An Investigation of the Association Between Newborn Screening Biomarkers and Total Serum Bilirubin Levels in Ontario Newborns

A Systematic Review and Secondary Data Analysis

Emeril Santander

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Master of Science in Epidemiology

School of Epidemiology and Public Health
Faculty of Medicine
University of Ottawa

ABSTRACT

The performance of biomarker-based screening tests may be improved by the identification of circumstances associated with biomarker differences aside from presence or absence of screen-targeted disease. We conducted a systematic review and database analysis investigating the association between total serum bilirubin (TSB) near the time of newborn screening for rare disease (NBS) and concentrations of NBS biomarkers, with implications for false positive (FP) results. Sixteen studies were identified reporting associations and lack thereof between NBS biomarkers and hyperbilirubinemia, evidencing a relatively unexplored subject. We conducted a cross-sectional analysis of NBS biomarkers and TSB using results from universal screening tests. Correlations revealed TSB level associated with most NBS biomarkers. Reassuringly, TSB was not significantly associated with FP risk for conditions with high FP rates such as congenital adrenal hyperplasia, congenital hypothyroidism, or phenylketonuria in regression analyses. Ascertainment of NBS specificity in neonates with common transient conditions such as hyperbilirubinemia may be a viable approach to improve the accuracy of NBS tests.

ACKNOWLEDGEMENTS

I would like to thank my co-supervisors, Drs. Steven Hawken and Beth Potter, without whom this thesis project would not have been possible. Since first collaborating with these outstanding epidemiologists in a workplace setting, I had been eagerly looking forward to learning from both of you in the context of graduate training. The experience was even more positive than I could have foreseen. Your consistent leadership and deep expertise are evident throughout this work, and I truly appreciate your contributions.

I would like to acknowledge the work of my Thesis Advisory Committee members, including their heartwarming willingness to review my research progress, for several hours at a time, over the course of multiple years. Dr. Erika Bariciak's expertise as an academic neonatologist was essential to properly understanding this thesis' subject matter, and the encouragement provided throughout was indispensable. I would like to thank Drs. Matthew Henderson and Pranesh Chakraborty for their irreplaceable guidance on the topics of clinical biochemistry and rare inherited disease. Although not a member of the committee, the systematic review component of this thesis benefitted greatly from the strategic advice of Dr. Lindsey Sikora, and I extend her my gratitude for her excellent input.

Throughout this project, I had the privilege of being able to consult my work colleagues at Newborn Screening Ontario on various technical and research-related topics that contributed to this thesis. I would like to thank Jennifer Milburn and Christine McRoberts for their unwavering support for this project as newborn screening leaders. Thank you, Michael Kowalski, for reviewing my thesis proposal; Janet Marcadier for discussions on diagnostic follow-up; and Matt and Pranesh for in-depth technical guidance and advice. More generally, I would thank the organization for entrusting me to investigate this research question. It is inspiring, on a daily basis, to be part of a team that is passionate about newborn screening.

I would like to thank my spouse, MacKenzie Stone, for her selfless and loving support as I navigated the graduate student experience. I could not have done this without you, Mac.

Finally, I would acknowledge that this project was enabled by the broad support of a loving network of family and friends. Thanks for believing this was possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF BIOMARKERS

Short Name	Full Name	Type
TSB	Total Serum Bilirubin	Heme Derivative
TYR	Tyrosine	Amino Acid
LEU + ILEU	Leucine with Isoleucine	Amino Acid
PHE	Phenylalanine	Amino Acid
ORN	Ornithine	Amino Acid
GLY	Glycine	Amino Acid
ARG	Arginine	Amino Acid
CIT	Citrulline	Amino Acid
VAL	Valine	Amino Acid
ASA	Argininosuccinic Acid	Amino Acid
MET	Methionine	Amino Acid
ALA	Alanine	Amino Acid
SUAC	Succinylacetone	Ketone
C0	Free Carnitine	Carnitine
C2	Acetylcarnitine	Acylcarnitine
C3	Propionylcarnitine	Acylcarnitine
C3DC	Malonylcarnitine	Acylcarnitine
C4	Butyrylcarnitine	Acylcarnitine
C4OH	3-Hydroxybutyrylcarnitine	Acylcarnitine
C4DC	Succinylcarnitine	Acylcarnitine
C5	Valerylcarnitine	Acylcarnitine
C5:1	Tiglylcarnitine	Acylcarnitine
C5OH	Methylmalonylcarnitine	Acylcarnitine
C5DC	Glutarylcarnitine	Acylcarnitine
C6	Hexanoylcarnitine	Acylcarnitine
C6DC	3-Methylglutarylcarnitine	Acylcarnitine
C8	Octanoylcarnitine	Acylcarnitine
C8:1	Octenoylcarnitine	Acylcarnitine
C10	Decanoylcarnitine	Acylcarnitine
C10:1	Decenoylcarnitine	Acylcarnitine
C12	Dodecanoylcarnitine	Acylcarnitine
C12:1	Dodecenoylcarnitine	Acylcarnitine
C14	Tetradecanoylcarnitine	Acylcarnitine
C14:1	Tetradecenoylcarnitine	Acylcarnitine
C14:2	Tetradecadienoylcarnitine	Acylcarnitine
C14OH	3-OH-tetradecanoylcarnitine	Acylcarnitine
C16	Hexadecanoylcarnitine	Acylcarnitine
C16OH	3-OH-hexadecanoylcarnitine	Acylcarnitine
C16:1OH	3-OH-hexadecenoylcarnitine	Acylcarnitine
C18	Octadecanoylcarnitine	Acylcarnitine
C18OH	3-OH-octadecanoylcarnitine	Acylcarnitine
C18:1	Octadecenoylcarnitine	Acylcarnitine
C18:1OH	3-OH-octadecenoylcarnitine	Acylcarnitine
C18:2	3-OH-octadecadienoylcarnitine	Acylcarnitine
GALT	Galactose-1-phosphate uridyl transferase	Enzyme
TSH	Thyroid Stimulating Hormone	Hormone
17-OHP	17-Hydroxyprogesterone	Hormone
BIOT	Biotinidase Enzyme Activity	Enzyme
IRT	Immunoreactive Trypsinogen	Enzyme
TREC	T-Cell Receptor Excision Circle	DNA fragment

Note: Acylcarnitines abbreviations contain the number of carbons (C) within the acyl group. DC denotes a dicarboxylic acid in the acyl group. OH denotes a hydroxyl group attached to the acyl group. The number after the colon represents the total number of double bonds in the acyl group.

1.0 Introduction

1.1 General Background

The neonatal transition to extrauterine life is a physiologically unique period of development during which many bodily systems begin to function independently from maternal support [1]. Although these rapid changes proceed uneventfully for most healthy neonates, those born with rare diseases can begin to experience impaired health soon after birth as disordered physical functions are performed independently for the first time. Congenital metabolic, endocrine, and immune abnormalities are particularly concerning to neonatal healthcare providers due to their pre-symptomaticity at birth and the accompanying possibilities of sudden health crises, death in early infancy, or irreversible disability.

Medical advances beginning in the 1930s identified that neonates with certain conditions exhibited markedly distinct biomarker concentrations in bodily fluids [2] that could be measured to identify and diagnose disease prior to the onset of symptoms. Given that these disorders are frequently treatable when identified early, over the subsequent decades many came to be targeted by biomarker-based public health screening initiatives — known today as newborn screening (NBS) programs [3]. These population-wide programs aim to identify and deliver treatment to newborns who require it by testing every infant born for a predefined panel of rare diseases.

Modern NBS using dried blood spot (DBS) samples is based on the measurement of approximately 50 or more disease-indicative biochemical markers [4] in blood samples, also known as 'metabolites', collected from newborns after the 24th hour of life. With analysis frequently taking place at centralized laboratories, the concentrations of a neonate's biomarkers are compared to pre-selected risk thresholds, also known as 'cut-offs' or decision limits, to yield a binary determination [5] of higher or lower disease risk. Neonates identified to be at higher risk for targeted disease are referred to specialized diagnostic care.

Within NBS, the minimization of risk determinations discordant with the true underlying disease status is a scientific priority. Previous NBS studies have identified a small number of neonatal factors that are associated with altered biomarker profile distributions, unrelated to the presence of target diseases. Biomarker variability studies have reported that female neonates [6], neonates born at preterm gestational age [7,8], those with lower birthweights [9,10] or those receiving parenteral nutrition [11] are more likely to have anomalous biomarker observations or screening results for one

or more diseases, at times including an increased risk of false positive findings [12]. Certain enzymatic biomarkers have been reported to vary across neonates born in different seasons [13]. Research that identifies these influential traits, factors, or circumstances is critical in that it enables i) the enhanced interpretation of neonatal biomarker profiles and ii) the implementation of 'augmented' screening approaches to reduce potential sources screening error. In a recent example, the identification of consistently elevated 17-hydroxyprogesterone among premature, sick, and stressed [14] neonates enabled the introduction of risk thresholds stratified by gestational age to account for these elevations [9]. These adjustments resulted in increased specificity without sacrificing sensitivity among premature neonates undergoing screening for congenital adrenal hyperplasia [15]. Other newborn screening augmentations include supplementary repeated screening to avoid transient physiologic or transfusion-related alterations in disease-indicative biomarkers [16], implementation of specificity-increasing second-tier tests [17], implementation of risk decision limits based on multiple biomarkers or biomarker ratios [18], and the development of more accurate adjunctive assay technologies [19]. Augmentative strategies are not unique to NBS and can also be found in other screening settings. For example, combinations of sensitive primary and specific secondary tests have been used to improve the detection of hearing loss [20] and autism [21] in children. Screening augmentations can improve the quality of screening for all screened individuals; the crucial first step toward enabling their implementation is the initial identification of factors influencing the biomarkers used to estimate targeted disease risk.

This thesis investigates relationships between neonatal total serum bilirubin (TSB), which is commonly elevated in newborn blood, and dried blood spot markers indicative of rare disease in order to inform potential improvements in NBS as a public health screening program. In the next sections, additional introductory details on neonatal hyperbilirubinemia, newborn dried blood spot screening, and screening performance are provided and recently available evidence suggesting that NBS markers may associate with TSB is reviewed. The first chapter ends with a formal statement of my two research objectives and describes how each is presented as a constituent article within this thesis.

1.2 Neonatal Hyperbilirubinemia and Pre-Discharge Hyperbilirubinemia Screening

Bilirubin is a product of the catabolism of heme complexes released during the breakdown of red blood cells and certain liver enzymes [22,23]. In human neonates, a variety of normal physiologic processes as well as diseases may cause elevations in circulating bilirubin ('hyperbilirubinemia')

leading to a characteristic yellowing of the skin and mucous membranes ('jaundice'). Although most newborns experience some degree of jaundice, severe or prolonged hyperbilirubinemia may lead to acute bilirubin encephalopathy or chronic bilirubin encephalopathy ('kernicterus'), neurologic disorders caused by bilirubin's toxic effects on neurons [24]. These potentially hazardous instances of hyperbilirubinemia are universally considered medical emergencies requiring prompt hospitalization and treatment. In Canada, nearly two-thirds of term newborns and almost all premature neonates experience jaundice that will resolve spontaneously in the first weeks of life; 2% of all births result in circulating bilirubin levels high enough to warrant treatment [25].

The causes of hyperbilirubinemia are broadly categorized as either non-pathologic or pathologic, based on whether the jaundice is caused by a disease process [26]. The most common reasons for neonatal hyperbilirubinemia are non-pathologic in origin and are mostly related to increased demands on the neonatal liver and immature liver capacities. The newborn liver's ability to conjugate bilirubin, a necessary step for excretion, is impaired by reduced activity of UDP-glucuronyltransferase which has a capacity that is only 1% of adult enzymatic function in term infants [27]. Increased liver

Box 1-1.The Measurement of Bilirubin

In humans, bilirubin circulates in different biochemical forms ('fractions') that can be measured from blood separately ('fractionated analysis') or in unison ('total analysis').

Bilirubin Fractions

Unconjugated Bilirubin (UCB)
This water-insoluble type of bilirubin is the immediate product of heme breakdown. For historical reasons related to early testing methods, this form is also known as *indirect* bilirubin. UCB may be attached or unattached to a companion protein in the blood called albumin. The UCB that is not attached to albumin is termed free bilirubin (B_f) and is the causative agent of bilirubin neurotoxicity.

Conjugated Bilirubin (CB)
During processing by the liver, UCB is attached ('conjugated') to glucuronic acid to make it water-soluble and easier to excrete as waste. This form cannot affect the brain. Elevations of CB may be termed 'conjugated hyperbilirubinemia' or cholestasis. Although cholestasis is a broader diagnosis than conjugated hyperbilirubinemia, the former is most frequently diagnosed based on the latter.

Delta Bilirubin (δ)
In blood, albumin may bind to CB. Delta bilirubin is albumin-bound conjugated bilirubin. The 'delta' refers to this bilirubin type being the difference between the simple addition of UCB and CB and total serum bilirubin:

$$\delta = TSB - (UCB + CB)$$

Combinations of Bilirubin Fractions

Total Serum Bilirubin (TSB)
This is the sum of all bilirubin fractions described above. It is used during hyperbilirubinemia screening and is the primary compound of interest for this thesis.

Direct Bilirubin (DB)
This type of bilirubin is the sum of conjugated bilirubin and delta bilirubin. At times, this name is used interchangeably with CB. This is not entirely accurate because CB does not necessarily include the albumin-bound portion of CB.

Other

Transcutaneous Bilirubin (TcB)
Interest in the measurement of bilirubin without the need for blood led to the development of transcutaneous Bilirubinometry (TCBR). These devices approximate TSB by shining light on the surface of the skin. TcB is neither a bilirubin fraction nor a combination thereof but rather a non-invasive point-of-care estimate of TSB.

Lumirubin
Bilirubin that is exposed to light is transformed in to non-toxic and water-soluble lumirubin, which is readily excreted in urine. This relieves the liver of the responsibility of processing bilirubin into a water-soluble form prior to elimination. Light-based conversion of bilirubin to lumirubin is the mechanism of action of phototherapy, the primary treatment for neonatal hyperbilirubinemia.

demands originate from higher erythrocyte volumes in neonates and fetal red blood cells with abbreviated lifespans relative to adult cells, leading to increased quantities of bilirubin from the breakdown of hemoglobin-derived heme [28]. The trauma of birth may also cause subcutaneous hemorrhage, further increasing the quantity of effete red blood cells requiring heme catabolism [29]. Among pre-term neonates, the mechanisms underlying non-pathologic hyperbilirubinemia are similar but increased in magnitude, leading to jaundice increased in severity and more likely to be prolonged [30]. Pathologic hyperbilirubinemia occurs in neonates who have infections [31], disorders affecting red blood cells [32–37] or the liver enzymes required to make bilirubin excretable [38,39] and other structural [40–43] or functional [44–46] impairments of the mechanisms that eliminate bilirubin from the body [47]. Notably, a number of NBS targets may also present with jaundice including cystic fibrosis [48], galactosemia [49], tyrosinemia [50], citrullinemia [51], congenital cytomegalovirus infection [52] and congenital hypothyroidism [53]. Severely hyperbilirubinemic neonates receive phototherapy as treatment, which converts toxic bilirubin in to non-toxic lumirubin [54]. Neonates with pathologic hyperbilirubinemia are also treated through management of the underlying disease.

Regardless of etiology, the potential for severe life-long harm faced by neonates experiencing hazardous levels of hyperbilirubinemia [29], together with the ready availability of safe, economical, and effective treatments [55,56], have prompted health experts to recommend universal pre-discharge neonatal screening for hyperbilirubinemia [57]. In 2004, the American Academy of Pediatrics recommended assessing all newborns for risk of hyperbilirubinemia using a total serum bilirubin (TSB) test in whole blood paired with analysis of clinical risk factors (see Box 1-1 for measures of bilirubin) [58]. The Canadian Pediatric Society issued similar screening recommendations in 2007 [25]. The primary objective of universal pre-discharge bilirubin screening remains controversial [59,60], because sparse evidence is currently available suggesting pre-discharge hyperbilirubinemia screening decreases the incidence of bilirubin-related brain damage [61] . Evidence from Ontario does confirm that hyperbilirubinemia-screened neonates are more likely to receive phototherapy and less likely to present emergently for a jaundice-related concern [62]. Proponents of universal pre-discharge hyperbilirubinemia screening assert the core objective is to enable the earlier administration of phototherapy so as to prevent the need for exchange transfusion, a procedure used in severe phototherapy-resistant hyperbilirubinemia and known to be invasive and higher in risk [60].

As pertinent to the objectives of this thesis, the implementation of universal pre-discharge hyperbilirubinemia screening in Ontario [63] has resulted in a synchronized and parallel configuration of universal laboratory-based screening initiatives sampling neonates in the first days of life. Since 2007, neonates born in Ontario have been recommended to receive hyperbilirubinemia screening between 24-48 hours after birth, either with a non-invasive transcutaneous bilirubin measurement or with a whole-blood based TSB test [63]. This interval is identical to the recommended sampling of the dried blood spot based NBS for rare disease, and both are additionally recommended to occur prior to the first separation from the birth setting to ensure universality. The result is a unique opportunity to investigate the association between the results of two universal biomarker-based screening initiatives. In the next subsection, an introduction is provided to NBS.

1.3 Neonatal Dried Blood Spot Screening for Rare Disease

Table 1-1. Ontario's Newborn Screening Panel and Biomarkers Derived from the Dried Blood Spot Sample

Disorder Group	Target Disorder	Indicative Biomarkers
Endocrine Disorders	Congenital Hypothyroidism	Thyroid Stimulating Hormone
	Congenital Adrenal Hyperplasia	17-Hydroxyprogesterone , Andostenedione, Cortisol, 11-Deoxycortisol, 21-Deoxycortisol
Amino Acid Disorders and Citrullinemias	Maple Syrup Urine Disease	Leucine, Alanine
	Phenylketonuria	Phenylalanine , Tyrosine
	Tyrosinemia	Succinylacetone, Tyrosine
	Homocystinuria	Methionine
	Citrullinemia or Argininosuccinic Aciduria	Citrulline, Arginin, Argininosuccinic Acid, Ornithine
Fatty Acid Oxidation Disorders	Medium-Chain Acyl-CoA Dehydrogenase Deficiency	C8, C10, C10:1, C2, C5DC, C12, C14:1
	Long Chain Hydroxyl Acyl Dehydrogenase or Trifunctional Protein Deficiencies	C16OH, C14OH, C16:1OH, C18OH, C16, C18, C18:1
	Very Long Chain Acyl Dehydrogenase Deficiency	C4, C2, C14:1, C12:1, C14, C14:1, C14:2, C16, C18, C18:1
	Carnitine Uptake Defect	C0, C2, C3, C16, C18
	Carnitine Palmitoyltransferase 1 Deficiency	C0, C16, C18
	Carnitine Palmitoyltransferase 2 Deficiency	C0, C16
Organic Acid Disorders	Propionic or Methylmalonic Acidemia	C2, C3, Methylcitrate
	Isovaleric Acidemia	C0, C2, C3, C4, C5, C16
	Glutaric Aciduria Type 1	C5DC, C2, C8, C5OH, C16
Immune Disorders	Severe Combined Immune Deficiency	T-Cell Receptor Excision Circles, Adenosine, Deoxyadenosine, Guanosine, Deoxyguanosine, Mutations
Hematological Disorders	Sickle Cell Disease	Normal and Variant Hemoglobins
Other Disorders	Cystic Fibrosis	Immunoreactive Trypsinogen, Mutation Panel
	Galactosemia	Galactose-1-phosphate uridylyltransferase activity
	Biotinidase Deficiency	Biotinidase activity

Note: The full names of biomarkers listed in this table are available in the front matter of this thesis.

Modern dried blood spot based NBS began in the 1960's in the northeastern United States. At the time, the rare disease phenylketonuria (PKU) was a major contributor to the prevalence of life-long mental retardation and a growing quantity of research indicated early administration of a specialized diet could prevent cognitive impacts [64,65]. Dr. Robert Guthrie, the relative of a child with PKU, developed the first bacterial inhibition assays in 1961 that could detect excess phenylalanine (PHE) produced in the urine of PKU-affected neonates [66]. In the context of widespread public enthusiasm for preventive health screening, Guthrie refined the test in 1963 to only require a minimal amount of blood spotted on to paper after puncture the neonate's heel [67]. By 1965, 27 of 50 United States had legislated that every neonate born within the jurisdiction receive the heelstick screening test for PKU [68]. In Canada, Prince Edward Island was the first province to begin NBS in 1963 and was followed by most provinces and territories by the end of that decade [69]. Facilitated in part by the introduction of new laboratory technologies [70], NBS progressively expanded to include testing for additional disorders using the Guthrie blood spot technique [71]. Over the next 50 years, dried blood spot-based NBS became the standard of care for identifying newborns with rare disorders [72].

The preanalytical phase of neonatal screening, which includes patient education, sample collection and sample transport, is described in Figure 1-1. In Ontario, dried bloodspot-based screening is performed at the Newborn Screening Ontario central laboratory located at the Children's Hospital of Eastern Ontario in Ottawa.

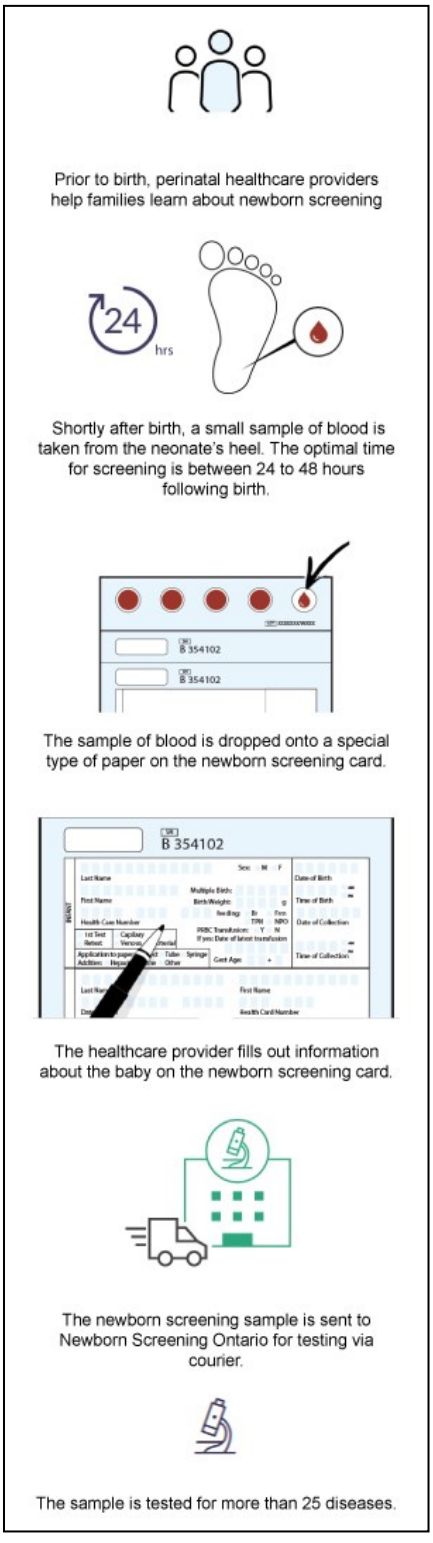


Figure 1-1. The initial newborn screening process as shown in patient education materials provided by Newborn Screening Ontario

Ontario's screening program is typical of most others implemented throughout North America in targeting disorders responsive to the early intervention that is enabled by timely identification in early life [72,73]. This unifying characteristic connects the diverse variety of predominantly monogenic rare disorders present on Ontario's screening panel (Table 1). Although NBS began as a public health screening initiative aiming to identify a single metabolic disease (PKU), current panel targets can also be endocrine, immune, or multi-system diseases.

Once the sample is received, blood-saturated sample cards are analyzed with an assortment of clinical laboratory methods to generate biomarker measurements. Each marker's concentration is compared to a pre-established disease-specific risk threshold ('cutoff') to develop risk determinations for one or more targeted disorders. At the conclusion of blood spot biomarker profile analysis, the neonatal care provider receives a risk report detailing whether the infant has screened positive as having a high risk for targeted rare diseases.

In the next section, factors influencing the diagnostic performance of newborn screening tests are described.

1.4 Screening Performance and Improvement Strategies

In NBS that biochemically measures the concentrations of disease markers in dried blood spots on a continuous scale, it is mainly the selection of biomarker cut-offs [74] (also termed 'decision limits' or 'thresholds') that alters the diagnostic performance characteristics of the screen, such as specificity and sensitivity. Most newborns have detectable concentrations of disease-indicative markers in their blood as part of normal physiological processes, however neonates who will ultimately be diagnosed with a screen-targeted disorder have abnormally high or low concentrations of the indicative biomarker [75]. The ideal indicative biomarker of targeted disease would consistently achieve complete separation between concentration levels observed in affected neonates relative to those unaffected, and in those cases a decision limit would entirely separate healthy neonates from those with disease. In practice, complete separation is almost never achieved with a biochemical threshold since healthy newborns may occasionally have abnormal biomarkers and infants truly affected by a targeted condition may not exhibit sufficiently abnormal values. Given that affected and unaffected infants have biomarker concentrations that originate from overlapping continuous distributions, establishing a biomarker cutoff requires the weighing of the relative impact of false positives versus false negatives in

terms of public health, economic, and individual patient considerations in order to yield a test with acceptable performance characteristics [76–79].

In particular, given the importance of early identification to initiate treatment in order to prevent severe morbidity and mortality for many diseases that are targets of NBS, cutoffs have typically been selected to maximize sensitivity and thus minimize the risk of a false negative screening result [80,81]. Thus, a major focus of research to further improve the accuracy of screening has been to reduce the occurrence of false positive findings (i.e., improve specificity) while maintaining the low risk of false negative findings [12,82–85]. False positive neonatal screening determinations have been well-established to be associated with patient and health system harms [86]. The psychosocial impacts of false positives primarily impact new parents who may experience emotional distress upon receipt of a positive risk determination [87,88]. False positive newborn screening tests have been reported to alter family health-seeking behaviours, prompting parents to seek more medical care for their infants relative to parents who received an accurate result and increasing their health resource utilization [89]. Research that can enhance screening performance not only benefits individual patients but also the broader health system by ensuring that health resources are utilized for newborns requiring treatment [90].

In section 1.5, literature suggesting interrelation between NBS biomarkers and neonatal hyperbilirubinemia is introduced followed by the formal hypothesis statement of this work.

1.5 Rare Disease Screening and the Hyperbilirubinemic Neonate

Previous studies have adduced that hyperbilirubinemia is a presenting feature in several disorders targeted by NBS [91,92]. The human liver is the primary site for bilirubin conjugation and metabolism of many macromolecules [93]; hyperbilirubinemia secondary to decreased liver function may also manifest as alterations in the markers indicative of metabolic disease in NBS. In adults, liver disease or injury has been shown to alter biomarkers of fatty acid beta oxidation [94]. In NBS for inherited tyrosinemias and other amino acid disorders, immaturity within enzymatic pathways of the liver have been identified as a source of transient elevations in these acids that can mimic rare disease [95]. Thus, due to the numerous physiological interconnections between liver function and NBS biomarkers, it appears biologically plausible that one or more of these markers may be altered among hyperbilirubinemic neonates.

During the data analysis phase of the investigations for this thesis, a study became available by McCarthy and colleagues that represented the first published analysis of dried blood spot NBS markers in hyperbilirubinemic infants [96]. This study was not identified by the systematic review presented as Article A because it was published after the last search date for our review. In a sample of approximately 1.7 million Californian neonates, McCarthy et al. ascertained hyperbilirubinemia through the presence of diagnostic codes in the neonatal health record and linked these with biomarker measurements obtained from universal dried blood spot NBS. The objectives of the McCarthy study were to predict incident hyperbilirubinemia and kernicterus using information from the dried blood spot markers, however, the study incidentally identified biomarker alterations within infants clinically diagnosed with hyperbilirubinemia. In multiple logistic models adjusting for several potential confounders, the study team reported that 31 of 42 NBS markers differed among apparently healthy infants with diagnosed hyperbilirubinemia relative to neonates without a hyperbilirubinemia diagnosis. The strongest associations were observed between hyperbilirubinemia status and increased phenylalanine, decreased ornithine, and decreased leucine/isoleucine, in infants without a diagnosis of metabolic disease. The authors concluded that the neonatal metabolic profiles of hyperbilirubinemic neonates were distinct from those of neonates without hyperbilirubinemia.

The McCarthy et al. [96] study provides additional early indications that NBS biomarkers may be altered among hyperbilirubinemic neonates, however the authors noted that lack of access to total serum bilirubin (TSB) results was the primary limitation of their study. The methodology implemented in Article B is a continuous analysis of TSB values and NBS biomarkers and thus addresses this limitation.

1.6 Research Objectives

The identification of neonatal factors that influence the concentrations of disease-indicative NBS biomarkers may enhance interpretation of newborn screening results and/or facilitate the implementation of screening augmentations that compensate for the errors attributable to the influential factor, thereby increasing screening performance.

The aim of this project is to investigate associations between infant total serum bilirubin levels and NBS biomarkers using linked data from laboratories conducting these two screening

activities. To that end, two specific objectives were each addressed using an appropriate epidemiologic approach:

i. The current state of the published evidence concerning associations between neonatal bilirubin and NBS biomarkers was systematically reviewed and summarized through a narrative synthesis. Methods and results for this component of the thesis are presented as Article A.

ii. Associations between NBS biomarkers as well as false positivity risk and newborn bilirubin levels were quantitatively assessed using descriptive, correlational, and logistic regression modelling analyses. This investigation is presented as Article B.

This thesis hypothesizes that biomarker levels measured during rare disease screening are associated with neonatal total serum bilirubin levels as obtained from universal pre-discharge hyperbilirubinemia screening.

Interface

The upcoming article describes a systematic review of associations between biomarkers used in neonatal screening and infant bilirubin or hyperbilirubinemia status. This systematic review was conducted to summarize published studies associating biochemical dried blood spot markers used in newborn screening and infant bilirubin or hyperbilirubinemia status.

Ethics

Systematic reviews do not involve human participants. Ethics board approval for this study was not required.

Contributions, Journal Style

All authors participated in the conception and design of this study. ES and JT participated in data acquisition and analysis. Interpretation of findings and manuscript drafting tasks were done by ES, with critical revisions by BP, SH, EB, MH, and PC.

Contributions, Detailed

The student (ES) was responsible for developing the systematic review search, screening citations and articles for eligibility, extraction of information from included studies, and evaluating risk of bias within included studies. The second author (JT) acted as second reviewer and verified inclusion decisions, extracted data, participated in risk of bias determinations, and participated in manuscript development. A health science librarian (LS) provided guidance on biomarker systematic review search methodology. MH, PC, and LS reviewed the search strategy and results of review for comprehensiveness and together with EB participated in manuscript development. SH and BP guided the design of the systematic review while also supervising information gathering, interpretation, and manuscript development.

Intended Journal Submission

This manuscript will be submitted to the International Journal of Neonatal Screening (IJNS) and has been formatted to meet the requirements of the publisher.

Article A. Associations between Infant Bilirubin and Screening Biomarkers of Rare Disease: A Systematic Review

Emeril Santander [1,2], Jessica Tao [1,3], Steven Hawken [2,4], Beth K. Potter [1,2], Matthew P.A. Henderson [1], Erika Bariciak [6,7], Lindsey Sikora [4] and Pranesh Chakraborty [1,6]

[1] Newborn Screening Ontario, Children's Hospital of Eastern Ontario, Ottawa, ON, K1H 8M5, Canada;
[2] School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, K1H 8L1, Canada
[3] Faculty of Medicine, University of Ottawa, Ottawa, ON, K1H 8L1, Canada
[4] Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[5] Health Science Libraries, University of Ottawa, Ottawa, ON, K1H 8L1, Canada
[6] Department of Pediatrics, Faculty of Medicine, University of Ottawa, Ottawa, ON K1H 8L1, Canada
[7] Department of Neonatology, Children's Hospital of Eastern Ontario, Ottawa, ON, K1H 8M5, Canada;

Abstract

Variation in neonatal screening biomarkers unattributable to screen-targeted disease can adversely impact screening accuracy. Neonatal jaundice is a common clinical finding that is frequently present during screening sampling. We conducted a systematic review to summarize peer-reviewed evidence associating bilirubin levels or hyperbilirubinemia status with an assortment of markers used to screen neonates for disorders on the Recommended Uniform Screening Panel, July 2018 version. Searches of MEDLINE, PUBMED, and Embase yielded 7,316 abstracts, and after two levels of screening, 16 studies were included. Reviewers extracted data on study characteristics, biomarkers measured, and biomarker-bilirubin association statistics. Results were summarized in tabular and narrative formats. Fifteen of 16 identified studies were observational, and one was a trial. We identified limited evidence of association between infant bilirubin and markers of thyroid function, markers of fatty acid oxidation, fetal hemoglobin, amino acid markers, enzymes, and diagnosed congenital heart defects. Studies typically reported no association between bilirubin and immune markers and certain amino acids. Most studies were at lower risk of bias. The identified literature suggests biomarkers used to screen neonates for rare disease could be altered among infants with hyperbilirubinemia. Additional investigation is warranted probing screening marker alterations and modified screening accuracy in this subpopulation.

