Estimating Screening Results Following the Introduction of Next-Generation Sequencing into Newborn Screening

Alvi Rahman

A thesis submitted in partial fulfillment of the requirements for the
Master's degree in Science

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

© Alvi Rahman, Ottawa, Canada, 2017

ABSTRACT

Objective: The objective of this thesis was to estimate the impact on newborn screening (NBS) results of changing screening technology from tandem mass spectrometry (MS/MS) to an approach using targeted next-generation sequencing (T-NGS) and MS/MS in parallel.

Methods: We integrated results of an analysis of MS/MS screening data for phenylketonuria (PKU) and medium-chain acyl-CoA dehydrogenase (MCAD) deficiency; and a query of genetic compendia for variants of genes associated with the two disorders.

Results: The introduction of T-NGS into NBS may reduce nearly 80% of false positives that are generated using the current screening approach. Based on estimated NBS results, T-NGS may be applied using a second-tier approach, which may improve specificity while maintaining sensitivity at its current level.

Discussion: T-NGS may enhance the performance of NBS for PKU by improving specificity when used as a second tier test, but may be limited by feasibility and cost under current circumstances. Future studies should consider the cost-effectiveness of T-NGS for all infants undergoing NBS.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervisor and mentor, Dr. Brenda Wilson, for her wisdom, unwavering support, and constant encouragement throughout the course of my thesis project. I am forever thankful to my co-supervisor, Dr. Beth Potter, who gave me the guidance to approach some of the toughest challenges I faced during my thesis project. I am also highly appreciative of the members of my thesis advisory committee, including Dr. Pranesh Chakraborty, Dr. Dennis Bulman, and Dr. Marie-Helene Roy-Gagnon, for helping me understand the clinical elements of my work, navigate through the complex world of genetics, and approach all of my statistical analyses. I would also like to acknowledge Dr. Lemuel Racacho for his guidance with understanding genetic compendia, as well as Silvia Visentin, Laure Tessier, Emeril Santander, and Monica Lamoureux for their all of their help in facing the administrative challenges of research. I would like to thank my family and friends for always encouraging me to keep going. Finally, I would like to thank all of my wonderful classmates in the MSc. Epidemiology program for being amazing friends and the best cohort I could have asked for.

TABLE OF CONTENTS

List of Tables vii
List of Figures x
Legend xi

Chapter 1: Introduction **1**
1.1 Overview of Screening Programs 1
1.2 Genomics in Screening 2

Chapter 2: Literature Review and Thesis Objectives **4**
2.1 History of NBS 4
2.2 Inborn Errors of Metabolism 5
2.3 The NBS System 6
2.31 Education 6
2.32 Screening 6
2.33 Follow-up and Management 6
2.34 Evaluation 7
2.4 Screening Test Performance 7
2.5 Limitations of Current Screening Technology 9
2.6 Newborn Screening in Canada 10
2.7 Expansion of NBS 11
2.8 Next-Generation Sequencing 12
2.81 Targeted Next-Generation Sequencing 13
2.9 Rationale for this Study 16
2.10 Research Question and Thesis Objectives 17

Chapter 3: Case Studies **19**
3.1 Phenylketonuria 19
3.2 MCAD Deficiency 20

Chapter 4: Methodology **23**
4.1 Study Design 23
4.2 Objective 1: Retrospective Analysis of MS/MS Data 23
4.21 Outcomes of MS/MS Screening 26
4.22 Data Cleaning 26
4.3 Objective 2: Analysis of Genetic Compendia to Estimate T-NGS Results 30
4.31 Genetic Compendia 30
4.32 Data Extraction 31
4.33 Estimating Genotype Proportions 33
4.4 Objective 3: Combining Observed MS/MS and Estimated T-NGS Results 36
4.41 Association between Genotype and MS/MS Screening Results 37
4.42 Association between Genotype and Diagnostic Follow-up Results 37
4.43 PKU: Benign Hyperphenylalaninemia 38
4.44 MCAD Deficiency: Carriers 38
4.45 False Positives and True Negatives 38

4.5 Sensitivity Analyses 39
4.51 Genotype Proportions in All Newborns 39
4.52 Genotype Proportions in True Positives 39
4.6 Ethical Considerations 39

Chapter 5: Results **41**
5.1 Retrospective Analysis of NSO Data 41
5.11 PKU Screening and Diagnosis 41
5.12 MCAD Deficiency Screening and Diagnosis 42
5.2 Analysis of Genetic Databases 43
5.21 Query of Genetic Variants 43
5.22 Characteristics of Genetic Variants 44
5.23 Allele Frequencies 44
5.24 Hardy-Weinberg Genotype Proportions 45
5.3 Primary Analyses: Estimating Results from a Parallel MS/MS and T-NGS Approach 46
5.31 Combined PKU Results 46
Step 1: Association between Genotype and MS/MS Results 47
Step 2: Association between Genotype and Diagnostic Follow-up Results 47
Step 3: Benign HPA 49
Step 4: Remaining False Positives and True Negatives 51
5.31 Combined MCAD Deficiency Results 52
Step 1: Association between Genotype and MS/MS Results 53
Step 2: Association between Genotype and Diagnostic Follow-up Results 54
Step 3: MCAD Deficiency Carriers 56
Step 4: Remaining False Positives and True Negatives 56
5.4 Sensitivity Analyses 58

