newborn screening fact sheets." Pediatrics **Sep;118**(3): e934-63.

Kim S, Lloyd-Puryear MA, Tonniges TF (2003). "Examination of the communication practices between state newborn screening programs and the medical home." Pediatrics **Feb;111**(2): E120-6.

Kingdon J (1995). Agendas, Alternatives, and Public Policies. New York, HarperCollins College Publishers.

Kladny B, Willaims A., Gupta A, Gettig EA, Krishnamurti L. (2011). "Genetic counseling following the detection of hemoglobinopathy trait on the newborn screen is well received, improves knowledge, and relieves anxiety." Genet Med. **Jul 13**(7): 658-61.

Koch J. (1997). Robert Guthrie – the PKU story: Crusade against mental retardation. Pasadena, CA, Hope Publishing House.

Kurtzer-White E., Luterman D. (2003). "Families and children with hearing loss: grief and coping." Ment Retard Dev Disabil Res Rev **9**(4): 232-5.

Kwon C., Farrell PM (2000). "The Magnitude and Challenge of False-Positive Newborn Screening Test Results." Arch Pedicatr Adolesc Med **Jul;154**(7): 714-718.

LaFranchi S. (2011). "Approach to the diagnosis and treatment of neonatal hypothyroidism." J Clin Endocrinol Metab **Oct;96**(10): 2959-67.

Lagoe E., Labella S, Arnold G, Rowley PT (2005). "Cystic fibrosis newborn screening: a pilot study to maximize carrier screening." Genet Test **Fall;9**(3): 255-60.

Lindee M. (2000). "Genetic Disease since 1945." Nat Rev Genet. **Dec 1**(3): 236-41.

Lipstein, E., Nabi, E, Perrin, JM, Luff, D, Browning, MF, Kuhlthau, KA (2010). "Parents' decision-making in newborn screening: opinions, choices, and information needs." Pediatrics **Oct;126**(4): 696-704.

Lipstein E, Perrin JM, Waisbren SE, Prosser LA (2009). "Impact of false-positive newborn metabolic screening results on early health care utilization." Genet Med **Oct;11**(10): 716-21.

Mandl K, Feit S, Larson C, Kohane IS (2002). "Newborn Screening program practices in the United States: notification, research, and consent." Pediatrics **Feb;109**(2): 269-73.

Marshall BC, Campbell PW 3rd (2009). "Improving the care of infants identified through Cystic Fibrosis newborn screening." J Pediatr **Dec;155**(6 Suppl): S71-2.

McCandless S. (2004). "A primer on expanded newborn screening by tandem mass spectrometry." Prim Care **Sep;31**(3): 583–604.

Miller F, Robert JS, Hayeems RZ (2009). "Questioning the consensus: managing carrier status results generated by newborn screening." Am J Public Health **Feb;99**(2): 210-5.

Morrison D, Clayton EW (2011). "False positive newborn screening results are not always benign." Public Health Genomics **14**(3): 173-7.

Motulsky A. (1997). "Screening for genetic diseases." N Engl J Med **May 1;336** (18): 1314-6.

Naouri M., Boisseau C, Bonicel P, Daudon P, Bonneau D, Chassaing N, Martin L (2009). "Manifestations of pseudoxanthoma elasticum in childhood." Br J Dermatol **Sep;161**(3): 635-9.

NAS (1975). Genetic screening: programs, principles, and research. National Academy of Sciences. Washington, DC.

NBS, Washington. (2012). "Washington State Department of Health, Office of Newborn Screening." from http://www.doh.wa.gov/YouandYourFamily/InfantsChildrenandTeens/NewbornScreening.aspx.

NNSGRC. (2012). "National Newborn Screening and Genetics Resource Center." Retrieved June, 2007-2012, from http://genes-r-us.uthscsa.edu/.

Parsons E., Bradley DM (2003). "Psychosocial issues in newborn screening for cystic fibrosis." Paediatr Respir Rev **Dec;4**(4): 285-92.

Paul D. (2000). "A double-edged sword." Nature Reviews **June (405)** (1): 515.

PCOB (2008). The Changing Moral Focus of Newborn Screening. Washington DC, President's Council on Bioethics.

Prosser L, Ladapo JA, Rusinak D, Waisbren SE (2008). "Parental Tolerance of False-positive Newborn Screening Results." Arch Pediatr Adolesc Med **162**(9): 870-876.