A 1.1 Introduction

Modern universal neonatal screening utilizes advanced laboratory technologies to screen for rare diseases that are asymptomatic in the newborn period but may cause severe harm or death when untreated in early life [1]. In many cases, screening enables earlier intervention to the extent that health outcomes for children diagnosed through screening are comparable to those of children without the rare disease [2].

The disorders targeted by neonatal screening programs are selected following consideration of their suitability for mass screening [3]. In the United States, the Secretary's Advisory Committee on Heritable Disorders in Newborns and Children (SACHDNC) recommends disorders that should be included on newborn screening panels [3]. The Secretary of Health Resources and Services may accept the recommendation and include the disorder on the Recommended Uniform Screening Panel (RUSP). The RUSP serves as a national standard for designated newborn screening laboratories in the United States [4].

Although core disorders present on the RUSP are primarily genetic in etiology, screening laboratories identify infants at elevated risk for disease by measuring biochemical markers present in dried blood whose altered concentrations are associated with presence of the target disorders [5]. Molecular genetic testing and/or other investigations are then conducted as part of confirmatory diagnostic care for those newborns with positive (high risk) biomarker screening results. Screening laboratories conducting tests for all disorders currently on RUSP may measure over 50 individual biomarkers when determining an infant's risk status.

Variability of screening biomarkers unrelated to the presence of target diseases has been identified within the samples of female [6], premature [7,8], and ill or stressed [9] neonates as well as those receiving parenteral nutrition [10]. These biomarker differences that are not attributable to screen-targeted diseases can alter the performance characteristics of newborn screening tests, increasing the likelihood of harmful screening outcomes such as false positive or false negative results [11]. Families of infants receiving false positive newborn screening determinations may experience adverse psychosocial impacts [12] and subsequently increase their use of health services, straining the healthcare system [13]. False negative risk determinations can be equally or more harmful, as neonates with targeted disease could miss time-sensitive opportunities for treatment [14]. Subgroup variation in biomarker distributions and diagnostic performance are properties inherent to most medical tests applied across diverse populations [15–17]. The identification of

factors that influence newborn screening biomarkers remains a priority for researchers, healthcare providers, and newborn screening laboratories seeking to minimize screening errors [18].

Hyperbilirubinemia is a common finding in neonates, occurring in 60% of term infants and almost all infants born premature [19–21]. This condition may be pathologic or physiologic in origin and occurs when bilirubin concentrations become elevated in blood. The infant's skin and mucous membranes may become yellowed, and in rare untreated instances of extreme bilirubin elevation, the infant may experience neurologic damage of variable permanence [22]. Despite the relatively high prevalence of hyperbilirubinemia among neonates undergoing newborn screening, evidence suggestive of relationships between infant bilirubin and biomarkers used to screen for rare diseases has not previously been systematically reviewed.

A 1.2 Objectives and PECOS Statement

In this review, we aimed to establish the current status of the published evidence regarding potential associations between biomarkers used to screen newborns for a standard panel of rare disorders and bilirubin levels or hyperbilirubinemia status among neonates.

The primary population of interest was human neonates 28 days of age or younger. The exposures of interest were biomarkers known to be indicative of a core disorder on the July 2018 edition of the Recommended Uniform Screening Panel. Our research question did not include a comparator group. The outcomes of interest were categorical hyperbilirubinemia status and/or continuous bilirubin level.

A 1.3 Methods

The review protocol was registered with PROSPERO (CRD 42019123350) on March 19, 2019 [23]. Two known protocol deviations are described in the Methods or Discussion sections as appropriate. Both the protocol and findings of this review adhere to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [24–26].

A 1.3.1 Literature Search

In consultation with an experienced health science librarian (LS), a systematic search was developed and performed on the MEDLINE, Embase, and Cochrane Central databases for studies

published prior to March 1st, 2019. The search strategy, available as Supplementary Material A, included both Medical Subject Headings (MeSH) and textword-based queries. These searches were reviewed by a clinical biochemist (MH) and a metabolic physician (PC) for comprehensiveness. The reference lists of included studies were also searched for potentially relevant studies.

A 1.3.2 Eligibility Criteria

Studies were eligible for inclusion in this review after meeting every inclusion criterion and avoiding each ground for exclusion (Table 1). Reviewers applying criteria had primarily methodological (ES) or primarily clinical (JT) expertise.

Table 1. Criteria for Study Selection

Inclusion	Exclusion
• Study is an observational or experimental design	• Language of publication was not English
• Participants are live neonates (≤ 28 days) or infants (≤ 12 months) [1]	• The study was not published in a peer-reviewed, indexed journal
• Study measured ≥ 1 biomarker associated with a RUSP core disorder, July 2018 version	• Study is an abstract-only publication or conference proceeding and a complete study publication is unavailable
• Study measured bilirubin level or determined hyperbilirubinemia status of participants	• Study was performed *in vitro* or *ex vivo*
• RUSP biomarker association with bilirubin level or hyperbilirubinemia status was reported	• Participant bilirubin level or hyperbilirubinemia status was not determined using objective laboratory techniques
	• The primary objective of the study was the validation of an assay or laboratory technique

[1] In recognition of the logistical challenges of studying participants within the brief neonatal period, this review considered studies on infants 1 year or less of age. The population of primary interest remains neonates 28 days or less of age.

A 1.3.3 Screening

Citation lists from each searched database were imported to Covidence software (Veritas Health Innovation, Melbourne, Australia) [27] which automated the implementation of a two-stage

screening process. Specifically, at one stage, we used a liberal accelerated selection process previously described by Khangura et al. [28]. Reviewer ES screened all citation titles and abstracts against eligibility criteria. Study abstracts deemed ineligible by ES were then screened by JT. At this stage, only references voted ineligible by both ES and JT were removed from the review; those identified as potentially eligible by either reviewer were advanced to the second stage of screening. Reasons for exclusion were not tracked at this stage.

In the second stage of eligibility screening, full texts of articles that passed the first screening stage were obtained via university access to academic journals or interlibrary loan. Both reviewers (ES and JT) independently reviewed the full text of each study and established adherence to inclusion criteria. These criteria were applied in an unranked manner, and studies excludable on multiple grounds were removed based on the first appreciable unmet criterion at review-time. After each reviewer contributed an inclusion vote to all studies identified, study-specific votes were compared and disagreements resolved through discussion. Cohen's kappa was calculated as a measure of inter-rater reliability at the full text review phase [29].

A 1.3.4 Data Extraction

Data from studies meeting all criteria at the full-text review stage were extracted using a standardized form generated by the study team in Excel software (Microsoft Corporation, Redmond, United States). This form is available as Supplementary Material B. In general, variables collected from each study related to characteristics of the article (author, publication year), infant demographics (number of participants, age, weight), comparator demographics, details on bilirubin tests performed and reported associations with newborn screening markers. We did not contact authors regarding data items not available in the published articles. Reviewer 1 (ES) completed data extraction and reviewer 2 (JT) reviewed each extracted element. Reviewers ES and JT met to resolve disagreements through discussion and re-review of full study texts.

A 1.3.5 Risk of Bias Assessment in Individual Studies

An appropriate standard risk-of-bias tool was selected based on the study design of each included study. One of the three possible variants of the Newcastle-Ottawa Scale (NOS) [30] was applied to observational studies based on study subtype: NOS-Cohort, NOS-Case Control, or NOS-Cross Sectional [31]. Experimental studies were assessed using the Cochrane Risk of Bias (RoB) 2.0 tool

[32]. Both ES and JT assessed each risk domain independently and met to compare determinations. Disagreements were resolved through discussion and re-examination of article text.

A 1.3.6 Summarization and Narrative Synthesis

Heterogeneity among screened disorders precluded the calculation of meta-analytic summary measures. The findings of included studies are summarized in tabular format and we report individual measures of association reported within studies. The findings are also summarized in a narrative synthesis by disorder groups.

A 1.4 Results

A 1.4.1 Study Selection

Database searches returned 7,316 individual citations (Figure 1) of which 1,926 were automatically recognized as duplicate and removed. Of the remaining 5,390 studies, 4,993 were excluded for irrelevance during initial screening. Among the remaining 397 studies, 71 met all inclusion criteria. Sixteen of these studies were non-hearing related and 55 were hearing-related (relating to protocol deviation explained below). Excellent inter-reviewer agreement was observed at the full-text review stage, with reviewers initially disagreeing on the inclusion of a single reference that was subsequently found includable (Cohen's k = 0.99).

In a deviation from the pre-published protocol, the study team elected to first report the findings of studies that do not pertain to hearing and report hearing-related studies in a second publication that is not part of this thesis document. This deviation aimed to preserve the thematic unity of the current review; studies finding associations between bilirubin and physiological markers of audition were deemed sufficiently conceptually distinct from studies of biochemical markers of disease that a discrete report was warranted. The most common ground for exclusion at the full-text review stage was non-quantitation of a biomarker associated with a RUSP disorder (n = 128 studies) followed by non-report of a measure of association (n = 60 studies). Studies excluded on the latter criterion were predominantly single-participant case studies that did not repeatedly measure biomarkers over time, which precluded the measurement of association between bilirubin and the target marker. A substantial quantity (n = 46) of studies were returned based on

English titles or abstracts and were subsequently excluded based on non-English text in the body of the report. A total of 16 studies were retained in our study for further analysis.

A 1.4.2 Studies Investigating Bilirubin Association with Markers of Endocrine Disorders

Seven included studies examined relationships between bilirubin and newborn screening

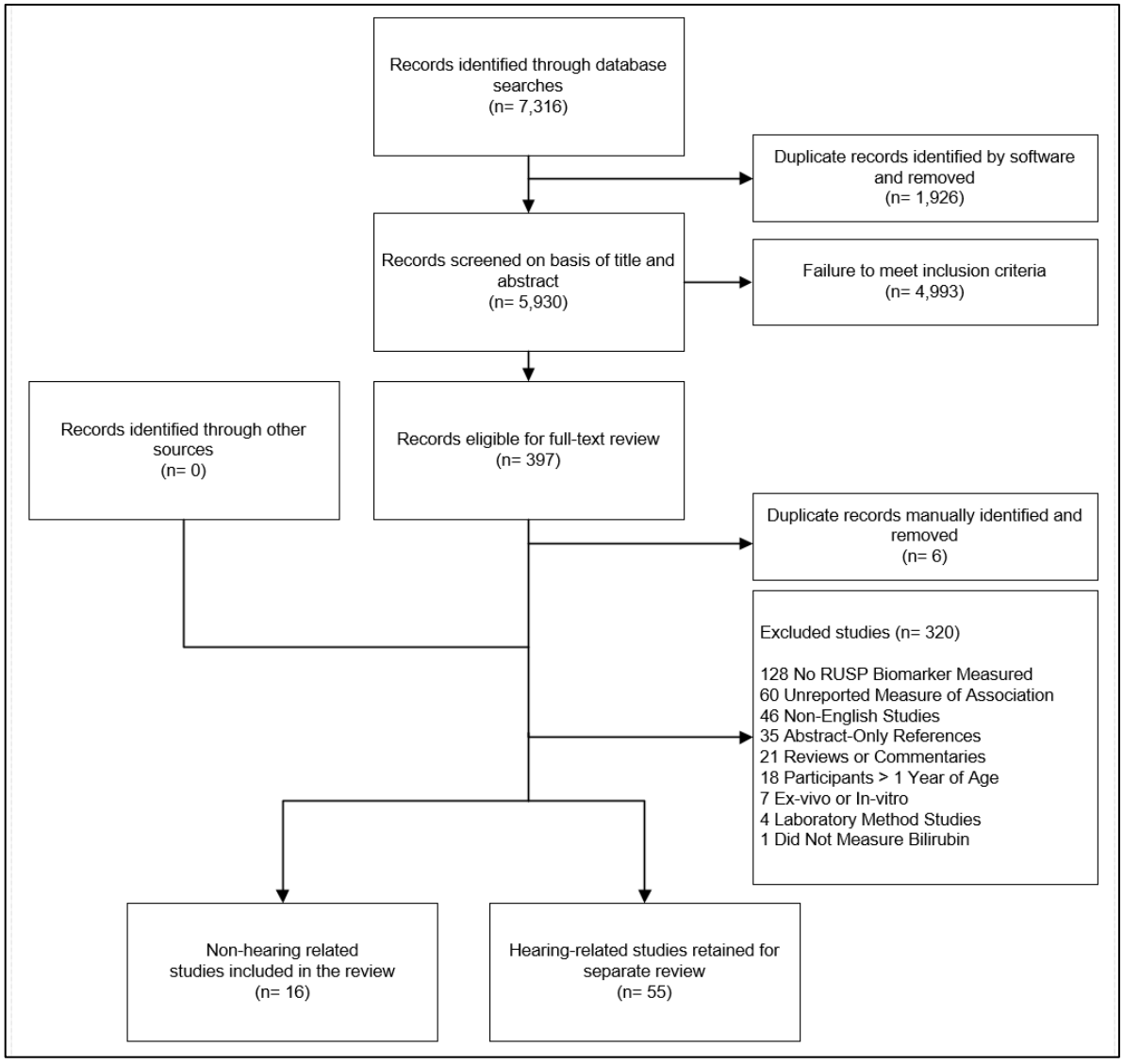


Figure 1. PRISMA flow diagram for the systematic review of the associations between infant bilirubin and neonatal screening biomarkers

biomarkers used during newborn screening to establish risk for endocrinopathies [33–39]. These articles were published between 1996 and 2018 in Denmark, France, India, Iran, Israel, Singapore, and the United States. Represented study designs included case-control (4 studies), analytic cross-sectional (1 study), randomized trial (1 study) and nested case control (1 study).

Studies varied widely in number of participating infants with confirmed hyperbilirubinemia, ranging from four [33] to 207 [34]; five of seven included studies investigated less than 100 confirmed hyperbilirubinemic participants and two of these studies investigated less than 10 confirmed hyperbilirubinemic participants.

Six studies reported associations between markers of thyroid function and bilirubin, such as thyroid stimulating hormone (TSH), triiodothyronine (T3), or thyroxine (T4). One study measured cortisol (Mauvais et al.) [38]. Three of four studies measuring TSH (Kvetny et al.[33], Mojtahedhi et al. [34], and Tan et al. [36]) reported decreased TSH among hyperbilirubinemic neonates, and Singh et al.[35] found unaltered TSH in hyperbilirubinemic neonates relative to healthy controls. Three of five studies (Mauvais et al.[38], Mojtahedhi et al. [34], Singh et al. [35]) measuring T4 found no association between T4 and bilirubin; Ulanovsky et al.[37] identified a strong association between hyperthyroxinemia and idiopathic hyperbilirubinemia but only among cesarean-delivered neonates. In a nested case-control study of premature neonates receiving total parenteral nutrition (TPN), Steinbach et al.[39] observed significantly decreased T4 levels among 13 infants who became cholestatic, which causes elevated bilirubin levels, in the first postnatal month.

Two studies [33,38] examined endocrine markers in infants with health states altered by disease. Kvetny et al.'s [33] study of neonates born to euthyroid mothers with autoimmune thyroid disease and non-autoimmune thyroid disease found a non-significantly elevated rate of severe hyperbilirubinemia among neonates born to mothers with thyroid disease of autoimmune origin. Mauvais et al.[38] studied conjugated hyperbilirubinemia in infants with Pituitary Stalk Interruption Syndrome (PSIS) and reported significantly decreased cortisol levels among PSIS infants who developed cholestasis.

A 1.4.3 Studies Investigating Bilirubin Association with Markers of Amino Acid Disorders

Three studies examined associations between infant bilirubin and biomarkers used to screen neonates for aminoacidopathies [39–41]. Dates of publication ranged between 2003 and 2018 with studies taking place in Turkey, China, and the United States. Included works were either retrospective cohort, analytic cross-sectional, or nested case control in design. Erdol et al.[40] studied serum homocysteine in 215 neonates with varying bilirubin levels, two of whom were deemed cholestatic, while Steinbach et al. [39] reported results of an amino acid panel performed on 13 cholestatic neonates. Thirty-seven participants in Gong et al.[41] were cholestatic when tested for

an amino acid panel. Participants with conjugated hyperbilirubinemia in the Gong et al.[42] study also had either biliary atresia or neonatal intrahepatic cholestasis due to citrin deficiency (NICCD).

In Erdol et al.'s [40] study of 215 neonates with varying levels of bilirubin, correlational analyses did not identify significant differences in plasma homocysteine as a function of total bilirubin or either bilirubin fraction. Steinbach et al.'s [39] study concluded that among premature neonates receiving TPN, those whom developed cholestasis had significantly elevated citrulline, histidine, and methionine; glutamate and serine were significantly decreased. Gong et al.[42] also reported significantly increased citrulline and methionine among cholestatic infants with the most marked elevations occurring among infants with NICCD. This study observed tyrosine was elevated among infants with NICCD but not among infants cholestatic secondary to biliary atresia. Cholestatic infants in this study also had elevated arginine and ornithine while alanine and glycine were significantly decreased relative to healthy controls.

All included studies reporting associations between bilirubin and markers of amino acid disorders studied in populations in altered states of health. Erdol et al.[40] studied a special cohort of infants referred to specialized metabolic care following a positive neonatal screen for PKU or biotinidase deficiency. Hyperbilirubinemic participants in Gong et al.[41] were also receiving metabolic or gastroenterological care for NICCD or biliary atresia. Participants in the Steinbach et al.[39] trial were neonates born after the 23rd week but prior to the 30th week of gestation who required intensive care and parenteral nutrition.

A 1.4.4 Studies Investigating Bilirubin Association with Markers of Fatty Acid Oxidation or Organic Acid Disorders

Four studies reported associations between bilirubin and markers used to screen neonates for fatty acid oxidation disorders [39,40,42,48]. Publication of these studies occurred between 2003 and 2018 in Turkey, Japan, China, and the United States. Included study designs were retrospective special cohort, case control, nested case control, and analytic cross-sectional. Three of four studies investigated acylcarnitine markers derived from dried blood in relation to infant bilirubin and the remaining study associated urine-derived methylmalonic acid (MMA) with bilirubin. The number of infants classified as hyperbilirubinemic included in these four studies ranged between two to

37. Erdol et al.[40], Gong et al.[42], and Steinbach et al.[39] measured acylcarnitines in addition to other biochemical markers.

Only a single included study (Lee et al.[48]) exclusively investigated acylcarnitines in relation to bilirubin.

Erdol et al.[40] investigated 215 neonates with varying levels of bilirubin and found no correlation between total or fractionated serum bilirubin and urinary levels of methylmalonic acid. When participants were dichotomized based on measured urinary MMA above or below the upper limit of normal (1.9 mmol/mmol creatinine), the two resulting subgroups did not differ in bilirubin level. In Lee et al.'s [48] case control study of 4 infants with conjugated hyperbilirubinemia secondary to NICCD, short-chain (C0, C2, and C3) as well as long and very-long chain (C10, C12, C14, C16, C18, C18:1) acylcarnitines were elevated relative to healthy controls. Cholestatic infants and healthy controls did not differ significantly in levels of medium-chain (C4, C5, C5OH, C6, C8) and some longer-chain (C12:1, C14:1, C14OH, C16:1, C16OH, C18:1OH) acylcarnitines. Gong et al.[42] reported similar findings in 37 neonates with NICCD or BA, with shorter-chain (C0, C2, C3, C4DC, C6) and longer-chain (C14, C16, C18, C18:2) acylcarnitines elevated in infants with conjugated hyperbilirubinemia relative to controls. Cholestatic infants in Gong et al.[42] did not differ in certain short and medium (C3DC, C4, C5, C5DC, C6, C6DC, C8) and long to very-long (C10, C12, C18) acylcarnitines relative to controls. Steinbach et al.'s [39] study of 14 TPN-alimented cholestatic infants only observed significant elevation of succinylcarnitine (C4DC), with all other measured acylcarnitines (C0, C2, C3, C4, C5DC, C5, C5OH, C5DC, C16, C16:1, C18, C18:1, C18:2) not differing significantly between cholestatic and non-cholestatic participants.

A 1.4.5 Studies Investigating Bilirubin Association with Hemoglobinopathy or Related Markers

Two included studies examined bilirubin in relation to SS disease, the most common form of sickle cell disease, or a marker used to determine risk for hemoglobinopathy [44,46]. United States-based Bainbridge et al.[44] compared the peak total serum bilirubin (TSB) values of 28 infants with SS disease identified by cord-blood electrophoresis and 39 matched controls. No statistically significant differences were identified between cases and controls in either peak TSB or the proportion of infants experiencing TSB elevations exceeding 5mg/dL.

Table 2. Designs, Participants, and Biomarker Sampling Methods of Included Studies

Reference [Design]	Study Population	Age at Marker Sampling [Source]	Age at Bilirubin Sampling [Bilirubin Type]
Arnell et al. 2012 [43] [Retrospective Cohort]	**206** Infants with Trisomy 21	**Not Reported** [Unreported Source]	**1 - 11 Weeks** [Conjugated]
Bainbridge et al. 1988 [44] [Case Control]	**28** Infants with Sickle Cell Disease **39** Matched Controls	**At Birth** [Cord Blood]	**1-3 Days** [Total]
Erdol et al. 2018 [40] [Retrospective Cohort]	**215** Neonates screen positive on neonatal screening for phenylketonuria or biotinidase deficiency	**Not Reported** [Plasma] [Methylmalonic Acid: Urine]	**Avg. 25.3 Days (SD: 1.48 Days)** [Total, Direct, and Indirect]
Eyada et al. 2017 [45] [Case Control]	**25** Neonates with Hyperbilirubinemia **25** Healthy Controls	**<= 14 Days** [Whole Blood]	**<= 14 Days** [Total]
Gong et al. 2014 [42] [Case Control]	**4,898** Healthy and Unwell Infants **8** Infants with NICCD **29** Infants with BA	**NICCD: 2.9 +- 1.2 months** **BA: 2.3 +- 0.7 months** **Controls: Not Reported** [Blood Spot]	**NICCD: 2.9 +- 1.2 months** **BA: 2.3 +- 0.7 months** **Controls: Not Collected** [Total and Direct]
Kanai et al. 2003 [46] [Cross-Sectional]	**69** Infants grouped by fetal hemoglobin proportion (HbF%) High (HbF > 90%) Moderate (HbF 80%-90%) Low (HbF < 80%)	**At Birth** [Cord Blood]	**1 - 5 Postnatal Days, Repeated Daily** [TCBR]
Kurt et al. 2009 [47] [Case Control]	**21** Neonates with Severe hyperbilirubinemia **16** Healthy Neonates	**Avg: 6.3 Days; SD: 2.2 Days** [Whole Blood]	**Avg: 6.3 Days; SD: 2.2 Days** [Total and Indirect]
Kvetny et al. 2006 [33] [Case Control]	**19** Neonates with AITD Mothers **12** Neonates with non-AI TD Mothers	**Not Reported** [Serum]	**Postnatal Day 5** [Total]
Lee et al. 2006 [48] [Case Control]	**4** Infants with NICCD **23** Healthy Infants	**2 to 5 Months** [Blood Spot]	**Various and Repeated** [Total and Direct]

➔ This table continues on the next page. Abbreviation are defined on the next page.

Table 2 (Continued). Designs, Participants, and Biomarker Sampling Methods of Included Studies

Reference [Design]	Study Population	Age at Marker Sampling [Source]	Age at Bilirubin Sampling [Bilirubin Type]
Mauvais et al. 2016 [38] [Case Control]	**11** Infants with PSIS **5** Infants with PSIS and Cholestasis	**<= 1 year** [Plasma]	**<= 1 Year** [Total and Conjugated]
Mojtahedhi et al. 2018 [34] [Cross-Sectional]	**207** Neonates Grouped by TSB **163** Group A (10 – 14.9 mg/dL) **31** Group B (15 – 19,9 mg/dL) **6** Group C (20 – 20.4 mg/dL)	**< 15 Days** [Unreported Source]	**< 15 Days** [Total and Direct]
Schulpis et al. 2003 [49] [Cross-Sectional]	**1,256** Term Neonates (Group 1) **246** Premature Neonates (Group 2) **156** Small for GA Neonates (Group 3)	**2-3 Postnatal Days** [Serum]	**2-3 Postnatal Days** [Total]
Singh et al. 2003 [35] [Case Control]	**100** Infants with Hyperbilirubinemia **100** Healthy Infants	**4-7 Postnatal Days** [Serum]	**4-7 Postnatal Days** [Total and Direct]
Steinbach et al 2003[39] [Case Control Nested in Randomized Trial]	**13** Neonates with Cholestasis **109** Neonates without Cholestasis	**Postnatal Day 7** [Blood Spot] [PDA: Unreported Source]	**< 28 Days** [Direct]
Tan et al. 1996 [36] [Randomized Trial]	**25** Healthy Control Neonates **123** Neonates with Hyperbilirubinemia **41** White Light Phototherapy **41** Fibre Optic Phototherapy **41** Combined Phototherapy	**1-2 Postnatal Days** [Whole Blood]	**1-2 Postnatal Days** [Direct]
Ulanovsky et al 2018 [37] [Case Control]	**3,446** Healthy Neonates **2,513** Vaginal Births **933** Caesarean Births **61** Neonates with Hyperbilirubinemia **44** Vaginal Births **17** Caesarean Births	**44 to 50 Postnatal Hours** [Blood Spot]	**<= 2 Postnatal Days** [Total]

Alphabetized Abbreviations: AITD = Autoimmune thyroid disease **; Avg** = average ; **BA =** Biliary Atresia ;**HbF%** = Proportion of Fetal Hemoglobin; **NICCD =** Neonatal Intrahepatic Cholestasis due to Citrin Deficiency; **Non-AI TD** = Non-Autoimmune Thyroid Disease; **PDA** = Patent Ductus Arteriosus ; **PSIS** = Pituitary Stalk Interruption Syndrome; **SD** = Standard Deviation**; TCBR** = **Transcutaneous Bilirubinometry; TSB** = Total Serum Bilirubin

Table 3. Summary of Main Findings

Reference [Design]	Relevant Disorder [Marker]	Quality Score	Confounder Management [Statistical Adjustment]	Main Findings
Arnell et al. 2012 [43] [Retrospective Cohort]	**CCHD** [Confirmed CHD]	**9 of 9**	**Restricted** by Trisomy 21 Status, Birthdate & Location [Unadjusted Statistical Comparison]	Among infants with Trisomy 21, neonatal cholestasis (cBil > 30 umol/L and cBil/TSB > 0.2) was significantly more common in children with CHD than in those without CHD (9 of 96 vs 0 of 110; $p = 0.0019$).
Bainbridge et al. 1988 [44] [Case Control]	**Hemoglobinopathies** [Variant Hemoglobins]	**7 of 9**	**Matched** for gestational age, birthweight, growth adequacy, race, sex, and birth year [Unadjusted Statistical Comparison]	The average peak TSB did not differ in neonates with Sickle Cell Disease (mean,SD: 145.35,61.58 umol/L) relative to healthy matched controls (mean,SD: 143.64,73.53 umol/L) ($p > 0.05$). The proportion of neonates experiencing TSB elevations in excess of 5mg/dL was also not significantly different (SCD: 48%; Controls: 51%; $p > 0.05$).
Erdol et al. 2018 [40] [Retrospective Cohort]	**Homocystinuria and PA/MMA** [Homocysteine] [Methylmalonic Acid]	**7 of 9**	**Restricted** to term infants at elevated risk for PKU or BIOT [Unadjusted Statistical Comparison]	No significant correlation was found between increased homocysteine and serum total, direct, or indirect bilirubin levels ($p > 0.05$). Infants divided in to two groups based on plasma homocysteine <> 10 uMol/L did not differ in total, direct, or indirect bilirubin levels ($p > 0.05$). No significant correlation identified between urinary methylmalonic acid levels and serum total, direct, or indirect bilirubin ($p > 0.05$).
Eyada et al. 2017 [45] [Case Control]	**SCID** [CD4+ & CD8+ T-Cells]	**6 of 9**	**Restricted** to exclude preterm, septic, transfused, hypoxic infants or infants with congenital anomalies **Matched** for gestational age [Unadjusted Statistical Comparison]	Prior to exposure to phototherapy, severely hyperbilirubinemic infants did not differ in their proportions of CD4+ (47.9%) or CD8+ (16.2%) lymphocytes relative to the CD4+ (49.8%) or CD8+ (17.3%) lymphocyte proportions of healthy controls ($p > 0.05$).
Gong et al. 2014 [42] [Case Control]	**Aminoacidopathies** **Fatty Acid Oxidation Disorders** **Organic Acid Conditions** [Amino Acids and Acylcarnitines]	**7 of 9**	Unreported Strategy [Unadjusted Statistical Comparison]	In patients with NICCD or BA, Cit, Arg, Met, Orn, C0, C2, C3, C6, C14, C16, C18, C18:2 acylcarnitines were higher than those in the control ($p < 0.05$), Ala and Gly were lower than those in the control (p 0.05). In patients with NICCD, the blood concentrations of Cit, Arg, Orn, Met, Tyr, and C18:2 were higher than those of patients with BA ($p < 0.002$).
Kanai et al. 2003 [46] [Cross-Sectional]	**Hemoglobinopathies** [Variant Hemoglobins]	**8 of 10**	**Restricted** to exclude BW <2500g, GA <36 wk, hemolytic anemia, neonatal asphyxia, maternal diabetes, CHD, intestinal malformations, infections, drug administration, TPN administration [Unadjusted Statistical Comparison]	Infants with higher HbF% trended toward having elevated TCBR in the first week of life, although statistically significant differences between the three strata were not observed. In pairwise comparisons, the High HbF% group had significantly higher TCBR readings than the Low HbF% group on postnatal days 4 and 5. Infants with moderate HbF% had significantly lower TCBR readings on the fifth day of life ($p < 0.05$).

➔ This table continues on the next two pages. Abbreviation are defined on the final page of this table.

Table 3 (Continued). Summary of Main Findings

Reference [Design]	Relevant Disorder [Marker]	Quality Score	Confounder Management [Statistical Adjustment]	Main Findings
Kurt et al. 2009 [47] [Case Control]	**SCID** [CD3+, CD4+. CD8+ T-Cells]	**6 of 9**	**Restricted** to exclude congenital malformations, TORCH infection, hypoxia, RD, HDN, transfusion, dBil > 2mg/dl **Matched** for GA, BW, and postnatal age [Unadjusted Statistical Comparison]	Relative to healthy newborns, neonates with severe hyperbilirubinemia (indirect bilirubin > 15 mg/dl) did not differ in their CD3+, CD4+ or CD8+ lymphocytes (% or cells/mm3) measured pre-phototherapy with flow cytometric immunophenotyping.
Kvetny et al. 2006 [33] [Case Control]	**Congenital Hypothyroidism** [TSH, T3, T4]	**8 of 9**	**Restricted** to euthyroid term neonates of mothers with thyroid disease **Matched** by birthweight and gestational age [Unadjusted Statistical Comparison]	The incidence of severe hyperbilirubinemia (> 300 umol/L) tended to be higher in neonates of women with autoimmune thyroid disease (4 of 19) relative to neonates born to mothers with non-autoimmune thyroid disease (0 of 12). This difference was not statistically significant.
Lee et al. 2006 [48] [Case Control]	**Fatty Acid Oxidation Disorders** **Organic Acid Conditions** [Acylcarnitines]	**8 of 9**	**Matched** for postnatal age [Unadjusted Statistical Comparison]	Relative to age-matched healthy controls, infants with neonatal intrahepatic cholestasis due to citrin deficiency had significantly elevated C0, C2, C3, C10, C12, C14, C16, C18, and C18:1 acylcarnitines. No significant differences in medium-chain acylcarnitines were observed.
Mauvais et al. 2016 [38] [Case Control]	**CH and CAH** [Cortisol and FT4]	**9 of 9**	**Matched** for PSIS status and age at diagnosis [Unadjusted Statistical Comparison]	Among 16 infants with Pituitary Stalk Interruption Syndrome, those who developed cholestasis (5 of 16) had significantly decreased cortisol levels (12.4 ng/mL) relative to the 11 infants who did not develop cholestasis (79.4 ng/mL). No significant difference was observed in FT4 (9.0 vs. 10.6 pmol/L)..
Mojtahedhi et al. 2018 [34] [Cross-Sectional]	**CH** [TSH and FT4]	**7 of 10**	Unreported Strategy [Unadjusted Statistical Comparison]	Hyperbilirubinemic neonates stratified by TSB differed in average TSH, with lowest TSH observed in the most hyperbilirubinemic infants (A: 4.96 mU/L; B: 3.88 mU/L; C: 1.80 mU/L; p = 0.003). No significant difference observed in FT4 (p = 0.10).
Schulpis et al. 2003 [49] [Cross-Sectional]	**Biotinidase Deficiency** [Biotinidase Activity]	**9 of 10**	**Stratified** by gestational age, growth adequacy, and bilirubin level [Unadjusted Statistical Comparison]	Moderate-to-moderately strong and significant inverse correlations were observed between TSB and biotinidase activity among term (r = -0.18 to -0.28, p < 0.001), preterm (r = -0.26 to -0.46, p < 0.01) and small-for-gestational age infants (r= -0.39 to -0.46, p < 0.01).

➔ This table continues on the next tpage. Abbreviation are defined on the next page.

Table 3 (Continued). Summary of Main Findings

Reference [Design]	Relevant Disorder [Marker]	Quality Score	Confounder Management [Statistical Adjustment]	Main Findings
Singh et al. 2003 [35] [Case Control]	**CH** [T3, T4, and TSH]	**7 of 9**	**Matched** for postnatal age, gestational age, birth weight [Unadjusted Statistical Comparison]	No significant correlations were observed between T3, T4, or TSH and bilirubin levels in hyperbilirubinemic infants nor healthy controls ($p > 0.01$).
Steinbach et al 2003 [39] [Case Control Nested in Randomized Trial]	**Aminoacidopathies** **Fatty Acid Oxidation Disorders** **Organic Acid Conditions** **CCHD** [Amino Acids, Acylcarnitines, and Confirmed Patent Ductus Arteriosus]	**7 of 9**	**Randomized** to TPN dosage **Restricted** to GA between 23.0 and 29.6 (Week.Days) **Matched** by age at dried blood spot collection [Unadjusted Statistical Comparison]	In a randomized trial of TPN dosages, 13 premature neonates who developed cholestasis had significantly elevated Cit, His, Met, and C4DC relative to those who did not become cholestatic ($p < 0.05$). Cholestatic infants had significantly decreased Glu, Ser, and thyroxine ($p < 0.05$). Neonates with cholestasis had a significantly higher prevalence of patent ductus arteriosus (76.9%) relative to non-cholestatic infants (42.4%) ($p = 0.04$).
Tan et al. 1996 [36] [Randomized Trial]	**CH** [TSH and FT4]	**Some Concern**	**Randomized** to phototherapy treatment group **Matched** for gestational age and birthweight [Unadjusted Statistical Comparison]	In a randomized trial of phototherapy light source efficacy, neonates randomized but not yet exposed to WL or FO phototherapies had significantly lower mean TSH (WL: 3.62 mU/L, FO: 3.86 mU/L) relative to healthy controls (5.77 mu/L) ($p < 0.05$). Neonates randomized to CP phototherapy had lower TSH (4.40 mU/L) than controls, but this was not statistically significant. Free T4 did not differ in any hyperbilirubinemic group relative to controls.
Ulanovsky et al 2018 [37] [Case Control]	**CH** [Total T4 (tT4)]	**8 of 9**	**Restricted** to exclude oxytocin infusion, preterm birth, hemolytic diseases, cephalohematoma, polycythemia, red blood cell membranopathies, exclusive breastfeeding **Stratified** by mode of delivery [Multivariable-Adjusted Comparison]	Among vaginally delivered neonates, tT4 above 13 ug/dL was not associated with increased risk of idiopathic hyperbilirubinemia (OR = 1.04, 95%CI: 0.61 - 1.73) relative to infants with tT4 below 13 ug/dL. Among caesarean delivered neonates, tT4 above 13 ug/dL was strongly associated with increased risk of idiopathic hyperbilirubinemia (OR= 5.49, 95%CI: 1.23 - 24.40) relative to caesarean-born neonates with tT4 below 13 ug/dL.

Alphabetized Abbreviations: AITD = Autoimmune thyroid disease **; Ala** = Alanine **; Arg** = Arginine **; Avg** = Average **; BA** = Biliary Atresia **; BIOT** = Biotinidase Deficiency **; BW** = Birthweight **; C4DC** = Succinylcarnitine **; CAH** = Congenital Adrenal Hyperplasia **; cBil** = Conjugated Bilirubin **; CCHD** = Critical Congenital Heart Disease **; CD*x*** = Cluster of Differentiation *x* **; CH** = Congenital Hypothyroidism **; CHD** = Congenital Heart Defects **; CI =** Confidence Interval; **Cit** = Citrulline **; CP** = Combined White Light and Fibre-Optic Phototherapy **; C*x*** = Acylcarnitine with *x* carbon atoms **; dBil** = Direct Bilirubin **; FO** = Fibre-Optic Phototherapy **; FT4** = Free Thyroxine **; GA** = Gestational Age **; Glu** = Glutamic Acid **; Gly** = Glycine **; HbF** = Fetal hemoglobin **; HBIL** = Hyperbilirubinemia **; HDN** = Hemolytic Disease of the Newborn **; His** = Histidine **; Met** = Methionine **; MMA** = Methylmalonic Acid **; NICCD** = Neonatal Intrahepatic Cholestasis due to Citrin Deficiency **; Non-AI TD** = Non-autoimmune thyroid disease **; OR =** Odds Ratio; **Orn** = Ornithine **; PA/MMA** = Propionic and Methylmalonic Acidemias **; PDA** = Patent Ductus Arteriosus **; PKU** = Phenylketonuria **; PSIS** = Pituitary Stalk Interruption Syndrome **; *r*** = Correlation Coefficient; **RD** = Respiratory Distress **; SCD** = Sickle Cell Disease **; SCID** = Severe Combined Immune Deficiency **; SD** = Standard Deviation **; Ser** = Serine **; T3** = Triiodothyronine **; T4** = Thyroxine **; TCBR =** Transcutaneous Bilirubinometry **; TORCH =** Toxoplasmosis, Other, Rubella, Cytomegalovirus **; TPN =** Total Parenteral Nutrition **; TSB** = Total Serum Bilirubin **; TSH =** Thyroid Stimulating Hormone **; tT4 =** Total Thyroxine; **Wk**= Weeks ; **WL** = White Light Phototherapy

Kanai et al.'s [46] Japan-based study team compared bilirubin readings measured transcutaneously (TCBR) among infants with high (>90%), moderate (80%-90%), and low (<80%) fractions of fetal hemoglobin (HbF%). Neonates with high HbF (>90%) had significantly higher TCBR readings on the fourth day of life relative to neonates with low HbF (<80%). The high HbF group had significantly elevated TCBR relative to the moderate HbF (80%-90%) on the fifth postnatal day but did not significantly differ in TCBR relative to the low HbF group. No significant differences in TCBR readings among the three strata of HbF were observed on postnatal days 0-3.

A 1.4.6 Studies Investigating Bilirubin Association with Markers of Immune Disease

Two studies were identified by this review associating markers of inherited immune disease with hyperbilirubinemia [45,47]. Both studies quantified lymphocyte subsets in hyperbilirubinemic neonates using flow cytometry and neither measured T-cell receptor excision circles (TRECs). Study teams based in Turkey (Kurt et al.[47]) and Egypt (Eyada et al. [45]) published their studies in 2009 and 2016, respectively. Both studies were case-control designs primarily probing for effects of phototherapy on neonatal immunity. Eyada et al.[45] immunophenotyped 21 hyperbilirubinemic neonates and 16 healthy controls; Kurt et al.[47] compared 25 hyperbilirubinemic neonates to 25 healthy controls.

Kurt et al.[47] reported that CD3+, CD4+ and CD8+ lymphocyte volumes (cells/mm^3) and fractions (%) were statistically comparable in term hyperbilirubinemic neonates pre-phototherapy relative to healthy controls. Eyada et al.[45] similarly concluded that CD4+ and CD8+ lymphocyte fractions (%) did not differ between pre-phototherapy hyperbilirubinemic neonates and healthy controls. In a secondary analysis, Eyada et al. [45] reported no statistically significant correlation between continuous total serum bilirubin level and CD4+% or CD8+% lymphocytes.

A 1.4.7 Studies Investigating Bilirubin in Association with Congenital Heart Disease

In a 2012 study of 206 infants with Trisomy 21, Sweden-based Arnell et al.[43] observed a significantly increased rate of conjugated hyperbilirubinemia among infants with comorbid congenital heart defects relative to those without heart defects [43]. The retrospective cohort study classified isolated atrial septal defects, ventricular septal defects, patent ductus arteriosus (PDA), atrioventricular septal defects, and tetralogy of Fallot as congenital heart defects. These findings in infants with Trisomy 21 mirrored Steinbach et al.'s [39] 2003 report observing a significantly

higher rate of PDA in 13 cholestatic infants receiving TPN compared to 109 infants who did not develop cholestasis [39].

A 1.4.8 Studies Investigating Bilirubin Association with Markers of Biotinidase Deficiency

In this review, one included cross-sectional study reported an association between neonatal bilirubin and serum biotinidase activity [49]. Schulpis et al. studied 1,658 term, premature, and small-for-gestational-age (SGA) neonates and reported statistically significant inverse correlations between total serum bilirubin levels (TSB) and biotinidase enzyme activity. Participants in the gestational age-based study groups were further stratified based on TSB levels and neonates in lower-TSB strata had significantly higher serum biotinidase activities than infants in higher-TSB strata.

A 1.4.9 Quality Assessments

i. Non-Randomized Studies

Fifteen of 16 included studies were deemed observational at the quality assessment stage of this review. Although Steinbach et al.'s [39] participants were randomized, the included study presented a nested case control analysis and was evaluated using a tool intended for observational designs. Quality scores derived using the Newcastle-Ottawa Scale are available in Table 4.

Case control studies overall ranged in quality from moderate to high. Most case control studies achieved the maximum quality score in the exposure domain (3/3) which indicated a low risk of bias. Two studies had unexplained differences in the rates of non-responding participants between cases and controls, reducing their scores on this domain [39,44]. The majority (6 of 10) of case control studies ensured study comparability by adjusting their association for at least two confounding factors, while three studies only controlled for a single confounder [45,47,48] and confounding control was unclear in one study [42]. In the selection bias domain, only three of 10 case control studies achieved the maximum score [35,45,47]. Studies had selection domain scores reduced for relying on hospitalized infants as controls (5 studies)[33,37,39,44,45], incomplete descriptions of the source of cases (3 studies)[35,45,47] and incomplete descriptions of the source of controls (2 studies)[35,47].

Table 4. Risk of Bias Appraisals of Included Observational Studies

Reference	Domain: Selection	Comparability	Exposure	Total Score
Case Control Studies	Maximum: 4	Maximum: 2	Maximum: 3	Maximum: 9
Bainbridge et al. 1988	★★★	★★	★★	7
Eyada et al. 2017	★★	★	★★★	6
Gong et al. 2014	★★★★		★★★	7
Kurt et al. 2009	★★	★	★★★	6
Kvetny et al. 2006	★★★	★★	★★★	8
Lee et al. 2006	★★★★	★	★★★	8
Mauvais et al. 2016	★★★★	★★	★★★	9
Singh et al. 2003	★★	★★	★★★	7
Steinbach et al. 2003	★★★	★★	★★	7
Ulanovsky et al. 2018	★★★	★★	★★★	8
Cohort Studies	Maximum: 4	Maximum: 2	Maximum: 3	Maximum: 9
Arnell et al. 2012	★★★★	★★	★★★	9
Erdol et al. 2018	★★	★★	★★★	7
Cross-Sectional Studies	Maximum: 5	Maximum: 2	Maximum: 3	Maximum: 10
Kanai et al. 2003	★★★★★	★	★★★	8
Mojtahedi et al. 2018	★★★★		★★★	7
Schulpis et al. 2003	★★★★	★★	★★★	9

The two included cohort studies were moderate to high in quality. Arnell et al.[43] achieved the highest possible overall score, indicating a low risk of bias in selection, comparability, and exposure domains. Similarly, Erdol et al. [40] achieved the top possible scores for the comparability and exposure domains. The latter study had a lowered selection score due to using a cohort subset for analysis without additional explanation and unclear ascertainment of whether participants were already hyperbilirubinemic at the study outset [40].

Three cross-sectional studies were assessed to have moderate-to-high quality [34,46,49]. All studies rated favorably in measurement of the exposure. Two studies were downrated in the comparability domain due to concerns about the selection of the control group [34,46] and unreported

control of confounding[34]. Cross-sectional studies were at low risk of selection-related bias, although two studies were downrated due to incomplete reporting of their sample selection strategies [34,49].

ii. Randomized Studies

One randomized study was included in this review [36]. Tan et al. achieved a composite score of 'some concerns' when assessed using the Cochrane RoB 2.0 tool [32]. The trial was at low risk of bias in the outcome measurement and missing outcome data domains, but some concerns were noted in the randomization and assignment-to-intervention domains. These concerns were due to the sequential nature of group allocation and incomplete reporting as to whether parents were blinded to their child's assigned intervention.

A 1.5 Discussion

This review sought to assess the current status of the published scientific literature regarding potential associations between biomarkers indicative of rare disease and neonatal bilirubin. Our study identified 16 studies investigating a range of potential associations between infant bilirubin and the biomarkers used to screen neonates for non-hearing related diseases on the Recommended Uniform Screening Panel (RUSP) [4]. These studies were mostly observational in design and most were assessed to be at lower-to-moderate risk of bias. However, many included studies were not adjusted for important confounders or designed to establish directionality, making causal arguments challenging. Despite substantial variety in study-level conclusions, the identified studies support the finding that certain indicative biomarkers used to screen neonates for rare disease may be altered in association with infant bilirubin levels or clinical hyperbilirubinemia status. To our knowledge, this is the first systematic review establishing hyperbilirubinemia as potentially related to the markers underlying newborn screening tests.

The largest subgroup of studies identified in this systematic review suggested an association between bilirubin and markers used to identify neonatal endocrinopathies, including TSH, T3, and T4 and cortisol. Although the direct relationship between the endocrinopathy, congenital hypothyroidism (CH), and hyperbilirubinemia has been previously established [50,51], primary or non-central forms of CH are characterized by TSH elevation [52] yet none of three studies that identifyied an inverse associations between TSH and bilirubin reported the inclusion of neonates

with central CH, which is typically characterized by lower TSH levels. Although it is possible that undiagnosed infants with central CH were included in these studies, the low prevalence of central CH makes it more plausible that there is an association between TSH and neonatal bilirubin level even in the absence of thyroid disease. A majority of studies investigating T4 reported no association with bilirubin, although the most recent and only confounder-adjusted study identified a strong positive association with hyperthyroxinemia in a small group of cesearean-born neonates. In the sole identified study investigating bilirubin in relation to cortisol [38], an inverse association was found with conjugated hyperbilirubinemia but the study only included participants with a rare endocrine disorder.

Studies of amino acid markers among hyperbilirubinemic infants provided a diversity of findings; these findings are difficult to extrapolate to non-diseased populations, given that in two of the three studies, participants were included based on TPN administration which is known to alter amino acid levels in non-diseased infants [39], diagnosed NICCD or BA[42], or following a positive screening result for a metabolic disorder [40]. Consistent elevations in citrulline and methionine were reported by two studies enrolling infants with conjugated hyperbilirubinemia. Alterations in the amino acid panel have been recently proposed as a potential non-invasive diagnostic for biliary atresia and other liver pathologies [53], and our findings concur that these markers may be altered in association to hyperbilirubinemia.

Findings from studies examining markers of fatty acid oxidation or organic acid disorders were similarly heterogenous, however there was some consistency between two identified studies performing acylcarnitine panels. Both Lee et al.[48] and Gong et al.[42] reported elevated shorter-chain and longer-chain acylcarnitine species (chain lengths below C8 and above C12) among neonates with hyperbilirubinemia secondary to liver disease, but no such association with mid-chain length acylcarnitines. Although Steinbach et al.[39] did not reproduce these findings in a population without NICCD or BA, the participants in that trial were born extremely preterm and received amino acid-containing TPN, potentially altering their amino acid profile independent of any disease status [54].

Our systematic review also identified a small number of additional studies with relevance to other groups of diseases targeted in newborn screening. With respect to hemoglobinopathies, Bainbridge et al.[44] found no association between maximal TSB among neonates with confirmed Hemoglobin SS disease, which is in agreement with previous findings suggesting that erythrocyte

sickling and lysis is uncommon in the neonatal period [55]. Kanai et al.'s [46] finding of association between fetal hemoglobin and transcutaneous bilirubin level is concordant with recent evidence identifying hemoglobin F as a predictor of gestational maturity [56], with premature neonates also being at higher risk of hyperbilirubinemia. Schulpis et al.[49] identified strong relationships between bilirubin levels and biotinidase enzyme activity.

With respect to immune deficiency, two independent flow-cytometric studies did not identify T-lymphocyte differences in hyperbilirubinemic neonates relative to controls, suggesting that hyperbilirubinemia may not be associated with screening markers of T-lymphopenia [45,47]. We note that inclusion of flow-cytometric T-cell studies were a second deviation from our pre-published protocol[23] in that these are not screening but rather diagnostic markers of Severe Combined Immune Deficiency. The study team reasoned that T-Cell Receptor Excision Circles (TREC) were screening markers used to approximate actual T-cells as measured by flow cytometry and therefore elected to include these studies.

We identified two studies associating congenital heart defects (CHD) with increased risk of hyperbilirubinemia [39,43]. There are multiple plausible biological mechanisms for this link, with the most direct being hypoxic liver injury due to hypoperfusion or venous congestion of the liver [57–59]. Polycythemia and abnormalities of coagulation have been reported in pediatric patients with hypoxemic CHD [60], and these are both recognized pre-hepatic causes of hyperbilirubinemia. Although our review identified studies associating hyperbilirubinemia with CHD, no evidence was found that the screening marker for this disease (SpO_2) was distributed differentially among hyperbilirubinemic neonates.

This review identified a limited number of studies (16), and several challenges exist when synthesizing the findings as a uniform whole. We note that the wide range of analytic techniques represented among these studies, and variability in the source biofluids may pose a substantial limitation to generalizing these associations to the dried blood spots used in most population-wide newborn screening programs [67]. The wide range of participant ages among identified studies also limits a unified interpretation of these findings, as rapid biochemical changes in metabolic markers have been extensively documented in the first year of life [68]. It may also have been a potential source of meta-bias to exclude non-English references, which resulted in the exclusion of 40 full-text articles. Although published evidence is available that English-language randomized controlled trials are also more likely to be significant [69], 15 of 16 studies in this

review were observational in design and more recent studies have been unable to identify evidence of a language-related meta-bias [70]. Lastly, the race/ethnicity of participants in the reviewed studies was not extracted.

The strengths of this review include the provision of an encompassing perspective of the documented associations between infant bilirubin and biomarkers associated with a comprehensive panel of rare diseases typically incorporated in population-based newborn screening programs. Our reliance on a national-standard disorder panel [4] as the starting point for a search strategy will further assist in ensuring our findings are relevant to newborn screening stakeholders across North America in particular. Given a vast and growing body of literature on neonatal hyperbilirubinemia, querying the evidence for the few studies with information applicable to rare disease screening can be a challenging task. Synthesizing the current state of the literature removes this barrier, enabling additional research on the association between bilirubin and neonatal screening biomarkers to proceed more efficiently.

The identification of neonatal attributes, demographic factors, environmental influences, and extant comorbidities at the time of screening sampling that may systematically alter the concentration of disease-indicative biomarkers is an effort to extend the most accurate possible screening to every neonate. Upon recognition of influential factors that alter screening performance, newborn screening laboratories have historically implemented additional assays [72], stratified risk thresholds [73], and international quality assurance programs [71] that function to reduce heterogeneous screening error. These responses rely on surveillance for these sources of influence, as included in this review.

Conclusions

Findings from our systematic review suggest that the current status of the published literature associating infant bilirubin with biomarkers of rare disease remains at an early stage of development. Among studies available, there is some evidence of association, but several studies found no significant relationships. Our results provide early indication that bilirubin level or hyperbilirubinemia status may associate with markers of endocrine function, short-chain and long-chain acylcarnitines, biotinidase activity, fetal hemoglobin, and the presence of congenital heart disease. The evidence did not strongly support associations of T-cells and homocysteine with infant bilirubin. Cumulatively, we find the evidence sufficient to justify a formal study associating neonatal bilirubin and the biomarkers used to screen newborns for rare disease.

A 1.6 References

1. Marsden D, Larson C, Levy HL. Newborn screening for metabolic disorders. *J Pediatr*. 2006;148(5):577-584. doi:10.1016/j.jpeds.2005.12.021

2. Thibodeau DI. Newborn Screening Systems. *Adv Neonatal Care*. 2011;11(1):71-72. doi:10.1097/ANC.0b013e318207953d

3. Kemper AR, Green NS, Calonge N, et al. Decision-making process for conditions nominated to the recommended uniform screening panel: Statement of the US department of health and human services secretary's advisory committee on heritable disorders in newborns and children. *Genet Med*. 2014;16(2):183-187. doi:10.1038/gim.2013.98

4. Secretary's Advisory Committee on Heritable Disorders in Newborns and Children (SACHDNC). Recommended Uniform Screening Panel [July 2018 Version]. https://www.hrsa.gov/advisory-committees/heritable-disorders/rusp/index.html. Published 2019. Accessed December 2, 2019.

5. Mak CM, Lee H-CH, Chan AY-W, Lam C-W. Inborn errors of metabolism and expanded newborn screening: review and update. *Crit Rev Clin Lab Sci*. 2013;50(6):142-162. doi:10.3109/10408363.2013.847896

6. Varness TS, Allen DB, Hoffman GL. Newborn screening for congenital adrenal hyperplasia has reduced sensitivity in girls. *J Pediatr*. 2005;147(4):493-498. doi:10.1016/j.jpeds.2005.04.035

7. Hustace T, Fleisher JM, Sanchez Varela AM, Podda A, Alvarez O. Increased prevalence of false positive hemoglobinopathy newborn screening in premature infants. *Pediatr Blood Cancer*. 2011;57(6):1039-1043. doi:10.1002/pbc.23173

8. Slaughter JL, Meinzen-Derr J, Rose SR, et al. The Effects of Gestational Age and Birth Weight on False-Positive Newborn-Screening Rates. *Pediatrics*. 2010;126(5):910-916. doi:10.1542/peds.2010-0943

9. Minutti CZ, Lacey JM, Magera MJ, et al. Steroid profiling by tandem mass spectrometry improves the positive predictive value of newborn screening for congenital adrenal hyperplasia. *J Clin Endocrinol Metab*. 2004;89(8):3687-3693. doi:10.1210/jc.2003-032235

10. Tim-Aroon T, Harmon HM, Nock ML, Viswanathan SK, McCandless SE. Stopping Parenteral Nutrition for 3 Hours Reduces False Positives in Newborn Screening. *J Pediatr*. 2015;167(2):312-316. doi:10.1016/j.jpeds.2015.04.063

11. Cheillan D, Vercherat M, Chevalier-Porst F, Charcosset M, Rolland MO, Dorche C. False-positive results in neonatal screening for cystic fibrosis based on a three-stage protocol (IRT/DNA/IRT): Should we adjust IRT cut-off to ethnic origin? *J Inherit Metab Dis*. 2005;28(6):813-818. doi:10.1007/s10545-005-0067-0

12. Morrison DR, Clayton EW. False positive newborn screening results are not always benign. *Public Health Genomics*. 2011;14(3):173-177. doi:10.1159/000322527

13. Karaceper MD, Chakraborty P, Coyle D, et al. The health system impact of false positive newborn screening results for medium-chain acyl-CoA dehydrogenase deficiency: A cohort study. *Orphanet J Rare Dis*. 2016;11(1):1-9. doi:10.1186/s13023-016-0391-5

14. Schreiner F, Brack C, Salzgeber K, Vorhoff W, Woelfle J, Gohlke B. False negative 17-hydroxyprogesterone screening in children with classical congenital adrenal hyperplasia. *Eur J Pediatr*. 2008;167(4):479-481. doi:10.1007/s00431-007-0505-0

15. Usher-Smith JA, Sharp SJ, Griffin SJ. The spectrum effect in tests for risk prediction, screening, and diagnosis. *BMJ*. 2016;353. doi:10.1136/bmj.i3139

16. Mulherin SA, Miller WC. Academia and Clinic Spectrum Bias or Spectrum Effect ? Subgroup Variation in Diagnostic. *Ann Intern Med*. 2002;137:598-602.

17. Elie C, Coste J. A methodological framework to distinguish spectrum effects from spectrum biases and to assess diagnostic and screening test accuracy for patient populations: Application to the Papanicolaou cervical cancer smear test. *BMC Med Res Methodol*. 2008;8:1-8. doi:10.1186/1471-2288-8-7

18. Ryckman KK, Berberich SL, Shchelochkov OA, Cook DE, Murray JC. Clinical and environmental influences on metabolic biomarkers collected for newborn screening. *Clin Biochem*. 2013;46(1-2):133-138. doi:10.1016/j.clinbiochem.2012.09.013

19. Kumar RK. Neonatal jaundice. An update for family physicians. *Aust Fam Physician*. 1999;28(7):679-682. http://www.ncbi.nlm.nih.gov/pubmed/10431423.

20. Almeida MFB De. Quando devemos iniciar a fototerapia. 2004;80:256-258.

21. Okwundu CI, Okoromah CA, Shah PS. Cochrane Review: Prophylactic phototherapy for preventing jaundice in preterm or low birth weight infants. *Evidence-Based Child Heal A Cochrane Rev J*. 2013;8(1):204-249. doi:10.1002/ebch.1898

22. McCandless DW. *Kernicterus*. Totowa, NJ: Humana Press; 2011. doi:10.1007/978-1-4419-6555-4

23. Santander E, Tao J, Sikora L, Hawken S, Potter B. Associations between infant bilirubin and neonatal screning biomarkers: a systematic review protocol. PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019123350. Published 2019. Accessed December 2, 2019.

24. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097

25. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: Elaboration and explanation. *BMJ*. 2015;349(January):1-25. doi:10.1136/bmj.g7647

26. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4:1. doi:10.1186/2046-4053-4-1

27. Veritas Health Innovation. Covidence Software. www.covidence.org.

28. Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: the evolution of a rapid review approach. *Syst Rev*. 2012;1(10):10. doi:10.1186/2046-4053-1-10

29. McHugh ML. Interrater reliability: the kappa statistic. *Biochem medica*. 2012;22(3):276-282. http://www.ncbi.nlm.nih.gov/pubmed/23092060.

30. Wells GA, Shea B, O'Connell D, Peterson J, Welch V et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed February 26, 2019.

31. Modesti PA, Reboldi G, Cappuccio FP, et al. Panethnic Differences in Blood Pressure in Europe: A Systematic Review and Meta-Analysis. *PLoS One*. 2016;11(1):e0147601. doi:10.1371/journal.pone.0147601

32. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj*. 2019:l4898. doi:10.1136/bmj.l4898

33. Kvetny J, Poulsen H. Transient hyperthyroxinemia in newborns from women with autoimmune thyroid disease and raised levels of thyroid peroxidase antibodies. *J Matern Neonatal Med*. 2006;19(12):809-814. doi:10.1080/14767050600927304

34. Mojtahedi SY, Izadi A, Seirafi G, Khedmat L, Tavakolizadeh R. Risk factors associated with neonatal jaundice: a cross-sectional study from iran. *Open Access Maced J Med Sci*. 2018;6(8):1387-1393. doi:10.3889/oamjms.2018.319

35. Singh B, Ezhilarasan R, Kumar P, Narang A. Neonatal hyperbilirubinemia and its association with thyroid hormone levels and urinary iodine excretion. *Indian J Pediatr*. 2003;70(4):311-315. doi:10.1007/BF02723587

36. Tan KL, Chirino-Barcelo Y, Aw TC, Joseph R. Effect of phototherapy on thyroid stimulatory hormone and free thyroxine levels. *J Paediatr Child Health*. 1996;32(6):508-511. doi:10.1111/j.1440-1754.1996.tb00964.x

37. Ulanovsky I, Smolkin T, Almashanu S, Mashiach T, Makhoul IR. Hyperthyroxinemia at birth: a cause of idiopathic neonatal hyperbilirubinemia? *World J Pediatr*. 2018;14(3):247-253. doi:10.1007/s12519-017-0113-7

38. Mauvais FX, Gonzales E, Davit-Spraul A, Jacquemin E, Brauner R. Cholestasis reveals severe cortisol deficiency in neonatal pituitary stalk interruption syndrome. *PLoS One*. 2016;11(2):1-12. doi:10.1371/journal.pone.0147750

39. Steinbach M, Clark RH, Kelleher AS, et al. Demographic and nutritional factors associated with prolonged cholestatic jaundice in the premature infant. *J Perinatol*. 2008;28(2):129-135. doi:10.1038/sj.jp.7211889

40. Erdol S, Ozgur T. Vitamin B12 deficiency associated with hyperbilirubinemia and cholestasis in infants. *Pakistan J Med Sci*. 2018;34(3):712-718. doi:10.12669/pjms.343.14564

41. Gong Z, Xu W-J, Tian G-L, Zhang T, Lv Z. Neonatal Intrahepatic Cholestasis Caused by Citrin Deficiency Differentiated from Biliary Atresia. *Eur J Pediatr Surg*. 2016;26(3):255-259. doi:10.1055/s-0035-1551566

42. Gong Z, Xu WJ, Tian GL, Zhang T, Lv Z. Neonatal Intrahepatic Cholestasis Caused by Citrin Deficiency Differentiated from Biliary Atresia. *Eur J Pediatr Surg*. 2016;26(3):255-259. doi:10.1055/s-0035-1551566

43. Arnell H, Fischler B. Population-based study of incidence and clinical outcome of neonatal cholestasis in patients with down syndrome. *J Pediatr*. 2012;161(5):899-902. doi:10.1016/j.jpeds.2012.04.037

44. Bainbridge R. Jaundice in Neonates With Sickle Cell Disease. *Am J Dis Child*. 1988;142(5):569. doi:10.1001/archpedi.1988.02150050107045

45. Eyada IK, El Saie AL, Ibrahem GA, Riad NM. Effect of phototherapy on B and T lymphocytes in Egyptian infants suffering from neonatal jaundice. *Allergol Immunopathol (Madr)*. 2017;45(3):290-296. doi:10.1016/j.aller.2016.09.003

46. Kanai M, Akaba K, Sasaki A, et al. Neonatal hyperbilirubinemia in Japanese neonates: Analysis of the heme oxygenase-1 gene and fetal hemoglobin composition in cord blood. *Pediatr Res*. 2003;54(2):165-171. doi:10.1203/01.PDR.0000072329.56635.35

47. Kurt A, Aygun AD, Kurt ANC, Godekmerdan A, Akarsu S, Yilmaz E. Use of phototherapy for neonatal hyperbilirubinemia affects cytokine production and lymphocyte subsets. *Neonatology*. 2009;95(3):262-266. doi:10.1159/000171216

48. Lee NC, Chien YH, Kobayashi K, et al. Time course of acylcarnitine elevation in neonatal intrahepatic cholestasis caused by citrin deficiency. *J Inherit Metab Dis*. 2006;29(4):551-555. doi:10.1007/s10545-006-0250-y

49. Schulpis KH, Gavrili S, Tjamouranis J, Karikas GA, Kapiki A, Costalos C. The effect of neonatal jaundice on biotinidase activity. *Early Hum Dev*. 2003;72(1):15-24. doi:10.1016/S0378-3782(02)00097-X

50. Tiker F, Gurakan B, Tarcan A, Kinik S. Congenital Hypothyroidism and Early Severe Hyperbilirubinemia. *Clin Pediatr (Phila)*. 2003;42(4):365-366. doi:10.1177/000992280304200411

51. Benchimol EI, Walsh CM, Ling SC. Early diagnosis of neonatal cholestatic jaundice: test at 2 weeks. *Can Fam Physician*. 2009;55(12):1184-1192. doi:10.1186/cc10021

52. Rose SR, Brown RS, Foley T, et al. Update of newborn screening and therapy for congenital hypothyroidism. *Pediatrics*. 2006;117(6):2290-2303. doi:10.1542/peds.2006-0915

53. Zhao D, Han L, He Z, Zhang J, Zhang Y. Identification of the plasma metabolomics as early diagnostic markers between biliary atresia and neonatal hepatitis syndrome. *PLoS One*. 2014;9(1). doi:10.1371/journal.pone.0085694

54. Gucciardi A, Zaramella P, Costa I, et al. Analysis and interpretation of acylcarnitine profiles in dried blood spot and plasma of preterm and full-term newborns. *Pediatr Res*. 2015;77(1):36-47. doi:10.1038/pr.2014.142

55. Hegyi T, Delphin ES, Bank A, Polin RA, Blanc WA. Sickle cell anemia in the newborn. *Pediatrics*. 1977;60(2):213-216. doi:10.1016/B978-0-7234-3571-6.00034-8

56. Wilson K, Hawken S, Potter BK, et al. Accurate prediction of gestational age using newborn screening analyte data. *Am J Obstet Gynecol*. 2016;214(4):513.e1-513.e9. doi:10.1016/j.ajog.2015.10.017

57. Balistreri WF. PEDIATRIC HEPATOLOGY. *Clin Liver Dis*. 2000;4(1):191-210. doi:10.1016/S1089-3261(05)70103-3

58. Weinberg AG. The Liver in Congenital Heart Disease. *Am J Dis Child*. 1970;119(5):390. doi:10.1001/archpedi.1970.02100050392002

59. Asrani SK, Asrani NS, Freese DK, et al. Congenital heart disease and the liver. *Hepatology*. 2012;56(3):1160-1169. doi:10.1002/hep.25692

60. Tempe DK, Virmani S. Coagulation abnormalities in patients with cyanotic congenital heart disease. *J Cardiothorac Vasc Anesth*. 2002;16(6):752-765. doi:10.1053/jcan.2002.128436

61. Milne AE. Hemolytic Anemia: An Unusual Presentation of Congenital Heart Disease. *Arch Pediatr Adolesc Med*. 1996;150(1):109. doi:10.1001/archpedi.1996.02170260113022

62. Ravenel SD, Johnson JD, Sigler AT. Intravascular hemolysis associated with coarctation of the aorta. *J Pediatr*. 1969;75(1):67-73. doi:10.1016/S0022-3476(69)80102-2

63. Kaplan M, Muraca M, Hammerman C, et al. Imbalance between production and conjugation of bilirubin: a fundamental concept in the mechanism of neonatal jaundice. *Pediatrics*. 2002;110(4):47-51. doi:10.1542/peds.110.4.e47

64. Barker SJ, Curry J, Redford D, Morgan S. Measurement of carboxyhemoglobin and methemoglobin by pulse oximetry: A human volunteer study. *Anesthesiology*. 2006;105(5):892-897. doi:10.1097/00000542-200611000-00008

65. Fouzas S, Priftis KN, Anthracopoulos MB. Pulse oximetry in pediatric practice. *Pediatrics*. 2011;128(4):740-752. doi:10.1542/peds.2011-0271

66. Murphy D, Pak Y, Cleary JP. Pulse Oximetry Overestimates Oxyhemoglobin in Neonates with Critical Congenital Heart Disease. *Neonatology*. 2016;109(3):213-218. doi:10.1159/000442379

67. White E. Measurement error in biomarkers: sources, assessment, and impact on studies. *IARC Sci Publ*. 2011;(163):143-161.