Riley C. (2007). Public Health Follow-up of Abnormal Newborn Screening Results. Epidemiology - Public Health Genetics, School of Public Health and Community Medicine. Seattle, WA, University of Washington. **Master of Public Health:** 1-104.

Riley C., Watts C (2011). The Role of Advocacy in Newborn Screening. Achieving Justice in Genomic Translation: Rethinking the Pathway to Benefit. W. B. e. al. New York, New York, Oxford University Press**:** 149-161.

Russ S, Hanna D, DesGeorges J, Forsman I (2010). "Improving follow-up to newborn hearing screening: a learning-collaborative experience." Pediatrics **Aug; 126**(Suppl 1): S59-69.

Sabatier P, Jenkins-Smith H (1993). Policy change and learning: An advocacy coalition approach. Boulder, CO, Westview Press.

Sabatier PA (2007). Theories of the Policy Process, Westview Press.

SACHDNC (2010). Newborn Screening for Severe Combined Immunodeficiency Disorder. Rockville, MD, Health Resources and Services Administration, Secretary's Advisory Committee on Heritable Disorders in Newborns and Children.

Scriver C. (2001). "Garrod's foresight; our hindsight." J Inherit Metab Dis **24**: p93-116.

Simopoulos A, Committee for the Study of Inborn Errors of Metabolism (2009). "Genetic screening: programs, principles, and research--thirty years later. Reviewing the recommendations of the Committee for the Study of Inborn Errors of Metabolism (SIEM)." Public Health Genomics. **12**(2): 105-11.

Skinner D, Choudhury S, Sideris J, Guarda S, Buansi A, Roche M, Powell C, Bailey Jr DB (2011). "Parents' decisions to screen newborns for FMR1 gene expansions in a pilot research project." Pediatrics **Jun;127**(6): e1455-63.

Skinner D, Sparkman KL, Bailey Jr DB (2003). "Screenign for fragile X syndrome: parent attitudes and perspectives." Genet Med **5**: 378-84.

Smith L, Oyeku SO, Homer C, Zuckerman B (2006). "Sickle cell disease: a question of equity and quality." Pediatrics **117**(5): 1763–1770.

Sorenson JR, Levy HL, Mangione TW, Sepe SJ (1984). "Parental Response to Repeat Testing of Infants with 'False-Positve' Results in a Newborn Screening Program." Pediatrics **73**: 183-187.

Tarini B, Burke W, Scott CR, Wilfond BS (2008). "Waiving informed consent in newborn screening research: balancing social value and respect." Am J Med Genet C Semin Med Genet **Feb 15;148C**(1): 23-30.

Tarini BA, Cristaksi DA, Welch HG. (2006). "State newborn screening in the tandem mass spectrometry era: more tests, more false-positive results." Pediatrics **Aug 118**(2): 448-56.

Terry S, Terry PF, Rauen KA, Uitto J, Bercovitch LG (2007). "Advocacy groups as research organizations: the PXE International example." Nat Rev Genet **Feb;8**(2): 157-64.

Timmermans S, Buchbinder M (2010). "Patients-in-waiting: Living between sickness and health in the genomics era." J Health Soc Behav **Dec;51**(4): 408-23.

Tluczek A, C., MA, Lynam PA (2010). "When the cystic fibrosis label does not fit: a modified uncertainty theory." Qual Health Res **Feb;20**(2): 209-23.

Tluczek A, Koscik RL, Farrell PM, Rock MJ (2005). "Psychosocial risk associated with newborn screening for cystic fibrosis: parents' experience while awaiting the sweat-test appointment." Pediatrics **Jun;115**(6): 692-703.

Tluczek A, Koscik RL, Modaff P, Pfeil D, Rock MJ, Farrell PM, Lifchez C, Freeman ME, Gershan W, Zaleski C, Sullivan B (2006). "Newborn screening for cystic fibrosis: parents' preferences regarding counseling at the time of infants' sweat test." J Genet Couns **Aug;15**(4): 277-91.

Tluczek A., Orland KM, Cavanagh L (2011). "Psychosocial consequences of false-positive newborn screens for cystic fibrosis." Qual Health Res **Feb;21**(2): 174-86.

Tluczek A, Orland KM, Nick SW, Brown RL (2009). "Newborn Screening An Appeal for Improved Parent Education." J Perinat Neonatal Nurs **Oct-Dec;23**(4): 326-34.

Vohr B., Jodoin-Krauzyk J, Tucker R, Johnson MJ, Topol D, Ahlgren M (2008). "Results of newborn screening for hearing loss: effects on the family in the first 2 years of life." Arch Pediatr Adolesc Med **Mar;162**(3): 205-11.