68. Teodoro-Morrison T, Kyriakopoulou L, Chen YK, et al. Dynamic biological changes in metabolic disease biomarkers in childhood and adolescence: A CALIPER study of healthy community children. *Clin Biochem*. 2015;48(13-14):828-836. doi:10.1016/j.clinbiochem.2015.05.005

69. Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet*. 1997;350(9074):326-329. doi:10.1016/S0140-6736(97)02419-7

70. Morrison A, Polisena J, Husereau D, et al. The effect of english-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *Int J Technol Assess Health Care*. 2012;28(2):138-144. doi:10.1017/S0266462312000086

71. McHugh DMS, Cameron CA, Abdenur JE, et al. Clinical validation of cutoff target ranges in newborn screening of metabolic disorders by tandem mass spectrometry: A worldwide collaborative project. *Genet Med*. 2011;13(3):230-254. doi:10.1097/GIM.0b013e31820d5e67

72. Minutti CZ, Lacey JM, Magera MJ, et al. Steroid profiling by tandem mass spectrometry improves the positive predictive value of newborn screening for congenital adrenal hyperplasia. *J Clin Endocrinol Metab*. 2004;89(8):3687-3693. doi:10.1210/jc.2003-032235

73. Hayashi GY, Carvalho DF, de Miranda MC, et al. Neonatal 17-hydroxyprogesterone levels adjusted according to age at sample collection and birthweight improve the efficacy of congenital adrenal hyperplasia newborn screening. *Clin Endocrinol (Oxf)*. 2017;86(4):480-487. doi:10.1111/cen.13292

A 1.7 Supplementary Materials

Table S1. Database Search Strategy

ID	Query
1	exp Infant/
2	(infant* or newborn* or neonat* or baby or babies).tw.
3	1 or 2
4	HYPERBILIRUBINEMIA, NEONATAL/ or HYPERBILIRUBINEMIA/ or HYPERBILIRUBINEMIA, HEREDITARY/
5	CHOLESTASIS, EXTRAHEPATIC/ or CHOLESTASIS/ or CHOLESTASIS, INTRAHEPATIC/
6	((neonat* adj1 jaundice) or hyperbilirubin* or hypobilirubin* or eubilirubin* or cholestas*).tw.
7	(kernicterus or icterus or icter*).tw.
8	4 or 5 or 6 or 7
9	Biomarkers/
10	exp inorganic chemicals/ or exp organic chemicals/ or exp heterocyclic compounds/ or exp polycyclic compounds/ or exp "hormones, hormone substitutes, and hormone antagonists"/ or exp "enzymes and coenzymes"/ or exp carbohydrates/ or exp lipids/ or exp "amino acids, peptides, and proteins"/ or exp "nucleic acids, nucleotides, and nucleosides"/ or exp biological factors/
11	exp "diagnostic techniques and procedures"/ or exp investigative techniques/
12	((bio* adj1 factor) or (bio* adj1 parameter) or biomarkers or biomarker or "biological marker" or "biological markers" or (enzyme*) or (cytochrome*) or (coenzyme*)).tw.
13	((amino adj1 sugar*) or carbasugar* or (deoxy adj1 sugar*) or (dietary adj1 carbohydrate) or glycoconjugate* or glycoside* or (imino adj1 sugar) or polysaccharide* or (sugar adj1 acid*) or (sugar adj1 alcohol*) or (sugar adj1 phospate*) or thiosugar*).tw.
14	(ceroid or fat* or (fatty adj1 acid*) or (fatty adj1 alcohol*) or glyceride* or glycolipid* or (lipid adj1 peroxide*) or lipofuscin or lipopeptide* or lipopolysaccharide* or lipoprotein* or (membrane adj1 lipid*) or polyhydroxyalkanoate*).tw.
15	(arabinonucleoside* or deoxyribonucleoside or ribonucleoside* or thionucleoside* or arabinonucleotide* or deoxyribonucleotide* or dideoxynucleotide* or (dinucleoside adj1 phospate*) or polynucleotide or ribonucleotide or thionucleotide).tw.
16	(biol* adj2 (panel* or assay* or test* or analys*)).tw.
17	((amino adj1 aci*) or acylcarn* or (acyl adj1 carn*) or (organic adj2 acid*) or acylglyci* or (acyl adj1 glyci*) or purine*).tw.
18	((ketone adj1 body adj2 (panel* or assay* or test* or analys*)) or ((ketone or ketones) adj2 (panel* or assay* or test* or analys*))).tw.
19	((liver adj2 (panel* or assay* or test* or analys*)) or ((liver adj1 function) adj2 (panel* or assay* or test* or analys*))).tw.
20	((basic or comprehensive) adj2 metabolic adj2 (panel* or assay* or test* or analys*)).tw.
21	((chem* adj2 (panel* or assay* or test* or analys*))).tw.
22	((blood adj2 gas*) adj2 (panel* or assay* or test* or analys*)).tw.
23	(electrolyt* adj2 (panel* or assay* or test* or analys*)).tw.
24	((fatty adj1 acid*) or (Hydroxy adj1 fatty adj1 acid*)).tw.
25	(lipid* or immunophenotyp* or interleukin* or IL* or cytokin*).tw.
26	((immun* or muta* or sequen* or genotyp* or steroid* or thyroid* or adrenal or glycosaminoglycan*) adj1 (panel* or assay* or test* or analys*)).tw.
27	((very adj1 long adj1 chain adj1 fatty adj1 acid) adj2 (panel* or assay* or test* or analys*)).tw.
28	(ammonia or lactate or lactic or glucose or glutaric or carnitine).tw.
29	((aspartate adj1 aminotransferase) or ast or (aspartate adj1 transaminase) or (serum adj1 glutamic oxaloacetic adj1 transaminase)).tw.
30	((alanine adj1 aminotransferase) or alt or (alanine adj1 transaminase) or (serum adj1 glutamic adj1 pyruvic adj1 transaminase)).tw.
31	(uric or acetone or acetoacetate).tw.
32	(hydroxybutyrate or (beta adj1 hydroxybut*) or (propion* adj1 carboxylase)).tw.
33	(hydroxypropionic or methylcitrate or methylcitric).tw.
34	(propionylglycine or tiglylglycine or isovalerylglycine or hydroxyisovaleric or hydroxyvaleric or methylcrotonylglycin* or (methylcrotonyl adj1 carboxylase)).tw.
35	(methylglutaric or methylglutaconic or hydroxybutyrate or (methyl adj1 acetoacetate) or hydroxyglutaric or glutaconic or orotic or phenylpyruvic or hydroxyphenylpyruvic or ("4" adj1 OH adj1 pyruvic) or hydroxyisocaropic or hydroxyisolvaric or (alpha adj1 hydroxy adj1 beta adj1 methylvaleric) or (alpha adj1 keto adj1 beta adj1 methylvaleric) or ketoisocaproic or ketoisovaleric).tw.

36	(cystathionine or phenylpyruvic or phenylacetic or phenyllactic or hydroxyphenylacetic or phenylacetylglutam* or succinylacetone or fetoprotein*).tw.
37	(alanine or arginosuccinic or citulline or cysteine or arginine or glutamine or glycine or methionine or phenylalanine or leucine or isoleucine or ornithine or tyrosine or valine).tw.
38	((creatine adj1 kinase) or (total adj1 carnitine) or myoglobin or dicarboxylic).tw.
39	(acylcarnitine or acetylcarnitine or propionylcarnitine or butyrylcarnitine or isovalerylcarnitine or butyrylcarnitine or glutarylcarnitine or hexanoylcarnitine or octanoylcarnitine or octenoylcarnitine or decanoylcarnitine or decenoylcarnitine or dodecanoylcarnitine or dodecenoylcarnitine or dodecanoylcarnitine or tetradecanoylcarnitine or tetradecenoylcarnitine or tetradecadienoylcarnitine or tetradecanoylcarnitine or tetradecenoylcarnitine or palmitoylcarnitine or palmitoleylcarnitine or palmitoylcarnitine or palmitoleylcarnitine or stearoylcarnitine or oleylcarnitine or linoleylcarnitine or stearoylcarnitine or oleylcarnitine).tw.
40	(ethylmalonic or methylsuccinic or glutaric or isobutyrylglycine or butyrylglycine or methylbutyrylglycine or isovalerylglycine or hexanoylglycine or octanoylglycine or phenylpropionylglycine or suberylglycine or cinnamoylglycine or dodecanedioic or tetradodecanedioic or hexadodecanedioic or (("12" or "14" or "16") adj1 DCA)).tw.
41	(sodium or potassium or tsh or thyrotropin or t4 or thyroxine or t3 or triiodothyronine).tw.
42	(("17" adj1 hydroxyprogesterone) or ("17" adj1 OHP) or androstenedione or cortisol or deoxycortisol).tw.
43	((hemoglobin or hb) adj1 (variant or S or F or F1 or A or A2 or C or E)).tw.
44	(biotinidase or (galactose adj1 "1" adj1 phosphate) or (galact* adj1 phosphate) or (phosphate adj1 uridylyltransferase) or galactitol or galactonate).tw.
45	(IRT or (immunoreactive adj1 trypsinogen) or (fecal adj1 elastase)).tw.
46	(TREC or (receptor adj1 excision adj1 circles) or adenosine or deoxyadenosine guanosine or deoxyguinosine or xanthine or hypoxanthine or inosine or deoxyinosine).tw.
47	((glucose adj1 tetrasaccharide) or glucosidase or uric).tw.
48	(sO2 or pO2 or SaO2 or (oxygen adj1 saturation)).tw.
49	((dermatan or heparan or chondroitin) adj1 sulfate).tw.
50	(phenylacetate or sotolone or isovaleric or (hydrogen adj1 sulphide) or trimethylamine or dimethylglycine).tw.
51	9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50
52	exp Bilirubin/
53	(3 and 8 and 51 and 52) not (ANIMALS/ or mice.tw. or mouse.tw. or rat.tw.)
54	limit 53 to humans
55	limit 54 to english language

Supplement S2. Data Extraction Form

➔ **Overall Study Characteristics**
First Author Surname
Study Year
Study Country
Funding Declared [Y/N]
If Declared, Funding Source:
Study Setting Declared [Y/N]
If Declared, Study Setting:
Study Setting, Add'l Notes
Study Design
➔ **Overall Participant Characteristics**
Number of Participants Intended to be Included in Study
Number of Participants Effectively Included in Study
Number and Rationale for Excluded Participants
Participants Age Details
Average Age
Median Age
Minimum Age
Maximum Age
Age Statistic, Other
Participants Gestational Age Details
Average GA
Median GA
Minimum GA
Maximum GA
GA Statistic, Other
Overall Participants Health Status
In narrative form, what were the inclusion criteria?
In narrative form, what were the exclusion criteria?
Hospital Admission Status [Inpatient/Outpatient]
At study outset, Hyperbilirubinemia treatment status:
At study outset, phototherapy status:
➔ **Characteristics of the 'target' or 'study' group** **(if more than 1 study group, use multiple columns)**
Number of Participants Intended to be Included in Study Group
Number of Participants *Actually* Included in Study Group
Number and Rationale for Excluded Participants
Age Details [Study Group Only]
Average Age
Median Age
Minimum Age
Maximum Age
Age Statistic, Other
Participants Gestational Age Details
Average GA
Median GA
Minimum GA
Maximum GA
GA Statistic, Other
In narrative form, what were the inclusion criteria for the study group?
Hospital Admission Status [Inpatient/Outpatient]
At study outset, Hyperbilirubinemia treatment status: [Treated/Untreated]
At study outset, phototherapy status: [Exposed/Unexposed]
➔ **Characteristics of the 'comparator' or 'control' group**
Number of Participants Intended to be Included in Control Group
Number of Participants *Actually* Included in Control Group

Number and Rationale for Excluded Participants
Age Details [CTRL Group Only]
Average Age
Median Age
Minimum Age
Maximum Age
Age Statistic, Other
Participants Gestational Age Details
Average GA
Median GA
Minimum GA
Maximum GA
GA Statistic, Other
Study Group Participants
Were the controls matched for any characteristics with the study group? [Y/N]
If Matched, Characteristics:
What was the primary 'distinguishing characteristic' of participants in the control group?
In narrative form, what were the inclusion criteria for the control group?
According to Table 1, does the control group significantly differ from the study group on any characteristic?
Hospital Admission Status [Inpatient/Outpatient]
At study outset, Hyperbilirubinemia treatment status: [Treated/Untreated/Not Reported]
At study outset, phototherapy status: [Exposed/Unexposed/Not Reported]
➔ **Bilirubin Details for Studies Measuring Continuous Bilirubin**
Did this study objectively measure bilirubin and obtain a continuous value that has both a normal and
Quantification Method/Technique
Source biofluid or tissue
Was bilirubin fractionated? [Y/N]
Bilirubin Fraction Used for Study Analyses [Bound/Unbound/Conj/Unconj]
Testing Model [Point of Care, Lab-based]
Unit of Measurement
Were bilirubin measurements repeated? [Y/N]
Number of repeated measurements:
At what age or time were the infants tested for bilirubin?
Was a threshold level used to identify hyperbilirubinemic infants?
If Yes,Threshold:
Was bilirubin measured prior to and after an exposure?
If Yes, Exposure:
Pre-Exposure Bilirubin Summary Value, Control Group:
Pre-Exposure Bilirubin Summary Value, Study Group:
Post-Exposure Bilirubin Summary Value, Control Group:
Post-Exposure Bilirubin Summary Value, Study Group:
➔ **RUSP Marker Details**
Did this study measure a single marker or a panel of markers?
Name of single biomarker:
Name of biomarker panel:
Source biofluid or tissue
Was the biomarker measured simultaneously, before or after the bilirubin?
What statistical technique was used to show bilirubin's association to the RUSP biomarker?
Is the reported measure of association adjusted for confounders or unadjusted?
➔ ***** Statements of Association *****
Measure of Association:
Was a confidence interval reported?
If yes, Confidence Interval:
Direction of association in hyperbilirubinemic infants **(↑, ↓, ↔, N/A)**
Did the study authors consider confounders when calculating the measure of association? (gestational age,
If yes, confounders:

Interface

The findings from our systematic review of the hyperbilirubinemia and newborn screening literature (Article A) demonstrated that relatively few studies have examined the possibility that biomarkers indicative of rare disease may be altered among hyperbilirubinemic neonates. Some studies were identified suggesting the potential that infants with hyperbilirubinemia may have differences in selected biomarkers relied upon to estimate risk for rare disease. Following this development, a cross-sectional database analysis was designed and executed.

Article B reports the results of this analysis of secondary data in which we investigated associations between 49 newborn screening biomarkers and total serum bilirubin levels sampled in the first week of life.

Ethics

Privacy and ethics approvals for this study were obtained from the ICES Privacy Office and the University of Ottawa Research Ethics Board. Approval documents are attached as Appendix A of this thesis. The Sunnybrook approval relates specifically to the use of coroner data, which was required for this study to identify neonates deceased in the first week of life.

Contributions, Journal Style

All authors participated in the conception and design of this study. ES was responsible for data acquisition and analysis with supervision from SH and BP. All authors participated in the interpretation off findings. ES drafted the manuscript and all authors contributed critical revisions.

Contributions, Detailed

The student (ES) was responsible for specifying the data request as well as cleaning, linking, and analyzing the data. SH and BP supervised the design and biostatistical analysis of the study as well as the interpretation of findings and the development of the manuscript. All authors assisted with result interpretation and development or revision of the manuscript.

Intended Journal Submission

We intend to submit this to *Biomarkers* and the manuscript is formatted pursuant to their requirements.

Article B. Bilirubin-Associated Variation in Biomarkers of Rare Disease and Lack of Association with False Positive Neonatal Screening Results

Emeril Santander [1,2], Steven Hawken [2,3], Beth K. Potter [1,2], Matthew P.A. Henderson [1], Erika Bariciak [5,6] and Pranesh Chakraborty [1,5]

[1] Newborn Screening Ontario, Children's Hospital of Eastern Ontario, Ottawa, ON, K1H 8M5, Canada;
[2] School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, K1H 8L1, Canada
[3] Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada
[4] Health Science Libraries, University of Ottawa, Ottawa, ON, K1H 8L1, Canada
[5] Department of Pediatrics, Faculty of Medicine, University of Ottawa, Ottawa, ON K1H 8L1, Canada
[6] Department of Neonatology, Children's Hospital of Eastern Ontario, Ottawa, ON, K1H 8M5, Canada;

Abstract

Purpose This study investigated associations between total serum bilirubin (TSB) levels and dried blood spot biomarkers measured for rare disease screening (NBS). The association between TSB and false positive NBS results was also examined.

Methods Our cross-sectional study of routinely collected data identified neonates born in Ontario between January 2007 and December 2017 who received NBS and TSB tests in the first postnatal week. Partial correlations were calculated between TSB, NBS biomarkers and clinical covariates. Multivariable logistic regression was used to investigate the association between TSB and false positive results for selected disorders.

Results Among 98,137 participants, levels 48 of 49 NBS biomarkers were correlated with TSB level. Partial correlations with TSB ranged between -0.12 for 17-hydroxyprogesterone (17-OHP) to 0.13 for tyrosine. There was no evidence that TSB was associated with false positive results for congenital adrenal hyperplasia (CAH), congenital hypothyroidism (CH), and phenylketonuria (PKU); three conditions with high false positive rates.

Discussion In a population sample of neonates undergoing NBS, screening markers were significantly associated with TSB levels but TSB was not significantly associated with odds of false positive CAH, CH, or PKU results. The NBS profile of neonates with elevated TSB was characterized by lower concentrations of 17-OHP and higher concentrations of tyrosine, galactose-1-phosphate uridylyltransferase activity, and certain short and long chain acylcarnitines.

B 1.1 Introduction

Public health screening initiatives to identify pre-symptomatic disease frequently rely on measurements of blood-derived biomarkers to estimate risk [1]. Concentrations of the ideal disease-indicative biomarker are distributed uniformly within individuals sharing the same disease status and differentially in the population with the targeted disorder relative to those not affected [2]. Variation in indicative biomarkers not attributable to screen-targeted disease can increase the risk of false positive or false negative screening results by shifting the measured marker away from concentrations otherwise typical of the underlying disease status [3].

Universal newborn screening programs [4] for rare diseases (NBS) are common examples of biomarker-based screening. These aim to detect treatment-responsive diseases, for example, congenital adrenal hyperplasia (CAH), congenital hypothyroidism (CH), and phenylketonuria (PKU). Biochemical and molecular markers are measured in blood spots collected after the 24th hour of life to guide screening for metabolic, endocrine, immune, and hematologic disorders [5–7]. Awareness is growing that fixed, non-disease related neonatal characteristics change NBS biomarker distributions [3,8–12], but few studies have examined whether transient conditions commonly prevalent at the time of sampling are associated with NBS marker differences. Modification of disease-indicative biomarker levels by a potentially unrelated common condition is relevant to neonatal clinicians interpreting screening results in the context of clinical presentations. Programs use this information to determine whether quality improvements are required to adapt screening to neonates hosting untargeted underlying illness at the time of NBS sampling.

Hyperbilirubinemia is a common finding manifesting transiently in the first week of life, with a 60% prevalence in term neonates [13]. Growth in the incidence of preventable complications ascribable to untreated hyperbilirubinemia has led to the introduction of universal pre-discharge neonatal bilirubin screening [14]. These efforts predominantly measure total serum bilirubin (TSB) near the 24th hour of life [15,16]. Within jurisdictions conducting both NBS and hyperbilirubinemia screening, the parallel configuration of universal biomarker-based screening interventions provides the opportunity to assess NBS biomarker variation and false positive result risk in association with the distribution of early life neonatal TSB.

Here, we investigated relationships between 49 NBS biomarkers and TSB levels as measured during universal NBS and hyperbilirubinemia screening in Ontario, Canada. We also assessed whether neonatal TSB levels were associated with risk of false positive results on newborn screening tests for rare diseases.

B 1.2 Methods

B 1.2.1 Overall Design

We conducted a cross-sectional study of Ontario preterm and term neonates born January 1st, 2007 to December 31st, 2017 who received NBS in the first six days of life as well as a TSB test in the first seven days of life. The Ontario Laboratories Information System (OLIS), a provincially centralized electronic repository of laboratory orders and results for clinical use, was the source for neonatal TSB values. Newborn Screening Ontario (NSO), the NBS program for the province, measured blood spot biomarker profiles. The NBS data, TSB results, and other administrative health datasets were deterministically linked using unique encoded identifiers and analyzed at ICES. This report conforms to RECORD guidelines [17]. Ethics approval was received from the University of Ottawa and the Sunnybrook Health Sciences Centre research ethics boards.

B 1.2.2 Study Population and Data Sources

Analytical results from blood spot screening, as well as short term follow-up data on clinical confirmation of screen positive results were securely linked to health administrative data at ICES. This included an encrypted unique identifier, maternal and neonatal demographic data, birth hospital, and age of the infant (in hours) at blood spot collection. Neonatal sex and vital status on the 7th day of life were identified using vital records databases (Registered Persons Database). Severely ill neonates have distinct biomarkers [18], therefore neonates deceased in the first week were excluded. Neonates whose earliest satisfactory NBS took place after the 6th postnatal day or who received a first TSB test after the 7th day of life were excluded. Gestational age (GA) and weight at birth (BW) were obtained from the ICES-derived MOMBABY dataset. Admissions to intensive care (NICU) in the first week of life were identified by a special care unit designation on the hospital record. As in previous studies of phototherapy administration [19], Canadian Classification of Health Interventions [20] code 1YZ12JADQ identified participants receiving inpatient (Discharge Abstract Database) or outpatient (National Ambulatory Care Reporting System) phototherapy administered in the first postnatal week. The final sample consisted of neonates who were alive 7 days following birth with complete covariate data who had valid NBS and a TSB linked within the repository, with one exception: participants missing markers non-contributory to disease risk determinations were permitted to re-enter the sample for regression analyses.

B 1.2.3 Neonatal Dried Blood Spot Screening for Rare Disease

NSO is responsible for rare disease screening in the province of Ontario. Maternal-neonate characteristics and blood spot collection details were noted by a healthcare provider at the birth site on a form attached to blood collection media (Supplementary Materials**, S1**). Laboratory methods used to measure dried blood biomarkers in the period of study have been previously reported and included mass spectrometry, immunoassays, and molecular techniques [21]. NBS biomarkers studied are available in Table 3. For neonates with multiple NBS samples, we selected the earliest satisfactory sample as the analytic sample for our study. Neonates with biomarker concentrations in excess of pre-specified thresholds were referred to the pediatric centre nearest their home for diagnostic assessment. Pediatricians forwarded diagnoses to NSO, which were classified as 'True Positive', 'False Positive', 'Variant Disease', 'Incidental Finding', or 'Other' using a matrix developed *a priori* by an expert committee (Supplementary Materials**, S2**). False negative screening outcomes were not available in the accessed data repositories.

B 1.2.4 Neonatal Total Serum Bilirubin

Licenced clinical laboratories in Ontario perform neonatal TSB testing pursuant to Canadian Pediatric Society guidelines first published in 2007 [22]. Results, units, reference intervals, abnormality flags, and testing date were forwarded to OLIS. All TSB results were reported in micromoles per liter ('umol/L') and included the Logical Observation Identifiers Names and Codes (LOINC) 14631-6. Bilirubin reported with this code may be measured in serum or plasma with spectophotometric, wet chemistry, or dry slide techniques. Transcutaneous bilirubinometry measurements were not included in this study. For neonates with a single TSB result within the first 7 days of life, we used that TSB value in our study. For neonates with multiple TSB test results, we used the result that was from a test taken ± 1 day from the NBS samples; when multiple TSB results met this criterion, we used the average of TSB values in this 72-hour window. For those who had multiple TSB results but none occurring ± 1 day from the NBS, we used the nearest TSB result to the NBS sample.

B 1.2.5 Biomarker Standardization and Other Covariates

After preliminary exclusions but prior to data linkage, biomarkers as well as GA and BW were winsorized [23]. Values exceeding the 99.99th centile or below the 0.01st centile were replaced with the 99.99th and 0.01st percentile value, respectively. The TSB and NBS values of each neonate were then standardized. Specifically, each biomarker value for each participant was standardized by subtracting the monthly provincial biomarker mean and dividing the result by the monthly provincial standard deviation (SD), such that the resulting transformed variable had a

mean of 0 and SD of 1. This rendered biomarker values comparable across measurement scales while also reducing variability attributable to external factors such as seasonal variation [18]. References to TSB or NBS biomarkers hereinafter refer to their standardized forms.

B 1.2.6 Analyses

Participant characteristics are presented as medians, quartiles, or proportions relative to the eligible population. Medians for NBS biomarkers were calculated across tercile strata of TSB within the sample ('*Lower TSB'*: 1st-33rd percentile TSB; '*Moderate TSB'*: 33rd-66th percentile TSB, '*Higher TSB*': 66th-99th percentile TSB). The difference between the median biomarker value in the Lower vs. Higher TSB groups was calculated and expressed as an absolute value. Crude Spearman correlations were calculated to identify monotonic relationship strengths between NBS biomarkers and TSB. Partial correlations were then obtained that incrementally adjusted for prematurity (BW and GA), postnatal age at NBS, and every other studied NBS marker concentration. Two-tailed 95% confidence intervals are provided [24]. Birth institution was obtained from hospital records of all births in the study period and the monthly rates of NBS and TSB testing were calculated on a hospital-specific basis. Participants were deemed to have received universal hyperbilirubinemia screening if $\geq 90\%$ of births occurring at the same institution on the same month had TSB results available. Correlation analyses were repeated in the subsample of infants who received universal screening as a sensitivity analysis to assess for influence by suspicion-based TSB testing.

Logistic regression was used on a disease-specific basis subject to sample size availability to investigate the association between TSB [25] and false positive versus screen negative NBS results. Multivariable logistic models were created adjusting for sex, postnatal age at sampling, birthweight, gestational age at birth, exclusive breastfeeding status, and TPN administration; well-established influencers of false positivity risk [3,12,26–29]. We pre-specified covariates for inclusion and did not conduct screening or selection. At least 10 false positive 'cases' were required [30] for every covariate degree of freedom included, but insufficient sample size with respect to false positives for most disorders precluded the inclusion of all potentially-important covariates. Continuous covariates were placed in multivariable models as restricted cubic splines [31] with three knots placed at the 5th, 50th, and 95th percentiles of the distribution for each covariate [32].

All statistical analyses were performed using the software SAS version 9.4 (SAS Institute, Cary, NC).

B 1.3 Results

B 1.3.1 Participants

The analytic sample was comprised of 98,137 preterm and term neonates, representing 7% of the population screened for rare disease (NBS) in the study period (Figure 1), with the vast majority of exclusions being due to the absence of a TSB test result during the neonatal period. Neonates excluded from the main analytic sample based on incomplete biomarker profiles (n = 6,461) were all eligible for false positivity risk modelling, raising the sample size available for logistic regression (n = 104,958). The provincial NBS participation rates were above 98% across all studied months (Supplementary Materials, **S4**). The monthly proportion of neonates receiving hyperbilirubinemia screening among births captured in the dataset ranged between < 1% in the earliest birth years eligible for the study to 29% among neonates born in the last month of the study eligibility period (**S4**).

Given the large proportion of neonates missing a TSB test result (due to the absence of universal hyperbilirubinemia screening at most birth hospitals during the study period), we investigated the characteristics of included participants (i.e., those with a TSB result) relative to neonates meeting all inclusion criteria except for observation of a valid TSB. Included participants on average had lower birth weight and gestational age and were more likely to have received phototherapy, total parenteral nutrition, and intensive neonatal care (Table 1).

The majority of included participants (n = 66,493; 68%) received hyperbilirubinemia screening due to the introduction of universal screening rather than for a diagnostic investigation. Most participants (66%) had a single eligible TSB test identified in the repository and 92% of these occurred within ± 1 day from the NBS sample collection (**S3**). Neonates with multiple TSB observations (n = 33,803, 34% of all participants) had 94% of TSB's reported within this time window. Neonates identified as having received universal TSB screening were similar with respect to the proportion receiving only a single TSB test (67%) relative to that observed in the main analytic sample, but nearly all of these TSB tests in the universally screened subsample occurred ± 1 day from the newborn screen (>99%). Among neonates receiving a TSB test in the context of universal screening, demographic characteristics were more similar to the broader population of neonates meeting all inclusion criteria except for possession of a valid TSB result, except for the proportion reported to be exclusively breastfed, which was markedly increased (**S5**).

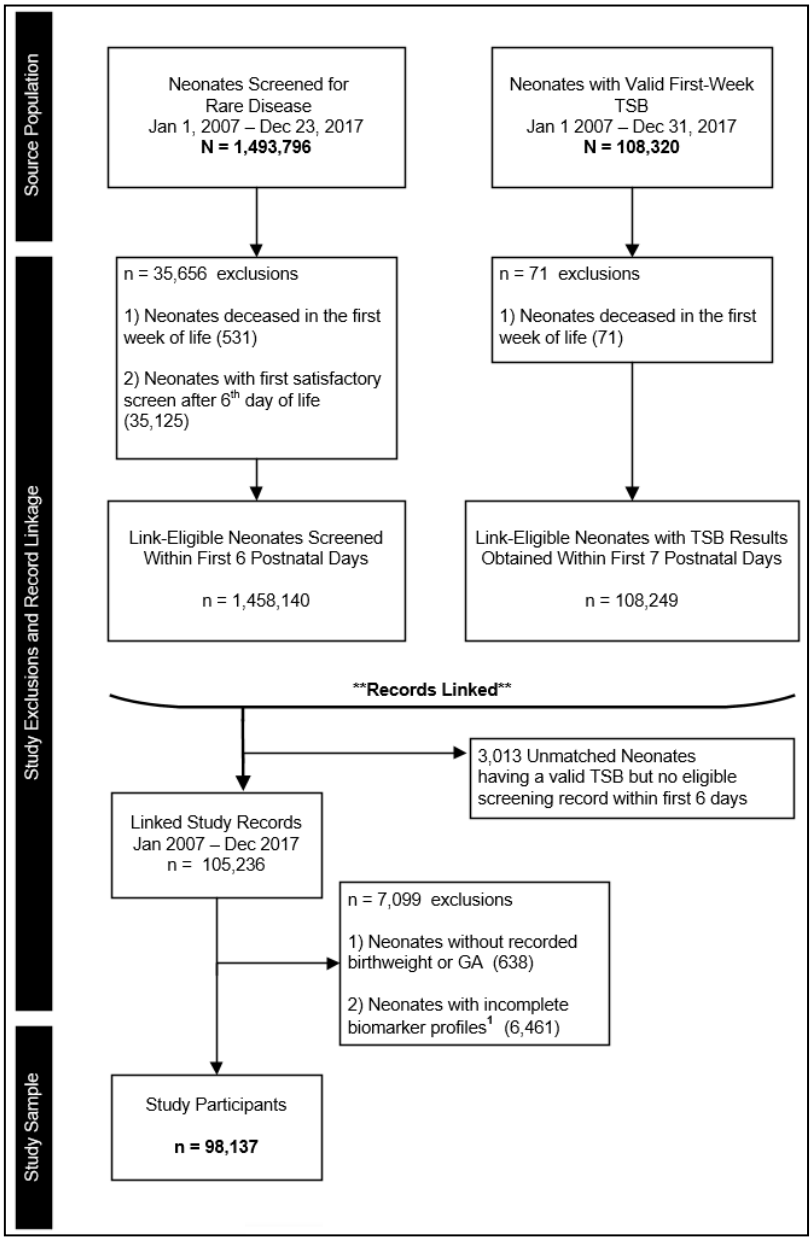


Figure 1. Flow Diagram for the Creation of the Study Sample

[1] Neonates with partially complete NBS biomarker profiles are excluded from descriptive and correlational analyses but are included in logistic regression models.

Table 1. Eligible Population and Study Sample Characteristics

	Screening Record Availability in Central Repository			
	NBS without TSB		NBS and TSB	
	n	%	n	%
Number of Neonates	**1,327,619**		**98,137**	
Sex				
Male	679,632	51.2%	50,948	51.9%
Female	647,987	48.8%	47,189	48.1%
Phototherapy in First Week				
Received	62,527	4.7%	7,239	7.4%
Not Received	1,265,092	95.3%	90,898	92.6%
NICU Admission in First Week				
Admitted	95,542	7.2%	14,765	15.0%
Not Admitted	1,232,077	92.8%	83,372	85.0%
Feeding Status				
Exclusive Breastfeeding	635,535	47.9%	63,098	64.3%
Other	692,084	52.1%	35,039	35.7%
TPN Administration				
Administered	6,626	0.5%	2,025	2.1%
Not Administered	1,320,993	99.5%	96,112	97.9%
Birthweight Category				
Normal Birthweight[>2,500g]	1,251,682	94.3%	89,307	91.0%
Low Birthweight[<2,500g]	67,829	5.1%	7,492	7.6%
Very Low Birthweight[<1,500g]	5,391	0.4%	811	0.8%
Extremely Low Birthweight[<1,000g]	2,717	0.2%	527	0.5%
Missing Weight	-	0.0%	0	0.0%
Gestational Age at Birth				
Term [>=38 Wk]	1,131,373	85.2%	79,040	80.5%
Moderate to Late Preterm [32.1-37.6 Wk]	186,847	14.1%	17,612	17.9%
Very Preterm[28-32 Wk]	6,583	0.5%	980	1.0%
Extreme Preterm [<28 Wk]	2,816	0.2%	505	0.5%
Missing Gestational Age	-	0.0%	0	0.0%
Proportion of Monthly Institution-Specific				
90-100%	Not Applicable: No TSB Test Reported		66,493	67.8%
80-89%			6,951	7.1%
70-79%			2,165	2.2%
60-69%			2,321	2.4%
50-59%			2,345	2.4%
40-49%			1,488	1.5%
30-39%			3,990	4.1%
20-29%			5,639	5.7%
10-19%			969	1.0%
< 9.9%			5,256	5.4%
Institution Not Available			520	0.5%
Age at Dried Blood Spot, Hours Median [25th-75th Centiles]	29 [25 – 42]		25 [24,32]	
Age at Total Serum Bilirubin, Days Median [25th-75th Centiles]	N/A		1 [1-2]	

Wk Weeks Gestation (Weeks.Days)

B 1.3.2 Dried Blood Biomarker Variability

Standardized NBS biomarker medians across standardized TSB strata are reported in Table 2. The median marker values of the Moderate TSB group were between the medians of the Lower and Higher TSB groups in all but seven NBS biomarkers (MET, C8:1, C16:1OH, C16-OH, C6DC, C10:1, BIOT). Comparison of the Higher and Lower TSB groups revealed C18:1OH had the most differing median across TSB terciles ($\Delta = 0.39$, interpretable as a proportion of a standard deviation), followed in ranked order of absolute difference by ii. C14 ($\Delta = 0.33$SD) iii. C18:1 and C16 ($\Delta = 0.32$SD) iv. GALT ($\Delta = 0.30$SD), and v. tyrosine ($\Delta = 0.29$SD). Twenty-four of 49 of biomarkers had medians in the Higher TSB group separated by more than a tenth of a standard deviation from the median of the Lowest TSB group.

Crude correlations ranged between -0.12 in 17-Hydroxyprogesterone (OHP) to 0.16 in tyrosine, with only three markers not significantly correlated to TSB (C6DC, C10:1, C-14:2) (Table 2). Five acylcarnitines at least 12 carbon atoms in length were significantly correlated with TSB with an absolute correlation ($|r|$) > 0.10 (C-12, C-14, C-14:1, C-16, C-18), joined at this correlational level by two short chain acylcarnitines (C2, C4OH). Tyrosine, leucine-isoleucine, GALT, thyroid stimulating hormone (TSH), and 17-OHP were the only non-acylcarnitine biomarkers with $|r| > 0.10$. Importantly, partial spearman correlation adjusting for GA, BW, and NBS sampling age demonstrated that biomarkers with crude $|r| > 0.10$ remained at this correlational level or increased in strength after adjustment for only GA, BW, and NBS sampling age; this was true for all biomarkers crudely exceeding $|r| > 0.10$ except C4OH, TSH and leucine (**S6**).

Following adjustment for BW, GA, postnatal age at NBS sampling, and the standardized concentration of every NBS marker other than the NBS marker being correlated to TSB, the most positively correlated marker with TSB was tyrosine ($r = 0.13$) and 17-OHP remained the strongest inverse correlation observed ($r = -0.12$) (Table 2). These were the only two markers correlating with TSB in excess of $|r| > 0.10$ in the fully adjusted spearman partial correlation analysis.

Table 2. Standardized NBS Biomarker Medians and Spearman Correlations to Total Serum Bilirubin

		Total Serum Bilirubin Group				Spearman Rho to TSB			
		LOWER TSB (*n* = 32,717)	MODERATE TSB (*n* = 32,708)	HIGHER TSB (*n* = 32,712)	Δ[1]	Crude (*n* = 98,137)		Adjusted [2] (*n* = 98,137)	
		Median	Median	Median		*r*	95% CI	*r*	95% CI
Amino Acids and Ketones	**TYR**	-0.31	-0.07	-0.02	0.29	0.16	(0.15,0.16)	0.13	(0.12,0.13)
	LEU + ILEU	-0.30	-0.23	-0.09	0.21	0.11	(0.10,0.11)	-0.05	(-0.05,-0.06)
	PHE	-0.07	0.05	0.09	0.16	0.08	(0.08,0.09)	0.05	(0.05,0.06)
	SUAC	-0.20	-0.14	-0.06	0.14	0.07	(0.06,0.08)	0.03	(0.02,0.03)
	GLY	-0.14	-0.07	0.00	0.14	0.06	(0.06,0.07)	0.03	(0.03,0.04)
	ARG	-0.21	-0.16	-0.09	0.12	0.06	(0.05,0.06)	0.06	(0.05,0.07)
	CIT	-0.02	-0.04	-0.12	0.10	-0.04	(-0.05,-0.04)	-0.03	(-0.03,-0.04)
	VAL	-0.22	-0.23	-0.13	0.09	0.04	(0.03,0.05)	0.04	(0.03,0.05)
	ASA	-0.15	-0.13	-0.06	0.09	0.04	(0.03,0.04)	0.02	(0.01,0.03)
	MET	-0.03	0.08	0.06	0.09	0.05	(0.05,0.06)	0.03	(0.03,0.04)
	ALA	-0.15	-0.13	-0.07	0.08	0.03	(0.03,0.04)	0.04	(0.04,0.05)
	ORN	-0.16	-0.21	-0.23	0.07	-0.03	(-0.03,-0.04)	-0.10	(-0.09,-0.11)
Carnitine and Acylcarnitines	**C18:1OH**	-0.37	-0.21	0.02	0.39	0.05	(0.05,0.06)	0.01	(0.00,0.01)
	C14	-0.27	-0.11	0.06	0.33	0.15	(0.14,0.15)	0.02	(0.01,0.03)
	C18:1	-0.30	-0.17	0.02	0.32	0.14	(0.13,0.15)	0.01	(0.00,0.01)
	C16	-0.28	-0.12	0.04	0.32	0.14	(0.13,0.14)	0.03	(0.02,0.03)
	C4OH	-0.28	-0.20	-0.02	0.26	0.12	(0.11,0.12)	0.02	(0.02-0.03)
	C2	-0.29	-0.22	-0.01	0.28	0.12	(0.11,0.13)	0.00	(0.00,-0.01)
	C18	-0.25	-0.11	0.00	0.25	0.12	(0.11,0.12)	0.00	(0.00,-0.01)
	C12	-0.20	-0.07	0.04	0.24	0.12	(0.11,0.12)	0.02	(0.02,0.03)
	C14:1	-0.22	-0.09	0.00	0.22	0.12	(0.11,0.12)	0.02	(0.01,0.02)
	C3	-0.26	-0.20	-0.08	0.18	0.09	(0.08,0.09)	0.04	(0.03,0.05)
	C18OH	0.01	0.13	0.17	0.16	0.05	(0.05,0.06)	0.01	(0.01,0.02)
	C4	-0.25	-0.20	-0.12	0.13	0.07	(0.07,0.08)	0.01	(0.00,0.01)
	C5OH	-0.17	-0.10	-0.04	0.13	0.08	(0.07,0.08)	0.01	(0.00,0.02)
	C0	-0.20	-0.18	-0.08	0.12	0.06	(0.05,0.06)	-0.02	(-0.01,-0.03)
	C8:1	-0.14	-0.26	-0.24	0.10	-0.06	(-0.05,-0.07)	-0.06	(-0.05,-0.06)
	C5	-0.20	-0.16	-0.11	0.09	0.06	(0.06,0.07)	-0.02	(-0.01,-0.03)
	C12:1	-0.17	-0.09	-0.07	0.10	0.05	(0.04,0.05)	-0.02	(-0.02,-0.03)
	C10	-0.13	-0.07	-0.06	0.07	0.05	(0.05,0.06)	0.00	(0.00,0.01)
	C4DC	-0.13	-0.15	-0.19	0.06	-0.03	(-0.02,-0.03)	-0.07	(-0.06,-0.07)
	C6	-0.12	-0.08	-0.05	0.07	0.03	(0.03,0.04)	0.00	(0.00,-0.01)
	C14OH	-0.16	-0.12	-0.09	0.07	0.05	(0.04,0.06)	0.01	(0.01,0.02)
	C18:2	-0.22	-0.18	-0.16	0.06	0.04	(0.04,0.05)	0.00	(0.00,0.00)
	C5DC	-0.08	-0.06	-0.03	0.05	0.04	(0.04,0.05)	0.00	(-0.01,0.01)
	C3DC	-0.07	-0.05	-0.02	0.05	0.04	(0.04,0.05)	0.02	(0.01,0.02)
	C14:2	-0.12	-0.10	-0.08	0.04	0.04	(0.04,0.05)	-0.01	(0.00,-0.01)
	C5:1	-0.04	0.00	0.01	0.05	0.01	(0.00,0.02)	0.00	(0.00,-0.01)
	C16:1OH	-0.13	-0.08	-0.09	0.04	0.03	(0.03,0.04)	-0.03	(-0.02,-0.03)
	C16OH	-0.20	-0.17	-0.17	0.03	0.06	(0.05,0.06)	0.01	(0.00,0.01)
	C8	-0.04	-0.03	-0.02	0.02	0.04	(0.03,0.04)	0.00	(0.00,-0.01)
	C6DC	-0.25	-0.27	-0.24	0.01	0.00	(0.00,-0.01)	-0.06	(-0.05,-0.07)
	C10:1	-0.09	-0.11	-0.09	0.00	0.00	(0.00,-0.01)	-0.01	(0.00,-0.02)
Other Markers	**GALT**	-0.20	-0.04	0.10	0.30	0.12	(0.11,0.12)	0.06	(0.06,0.07)
	TSH	0.07	0.00	-0.11	0.18	-0.12	(-0.12,-0.13)	-0.03	(-0.02,-0.04)
	17-OHP	-0.04	-0.10	-0.23	0.19	-0.12	(-0.11,-0.13)	-0.12	(-0.11,-0.12)
	BIOT	-0.08	-0.20	-0.15	0.07	-0.05	(-0.04,-0.05)	-0.06	(-0.06,-0.07)
	IRT	-0.18	-0.25	-0.26	0.08	-0.05	(-0.05,-0.06)	-0.02	(-0.01,-0.03)
	TREC Count	-0.19	-0.20	-0.25	0.06	-0.03	(-0.03,-0.04)	-0.04	(-0.03,-0.04)

Note: The full names of listed biomarkers are available in the front matter of this thesis.

[1] Absolute difference of median between *Lower* and *Higher* TSB groups in SD units
[2] Partial Spearman correlation adjusted for gestational age, birthweight, postnatal age at sampling, and all other measured NBS markers

Correlations adjusted in incremental covariate sequence are provided as a supplementary material (**S6**). As a sensitivity analysis, the correlations presented in Table 2 were reproduced with solely neonates receiving TSB in a universal screening setting (n = 66,493, 68% of overall sample). The few correlations that differed in the subgroup changed by ≤ 0.03 SD or less, except for GALT whose crude correlation with TSB increased (to $r_{\text{universal}} = 0.16$, 95% CI: 0.15, 0.17). We were satisfied that the incidental inclusion of neonates tested for TSB in settings we could not verify were universal did not alter our biomarker variability analysis. The correlations specific to this group are provided as supplement (**S7**).

B 1.3.3 False Positivity Risk

Three disorders (Congenital Adrenal Hyperplasia [CAH], Congenital Hypothyroidism [CH] and Phenylketonuria [PKU]) had a sufficient number of neonates with false positive newborn screening results to investigate associations between TSB levels and false positive versus fully screen negative results. The referent group for all models was comprised of the same 103,794 participants who were found at low risk (i.e., negative screening result) for every disorder on the Ontario panel. Participant demographics across disorder-specific categories of screening outcome are reported in Table 3.

TSB levels were not significantly associated with risk of false positivity for CAH, CH, or PKU in either bivariate (Table 4) or multivariable logistic regression analyses, adjusted for important covariates (Table 5). However, there were notable associations of included covariates with false positive results. Independent variables other than standardized TSB were all highly significantly associated with false positive risk for CAH in univariable logistic regression, including an increase in the odds of false positive CAH results in male neonates.

Table 3. Characteristics of Participants Included in Logistic Regression Models (n = 104,598)[1-2]

	Screening Outcome							
	SN Fully Screen Negative		**CAH** False Positives		**CH** False Positives		**PKU** False Positives	
Number of Neonates	103,794		91		140		84	
Total Serum Bilirubin, Standardized Median [25th-75th Centiles]	**-0.21 [-0.63 – 0.45]**		**-0.18 [-0.60 – 0.29]**		**-0.27 [-0.73 – 0.46]**		**-0.21 [-0.72 – 0.20]**	
Total Serum Bilirubin, µmol/L Median [25th-75th Centiles]	112 [91,150]		115 [97,160]		108 [86,155]		111 [87,136]	
Age at Dried Blood Spot, Hours Median [25th-75th Centiles]	25 [24 – 32]		29 [25 – 50]		25 [24 – 26]		25 [24 – 28]	
	n	%	*n*	%	*n*	%	*n*	%
Sex								
Male	53,950	52.0%	66	72.5%	72	51.4%	49	58.3%
Female	49,764	48.0%	25	27.5%	68	48.6%	45	41.7%
Phototherapy in First Week								
Administered	7,503	7.2%	36	39.6%	18	12.9%	51	60.7%
Not Administered	96,211	92.8%	55	60.4%	122	87.1%	33	39.3%
NICU Admission in First Week								
Admitted	95,542	7.2%	81	89.0%	26	18.6%	> 79	-
Not Admitted	1,232,077	92.8%	10	11.0%	114	81.4%	≤ 5	-
Breastfeeding Status								
Exclusive Breastfeeding	66,245	63.9%	13	14.3%	86	61.4%	7	8.3%
Other	37,594	36.1%	78	85.7%	54	38.6%	77	91.6%
TPN Administration								
Administered	1,962	1.9%	24	26.4%	≤ 5	-	49	58.3%
Not Administered	101,832	98.1%	67	73.6%	> 135	-	35	41.7%
Birthweight Category								
Normal Birthweight[>2,500g]	94,717	91.3%	48	52.7%	130	92.9%	12	14.3%
Low Birthweight[<2,500g]	7,770	7.5%	29	31.9%	9	6.4%	24	28.6%
Very Low Birthweight[<1,500g]	757	0.7%	7	7.7%	≤ 5	-	34	40.5%
Extremely Low Birthweight[<1,000g]	470	0.5%	7	7.7%	≤ 5	-	14	16.7%
Gestational Age at Birth								
Term [>=38 Wk]	83,749	80.7%	≤ 5	-	112	80.0%	9	10.7%
Moderate to Late Preterm[32.1-37.6Wk]	18.577	17.9%	64	70.3%	28	20.0%	28	33.3%
Very Preterm[28-32Wk]	942	0.9%	12	13.2%	≤ 5	-	33	39.3%
Extreme Preterm [<28Wk]	446	0.4%	< 12	-	≤ 5	-	14	16.7%

Note: Table cells have been made uncertain as contractually required to prevent health disclosures concerning five or fewer neonates.

[1] Neonates with partially complete biomarker profiles (*n* = 6,461,see Figure 1) have been reincorporated as participants for this analysis.
[2] Neonates categorized True Positive, Variant, Incidental, Other, or False Positive for a non-tabulated disorder are not shown (*n* = 489)

Table 4. Crude Odds Ratios

	False Positive Disorder (vs. Fully Screen Negative)					
	CAH		**CH**		**PKU**	
	OR	95% CI	OR	95% CI	OR	95% CI
Standardized TSB	**0.98**	**0.80 – 1.21**	**0.95**	**0.80 – 1.12**	**0.83**	**0.65 – 1.06**
Birthweight, grams	0.998	0.998 – 0.999	1.00	1.00 – 1.00	0.997	0.997 – 0.998
Gestational Age, weeks	0.68	0.66 – 0.71	1.06	0.97 – 1.17	0.63	0.61 – 0.66
Postnatal Age at NBS Sampling (hours)	1.02	1.01 – 1.03	0.91	0.87 – 0.94	1.01	1.00 – 1.02
Male Sex	2.44	1.54 – 3.86	0.98	0.70 – 1.36	1.29	0.84 – 1.99
Exclusive Breastfeeding	0.09	0.05 – 0.17	0.90	0.64 – 1.27	0.05	0.02 – 0.11
TPN Administered	18.58	11.63 – 29.68	0.37	0.05 – 2.67	72.61	46.95 – 112.89

The only variable crudely associated with screening falsely positive for CH was postnatal age at NBS sampling, with each hourly increase in age corresponding to a 9% decrease in the odds of a false positive result if linearity of effect is assumed. Variables other than TSB and sex were significantly associated with false positive outcomes for PKU in the unadjusted analysis.

Table 5. Multivariable Odds Ratios of False Positivity

	False Positive Disorder (vs. Fully Screen Negative)					
	CAH		**CH**		**PKU**	
	OR	95% CI	OR	95% CI	OR	95% CI
Standardized TSB	**0.87**	**0.68 –1.13**	**1.06**	**0.90 -1.26**	**1.20**	**0.88 – 1.63**
Male Sex	2.05	1.29 –3.26	1.07	0.76 – 1.49	1.33	0.85 – 2.08
Exclusive Breastfeeding	0.33	0.18 –0.61	0.79	0.56 – 1.13	0.30	0.12 – 0.71
TPN Administered	1.37	0.73 –2.54	0.51	0.06 -4.31	3.29	1.85 – 5.87
Birthweight Spline	$p < 0.001$		$p < 0.01$		$p < 0.0001$	
Gestational Age Spline	$p < 0.0001$		$p < 0.01$		$p = 0.20$	
Postnatal Age at NBS Spline	$p = 0.08$		$p < 0.0001$		$p < 0.01$	

In final multivariable models, standardized TSB was not significantly associated with increased odds of false positive results for any of the modelled disorders (Table 5). Male sex remained significantly associated with false positive results for CAH following adjustment. Administration of TPN was associated with a greater than three-fold increase in the odds of false positive PKU results.

B 1.4 Discussion

This is the second comprehensively adjusted study to find nearly profile-wide NBS biomarker differences associated with neonatal bilirubin level or hyperbilirubinemia status,

independently sourced from two of the largest and most diverse neonatal screening jurisdictions in North America [33]. We reported differences in median concentrations of most NBS biomarkers across tercile strata of neonatal TSB in the first days of life. In adjusted correlational analyses, modest but statistically significant monotonic associations between NBS biomarkers and TSB were found in 33 of 49 markers after removing the influences of prematurity, postnatal age at blood spot sampling, and other studied screening markers. The strongest positive and inverse correlations with TSB were observed in tyrosine and 17-OHP, respectively. In multivariable regression models adjusted for confounding factors, standardized TSB levels were not significantly associated with the risk of false positive CAH, CH, or PKU results in comparison to a reference group of neonates who screened negative for all disorders on the Ontario panel. Our study did not find evidence that the specificity of neonatal CAH, CH, and PKU screening in the province could benefit from augmentations compensating for the altered biochemical state of neonates who have elevated TSB near the time of rare disease screening.

Our findings associating NBS biomarkers and TSB are in agreement with the preceding study that associated NBS markers to clinically diagnosed hyperbilirubinemia among 1.2 million infants in California [33]. We concur in reporting a positive adjusted TSB association with tyrosine, saturated acylcarnitines 14 carbons in length or longer, and Galactose-1-phosphate uridyl transferase (GALT) activity as well as inverse association with 17-Hydroxyprogesterone (17-OHP). Although both studies report an inverse TSB association with TSH, our partial correlation results suggested this relationship was primarily attributable to the timing of NBS collection (Supplementary Materials, **S6**). Reproduction of the earlier findings across a broad range of TSB values obtained mainly from parallel universal screening adds support for an association with blood spot biomarkers that extends to neonates without a clinical diagnosis of hyperbilirubinemia.

The current and California studies question whether neonates should continue to be conceptualized as a single biochemical population whose screening markers vary in isolated relation to panel-targeted disease status. Our demonstration of shifts in the standardized location of the 50th percentile of biomarkers across neonatal TSB strata suggests the distributional position of a single population-wide abnormality cut-off (e.g. 99th percentile) will likely fail to translate to an equivalent percentile in some TSB-stratified biomarker distributions, assuming comparable variance. This increases the risk that screening specificity differs for neonates with varying TSB levels at NBS sampling. Also known as spectrum effect [34], the potential for this phenomenon led us to model false positivity odds as a function of TSB. Although our regression results show lack of TSB association with false positive results for three disorders, verification of homogenous specificity across the TSB spectrum remains outstanding for other disorders whose TSB-related

shift places larger proportions of TSB-defined subpopulations in closer proximity to applicable cut-offs (e.g. Long Chain Acyl CoA Deficiencies). Biomarker variability in NBS associated with untargeted conditions such as hyperbilirubinemia support screening definitions of abnormality that are multidimensional and personalized to consider any relevant aspect of the neonatal phenotype influential to screening marker measurements prevalent at the time of sampling [35].

Our results confirm a population-wide association between increased neonatal TSB and decreased 17-OHP, first reported by McCarthy and colleagues [33]. Documented assay interference from both endogenous and exogenous steroids leads us to use caution when interpreting our finding of an inverse relationship between TSB and 17-OHP levels [36–38]. Central endocrine disease is a known risk factor for cholestasis [39–42], but we are the first to report that lower concentrations of an adrenal steroid or cross-reactants associate dose-responsively with TSB in the first days of life among neonates predominantly lacking endocrinopathy. Additional research on the relationship between neonatal adrenal function and early TSB levels may be required to further contextualize these observations.

Several variable biomarkers across TSB strata in this study were erythrocyte-related (e.g. GALT [43], C14-18 [44–47]). Insofar as elevated TSB can imply an increased heme-catabolic load on the neonatal liver [48], polycythemia could be a shared cause leading to TSB elevation and hyperconcentration of erythrocyte-related markers in NBS. Neonates in our study were mostly tested for TSB and NBS near the 24th hour of life and bilirubin elevations appreciable at this age are likely due to *in vivo* hemolysis [49]. Erythrocyte marker associations with early TSB levels support the hypothesis that NBS tests reliant on erythrocyte-related biomarkers may perform differently in subpopulations of neonates characterized by polycythemia or *in vivo* hemolysis. Surveillance for modified screening accuracy in disorders indicated by these markers may be of use among extensively hemolyzed neonates or those with atypical erythrocyte loads at the time of NBS sampling, including newborns treated with delayed cord clamping at birth which has become the standard practice in many neonatal settings [50].

Neonates who met eligibility criteria for our analytic sample differed slightly from those excluded from the study due to the absence of a TSB test (Table 1). The observed differences of participant characteristics would manifest if hospitals providing more complex care began centralized TSB reporting earlier in the study period, or if more vulnerable neonates were more likely to receive a test outside the context of universal screening. Participant neonates were more likely to have been born later in the study period and therefore these differences may also relate to changes in underlying healthcare practices. Implementation of hyperbilirubinemia screening has been associated with increased phototherapy utilization in Ontario [19]. Our study spanned the

introduction of hospital-based breastfeeding support initiatives [51] that increased the provincial rates of exclusive breastfeeding [52]; a risk factor for early hyperbilirubinemia [53,54] that we were unable to reproduce in our study. These potential selection biases underscore the importance of our sensitivity analysis considering only neonates who received a TSB test in the context of confirmed universal screening for hyperbilirubinemia.

The principal strength of this study is our reliance on the alignment of two population-based screening initiatives to form the basis of our study. This design supports generalizable inferential statements about biomarkers in neonates that are neither limited to individuals sufficiently ill to indicate the requisite tests [55] nor to individuals sufficiently well for research blood draws [56]. The use of biomarker standardization methods also reduced the risk that changing laboratory practices and other analytical considerations introduced bias, as has been reported in other biomarker studies [57]. The main limitation of our study is that use of TSB values from a clinical repository prevented analysis of error introduced by differing TSB analytic methods [58,59] which, if present, would bias our associations toward the null [60]. The comparison groups in our logistic regression analyses (false positive vs. fully screen negative) were also two of six possible screening outcome categories (**S6**), requiring specific interpretation of associations. We believe this outcome configuration best informed our research goals related to false positive results, however it is acknowledged that if hyperbilirubinemic neonates lacking targeted disease were differentially likely to be miscategorized (e.g. 'Incidental') our false positive risk associations would be erroneously attenuated. Our study was unable to report conventional diagnostic accuracy metrics such as specificity and sensitivity because false negative screening outcomes were not available. Obtaining data on children with false negative newborn screening results is challenging, and in Ontario, would have required a survey of active practitioners for information on patients with screen-targeted disease identified clinically. Primary data collection for false negative screening outcomes was outside the scope of this routinely-collected data study investigating false positive screening results. If such a survey were conducted, the high sensitivity of newborn screening tests for CAH [61], CH [62] and PKU [63] together with the extremely low birth prevalence of these disorders would suggest almost all screen negative participants in our study were truly negative. Maternal smoking [64] and mode of birth could not be included in false positive risk models due to sample size. Race and ethnicity are associated with variation in TSB levels [65] as well as the levels of many NBS biomarkers [65] and may residually confound our reported associations. Data on race and ethnicity are not currently collected in most Canadian healthcare settings [66] and were not available for this study, although proposed standards for the collection of race-based information and indigenous identity may permit adjustment for the influence of these

factors in future studies [67]. Lastly, the associations in our study may not be reproducible in jurisdictions where the timing of screening differs from that in Ontario. This is because the circulating concentrations of TSB as well as those of many NBS markers are age dependent, and we had the benefit of a large cohort of newborns with TSB measurements very close to the time of NBS screening. Neonates with higher levels of TSB produced NBS blood spot samples that were biochemically different from neonates with lower levels of TSB near the time of NBS sampling. For certain NBS biomarkers, awareness of the pre-discharge TSB value may add valuable context to the interpretation of altered screening profiles. Reassuringly, for those diseases for which we had sufficient cases to investigate, we did not strong evidence that the association between TSB and biomarkers translated to a higher risk of false positive NBS results for neonates with higher or lower TSB values.

B 1.5 Conclusions

Most NBS biomarkers vary modestly but statistically significantly in association with TSB levels both univariately and after adjustment for important confounders including other NBS analytes. Increased TSB was associated with elevated tyrosine, several acylcarnitines, GALT enzyme, and decreased 17-OHP. TSB near the time of NBS sampling was not significantly associated with increased risk of false positive results for CAH, CH, or PKU. Therefore, our study did not provide evidence that TSB levels should be considered in defining screening protocols.