Waisbren S, Albers S, Amato S, et al. (2003). "Effect of Expanded Newborn Screening for Biochemical Genetic Disorders on Child Outcomes and Parental Stress." JAMA **290**(19): 2564-2571.

Wang G., Watts C (2007). "The role of genetics in the provision of essential public health services." Am J Public Health. **Apr;97**(4): 620-5.

Weible CM, Sabatier, PA. (2007). A Guide to the Advocacy Coalition Framework. Handbook of Public Policy Analysis: Theory, Politics, and Methods. F. M. Fischer, Gerald. Sidney, Mara. Boca Raton, CRC/Taylor & Francis, Marcel Dekker, Inc.

Wilfond B, Parad RB, Fost N (2005). "Balancing benefits and risks for cystic fibrosis newborn screening: implications for policy decisions." J Pediatr **Sep;147**(3 Suppl): S109-13.

Wilson J., Jungner G (1968). Principles and Practice of Screening for Disease. Geneva, World Health Organization.

Appendix A: PubMed Search Results for Parents and Newborn Screening

PUBMED SEARCH TERMS	Initial Search Result	# of articles excluded based on title/ short description: not on topic or before 2000	# of duplicate articles excluded	# of articles excluded based on abstract review			# of articles excluded based on full article review				Total # of articles in Review	# of articles with original data
				not on topic	not in the US	no parent component	not on topic	not in the US	no parent component	article unavai-lable		
newborn screening, parents, experience	108	76	0	0	8	3	0	2	4	0	**15**	9
newborn screening, parents, views	32	25	0	0	5	0	0	0	0	0	**2**	2
newborn screening, parents, perspective	25	14	1	3	4	1	2	0	0	0	**0**	0
newborn screening, false positives, parents	19	6	0	5	7	0	0	0	0	0	**1**	1
false positive, newborn screening, impact	48	22	5	0	10	5	0	0	2	0	**3**	3
psychosocial, false positive, newborn screening	14	3	1	3	3	0	0	0	2	0	**2**	2
Newborn screening, parents, decision-making	56	24	3	0	13	3		1	1	1	**10**	6
Articles added in based on Reference Review	12	0	0	0	0	0	0	0	3	0	**9**	7
TOTAL	**314**	**170**	**10**	**11**	**50**	**12**	**2**	**3**	**12**	**1**	**42**	**30**

Appendix B: HSD reference #42995, "Quality Assurance Analysis to Revise Newborn Screening Algorithm for Congenital Hypothyroidism"

UNIVERSITY of WASHINGTON
HUMAN SUBJECTS DIVISION

5/11/2012

To: Ms. Catherine Riley
Doctoral Student, Public Health

Re: HSD reference #42995, "Quality Assurance Analysis to Revise Newborn Screening Algorithm for Congenital Hypothyroidism"

Dear Ms. Riley,

The Human Subjects Division received your determination request on 5/10/2012. Your research activity described in the above-referenced determination request has been by screened by Subcommittee EG.

As outlined in your determination request, the research activity will only involve the receipt and analysis of data that is not individually identifiable, as the data cannot be linked to specific individuals by the investigators either directly or indirectly through coding systems. Information in your determination request indicates that:

(1)the private information or specimens were not collected specifically for the currently proposed research project through an interaction or intervention with living individuals; and

(2) if the information is coded, the investigators cannot readily ascertain the identity of the individuals to whom the coded private information or specimens pertain because the investigators and the holder of the key have entered into an agreement prohibiting the release of the key to the investigators under any circumstances.

Given this information and the definition of "human subject" under 45 CFR 46.102(f), the research has been determined to not meet the federal regulatory definition of "human subjects research". Therefore, you do not need IRB review and approval to perform your activities. Please keep this memo and a copy of your returned determination request for your records.

If you have further questions or concerns, feel free to contact me.