B 1.6 Required Disclosures

This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). This study was supported by Newborn Screening Ontario. Parts of this material are based on data and information compiled and provided by the MOHLTC and the Canadian Institute of Health Information (CIHI). The analyses, conclusions, opinions, and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

B 1.7 References

1. Schulte PA. The use of biomarkers in surveillance, medical screening, and intervention. *Mutat Res - Fundam Mol Mech Mutagen*. 2005;592(1-2):155-163. doi:10.1016/j.mrfmmm.2005.06.019

2. Sahai I, Marsden D. Newborn screening. *Crit Rev Clin Lab Sci*. 2009;46(2):55-82. doi:10.1080/10408360802485305

3. Cheillan D, Vercherat M, Chevalier-Porst F, Charcosset M, Rolland MO, Dorche C. False-positive results in neonatal screening for cystic fibrosis based on a three-stage protocol (IRT/DNA/IRT): Should we adjust IRT cut-off to ethnic origin? *J Inherit Metab Dis*. 2005;28(6):813-818. doi:10.1007/s10545-005-0067-0

4. Mak CM, Lee H-CH, Chan AY-W, Lam C-W. Inborn errors of metabolism and expanded newborn screening: review and update. *Crit Rev Clin Lab Sci*. 2013;50(6):142-162. doi:10.3109/10408363.2013.847896

5. Secretary's Advisory Committee on Heritable Disorders in Newborns and Children (SACHDNC). Recommended Uniform Screening Panel [July 2018 Version]. Published 2019. Accessed December 2, 2019. https://www.hrsa.gov/advisory-committees/heritable-disorders/rusp/index.html

6. Levy PA. An overview of newborn screening. *J Dev Behav Pediatr*. 2010;31(7):622-631. doi:10.1097/DBP.0b013e3181eedf01

7. Wilson K, Kennedy S, Potter B, Geraghty M, Chakraborty P. Developing a national newborn screening strategy for Canada. *Health Law Rev*. 2010;18(2):31.

8. Varness TS, Allen DB, Hoffman GL. Newborn screening for congenital adrenal hyperplasia has reduced sensitivity in girls. *J Pediatr*. 2005;147(4):493-498. doi:10.1016/j.jpeds.2005.04.035

9. Chace DH, De Jesús VR, Lim TH, Hannon WH, Clark RH, Spitzer AR. Detection of TPN contamination of dried blood spots used in newborn and metabolic screening and its impact on quantitative measurement of amino acids. *Clin Chim Acta*. 2011;412(15-16):1385-1390. doi:10.1016/j.cca.2011.04.009

10. Ryckman KK, Berberich SL, Shchelochkov OA, Cook DE, Murray JC. Clinical and environmental influences on metabolic biomarkers collected for newborn screening. *Clin Biochem*. 2013;46(1-2):133-138. doi:10.1016/j.clinbiochem.2012.09.013

11. Ryckman KK, Dagle JM, Shchelochkov OA, et al. Association of amino acids with common complications of prematurity. *Pediatr Res*. 2013;73(6):700-705. doi:10.1038/pr.2013.43

12. Holtzman NA, Meek AG, Mellits ED. Neonatal screening for phenylketonuria. IV. Factors influencing the occurrence of false positives. *Am J Public Health*. 1974;64(8):775-779. doi:10.2105/ajph.64.8.755

13. Cai A, Qi S, Su Z, et al. A Pilot Metabolic Profiling Study of Patients With Neonatal Jaundice and Response to Phototherapy. *Clin Transl Sci*. 2016;9(4):216-220. doi:10.1111/cts.12401

14. Lauer BJ, Spector ND. Hyperbilirubinemia in the Newborn. *Pediatr Rev*. 2011;32(8):341-349. doi:10.1542/pir.32-8-341

15. Darling EK, Guttmann A, Sprague AE, Ramsay T, Walker MC. Implementation of the Canadian Paediatric Society's hyperbilirubinemia guidelines: A survey of Ontario hospitals. *Paediatr Child Heal*. 2014;19(3):133-137. doi:10.1093/pch/19.3.133

16. PCMCH. Clinical Pathway Handbook for Hyperbilirubinemia in Term and Late Pre-Term. Published online 2017.

17. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med*. 2015;12(10):1-22. doi:10.1371/journal.pmed.1001885

18. Fell DB, Hawken S, Wong CA, et al. Using newborn screening analytes to identify cases of neonatal sepsis. *Sci Rep*. 2017;7(1):18020. doi:10.1038/s41598-017-18371-1

19. Darling EK, Ramsay T, Sprague AE, Walker MC, Guttmann A. Universal bilirubin screening and health care utilization. *Pediatrics*. 2014;134(4):e1017-24. doi:10.1542/peds.2014-1146

20. Lalonde AN, Taylor E. Medical classification systems in Canada: Moving toward the year 2000. *Cmaj*. 1997;157(11):1561-1565.

21. Wilson K, Hawken S, Potter BK, et al. Accurate prediction of gestational age using newborn screening analyte data. *Am J Obstet Gynecol*. 2016;214(4):513.e1-513.e9. doi:10.1016/j.ajog.2015.10.017

22. Canadian Pediatric Society. Guidelines for detection, management and prevention of hyperbilirubinemia in term and late preterm newborn infants (35 or more weeks' gestation) - Summary. *Paediatr Child Health*. 2007;12(5):401-418. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2528770&tool=pmcentrez&rendertype=abstract

23. Wilcox R. Trimming and Winsorization. In: *Encyclopedia of Biostatistics*. John Wiley & Sons, Ltd; 2005. doi:10.1002/0470011815.b2a15165

24. Fisher RA. Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population. *Biometrika*. 1915;10(4):507. doi:10.2307/2331838

25. Huang Y, Pepe MS, Feng Z. Logistic regression analysis with standardized markers. *Ann Appl Stat*. 2013;7(3):1640-1662. doi:10.1214/13-AOAS634

26. Hayashi G, Faure C, Brondi MF, et al. Weight-adjusted neonatal 17OH-progesterone cutoff levels improve the efficiency of newborn screening for congenital adrenal hyperplasia. *Arq Bras Endocrinol Metabol*. 2011;55(8):632-637. Accessed March 14, 2016. http://www.ncbi.nlm.nih.gov/pubmed/22218447

27. Kwon C, Farrell PM. The magnitude and challenge of false-positive newborn screening test results. *Arch Pediatr Adolesc Med*. 2000;154(7):714-718. doi:10.1001/archpedi.154.7.714

28. Merritt RJ. Cholestasis associated with total parenteral nutrition. *J Pediatr Gastroenterol Nutr*. 1986;5(1):9-22. http://www.ncbi.nlm.nih.gov/pubmed/3080580

29. Hayashi GY, Carvalho DF, de Miranda MC, et al. Neonatal 17-hydroxyprogesterone levels adjusted according to age at sample collection and birthweight improve the efficacy of congenital adrenal hyperplasia newborn screening. *Clin Endocrinol (Oxf)*. 2017;86(4):480-487. doi:10.1111/cen.13292

30. Peng CYJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. *J Educ Res*. 2002;96(1):3-14. doi:10.1080/00220670209598786

31. Gauthier J, Wu Q V., Gooley TA. Cubic splines to model relationships between continuous variables and outcomes: a guide for clinicians. *Bone Marrow Transplant*. 2020;55(4):675-680. doi:10.1038/s41409-019-0679-x

32. Harrell FE. *Regression Modeling Strategies*. Springer International Publishing; 2015. doi:10.1007/978-3-319-19425-7

33. McCarthy ME, Oltman SP, Baer RJ, et al. Newborn Metabolic Profile Associated with Hyperbilirubinemia With and Without Kernicterus. *Clin Transl Sci*. 2019;12(1):28-38. doi:10.1111/cts.12590

34. Usher-Smith JA, Sharp SJ, Griffin SJ. The spectrum effect in tests for risk prediction, screening, and diagnosis. *BMJ*. 2016;353. doi:10.1136/bmj.i3139

35. Janes H, Pepe MS. Adjusting for covariates in studies of diagnostic, screening, or prognostic markers: An old concept in a new setting. *Am J Epidemiol*. 2008;168(1):89-97. doi:10.1093/aje/kwn099

36. Han L, Tavakoli NP, Morrissey M, Spink DC, Cao Z (Tim). Liquid chromatography-tandem mass spectrometry analysis of 17-hydroxyprogesterone in dried blood spots revealed matrix effect on immunoassay. *Anal Bioanal Chem*. 2019;411(2):395-402. doi:10.1007/s00216-018-1449-0

37. Wong T, Shackleton CH, Covey TR, Ellis G. Identification of the steroids in neonatal plasma that interfere with 17 alpha-hydroxyprogesterone radioimmunoassays. *Clin Chem*. 1992;38(9):1830-1837. http://www.ncbi.nlm.nih.gov/pubmed/1526021

38. Makela SK, Ellis G. Nonspecificity of a direct 17 alpha-hydroxyprogesterone radioimmunoassay kit when used with samples from neonates. *Clin Chem*. 1988;34(10):2070-2075. doi:10.1093/clinchem/34.10.2070

39. Mauvais FX, Gonzales E, Davit-Spraul A, Jacquemin E, Brauner R. Cholestasis reveals severe cortisol deficiency in neonatal pituitary stalk interruption syndrome. *PLoS One*. 2016;11(2):1-12. doi:10.1371/journal.pone.0147750

40. Lee WS, Lum LCS, Harun F. Addisonian-like crisis in congenital hypopituitarism and cholestatic jaundice. *Med J Malaysia*. 2003;58(2):279-281.

41. Al-Hussaini A, Almutairi A, Mursi A, Alghofely M, Asery A. Isolated cortisol deficiency: A rare cause of neonatal cholestasis. *Saudi J Gastroenterol*. 2012;18(5):339-341. doi:10.4103/1319-3767.101137

42. Ellaway CJ, Silink M, Cowell CT, et al. Cholestatic jaundice and congenital hypopituitarism. *J Paediatr Child Health*. 1995;31(1):51-53. doi:10.1111/j.1440-1754.1995.tb02914.x

43. Berry GT. *Disorders of Galactose Metabolism*. Fifth Edit. Elsevier Inc.; 2014. doi:10.1016/B978-0-12-410529-4.00055-3

44. Teodoro-Morrison T, Kyriakopoulou L, Chen YK, et al. Dynamic biological changes in metabolic disease biomarkers in childhood and adolescence: A CALIPER study of healthy community children. *Clin Biochem*. 2015;48(13-14):828-836. doi:10.1016/j.clinbiochem.2015.05.005

45. Smith EH, Matern D. Acylcarnitine Analysis by Tandem Mass Spectrometry. *Curr Protoc Hum Genet*. 2010;64(1):1-20. doi:10.1002/0471142905.hg1708s64

46. Cavedon CT, Bourdoux P, Mertens K, et al. Age-related variations in acylcarnitine and free carnitine concentrations measured by tandem mass spectrometry. *Clin Chem*. 2005;51(4):745-752. doi:10.1373/clinchem.2004.043646

47. Holub M, Tuschl K, Ratschmann R, et al. Influence of hematocrit and localisation of punch in dried blood spots on levels of amino acids and acylcarnitines measured by tandem mass spectrometry. *Clin Chim Acta*. 2006;373(1-2):27-31. doi:10.1016/j.cca.2006.04.013

48. Kaplan M, Muraca M, Hammerman C, et al. Imbalance between production and conjugation of bilirubin: a fundamental concept in the mechanism of neonatal jaundice. *Pediatrics*. 2002;110(4):47-51. doi:10.1542/peds.110.4.e47

49. Cohen RS, Wong RJ, Stevenson DK. Understanding neonatal jaundice: A perspective on causation. *Pediatr Neonatol*. 2010;51(3):143-148. doi:10.1016/S1875-9572(10)60027-7

50. Chen X, Li X, Chang Y, Li W, Cui H. Effect and safety of timing of cord clamping on neonatal hematocrit values and clinical outcomes in term infants: A randomized controlled trial. *J Perinatol*. 2018;38(3):251-257. doi:10.1038/s41372-017-0001-y

51. Ministry of Health. Moms and Babies to Benefit From Breastfeeding Supports. Published 2013. Accessed October 4, 2020. https://news.ontario.ca/mohltc/en/2013/09/moms-and-babies-to-benefit-from-new-breastfeeding-supports.html

52. Baby-Friendly Initiative Ontario. Baby-Friendly Initiative Report for Ontario. Published 2019. Accessed October 4, 2020. http://www.bfiontario.ca/wp-content/uploads/2019/09/BFION-BFI-Report-2019-April-FINAL.pdf

53. Gourley GR. Breast-feeding, neonatal jaundice and kernicterus. *Semin Neonatol*. 2002;7(2):135-141. doi:10.1053/siny.2002.0101

54. Bertini G, Dani C, Tronchin M, Rubaltelli FF. Is breastfeeding really favoring early neonatal jaundice? *Pediatrics*. 2001;107(3). doi:10.1542/peds.107.3.e41

55. Kyriacou DN, Lewis RJ. Confounding by Indication in Clinical Research. *JAMA*. 2016;316(17):1818. doi:10.1001/jama.2016.16435

56. Lin JC, Strauss RG, Kulhavy JC, et al. Phlebotomy overdraw in the neonatal intensive care nursery. *Pediatrics*. 2000;106(2). doi:10.1542/peds.106.2.e19

57. Potter DM, Butterfield LH, Divito SJ, Sander CA, Kirkwood JM. Pitfalls in retrospective analyses of biomarkers: A case study with metastatic melanoma patients. *J Immunol Methods*. 2012;376(1-2):108-112. doi:10.1016/j.jim.2011.12.003

58. Van Imhoff DE, Dijk PH, Weykamp CW, Cobbaert CM, Hulzebos C V. Measurements of neonatal bilirubin and albumin concentrations: A need for improvement and quality control. *Eur J Pediatr*. 2011;170(8):977-982. doi:10.1007/s00431-010-1383-4

59. Schreiner RL, Glick MR. Interlaboratory Variability. *Pediatrics*. 1982;69(3):277-281.

60. White E. Measurement error in biomarkers: sources, assessment, and impact on studies. *IARC Sci Publ*. 2011;(163):143-161.

61. Zetterström RH, Karlsson L, Falhammar H, Lajic S, Nordenström A. Update on the Swedish newborn screening for congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Int J Neonatal Screen*. 2020;6(3):2-8. doi:10.3390/IJNS6030071

62. Knowles RL, Oerton J, Cheetham T, et al. Newborn screening for primary congenital hypothyroidism: Estimating test performance at different tsh thresholds. *J Clin Endocrinol Metab*. 2018;103(10):3720-3728. doi:10.1210/jc.2018-00658

63. Chace DH, Sherwin JE, Hillman SL, Lorey F, Cunningham GC. Use of phenylalanine-to-tyrosine ratio determined by tandem mass spectrometry to improve newborn screening for phenylketonuria of early discharge specimens collected in the first 24 hours. *Clin Chem*. 1998;44(12):2405-2409. http://www.ncbi.nlm.nih.gov/pubmed/9836704

64. Gloria-Bottini F, Magrini A, Cozzoli E, Bergamaschi A, Bottini E. ADA genetic polymorphism and the effect of smoking on neonatal bilirubinemia and developmental parameters. *Early Hum Dev*. 2008;84(11):739-743. doi:10.1016/j.earlhumdev.2008.05.001

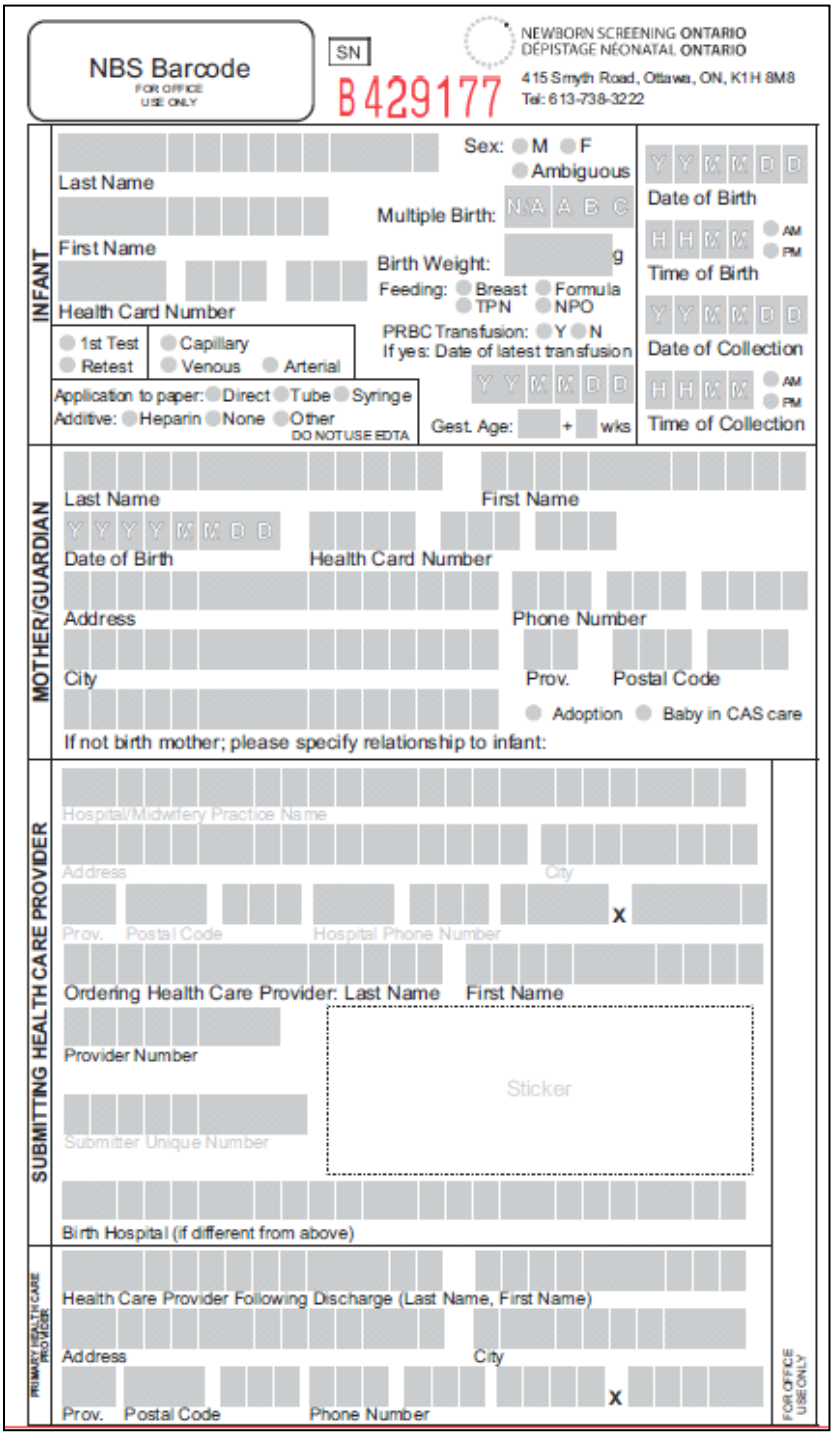
65. Watchko JF. Identification of Neonates at Risk for Hazardous Hyperbilirubinemia: Emerging Clinical Insights. *Pediatr Clin North Am*. 2009;56(3):671-687. doi:10.1016/j.pcl.2009.04.005

66. Nnorom O, Findlay N, Lee-Foon NK, et al. Dying to learn: A scoping review of breast and cervical cancer studies focusing on black canadian women. *J Health Care Poor Underserved*. 2019;30(4):1331-1359. doi:10.1353/hpu.2019.0100

67. Canadian Institute for Health Information. *Proposed Standards for Race-Based and Indigenous Identity Data Collection and Health Reporting in Canada*. CIHI; 2020. https://www.cihi.ca/en/proposed-standards-for-race-based-and-indigenous-identity-data

B 1.8 Supplementary Materials

S1. Newborn Screening for Rare Disease Requisition



S2. Pediatrician-Provided Diagnosis Categorization Table

	Newborn Screening Outcome				
Referred for Diagnostics on Suspicion of:	**TP** True Positive	**VD** Variant Disease	**FP** False Positive	**IF** Incidental Finding	**OT** Other
CAH Congenital Adrenal Hyperplasia	• 21-OH Deficiency: Salt-Wasting • 21-OH Deficiency: Simple Virilizing	Deficiencies of: • 11-Hydroxylase • 3β-hydroxysteroid dehydrogenase	• Not Affected	• Any panel-targeted disorder other than that initially suspected • CAH Carriers • Persistent Laboratory Abnormalities Inconsistent with CAH (Neonate *or* Mother)	↓ Applicable to All ↓
CH Congenital Hypothyroidism	• Dysgenesis: Ectopic • Dysgenesis: Athyrotic • Dysgenesis: Hypoplastic • Dyshormogenesis • Presumed dysgenesis or dyshormogenesis but no imaging done • Undetermined/Idiopathic with either: i) TSH ≥ 30 ; *or* ii) FT4 ≤ 10; *or* iii) abnormal thyroid imaging	Undetermined Idiopathic • TSH ≥ 30 or FT4 ≤ 10 with normal scan Undetermined Idiopathic Subclinical • TSH ≥30 and normal FT4 and normal scan	• Not Affected	• Any panel-targeted disorder other than that initially suspected • Iodine exposure • Prematurity • Maternal propylthiouracil exposure • Maternal autoimmune thyroid disease	• Parental non-consent to confirmatory diagnostics • Neonate deceased prior to confirmatory diagnostics • Interprovincial transfer of care: confirmatory diagnostics performed in another province • International transfer of care: confirmatory diagnostics performed in foreign country • Lost to follow-up
PKU Phenylketonuria	• Biopterin Responsive PKU • Classical PKU (Phe ≥ 1200 µmol/L) • Moderate PKU (Phe 900 – 1199 µmol/L) • Mild PKU (Phe 600 – 899 µmol/L) • Mild Hyperphenylalaninemia (Phe 360 – 599 µmol/L)	• Mild Hyperphenylalaninemia (Phe 120 – 359 µmol/L)	• Not Affected	• Any panel-targeted disorder other than that initially suspected • Liver disease or dysfunction • Persistent Laboratory Abnormalities Inconsistent with PKU (Neonate *or* Mother) • Maternal PKU Tetrahydrobiopterin Deficiencies or Variant Deficiencies of: • Dihydropteridine reductase • GTP Cyclohydrolyase • Pterin-4α-carbinolamine dehydratase • Sepiapterin reductase • 6-Pyruvoyl-Tetrahydropterin Synthase	

S3. Repository-Reported Screening Participation Rates among Ontario-Resident Newborns

		Birth Year										
	Birth Month	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Total Serum Bilirubin Testing	January	<1%	<1%	<1%	1%	1%	1%	2%	10%	16%	16%	25%
	February	<1%	<1%	<1%	1%	0%	4%	3%	11%	15%	16%	25%
	March	<1%	<1%	<1%	1%	1%	3%	3%	11%	15%	17%	25%
	April	<1%	<1%	<1%	1%	1%	2%	3%	10%	14%	16%	27%
	May	<1%	<1%	<1%	1%	0%	2%	3%	11%	15%	17%	27%
	June	<1%	<1%	<1%	1%	1%	2%	3%	12%	14%	17%	27%
	July	<1%	<1%	<1%	1%	1%	2%	3%	16%	15%	19%	27%
	August	<1%	<1%	<1%	1%	1%	1%	3%	15%	15%	19%	28%
	September	<1%	<1%	<1%	1%	1%	1%	3%	15%	15%	20%	28%
	October	<1%	<1%	<1%	1%	1%	2%	4%	16%	15%	21%	28%
	November	<1%	<1%	1%	1%	1%	2%	8%	16%	15%	22%	29%
	December	<1%	<1%	1%	1%	1%	2%	9%	16%	16%	25%	29%
Dried Blood Spot NBS	January	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	February	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	March	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	April	>99%	98%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	May	>99%	98%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	June	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	July	>99%	99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	August	>99%	99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	September	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	October	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	November	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%
	December	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%	>99%

Note: The denominator for each percentage is Ontario-resident neonates with a birth record in the Discharge Abstract Database [DAD]
The numerator for each percentage is Ontario-resident neonates with a screening record

S4. Intervals between Dried Blood Spot and Total Serum Bilirubin Screening

	Days Between NBS and TSB	Total Serum Bilirubin Observations in First Postnatal Week					
		Single TSB Neonates		≥ 2 TSB Neonates ± 1 Day from NBS		≥ 2 TSB Neonates ≥ 2 Days from NBS	
		(*n* = 64,334)		(*n* = 31,894)		(*n* = 1,909)	
		n	%	n	%	n	%
All Participants	**0**	40,594	63.1%	23,354	73.2%	N/A	
	1	18,824	29.3%	8,540	26.8%		
(N = **98,137**)	**2**	1,371	2.1%	N/A		782	41.0%
	3	1,284	2.0%			584	30.6%
	4	969	1.5%			350	18.3%
	5	742	1.2%			168	8.8%
	6	547	0.9%			25	1.3%
	7 *	≤ 5	0.0%			-	-
		(*n* = 44,862)		(*n* = 21,466)		(*n* = 165)	
		n	%	n	%	n	%
Participants Tested for TSB in Universal Context	**0**	31,886	71.1%	16,133	75.2%	N/A	
	1	12,697	28.3%	5,333	24.8%		
(N = **66,493**)	**2**	121	0.3%	N/A		108	65.5%
	3	67	0.1%			37	22.4%
	4	49	0.1%			11	6.7%
	5	31	0.1%			8	4.8%
	6	11	0.0%			1	0.6%
	7	-				-	-

Note: Each column corresponds to a branch of the TSB summarization algorithm described in the Methods section. Multi-TSB neonates are listed based on TSB test nearest in time to the NBS.

* In rare circumstances such as prior to a transfusion, a sample may be satisfactory even if collected on the 0th day of life. Neonates with a satisfactory 0-day NBS with a TSB taken at the upper limit of TSB inclusion (7 postnatal days) will have a 7-day difference between tests.

S5. Demographic Characteristics of Neonates Verified Screened in Universal Setting

Number of Neonates	**66,493**	
	n	**%**
Sex		
Male	34,239	51.5%
Female	32,254	48.5%
Phototherapy in First Week		
Received	3,974	6.0%
Not Received	62,519	94.0%
NICU Admission in First Week		
Admitted	7,238	10.9%
Not Admitted	59,255	89.1%
Feeding Status		
Exclusive Breastfeeding	43,387	65.3%
Other	23,106	34.8%
Any TPN Administered?		
Yes	792	1.2%
No	65,701	98.8%
Birthweight Category		
Normal Birthweight[>2,500g]	61,530	92.5%
Low Birthweight[<2,500g]	4,133	6.2%
Very Low Birthweight[<1,500g]	493	0.7%
Extremely Low Birthweight[<1,000g]	337	0.5%
Missing BW	-	0.0%
Gestational Age at Birth		
Term [>=38 Wk]	55,352	83.2%
Moderate to Late Preterm[32.1-37.6Wk]	10,234	15.4%
Very Preterm[28-32Wk]	581	0.9%
Extreme Preterm [<28Wk]	326	0.5%
Missing Gestational Age	-	0.0%
Proportion of Monthly Institution-Specific Resident Births Tested for TSB (%)		
90-100%	66,493	100%
80-89%	Not Applicable: Subgroup Eligibility Requirements	
70-79%		
60-69%		
50-59%		
40-49%		
30-39%		
20-29%		
10-19%		
< 9.9%		
Institution Not Available		
Age at Dried Blood Spot, Hours Median [25th-75th Centiles]	25 [24 - 28]	
Age at Total Serum Bilirubin, Days Median [25th-75th Centiles]	1 [1 - 2]	

S6. Sequentially Adjusted Spearman Correlations in All Study Participants **(*n* =98,137)**

NBS Biomarker	Spearman *r*							
	Direct		Partial					
	Crude		Prematurity [1] Adjusted		Prematurity [1] and NBS Sampling Age Adjusted		Prematurity, NBS Sampling Age, and Other Biomarker [2] Adjusted	
	r	95 CI	*r*	95 CI	*r*	95 CI	*r*	95 CI
TYR	0.16	(0.15,0.16)	0.14	(0.14,0.15)	0.16	(0.15,0.16)	0.13	(0.12,0.13)
C14	0.15	(0.14,0.15)	0.14	(0.14,0.15)	0.15	(0.14,0.16)	0.02	(0.01,0.03)
C16	0.14	(0.13,0.14)	0.15	(0.14,0.15)	0.15	(0.14,0.15)	0.03	(0.02,0.03)
C18:1	0.14	(0.13,0.15)	0.14	(0.13,0.15)	0.12	(0.11,0.12)	0.01	(0.00,0.01)
GALT	0.12	(0.11,0.12)	0.13	(0.12,0.13)	0.13	(0.13,0.14)	0.06	(0.06,0.07)
C18	0.12	(0.11,0.12)	0.12	(0.11,0.12)	0.12	(0.11,0.12)	0.00	(0.00,-0.01)
C2	0.12	(0.11,0.13)	0.12	(0.12,0.13)	0.11	(0.10,0.11)	0.00	(0.00,-0.01)
C12	0.12	(0.11,0.12)	0.12	(0.11,0.13)	0.13	(0.13,0.14)	0.02	(0.02,0.03)
C14:1	0.12	(0.11,0.12)	0.12	(0.11,0.12)	0.12	(0.12,0.13)	0.02	(0.01,0.02)
C4OH	0.12	(0.11,0.12)	0.12	(0.12,0.13)	0.10	(0.10,0.11)	0.02	(0.02-0.03)
LEU	0.11	(0.10,0.11)	0.09	(0.08,0.10)	0.05	(0.04,0.05)	-0.05	(-0.05, -0.06)
C3	0.09	(0.08,0.09)	0.08	(0.07,0.08)	0.09	(0.08,0.10)	0.04	(0.03,0.05)
PHE	0.08	(0.08,0.09)	0.08	(0.07,0.08)	0.12	(0.12,0.13)	0.05	(0.05,0.06)
C5OH	0.08	(0.07,0.08)	0.07	(0.06,0.08)	0.07	(0.07,0.08)	0.01	(0.00,0.02)
C4	0.07	(0.07,0.08)	0.06	(0.06,0.07)	0.07	(0.07,0.08)	0.01	(0.01,0.02)
SUAC	0.07	(0.06,0.08)	0.07	(0.06,0.08)	0.06	(0.05,0.06)	0.03	(0.02,0.03)
GLY	0.06	(0.06,0.07)	0.07	(0.07,0.08)	0.09	(0.08,0.09)	0.03	(0.03,0.04)
C5	0.06	(0.06,0.07)	0.04	(0.03,0.04)	0.04	(0.03,0.05)	-0.02	(-0.01, -0.03)
ARG	0.06	(0.05,0.06)	0.05	(0.05,0.06)	0.04	(0.03,0.05)	0.06	(0.05,0.07)
C0	0.06	(0.05,0.06)	0.05	(0.05,0.06)	0.06	(0.06,0.07)	-0.02	(-0.01, -0.03)
C16OH	0.06	(0.05,0.06)	0.06	(0.06,0.07)	0.07	(0.07,0.08)	0.01	(0.00,0.01)
C14OH	0.05	(0.04,0.06)	0.05	(0.05,0.06)	0.07	(0.06,0.08)	0.01	(0.01,0.02)
C18OH	0.05	(0.05,0.06)	0.05	(0.05,0.06)	0.05	(0.05,0.06)	0.01	(0.01,0.02)
MET	0.05	(0.05,0.06)	0.05	(0.04,0.06)	0.10	(0.09,0.10)	0.03	(0.03,0.04)
C18:1OH	0.05	(0.05,0.06)	0.05	(0.05,0.06)	0.06	(0.05,0.06)	0.01	(0.00,0.01)
C10	0.05	(0.05,0.06)	0.06	(0.05,0.07)	0.06	(0.06,0.07)	0.00	(0.00,0.01)
C8	0.04	(0.03,0.04)	0.04	(0.03,0.04)	0.04	(0.03,0.04)	0.00	(0.00,-0.01)
VAL	0.04	(0.03,0.05)	0.04	(0.03,0.05)	0.04	(0.03,0.05)	0.04	(0.03,0.05)
C14:2	0.04	(0.04,0.05)	0.03	(0.03,0.04)	0.03	(0.03,0.04)	-0.01	(0.00,-0.01)
C18:2	0.04	(0.04,0.05)	0.03	(0.03,0.04)	0.03	(0.02,0.03)	0.00	(0.00,0.00)
C3DC	0.04	(0.04,0.05)	0.05	(0.04,0.05)	0.06	(0.05,0.06)	0.02	(0.01,0.02)
C12:1	0.05	(0.04,0.05)	0.06	(0.05,0.06)	0.08	(0.07,0.08)	-0.02	(-0.02, -0.03)
C5DC	0.04	(0.04,0.05)	0.04	(0.03,0.04)	0.05	(0.04,0.05)	0.00	(-0.01,0.01)
ASA	0.04	(0.03,0.04)	0.03	(0.03,0.04)	0.03	(0.02,0.03)	0.02	(0.01,0.03)
C16:1-OH	0.03	(0.03,0.04)	0.04	(0.03,0.04)	0.05	(0.05,0.06)	-0.03	(-0.02, -0.03)
C6	0.03	(0.03,0.04)	0.04	(0.03,0.04)	0.04	(0.04,0.05)	0.00	(0.00,-0.01)
ALA	0.03	(0.03,0.04)	0.05	(0.04,0.06)	0.07	(0.06,0.07)	0.04	(0.04,0.05)
C5:1	0.01	(0.00,0.02)	0.01	(0.00,0.01)	0.02	(0.01,0.02)	0.00	(0.00,-0.01)
C10:1	0.00	(0.00,-0.01)	-0.01	(0.00,-0.01)	-0.02	(-0.01,-0.03)	-0.01	(0.00,-0.02)
C6DC	0.00	(0.00,-0.01)	0.01	(0.00,0.01)	-0.02	(-0.01,-0.03)	-0.06	(-0.05, -0.07)
C4DC	-0.03	(-0.02, -0.03)	0.00	(0.00,-0.01)	0.01	(0.00,0.01)	-0.07	(-0.06, -0.07)
ORN	-0.03	(-0.03, -0.04)	-0.03	(-0.02, -0.04)	-0.05	(-0.05, -0.06)	-0.10	(-0.09, -0.11)
TREC	-0.03	(-0.03, -0.04)	-0.02	(-0.01, -0.02)	0.00	(0.00,-0.01)	-0.04	(-0.03, -0.04)
CIT	-0.04	(-0.04, -0.05)	-0.03	(-0.03, -0.04)	-0.01	(0.00,-0.02)	-0.03	(-0.03, -0.04)
IRT	-0.05	(-0.05, -0.06)	-0.05	(-0.05, -0.06)	-0.05	(-0.04, -0.05)	-0.02	(-0.01, -0.03)
BIOT	-0.05	(-0.04, -0.05)	-0.04	(-0.04, -0.05)	-0.07	(-0.07, -0.08)	-0.06	(-0.06, -0.07)
C8:1	-0.06	(-0.05, -0.07)	-0.06	(-0.06, -0.07)	-0.07	(-0.07, -0.08)	-0.06	(-0.05, -0.06)
TSH	-0.12	(-0.12, -0.13)	-0.12	(-0.11, -0.12)	-0.04	(-0.03, -0.05)	-0.03	(-0.02, -0.04)
17-OHP	-0.12	(-0.11, -0.13)	-0.16	(-0.15, -0.17)	-0.10	(-0.09, -0.11)	-0.12	(-0.11, -0.12)

Biomarkers listed in the **same order as in the main article text**

[1] Effect of birthweight and gestational age removed
[2] Effect of birthweight, gestational age, NBS sampling age, and all other NBS markers removed

S7. Sequentially Adjusted Spearman Correlations in Universal Context Participants Only (*n* =66,493)

NBS Biomarker	Spearman *r*							
	Direct		Partial					
	Crude		Prematurity [1] Adjusted		Prematurity [1] and NBS Sampling Age Adjusted		Prematurity, NBS Sampling Age, and Other Biomarker [2] Adjusted	
	r	95 CI	*r*	95 CI	*r*	95 CI	*r*	95 CI
TYR	0.15	(0.14,0.15)	0.14	(0.13,0.14)	0.14	(0.14,0.15)	0.11	(0.11,0.12)
C14	0.16	(0.15,0.16)	0.15	(0.15,0.16)	0.15	(0.14,0.16)	0.02	(0.01,0.02)
C16	0.16	(0.15,0.17)	0.17	(0.16,0.18)	0.16	(0.15,0.17)	0.03	(0.02,0.04)
C18:1	0.16	(0.15,0.16)	0.16	(0.15,0.17)	0.13	(0.13,0.14)	0.01	(0.00,0.02)
GALT	0.16	(0.15,0.17)	0.17	(0.16,0.17)	0.17	(0.16,0.17)	0.09	(0.09,0.10)
C18	0.13	(0.13,0.14)	0.14	(0.13,0.14)	0.13	(0.12,0.14)	-0.01	(0.00,-0.02)
C2	0.15	(0.14,0.15)	0.15	(0.14,0.16)	0.13	(0.12,0.14)	0.00	(0.00,0.01)
C12	0.13	(0.12,0.14)	0.13	(0.13,0.14)	0.13	(0.12,0.14)	0.03	(0.03,0.04)
C14:1	0.11	(0.11,0.12)	0.11	(0.11,0.12)	0.11	(0.10,0.12)	0.02	(0.01,0.03)
C4OH	0.11	(0.11,0.12)	0.12	(0.11,0.12)	0.09	(0.09,0.10)	0.01	(0.00,0.02)
LEU	0.11	(0.10,0.12)	0.10	(0.09,0.11)	0.07	(0.06,0.07)	-0.05	(-0.04, -0.06)
C3	0.10	(0.09,0.11)	0.09	(0.09,0.10)	0.11	(0.10,0.11)	0.04	(0.03,0.05)
PHE	0.11	(0.10,0.11)	0.10	(0.09,0.11)	0.14	(0.14,0.15)	0.06	(0.05,0.07)
C5OH	0.10	(0.09,0.11)	0.09	(0.09,0.10)	0.10	(0.09,0.10)	0.01	(0.01,0.02)
C4	0.08	(0.07,0.09)	0.07	(0.06,0.08)	0.08	(0.07,0.09)	0.01	(0.00,0.01)
SUAC	0.07	(0.07,0.08)	0.07	(0.07,0.08)	0.06	(0.05,0.07)	0.02	(0.02,0.03)
GLY	0.09	(0.08,0.09)	0.09	(0.08,0.10)	0.10	(0.09,0.10)	0.03	(0.03,0.04)
C5	0.07	(0.06,0.07)	0.05	(0.04,0.06)	0.06	(0.05,0.06)	-0.02	(-0.01, -0.03)
ARG	0.06	(0.05,0.07)	0.05	(0.05,0.06)	0.05	(0.04,0.05)	0.07	(0.06,0.08)
C0	0.08	(0.07,0.09)	0.08	(0.07,0.08)	0.09	(0.08,0.09)	-0.02	(-0.01, -0.03)
C16OH	0.06	(0.05,0.07)	0.06	(0.05,0.07)	0.06	(0.06,0.07)	0.01	(0.00,0.01)
C14OH	0.05	(0.04,0.06)	0.05	(0.04,0.06)	0.06	(0.05,0.07)	0.01	(0.01,0.02)
C18OH	0.05	(0.04,0.06)	0.05	(0.04,0.06)	0.05	(0.04,0.06)	0.01	(0.00,0.02)
MET	0.07	(0.06,0.08)	0.07	(0.06,0.08)	0.11	(0.11,0.12)	0.03	(0.02,0.04)
C-18:1OH	0.06	(0.05,0.06)	0.06	(0.05,0.06)	0.05	(0.05,0.06)	0.01	(0.00,0.01)
C10	0.06	(0.05,0.06)	0.06	(0.05,0.07)	0.06	(0.05,0.06)	0.01	(0.00,0.02)
C8	0.04	(0.03,0.05)	0.04	(0.03,0.05)	0.04	(0.03,0.04)	0.00	(0.01,-0.01)
VAL	0.04	(0.03,0.05)	0.04	(0.03,0.05)	0.04	(0.03,0.05)	0.00	(0.00,-0.01)
C14:2	0.04	(0.03,0.05)	0.03	(0.02,0.04)	0.03	(0.02,0.03)	-0.01	(0.00,-0.02)
C18:2	0.06	(0.06,0.07)	0.05	(0.05,0.06)	0.05	(0.05,0.06)	-0.01	(0.00,-0.01)
C3DC	0.05	(0.04,0.05)	0.04	(0.04,0.05)	0.05	(0.04,0.05)	0.02	(0.01,0.03)
C12:1	0.04	(0.03,0.05)	0.04	(0.03,0.05)	0.05	(0.04,0.06)	-0.03	(-0.03, -0.04)
C5DC	0.03	(0.02,0.04)	0.02	(0.02,0.03)	0.03	(0.02,0.03)	-0.01	(0.00,-0.02)
ASA	0.03	(0.03,0.04)	0.03	(0.02,0.04)	0.03	(0.02,0.03)	0.02	(0.01,0.03)
C16:1-OH	0.04	(0.04,0.05)	0.05	(0.04,0.05)	0.06	(0.05,0.07)	-0.03	(-0.02, -0.04)
C6	0.05	(0.04,0.05)	0.05	(0.04,0.06)	0.05	(0.04,0.06)	0.00	(0.01,-0.01)
ALA	0.07	(0.06,0.08)	0.08	(0.07,0.09)	0.10	(0.09,0.11)	0.06	(0.05,0.07)
C5:1	0.01	(0.00,0.02)	0.01	(0.00,0.01)	0.01	(0.01,0.02)	-0.01	(0.00,-0.01)
C10:1	0.00	(0.00,-0.01)	-0.01	(0.00,-0.02)	-0.02	(-0.01, -0.03)	-0.01	(0.00,-0.02)
C6DC	-0.02	(-0.01, -0.02)	-0.01	(0.00,-0.02)	-0.04	(-0.03, -0.05)	-0.06	(-0.05, -0.07)
C4DC	0.00	(0.00,0.01)	0.02	(0.01,0.03)	0.02	(0.02,0.03)	-0.06	(-0.05, -0.07)
ORN	-0.03	(-0.03, -0.04)	-0.03	(-0.03, -0.04)	-0.05	(-0.05, -0.06)	-0.11	(-0.11, -0.12)
TREC	-0.02	(-0.01, -0.03)	-0.01	(0.00,-0.02)	0.00	(0.00,-0.01)	-0.03	(-0.03, -0.04)
CIT	-0.03	(-0.02, -0.03)	-0.02	(-0.02, -0.03)	-0.01	(0.00,-0.01)	-0.03	(-0.02, -0.04)
IRT	-0.06	(-0.06, -0.07)	-0.06	(-0.06, -0.07)	-0.06	(-0.05, -0.06)	-0.03	(-0.02, -0.03)
BIOT	-0.06	(-0.06, -0.07)	-0.06	(-0.05, -0.07)	-0.09	(-0.08, -0.09)	-0.08	(-0.07, -0.09)
C8:1	-0.07	(-0.06, -0.07)	-0.07	(-0.06, -0.07)	-0.08	(-0.07, -0.08)	-0.06	(-0.05, -0.07)
TSH	-0.12	(-0.11, -0.13)	-0.12	(-0.11, -0.12)	-0.05	(-0.04, -0.06)	-0.03	(-0.02, -0.04)
17-OHP	-0.14	(-0.14, -0.15)	-0.17	(-0.17, -0.18)	-0.12	(-0.11, -0.13)	-0.14	(-0.13, -0.14)

Biomarkers listed in the **same order as in the main article text**

[1] Effect of birthweight and gestational age removed
[2] Effect of birthweight, gestational age, NBS sampling age, and all other NBS markers removed

2.0 Discussion

2.1 Principal Findings

This two-article thesis began with a systematic review aiming to establish the current state of the published literature regarding potential associations between biomarkers of rare disease and bilirubin or hyperbilirubinemia status (Figure 2-1). A small number (n = 16) of studies, most of which were deemed individually at low risk of bias within the inherent constraints of their designs, were identified describing predominantly unadjusted associations as well as lack of association between bilirubin levels/hyperbilirubinemia status and certain hormones, amino acids, acylcarnitines, and the presence of congenital heart disease. Despite substantial heterogeneity among the samples, methods, and main findings of these studies, the review concluded that associations between bilirubin and neonatal biomarkers indicative of rare disease have been reported in the past, particularly among markers used to screen for endocrine diseases. The review speculated on potential mechanisms of association observed in the included studies and concluded additional data analysis associating neonatal total serum bilirubin and markers of rare disease was indicated.

A cross-sectional study was subsequently developed to investigate the hypothesis that newborn screening biomarkers associate with total serum bilirubin (TSB), as measured for the purposes of universal pre-discharge hyperbilirubinemia screening. Enabled in part by the implementation of a provincially centralized laboratory result repository in 2007, our study identified 98,137 neonates undergoing contemporaneous, and in a majority of cases, universal screening for hyperbilirubinemia and rare disease. In correlation analyses, statistically significant Spearman correlations were observed across 33 of 49 screening markers after adjustment for factors known to influence dried blood spot marker measurements such as birthweight, gestational age, and postnatal age at blood sampling. Substantial differences in standardized analyte medians were observed across tercile-based bilirubin strata in screening biomarkers with higher correlation to TSB, corroborating the findings from correlation analysis. The most notable findings among bilirubin-associated screening biomarkers included elevations of tyrosine, certain short-chain and longer-chain acylcarnitines, GALT enzyme, and decreases in 17-OHP. In a sensitivity analysis restricted only to participants we could confirm received TSB tests in a universal screening setting, subtle differences of correlation coefficients were reported but these did not alter the interpretation of the results. This report is the second study to report NBS biomarker differences in association with hyperbilirubinemia status or TSB level.

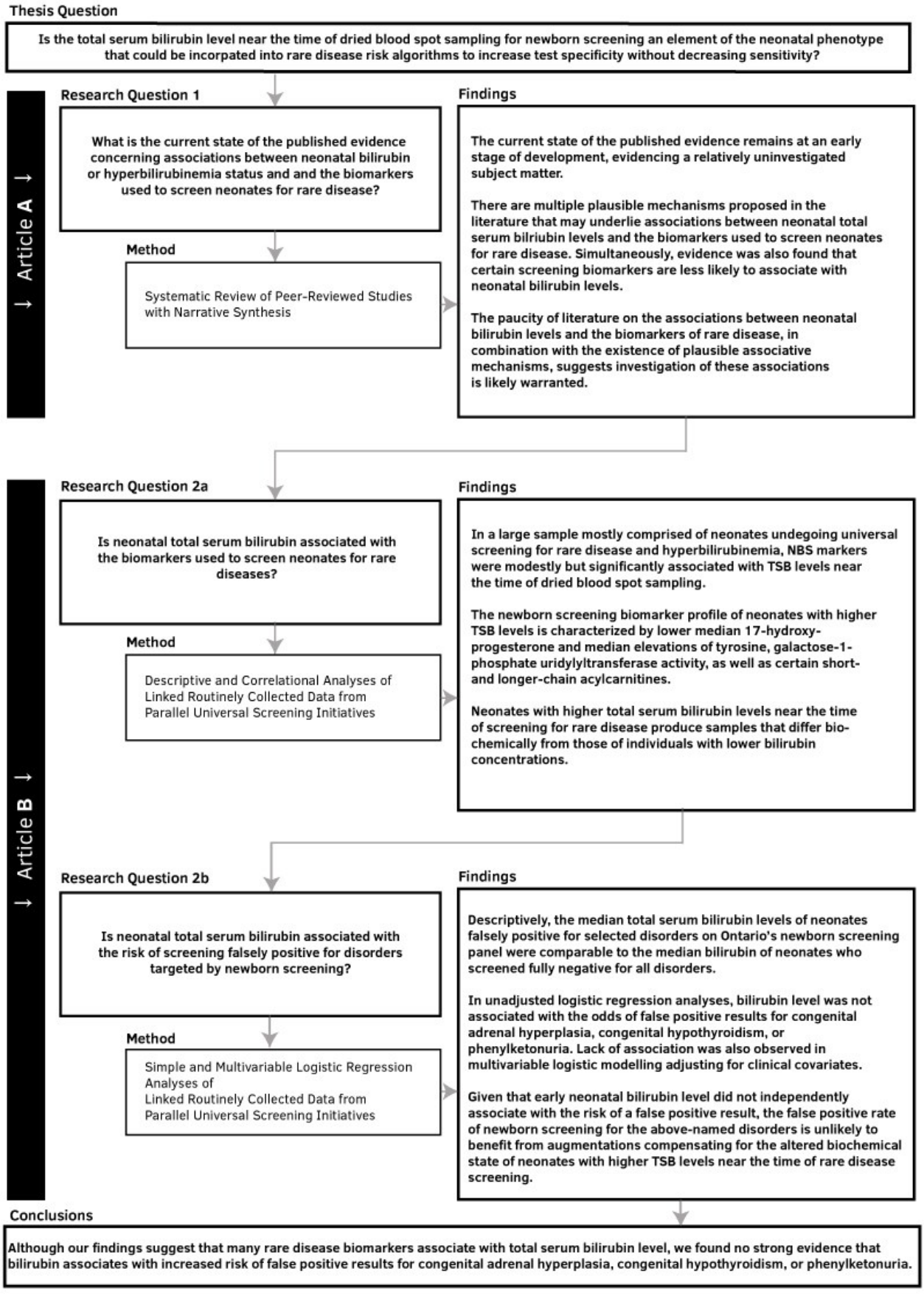


Figure 2-1. Visual Summary of Thesis Investigation

Multivariable logistic regression models were developed to estimate the independent association of total serum bilirubin (exposure) with the odds of a false positive screening determination (outcome present) versus a screening determination negative for all screened diseases (outcome absent) across three disorders with high rates of false positive findings. Each model included sex, breastfeeding status, parenteral nutrition administration status, birthweight, gestational age, and postnatal age at sampling as covariates with the latter three variables modelled as splines. We found that TSB was not significantly associated with the odds of a false positive screening result for congenital adrenal hyperplasia, congenital hypothyroidism, or phenylketonuria. Given that the risk of false positive results was not observed to differ across TSB levels in our sample, we suggest that the incorporation of TSB levels into newborn screening risk algorithms may not be an effective strategy to improve screening test specificity for these diseases.

2.2 Integrative Synthesis

2.2.1 Enabling Precision in Newborn Screening

In the past decade, enthusiasm within the healthcare community has grown for all aspects of precision medicine [97]. The central theme of precision health initiatives is the personalization of prevention and treatment strategies in a manner that considers individual variability [98]. Precision approaches to neonatal rare disease detection have been termed ‘precision newborn screening’. A consensus definition of properties characteristic of precision screening is not yet available, however researchers have attached this term to disease screening methods that are extremely specific and feature near-zero false positive rates while retaining near-perfect sensitivity (99%-100%) and thus also missing few cases of rare treatable diseases [99]. It could be inferred from this that a hallmark of precision newborn screening is high specificity alongside high sensitivity, i.e., screening that precisely directs additional healthcare to those who need it and only to those who need it, therefore sparing individuals without targeted illness unnecessary diagnostics or health professional encounters. As new treatable disorders are added to newborn screening panels, there is potential for important clinical benefits for children with the newly targeted rare diseases but also an inevitable increase in the yearly number of neonates receiving false positive determinations [100]. Thus, the minimization of false positive screening results is both an activity of immediate relevance to reducing screening-related psychosocial harms [101] as well as an effort to preserve the pediatric health system capacity required for further panel expansion.

Proposed approaches to achieving improved screening precision are now becoming available in the newborn screening literature [102]. Generally, healthcare providers caring for neonates born in a universal newborn screening jurisdiction submit two main sources of biological information to centralized newborn screening programs for rare disease detection purposes: 1) a blood spot sample and; 2) a completed screening information card. Consistent with newborn screening's laboratory-based origins [66], many research initiatives working to improve screening performance have focused on the blood spot sample as a source of more specific disease risk signals. Presently, genetic analysis from neonatal dried blood spot is among the most actively-researched precision screening strategies and most genetic professionals anticipate the imminent incorporation of sequencing technologies to newborn screening [103]. The substantial gains in precision anticipated from the application of genomics to newborn screening may have resulted in perceptions that sequencing DNA is the only technique that can facilitate increased screening precision. For example, a 2017 academic book chapter titled 'Newborn Screening in the Era of Precision Medicine' discusses only whole genome and exome sequencing as routes to increased screening precision [104]. Although future NBS will almost certainly utilize genomic technologies for screening, several challenges exist to the application of sequencing to newborn screening activities. Notable outstanding issues include the prolongation of screening turnaround times, as well as the detection of variants of unknown significance or untargeted later-onset disorders [105]. Until genetic technologies can be refined to meet the time-sensitive and targeted needs of newborn screening processes, it may be worthwhile to consider complementary precision strategies that improve the specificity of biochemical newborn screening without extending screening completion times or exposing excess genetic information.

Newborn screening improvement strategies used to increase specificity without decreasing sensitivity were first described in Chapter 1 (pp. 2). It was highlighted that these augmentations are not always technological interventions (e.g. new second-tier assays) but may also take the form of policies or practices designed to prevent diagnostic referral of neonates at lower risk of targeted disease. For example, a newborn screening laboratory may establish a policy to not refer positive galactosemia results on infants older than six months, based on awareness that this aggressive enzyme defect almost always manifests symptomatically within the neonatal period when untreated [106]. Postnatal age at sampling also influences typical circulating biomarker concentrations and screening laboratories use different thresholds to identify risk in samples drawn from neonates after the first week of life [107]. The element common to most specificity-improving policies or interpretation practices is the incorporation of additional clinical

information to better contextualize the biomarker analysis and enable the delivery of more accurate risk estimates.

The augmentative screening strategy that may have been enabled by this thesis investigation was covariate-adjusted screening abnormality thresholds ('cut-offs'). Our work aimed to identify neonatal factors associated with systematically altered screening biomarkers in order to facilitate the customization of screening risk algorithms adjusting for the influential factor, if such personalization was found to be necessary. The investigated neonatal factor (TSB) was selected based on the prevalence of TSB elevation among screened individuals [108], reports identified during preliminary review suggestive of association with newborn screening markers [109], and the availability of universal bilirubin measurements taken for pre-discharge hyperbilirubinemia screening [63]. Even though several biomarkers were statistically associated with TSB levels, we did not find evidence that TSB levels were associated with false positive screening results for selected disorders (Article B).

Covariate-adjusted biomarker cut-offs are not uncommon in biomarker-based newborn screening. A prototypical example is gestational age-stratified cut-offs for 17-hydroxyprogesterone (17-OHP) levels in screening algorithms for congenital adrenal hyperplasia (CAH) [110]. Circulating levels of 17-OHP are increased in preterm, ill, or stressed neonates and premature babies comprise 70% of all false positive CAH determinations [111]. The application of a single population-wide abnormal 17-OHP level to all neonates receiving CAH screening would asymmetrically deposit false positive screening results among preterm neonates relative to those who are born closer to term, as 17-OHP levels in premature neonates are more likely to be elevated in the absence of CAH [9]. A uniform abnormality threshold would therefore also result in differential screening specificity among preterm subgroups of the CAH-screened population, a phenomenon termed spectrum effect [112]. Thus, higher abnormality thresholds (screening cut-offs) for 17-OHP as an NBS screening marker are often used to identify CAH for neonates born early. This use of different cut-offs within different gestational age strata increases overall screening specificity by improving specificity among pre-term neonates. Thus, programs can account for the spectrum effect that occurs due to factors (in this example, gestational age) that are influential to screening markers by adjusting cut-offs to account for those factors [113].

To date, neonatal characteristics used to adjust screening abnormality thresholds have been mainly limited to simple demographics routinely noted on the screening information card (e.g.

postnatal age at collection, sex, weight, gestational age). One precision screening tool used by many screening laboratories worldwide, Collaborative Laboratory Integrated Reports (CLIR) by the Mayo Clinic, can produce continuous rare disease risk scores from uploaded screening biomarker profiles that are adjusted for postnatal age at DBS sampling, birthweight, and stratified by sex [114]. Considered state-of-the-art within the newborn screening literature, this tool considers only demographic characteristics (sex, birthweight) and a sampling characteristic (postnatal age at sampling) as covariates.

The investigation reported in this thesis did not assess biomarker variability due to screening participant demographics, instead focusing on associations with a second biomarker (TSB) measured universally at the site of birth. Our study recognized the possibility that the biomarkers used to determine neonatal risk for rare disease may not only be affected by immutable demographic characteristics (e.g. birthweight, gestational age at birth) and sampling characteristics, but also transient conditions such as neonatal hyperbilirubinemia. There is precedent for recognizing that clinical status and associated interventions can influence NBS biomarker levels. For example, preterm or ill neonates frequently receive total parenteral nutrition (TPN) to deliver several nutrients necessary for growth, including amino acids [115]. The high concentrations of amino acids delivered by TPN directly to the neonatal bloodstream can increase the amino acid levels detected by NBS laboratories in DBS samples, mimicking certain screen-targeted amino acid disorders [11]. The identification of this intervention-associated biomarker alteration enabled research that subsequently found that a brief 3-hour interruption of TPN prior to DBS sampling could reduce false positive amino acid disorder findings by 74% among NICU neonates receiving NBS [116]. This example illustrates a recognition among NBS programs that temporary interventions or transient health states may need to be considered when developing strategies to increase screening specificity.

To that end, this project investigated whether a transiently variable aspect of the neonatal phenotype (TSB level) measured by a parallel universal screening initiative was associated with false positive risk for selected newborn screening tests. As bilirubin level was not independently linked to increased risk of false positive results for the diseases we specifically studied, it was concluded that consideration of this element is unlikely to improve screening specificity for those diseases. However, our study (Article B) provides support for the perspective that demographic and sampling characteristics are not the only factors that can be systematically associated with NBS biomarker levels.

2.2.3 Increasing Specificity with Routinely Collected Data

This thesis used linked administrative databases to investigate associations between DBS biomarkers and a neonatal clinical factor (TSB). Although false positives can be numerous and concerning at population-scale, false positive findings are rare events at the individual level and the identification of risk factors for these results shares parallels with the epidemiologic study of other rare outcomes. In particular, studying false positive screening outcomes prospectively in the context of a cohort study or a trial would be complex and resource intensive. Participant enrollment would need to take place at multiple birthing centres over several years for an analyzable false positive sample size to accrue. Resultantly, many NBS false positivity studies are extremely small-scale, retrospective, and sparse in covariates. Of these three methodological challenges, covariate sparseness is a notable barrier to the discovery of risk factors for false positive results. In most instances and absent enrichment from other sources, the upper limit of covariates available for blood spot biomarker profile variability analysis is represented by the fields reported on the screening requisition (Figure 2-2).

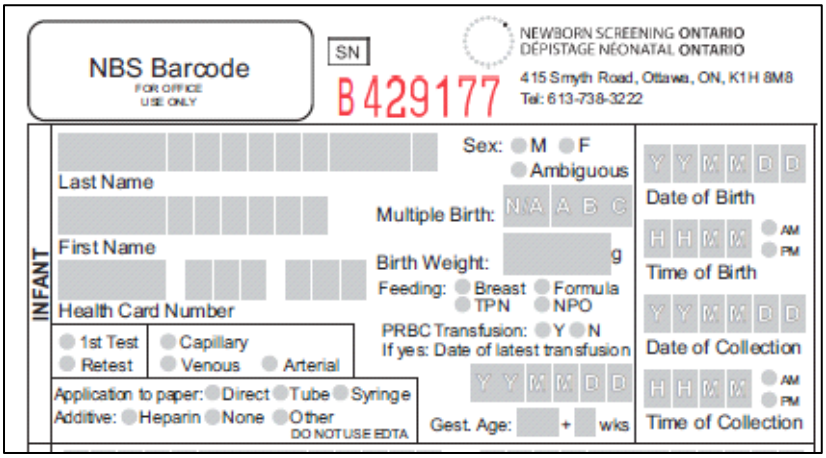


Figure 2-2. Neonatal Data Elements Transmitted to Ontario's Newborn Screening Program

In this work, the biomarker results and the requisition data were linked with a centralized laboratory result repository and other administrative records to investigate biomarker variability along the neonatal TSB spectrum. To our knowledge, our large database techniques facilitated the first report of phototherapy administration rates by screening false positive status, as well as the first reports of comparable median TSB levels in screen negative and false positive neonates (Article B, Table 3). Our principal finding of non-association between TSB level and false positive risk for three disorders with high false positive rates further supports the value routinely collected data methods to cost-effectively and rapidly identify or rule out potential sources of NBS biomarker influence.

In future studies, routinely-collected data methods like ours could be used to test hypotheses about interventions (e.g. TPN administration, phototherapy), demographics (e.g. birthweight), and conditions (e.g. hyperbilirubinemia) that are suspected to associate with NBS biomarkers, or may be risk factors for receiving a false positive newborn screening result, provided the hypothesized risk factor is recorded in administrative data. Through these methods, new sources of biomarker variability could be identified and added as required fields on the newborn screening requisition (Figure 2-2), enabling the creation of covariate-adjusted cut-offs. The increasing availability of electronic health records in centralized repositories, paired with routinely collected data studies as reported here, could in this way enable newborn algorithms better customized to providing sensitive and specific screening in biochemically diverse neonatal populations.

Thesis References

1. Morton SU, Brodsky D. Fetal Physiology and the Transition to Extrauterine Life. *Clin Perinatol*. 2016;43(3):395-407. doi:10.1016/j.clp.2016.04.001

2. Jervis GA. Phenylpyruvic Oligophrenia. *Arch Neurol Psychiatry*. 1937;38(5):944. doi:10.1001/archneurpsyc.1937.02260230042003

3. Sahai I, Marsden D. Newborn screening. *Crit Rev Clin Lab Sci*. 2009;46(2):55-82. doi:10.1080/10408360802485305

4. Garg U, Smith LD. Introduction to laboratory diagnosis and biomarkers in inborn error of metabolism. In: *Biomarkers in Inborn Errors of Metabolism*. Elsevier; 2017:1-24. doi:10.1016/B978-0-12-802896-4.00001-8

5. White E. Measurement error in biomarkers: sources, assessment, and impact on studies. *IARC Sci Publ*. 2011;(163):143-161.

6. Varness TS, Allen DB, Hoffman GL. Newborn screening for congenital adrenal hyperplasia has reduced sensitivity in girls. *J Pediatr*. 2005;147(4):493-498. doi:10.1016/j.jpeds.2005.04.035

7. Tarini BA, Clark SJ, Pilli S, et al. False-positive newborn screening result and future health care use in a state medicaid cohort. *Pediatrics*. 2011;128(4):715-722. doi:10.1542/peds.2010-2448

8. Gucciardi A, Zaramella P, Costa I, et al. Analysis and interpretation of acylcarnitine profiles in dried blood spot and plasma of preterm and full-term newborns. *Pediatr Res*. 2015;77(1):36-47. doi:10.1038/pr.2014.142

9. Hayashi G, Faure C, Brondi MF, et al. Weight-adjusted neonatal 17OH-progesterone cutoff levels improve the efficiency of newborn screening for congenital adrenal hyperplasia. *Arq Bras Endocrinol Metabol*. 2011;55(8):632-637. http://www.ncbi.nlm.nih.gov/pubmed/22218447. Accessed March 14, 2016.

10. Slaughter JL, Meinzen-Derr J, Rose SR, et al. The Effects of Gestational Age and Birth Weight on False-Positive Newborn-Screening Rates. *Pediatrics*. 2010;126(5):910-916. doi:10.1542/peds.2010-0943

11. Chace DH, De Jesús VR, Lim TH, Hannon WH, Clark RH, Spitzer AR. Detection of TPN contamination of dried blood spots used in newborn and metabolic screening and its impact on quantitative measurement of amino acids. *Clin Chim Acta*. 2011;412(15-16):1385-1390. doi:10.1016/j.cca.2011.04.009

12. Kwon C, Farrell PM. The magnitude and challenge of false-positive newborn screening test results. *Arch Pediatr Adolesc Med*. 2000;154(7):714-718. doi:10.1001/archpedi.154.7.714

13. Ryckman KK, Berberich SL, Shchelochkov OA, Cook DE, Murray JC. Clinical and environmental influences on metabolic biomarkers collected for newborn screening. *Clin Biochem*. 2013;46(1-2):133-138. doi:10.1016/j.clinbiochem.2012.09.013

14. White PC. Optimizing newborn screening for congenital adrenal hyperplasia. *J Pediatr*. 2013;163(1):10-12. doi:10.1016/j.jpeds.2013.02.008

15. Pang S, Shook MK. Current status of neonatal screening for congenital adrenal hyperplasia. *Curr Opin Pediatr*. 1997;9(4):419-423. http://www.ncbi.nlm.nih.gov/pubmed/9300201. Accessed March 14, 2016.