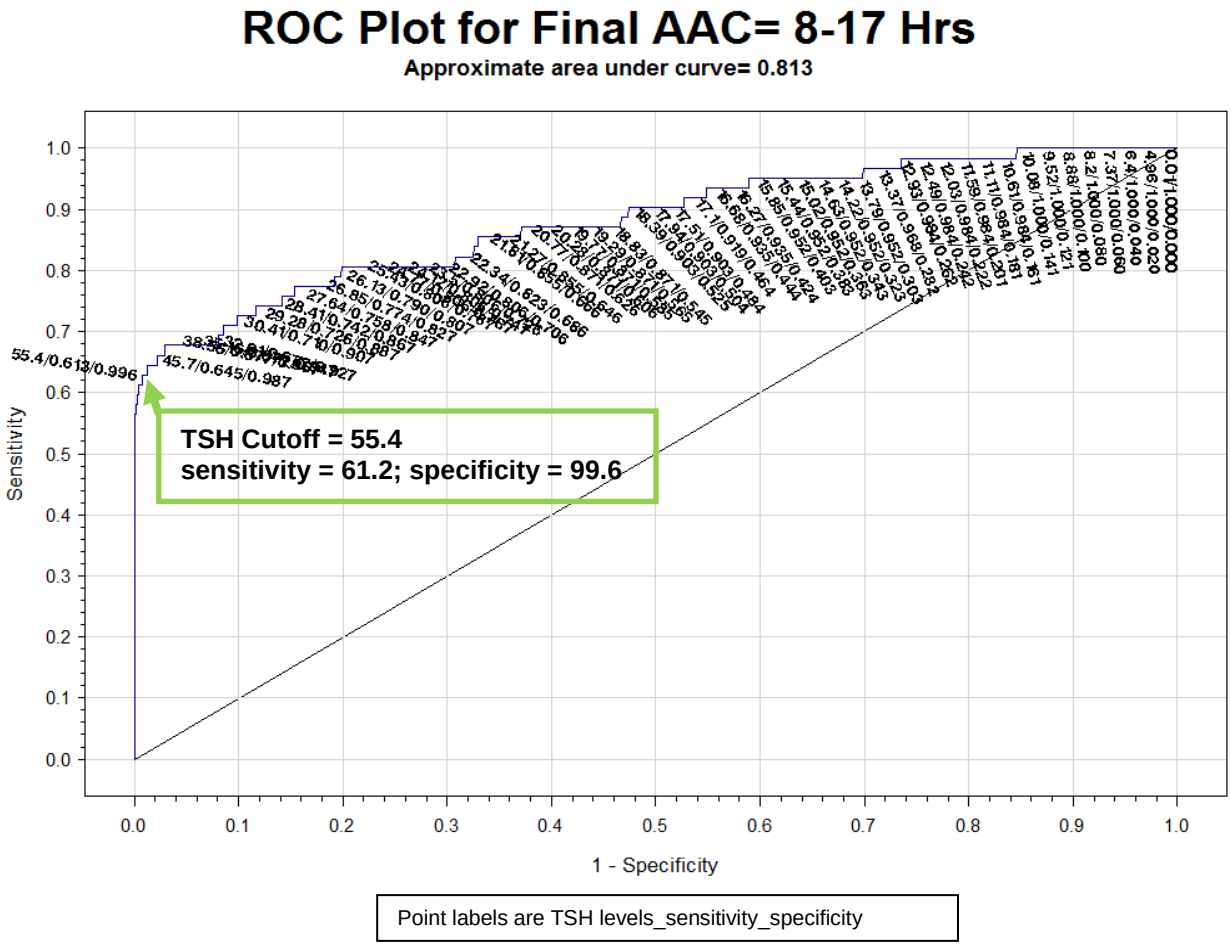
Best regards,

Laura Busch
Human Subjects Review Administrator
Minimal Risk Subcommittee EG
(206) 543-1926
lbusch@u.washington.edu

4333 Brooklyn Ave. NE, Box 359470 Seattle, WA 98195-9470

(206) 543-0098 fax 206.543.9219 hsdinfo@uw.edu www.washington.edu/research/hsd

Appendix C: ROC curve example



VITA

Catharine Riley earned her Bachelor of Science in Molecular and Cellular Biology from the University of Arizona and went on to complete a Master of Public Health Administration and Policy from the University of Arizona and a Master of Public Health Genetics from the University of Washington. In 2012 she earned her Doctor of Philosophy in Public Health Genetics at the University of Washington. Dr. Riley is currently a follow-up specialist for the Washington State Newborn Screening Program. Prior to her doctorate work Dr. Riley served as the Director of the Arizona Turning Point Project, a national initiative funded by the Robert Wood Johnson Foundation to strengthen and improve the public health infrastructure. She also held the positions of staff researcher on a project to enumerate the public health workforce in Arizona and Continuing Education Coordinator for the Mel and Enid Zuckerman Arizona College of Public Health.