16. Kugelman A, Riskin A, Bader D, Koren I. Pitfalls in screening programs for congenital hypothyroidism in premature newborns. *Am J Perinatol*. 2009;26(5):383-385. doi:10.1055/s-0028-1110091

17. Minutti CZ, Lacey JM, Magera MJ, et al. Steroid profiling by tandem mass spectrometry improves the positive predictive value of newborn screening for congenital adrenal hyperplasia. *J Clin Endocrinol Metab*. 2004;89(8):3687-3693. doi:10.1210/jc.2003-032235

18. Eastman JW, Sherwin JE, Wong R, et al. Use of the phenylalanine: Tyrosine ratio to test newborns for phenylketonuria in a large public health screening programme. *J Med Screen*. 2000;7(3):131-135. doi:10.1136/jms.7.3.131

19. Han L, Tavakoli NP, Morrissey M, Spink DC, Cao Z (Tim). Liquid chromatography-tandem mass spectrometry analysis of 17-hydroxyprogesterone in dried blood spots revealed matrix effect on immunoassay. *Anal Bioanal Chem*. 2019;411(2):395-402. doi:10.1007/s00216-018-1449-0

20. Benito-Orejas JI, Ramírez B, Morais D, Almaraz A, Fernández-Calvo JL. Comparison of two-step transient evoked otoacoustic emissions (TEOAE) and automated auditory brainstem response (AABR) for universal newborn hearing screening programs. *Int J Pediatr Otorhinolaryngol*. 2008;72(8):1193-1201. doi:10.1016/j.ijporl.2008.04.011

21. Srisinghasongkram P, Pruksananonda C, Chonchaiya W. Two-Step Screening of the Modified Checklist for Autism in Toddlers in Thai Children with Language Delay and Typically Developing Children. *J Autism Dev Disord*. 2016;46(10):3317-3329. doi:10.1007/s10803-016-2876-4

22. Maisels J, Watchko J. Jaundice. In: MacDonald MG, Seshia M, eds. *Avery's Neonatology: Pathophysiology and Management of the Newborn*. 7th ed. Philadelphia: Wolters Kluwer; 2016:587-645.

23. Tenhunen R. The enzymatic conversion of heme to bilirubin in vivo. *Ann Clin Res*. 1976;8 Suppl 17(2):2-9. http://www.ncbi.nlm.nih.gov/pubmed/1008491.

24. Hansen TWR. Prevention of neurodevelopmental sequelae of jaundice in the newborn. *Dev Med Child Neurol*. 2011;53(SUPPL.4):24-28. doi:10.1111/j.1469-8749.2011.04059.x

25. Canadian Pediatric Society. Guidelines for detection, management and prevention of hyperbilirubinemia in term and late preterm newborn infants (35 or more weeks' gestation) - Summary. *Paediatr Child Health*. 2007;12(5):401-418. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2528770&tool=pmcentrez&rendertype=abstract.

26. Lauer BJ, Spector ND. Hyperbilirubinemia in the Newborn. *Pediatr Rev*. 2011;32(8):341-349. doi:10.1542/pir.32-8-341

27. Kawade N, Onishi S. The prenatal and postnatal development of UDP-glucuronyltransferase activity towards bilirubin and the effect of premature birth on this activity in the human liver. *Biochem J*. 1981;196(1):257-260. doi:10.1042/bj1960257

28. Maisels MJ. What's in a Name? Physiologic and Pathologic Jaundice: The Conundrum of Defining Normal Bilirubin Levels in the Newborn. *Pediatrics*. 2006;118(2):805-807. doi:10.1542/peds.2006-0675

29. Watchko JF. Identification of Neonates at Risk for Hazardous Hyperbilirubinemia: Emerging Clinical Insights. *Pediatr Clin North Am*. 2009;56(3):671-687. doi:10.1016/j.pcl.2009.04.005

30. Billing BH, Cole PG, Lathe GH. Increased plasma bilirubin in newborn infants in relation to birth weight. *Br Med J*. 1954;2(4899):1263-1265. doi:10.1136/bmj.2.4899.1263

31. Maamouri G, Khatami F, Mohammadzadeh A. Hyperbilirubinemia and Neonatal Infection. *Int J Pediatr*. 2013;1(1):5-12.

32. Iolascon A, Perrotta S, Stewart GW. Red blood cell membrane defects. *Rev Clin Exp Hematol*. 2003;7(1):22-56. doi:10.1177/017084068800900203

33. Oort M, Heerspink W, Roos D, Flavell RA, Bernini LF. Haemolytic Disease of the Newborn and Chronic Anaemia Induced by γβ Thalassaemia in a Dutch Family. *Br J Haematol*. 1981;48(2):251-262. doi:10.1111/j.1365-2141.1981.tb02710.x

34. Kedar PS, Warang P, Colah RB, Mohanty D. Red cell pyruvate kinase deficiency in neonatal jaundice cases in India. *Indian J Pediatr*. 2006;73(11):985-988. doi:10.1007/BF02758302

35. Gallagher PG. Abnormalities of the erythrocyte membrane. *Pediatr Clin North Am*. 2013;60(6):1349-1362. doi:10.1016/j.pcl.2013.09.001

36. Berkowitz FE. Hemolysis and infection: categories and mechanisms of their interrelationship. *Rev Infect Dis*. 2018;13(6):1151-1162. http://www.ncbi.nlm.nih.gov/pubmed/1775848.

37. Bainbridge R. Jaundice in Neonates With Sickle Cell Disease. *Am J Dis Child*. 1988;142(5):569. doi:10.1001/archpedi.1988.02150050107045

38. Travan L, Lega S, Crovella S, Montico M, Panontin E, Demarini S. Severe neonatal hyperbilirubinemia and UGT1A1 promoter polymorphism. *J Pediatr*. 2014;165(1):42-45. doi:10.1016/j.jpeds.2014.03.013

39. Fretzayas A, Moustaki M, Liapi O, Karpathios T. Eponym: Gilbert syndrome. *Eur J Pediatr*. 2012;171(1):11-15. doi:10.1007/s00431-011-1641-0

40. Lane E, Murray KF. Neonatal Cholestasis. *Pediatr Clin North Am*. 2017;64(3):621-639. doi:10.1016/j.pcl.2017.01.006

41. Lakshminarayanan B, Davenport M. Biliary atresia: A comprehensive review. *J Autoimmun*. 2016;73:1-9. doi:10.1016/j.jaut.2016.06.005

42. Craig JM, Landing BH. Form of hepatitis in neonatal period simulating biliary atresia. *AMA Arch Pathol*. 1952;54(4):321-333. http://www.ncbi.nlm.nih.gov/pubmed/12984938.

43. Colombo C, Battezzati PM, Strazzabosco M, Podda M. Liver and biliary problems in cystic fibrosis. *Semin Liver Dis*. 1998;18(3):227-235. doi:10.1055/s-2007-1007159

44. Quigley EMM, Marsh MN, Shaffer JL, Markin RS. Hepatobiliary complications of total parenteral nutrition. *Gastroenterology*. 1993;104(1):286-301. doi:10.1016/0016-5085(93)90864-9

45. Merritt RJ. Cholestasis associated with total parenteral nutrition. *J Pediatr Gastroenterol Nutr*. 1986;5(1):9-22. http://www.ncbi.nlm.nih.gov/pubmed/3080580.

46. Roy CC, Belli DC. Hepatobiliary complications associated with tpn: An enigma. *J Am Coll Nutr*. 1985;4(6):651-660. doi:10.1080/07315724.1985.10720107

47. Memon N, Weinberger BI, Hegyi T, Aleksunes LM. Inherited disorders of bilirubin clearance. *Pediatr Res*. 2016;79(3):378-386. doi:10.1038/pr.2015.247

48. Lykavieris P, Bernard O, Hadchouel M. Neonatal cholestasis as the presenting feature in cystic fibrosis. *Arch Dis Child*. 1996;75(1):67-70. doi:10.1136/adc.75.1.67

49. Applebaum MN, Thaler MM. Reversibility of extensive liver damage in galactosemia. *Gastroenterology*. 1975;69(2):496-502. http://www.ncbi.nlm.nih.gov/pubmed/1150052.

50. Endo F, Sun MS. Tyrosinaemia type I and apoptosis of hepatocytes and renal tubular cells. *J Inherit Metab Dis*. 2002;25(3):227-234. doi:10.1023/A:1015646400182

51. Ohura T, Kobayashi K, Tazawa Y, et al. Clinical pictures of 75 patients with neonatal intrahepatic cholestasis caused by citrin deficiency (NICCD). *J Inherit Metab Dis*. 2007;30(2):139-144. doi:10.1007/s10545-007-0506-1

52. Ozkan TB, Mistik R, Dikici B, Nazlioglu HO. Antiviral therapy in neonatal cholestatic cytomegalovirus hepatitis. *BMC Gastroenterol*. 2007;7:5-9. doi:10.1186/1471-230X-7-9

53. Laukkarinen J, Sand J, Saaristo R, et al. Is bile flow reduced in patients with hypothyroidism? *Surgery*. 2003;133(3):288-293. doi:10.1067/msy.2003.77

54. Kurt A, Aygun AD, Kurt ANC, Godekmerdan A, Akarsu S, Yilmaz E. Use of phototherapy for neonatal hyperbilirubinemia affects cytokine production and lymphocyte subsets. *Neonatology*. 2009;95(3):262-266. doi:10.1159/000171216

55. Watchko JF. Recent Advances in the Management of Neonatal Jaundice. *Res Reports Neonatol*. 2014;4:183-193.

56. Okwundu CI, Okoromah CAN, Shah PS. Prophylactic phototherapy for preventing jaundice in preterm or low birth weight infants. *Cochrane database Syst Rev*. 2012;1:CD007966. doi:10.1002/14651858.CD007966.pub2

57. Keren R, Bhutani VK. Predischarge risk assessment for severe neonatal hyperbilirubinemia. *Neoreviews*. 2007;8(2):e68-e76.

58. American Academy of Pediatrics Subcommittee on Hyperbilirubinemia. Management of hyperbilirubinemia in the newborn infant 35 or more weeks of gestation. *Pediatrics*. 2004;114(1):297-316. doi:10.1542/peds.114.1.297

59. Grosse SD, Prosser LA, Botkin JR. Screening for Neonatal Hyperbilirubinemia—First Do No Harm? *JAMA Pediatr*. 2019;173(7):617. doi:10.1001/jamapediatrics.2019.1194

60. Watchko JF, Maisels MJ. Avoiding Harm From Hyperbilirubinemia Screening. *JAMA Pediatr*. 2019;29:S25-S45. doi:10.1001/jamapediatrics.2019.3765

61. Trikalinos TA, Chung M, Lau J, Ip S. Systematic Review of Screening for Bilirubin Encephalopathy in Neonates. *Pediatrics*. 2009;124(4):1162-1171. doi:10.1542/peds.2008-3545

62. Darling EK, Ramsay T, Sprague AE, Walker MC, Guttmann A. Universal bilirubin screening and health care utilization. *Pediatrics*. 2014;134(4):e1017-24. doi:10.1542/peds.2014-1146

63. Darling EK, Guttmann A, Sprague AE, Ramsay T, Walker MC. Implementation of the Canadian Paediatric Society's hyperbilirubinemia guidelines: A survey of Ontario hospitals. *Paediatr Child Heal*. 2014;19(3):133-137. doi:10.1093/pch/19.3.133

64. Blainey JD, Gulliford R. Phenylalanine-restricted diets in the treatment of phenylketonuria. *Arch Dis Child*. 1956;31(160):452-466. doi:10.1136/adc.31.160.452

65. Bickel H, Gerrard J, Hickmans EM. Influence of phenylalanine intake on phenylketonuria. *Lancet (London, England)*. 1953;265(6790):812-813. http://www.ncbi.nlm.nih.gov/pubmed/13098090.

66. Guthrie R. The origin of newborn screening. *Screening*. 1992;1:5-15. doi:10.1016/0925-6164(92)90025-Z

67. Guthrie R, Susi A. A Simple Phenylalanine Method for Detecting Phenylketonuria in Large Populations of Infants. *Pediatrics*. 1963;32(September):338-343. http://www.ncbi.nlm.nih.gov/pubmed/14063511.

68. Brosco JP, Paul DB. The political history of PKU: Reflections on 50 years of newborn screening. *Pediatrics*. 2013;132(6):987-989. doi:10.1542/peds.2013-1441

69. Therrell BL, Adams J. Newborn screening in North America. *J Inherit Metab Dis*. 2007;30(4):447-465. doi:10.1007/s10545-007-0690-z

70. Chace DH, Millington DS, Terada N, Kahler SG, Roe CR, Hofman LF. Rapid diagnosis of phenylketonuria by quantitative analysis for phenylalanine and tyrosine in neonatal blood spots by tandem mass spectrometry. *Clin Chem*. 1993;39(1):66-71. doi:10.1093/clinchem/39.1.66

71. Schulze A, Lindner M, Kohlmüller D, Olgemöller K, Mayatepek E, Hoffmann GF. Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics*. 2003;111(6 Pt 1):1399-1406. http://www.ncbi.nlm.nih.gov/pubmed/12777559.

72. Caggana M, Jones EA, Shahied SI, Tanksley S, Hermerath CA, Lubin IM. Newborn screening: From Guthrie to whole genome sequencing. *Public Health Rep*. 2013;128(SUPPL. 2):14-19. doi:10.1177/00333549131280s204

73. Brown L, Xu-Bayford J, Allwood Z, et al. Neonatal diagnosis of severe combined immunodeficiency leads to significantly improved survival outcome: The case for newborn screening. *Blood*. 2011;117(11):3243-3246. doi:10.1182/blood-2010-08-300384

74. Marquardt G, Currier R, McHugh DMS, et al. Enhanced interpretation of newborn screening results without analyte cutoff values. *Genet Med*. 2012;14(7):648-655. doi:10.1038/gim.2012.2

75. Shlomi T, Cabili MN, Ruppin E. Predicting metabolic biomarkers of human inborn errors of metabolism. *Mol Syst Biol*. 2009;5(263):1-8. doi:10.1038/msb.2009.22

76. Therrell BL, Williams D, Johnson K, Lloyd-Puryear MA, Mann MY, Ramos LR. Financing newborn screening: Sources, issues, and future considerations. *J Public Heal Manag Pract*. 2007;13(2):207-213. doi:10.1097/00124784-200703000-00020

77. Orzalesi M, Danhaive O. Ethical problems with neonatal screening. *Ann Ist Super Sanita*. 2009;45(3):325-330. http://www.ncbi.nlm.nih.gov/pubmed/19861738.

78. The President's Council on Bioethics. *The Changing Moral Focus of Newborn Screening*.; 2008. doi:10.1017/CBO9781107415324.004

79. Schoen EJ, Baker JC, Colby CJ, To TT. Cost-benefit analysis of universal tandem mass spectrometry for newborn screening. *Pediatrics*. 2002;110(4):781-786. http://www.ncbi.nlm.nih.gov/pubmed/12359795.

80. Fingerhut R, Olgemöller B. Newborn screening for inborn errors of metabolism and endocrinopathies: An update. *Anal Bioanal Chem*. 2009;393(5):1481-1497. doi:10.1007/s00216-008-2505-y

81. Mak CM, Lee H-CH, Chan AY-W, Lam C-W. Inborn errors of metabolism and expanded newborn screening: review and update. *Crit Rev Clin Lab Sci*. 2013;50(6):142-162. doi:10.3109/10408363.2013.847896

82. Centers for Disease Control and Prevention (CDC). Good laboratory practices for biochemical genetic testing and newborn screening for inherited metabolic disorders. *MMWR Recomm reports Morb Mortal Wkly report Recomm reports*. 2012;61(RR-2):1-44. doi:10.1016/j.jpeds.2017.11.012

83. Stinton C, Geppert J, Freeman K, et al. Newborn screening for Tyrosinemia type 1 using succinylacetone - a systematic review of test accuracy. *Orphanet J Rare Dis*. 2017;12(1):1-8. doi:10.1186/s13023-017-0599-z

84. Matern D, Tortorelli S, Oglesbee D, Gavrilov D, Rinaldo P. Reduction of the false-positive rate in newborn screening by implementation of MS/MS-based second-tier tests: The Mayo Clinic experience (2004-2007). *J Inherit Metab Dis*. 2007;30(4):585-592. doi:10.1007/s10545-007-0691-y

85. De Jesús VR, Mei J V., Bell CJ, Hannon WH. Improving and Assuring Newborn Screening Laboratory Quality Worldwide: 30-Year Experience at the Centers for Disease Control and Prevention. *Semin Perinatol*. 2010;34(2):125-133. doi:10.1053/j.semperi.2009.12.003

86. Morrison DR, Clayton EW. False positive newborn screening results are not always benign. *Public Health Genomics*. 2011;14(3):173-177. doi:10.1159/000322527

87. Van Der Ploeg CPB, Lanting CI, Kauffman-de Boer MA, Uilenburg NN, De Ridder-Sluiter JG, Verkerk PH. Examination of long-lasting parental concern after false-positive results of neonatal hearing screening. *Arch Dis Child*. 2008;93(6):508-511. doi:10.1136/adc.2007.129320

88. Schmidt JL, Castellanos-Brown K, Childress S, et al. The impact of false-positive newborn screening results on families: A qualitative study. *Genet Med*. 2012;14(1):76-80. doi:10.1038/gim.2011.5

89. Karaceper MD, Chakraborty P, Coyle D, et al. The health system impact of false positive newborn screening results for medium-chain acyl-CoA dehydrogenase deficiency: A cohort study. *Orphanet J Rare Dis*. 2016;11(1):1-9. doi:10.1186/s13023-016-0391-5

90. Webster D. Quality performance of newborn screening systems: Strategies for improvement. *J Inherit Metab Dis*. 2007;30(4):576-584. doi:10.1007/s10545-007-0639-2

91. Gong Z, Xu W-J, Tian G-L, Zhang T, Lv Z. Neonatal Intrahepatic Cholestasis Caused by Citrin Deficiency Differentiated from Biliary Atresia. *Eur J Pediatr Surg*. 2016;26(3):255-259. doi:10.1055/s-0035-1551566

92. Tamamori A, Okano Y, Ozaki H, et al. Neonatal intrahepatic cholestasis caused by citrin deficiency: Severe hepatic dysfunction in an infant requiring liver transplantation. *Eur J Pediatr*. 2002;161(11):609-613. doi:10.1007/s00431-002-1045-2

93. Balistreri WF. PEDIATRIC HEPATOLOGY. *Clin Liver Dis*. 2000;4(1):191-210. doi:10.1016/S1089-3261(05)70103-3

94. Beyoğlu D, Idle JR. The metabolomic window into hepatobiliary disease. *J Hepatol*. 2013;59(4):842-858. doi:10.1016/j.jhep.2013.05.030

95. Russo PA, Mitchell GA, Tanguay RM. Tyrosinemia: A review. *Pediatr Dev Pathol*. 2001;4(3):212-221. doi:10.1007/s100240010146

96. McCarthy ME, Oltman SP, Baer RJ, et al. Newborn Metabolic Profile Associated with Hyperbilirubinemia With and Without Kernicterus. *Clin Transl Sci*. 2019;12(1):28-38. doi:10.1111/cts.12590

97. Twilt M. Precision Medicine: The new era in medicine. *EBioMedicine*. 2016;4:24-25. doi:10.1016/j.ebiom.2016.02.009

98. Collins FS, Varmus H. A New Initiative on Precision Medicine. *N Engl J Med*. 2015;372(9):793-795. doi:10.1056/NEJMp1500523

99. Minter Baerg MM, Stoway SD, Hart J, et al. Precision newborn screening for lysosomal disorders. *Genet Med*. 2018;20(8):847-854. doi:10.1038/gim.2017.194

100. Tarini BA, Christakis DA, Welch HG. False-positive Results in Expanded Newborn Screening: In Reply. *Pediatrics*. 2007;119(2):415-415. doi:10.1542/peds.2006-3310

101. Tu WJ, He J, Chen H, Shi XD, Li Y. Psychological effects of false-positive results in expanded newborn screening in China. *PLoS One*. 2012;7(4). doi:10.1371/journal.pone.0036235

102. Hall PL, Marquardt G, McHugh DMS, et al. Postanalytical tools improve performance of newborn screening by tandem mass spectrometry. *Genet Med*. 2014;16(12):889-895. doi:10.1038/gim.2014.62

103. Ulm E, Feero WG, Dineen R, Charrow J, Wicklund C. Genetics Professionals' Opinions of Whole-Genome Sequencing in the Newborn Period. *J Genet Couns*. 2015;24(3):452-463. doi:10.1007/s10897-014-9779-3

104. Yang L, Chen J, Shen B. Newborn Screening in the Era of Precision Medicine. In: *Translational Informatics in Smart Healthcare - Advances in Experimental Medicine and Biology.* Vol 1005. ; 2017:47-61. doi:10.1007/978-981-10-5717-5_3

105. Almannai M, Marom R, Reid Sutton V. Newborn screening: A review of history, recent advancements, and future perspectives in the era of next generation sequencing. *Curr Opin Pediatr*. 2016;28(6):694-699. doi:10.1097/MOP.0000000000000414

106. Demirbas D, Coelho AI, Rubio-Gozalbo ME, Berry GT. Hereditary galactosemia. *Metabolism*. 2018;83:188-196. doi:10.1016/j.metabol.2018.01.025

107. Schymik I, Liebig M, Mueller M, et al. Pitfalls of neonatal screening for very-long-chain acyl-CoA dehydrogenase deficiency using tandem mass spectrometry. *J Pediatr*. 2006;149(1):128-130. doi:10.1016/j.jpeds.2006.02.037

108. Cohen RS, Wong RJ, Stevenson DK. Understanding neonatal jaundice: A perspective on causation. *Pediatr Neonatol*. 2010;51(3):143-148. doi:10.1016/S1875-9572(10)60027-7

109. Cai A, Qi S, Su Z, et al. A Pilot Metabolic Profiling Study of Patients With Neonatal Jaundice and Response to Phototherapy. *Clin Transl Sci*. 2016;9(4):216-220. doi:10.1111/cts.12401

110. Allen DB, Hoffman GL, Fitzpatrick P, Laessig R, Maby S, Slyper A. Improved precision of newborn screening for congenital adrenal hyperplasia using weight-adjusted criteria for 17-hydroxyprogesterone levels. *J Pediatr*. 1997;130(1):128-133. doi:10.1016/S0022-3476(97)70321-4

111. Hayashi GY, Carvalho DF, de Miranda MC, et al. Neonatal 17-hydroxyprogesterone levels adjusted according to age at sample collection and birthweight improve the efficacy of congenital adrenal hyperplasia newborn screening. *Clin Endocrinol (Oxf)*. 2017;86(4):480-487. doi:10.1111/cen.13292

112. Usher-Smith JA, Sharp SJ, Griffin SJ. The spectrum effect in tests for risk prediction, screening, and diagnosis. *BMJ*. 2016;353. doi:10.1136/bmj.i3139

113. Mulherin SA, Miller WC. Academia and Clinic Spectrum Bias or Spectrum Effect ? Subgroup Variation in Diagnostic. *Ann Intern Med*. 2002;137:598-602.

114. Sörensen L, von Döbeln U, Åhlman H, et al. Expanded Screening of One Million Swedish Babies with R4S and CLIR for Post-Analytical Evaluation of Data. *Int J Neonatal Screen*. 2020;6(2):42. doi:10.3390/ijns6020042

115. Patel P, Bhatia J. Total parenteral nutrition for the very low birth weight infant. *Semin Fetal Neonatal Med*. 2017;22(1):2-7. doi:10.1016/j.siny.2016.08.002

116. Morris M, Fischer K, Leydiker K, Elliott L, Newby J, Abdenur JE. Reduction in newborn screening metabolic false-positive results following a new collection protocol. *Genet Med*. 2014;16(6):477-483. doi:10.1038/gim.2013.171

Appendix A: Certificates of Ethical Approval

A1. Institute of Clinical Evaluative Sciences Letter

Wednesday, July 29, 2020

ICES Central
G1 06, 2075 Bayview Avenue
Toronto, Ontario M4N 3M5
www.ices.on.ca

NOTICE re: ICES Projects Conducted Under Section 45 of PHIPA

ICES is a prescribed entity under section 45 of *Ontario's Personal Health Information Protection Act* (PHIPA). Section 45 is the provision that enables analysis and compilation of statistical information related to the management, evaluation and monitoring of, allocation of resources to, and planning for the health system. Section 45 authorizes health information custodians to disclose personal health information to a prescribed entity, like ICES, without consent for such purposes.

Projects conducted wholly under section 45, by definition, do not require review by a Research Ethics Board. This is confirmed in a letter from the REB of Sunnybrook Health Sciences Centre, ICES' Research Ethics Board of Record (See Appendix A).

As a prescribed entity, ICES must submit to tri-annual review and approval of its privacy and security policies, procedures and practices by Ontario's Information and Privacy Commissioner. These include policies, practices and procedures that require internal review and approval of every project by ICES' Privacy and Compliance Office. ICES was approved by the Commissioner for a fifth time in 2017.

Please do not hesitate to contact us if you require further information.

Sincerely,

ICES
Privacy & Legal Office
G-106, 2075 Bayview Avenue, Toronto, ON M4N 3M5
T: 416-480-4055
F: 416-480-6048
plo@ices.on.ca

A2. Sunnybrook Health Science Centre Research Ethics Board Memo Concerning Exemption for Section 45 Research

Special Note: The document below is referenced as "Appendix A" on the letter included on the previous page.

Sunnybrook
HEALTH SCIENCES CENTRE

Research Ethics Office, Room C819
2075 Bayview Avenue
Toronto, ON Canada M4N 3M5
t: 416-480-6100 ext. 4276 or 88144
www.sunnybrook.ca/reo

To: Institute for Clinical Evaluative Sciences (ICES)

From: Dr. Brian J. Murray
Chair, Research Ethics Board
Sunnybrook Health Sciences Centre

Date: January 24, 2017

Subject: Research Ethics Board Review of PHIPA s.45 projects

The Sunnybrook Health Sciences Centre (Sunnybrook) Research Ethics Board (REB) acts as the REB of record for the Institute for Clinical Evaluative Sciences (ICES) for some studies.

It has been determined that studies/projects that fall under section 45[1] of the *Personal Health Information Protection Act: Disclosure for planning and management of health system* do not require REB review and approval.

[Signature]

Brian J. Murray, MD FRCP(C) D,ABSM
Chair, Research Ethics Board

[1] Personal Health Information Protection Act, 2004, S.O. 2004, c. 3, Sched. A s.45

A3. University of Ottawa Determination of Exemption for Study Reported as "Article B" in this Thesis

23/05/2019

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Lettre d'approbation administrative | Letter of administrative approval

Numéro de dossier / Ethics File Number	H-05-19-4017
Titre du projet / Project Title	An Investigation of the Association between Metabolic Screening Analytes and Bilirubin Levels in Ontario Newborns
Type de projet / Project Type	Thèse de maîtrise / Master's thesis
CÉR primaire / Primary REB	
Statut du projet / Project Status	Approuvé / Approved
Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)	23/05/2019
Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)	06/01/2020

Équipe de recherche / Research Team

Chercheur / Researcher	Affiliation	Role
Emeril SANTANDER	Département d'épidémiologie et santé publique / Department of Epidemiology and Public Health	Chercheur Principal / Principal Investigator
Steven HAWKEN	Département d'épidémiologie et santé publique / Department of Epidemiology and Public Health	Superviseur / Supervisor
Elizabeth POTTER	Département d'épidémiologie et santé publique / Department of Epidemiology and Public Health	Co-superviseur / Co-supervisor

Conditions spéciales ou commentaires / Special conditions or comments:

This project uses data from a prescribed entity under section 45 of Ontario's Personal Health Information Protection Act (PHIPA), i.e. The Institute for Clinical Evaluative Sciences (ICES).

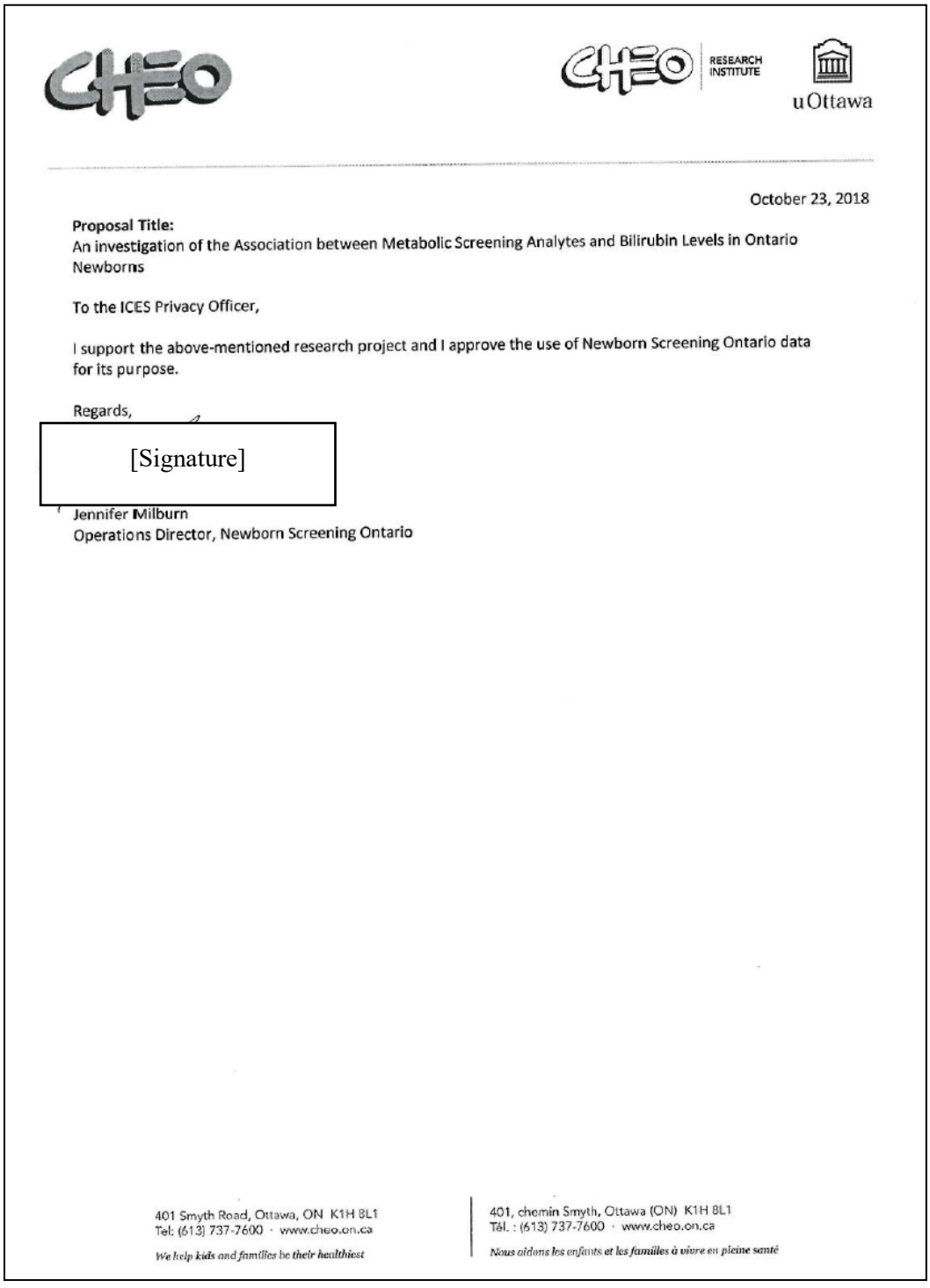
Ethics review is therefore not required but the University will keep track of the project via this administrative review process.

550, rue Cumberland, pièce 154
Ottawa (Ontario) K1N 6N5 Canada

550 Cumberland Street, Room 154
Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

Appendix B: Letters of Support



October 23, 2018

Proposal Title:
An investigation of the Association between Metabolic Screening Analytes and Bilirubin Levels in Ontario Newborns

To the ICES Privacy Officer,

I support the above-mentioned research project and I approve the use of Newborn Screening Ontario data for its purpose.

Regards,

[Signature]

Jennifer Milburn
Operations Director, Newborn Screening Ontario

401 Smyth Road, Ottawa, ON K1H 8L1
Tel: (613) 737-7600 · www.cheo.on.ca
We help kids and families be their healthiest

401, chemin Smyth, Ottawa (ON) K1H 8L1
Tél. : (613) 737-7600 · www.cheo.on.ca
Nous aidons les enfants et les familles à vivre en pleine santé