Chapter 6: Discussion **61**
6.1 Performance of MS/MS Screening 61
6.2 A Combined T-NGS and MS/MS Approach 62
6.21 Primary Analyses: PKU 62
6.22 Primary Analyses: MCAD Deficiency 64
6.3 Sensitivity Analyses 65
6.31 Sensitivity Analyses: PKU 65
6.31 Sensitivity Analyses: MCAD Deficiency 66
6.4 Models for Implementing T-NGS in NBS 66
6.41 Second-tier NGS 67
6.42 Universal T-NGS 68
6.5 Issues to Consider for Universal T-NGS 69
6.51 Storage 69
6.52 Consent 70
6.53 Evidence Generation and Feasibility 70
6.54 Unsolicited Findings 71
6.6 Strengths 72
6.7 Limitations 73

6.8 Conclusion 76

Appendices **77**
Appendix A: Variable Missing Proportions 77
Appendix B: Parent Information Form 78
Appendix C: Policy Regarding Storage and Secondary Use of NBS Samples 81
Appendix D: REB Approval Letters 83
Appendix E: Variants with Conflicting Interpretations 86
Appendix F: Complete Sensitivity Analyses 89

References **97**

LIST OF TABLES

Chapter 1

Table 1.1: Wilson and Jungner's principles of screening for disease 1

Chapter 2

Table 2.1: Types of alleles classified by phenotypic manifestation 14

Table 2.2: Description of possible results using the current NBS system 16

Chapter 3

Table 3.1: Classification of PKU according to severity 20

Chapter 4

Table 4.1: MS/MS screening criteria for PKU 24

Table 4.2: MS/MS screening criteria for MCAD deficiency 24

Table 4.3: List of variables in NSO data set for MS/MS screening and diagnostic follow-up 25

Table 4.4: Data fields extracted from genetic compendia 33

Table 4.5: Hardy-Weinberg variables and their description 35

Table 4.6: Genotype proportions of infants in the NSO population 36

Chapter 5

Table 5.1: Summary of MS/MS screening results at NSO 41

Table 5.2: Summary of diagnostic testing results for PKU 41

Table 5.3: Summary of diagnostic testing results for MCAD deficiency 42

Table 5.4: Clinical significance of variants that were included and excluded from ClinVar and gnomAD 44

Table 5.5: Population frequencies of each type of allelic variant for *PAH* and *ACADM* 45

Table 5.6: Estimated genotype proportions for *PAH* and *ACADM* 45

Table 5.7: Estimated genotype frequencies of infants in the NSO population 46

Table 5.8A: MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. 46
Table 5.8B: MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated true negatives with *P or LP || P or LP* genotype 47
Table 5.9: Genotype proportions of infants diagnosed with PKU 48
Table 5.8C: MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated infants who were true positives 49
Table 5.10 Genotype proportions of infants diagnosed with benign HPA 50
Table 5.8D: MS/MS results of infants in the NSO population, categorized by estimated genotype. Allocated remaining infants diagnosed with benign HPA 51
Table 5.8E: MS/MS results of infants in the NSO population, categorized by estimated genotype. Allocated remaining false positives 51
Table 5.8F: MS/MS results of infants in the NSO population, categorized by estimated genotype. Allocated infants who were true negatives 52
Table 5.11A MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. 53
Table 5.11B MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated false negatives and true negatives who were *P or LP || P or LP* 54
Table 5.12 Genotypes of patients diagnosed with MCAD deficiency through biochemical testing 54
Table 5.11C MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were true positives 55
Table 5.11D MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were MCAD deficiency carriers. 56
Table 5.11E MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were false positives. 57
Table 5.11F MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were true negatives. 57

Table 5.13A Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || P or LP* to the lower limit of the 95% confidence interval of the estimate. 58

Table 5.13B Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || P or LP* to the upper limit of the 95% confidence interval of the estimate. 59

Table 5.13C Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || VUS* to the lower limit of the 95% confidence interval of the estimate. 92

Table 5.13D Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || VUS* to the upper limit of the 95% confidence interval of the estimate. 92

Table 5.13E Sensitivity analysis for PKU. Varied proportion of infants who were *VUS ||VUS* to the lower limit of the 95% confidence interval of the estimate. 93

Table 5.13F Sensitivity analysis for PKU. Varied proportion of infants who were *VUS || VUS* to the upper limit of the 95% confidence interval of the estimate. 93

Table 5.13G Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || P or LP* genotype to the lower limit of the 95% confidence interval. 94

Table 5.13H Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || P or LP* genotype to the lower limit of the 95% confidence interval. 94

Table 5.13I Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || VUS* genotype to the lower limit of the 95% confidence interval. 95

Table 5.13J Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || VUS* genotype to the lower limit of the 95% confidence interval. 95

Table 5.14A Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || P or LP* to the lower limit of the 95% confidence interval of the estimate. 60

Table 5.14B Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || P or LP* to the upper limit of the 95% confidence interval of the estimate. 60

Table 5.14C Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP // VUS* to the lower limit of the 95% confidence interval of the estimate. 97

Table 5.14D Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP // VUS* to the upper limit of the 95% confidence interval of the estimate. 97

Table 5.14E Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *VUS // VUS* to the lower limit of the 95% confidence interval of the estimate. 98

Table 5.14F Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *VUS // VUS* to the upper limit of the 95% confidence interval of the estimate. 98

Chapter 1

Table 1.1: Wilson and Jungner's principles of screening for disease 1

Chapter 2

Table 2.1: Types of alleles classified by phenotypic manifestation 14

Table 2.2: Description of possible results using the current NBS system 16

Chapter 3

Table 3.1: Classification of PKU according to severity 20

Chapter 4

Table 4.1: MS/MS screening criteria for PKU 24

Table 4.2: MS/MS screening criteria for MCAD deficiency 24

Table 4.3: List of variables in NSO data set for MS/MS screening and diagnostic follow-up 25

Table 4.4: Data fields extracted from genetic compendia 33

Table 4.5: Hardy-Weinberg variables and their description 35

Table 4.6: Genotype proportions of infants in the NSO population 36

Chapter 5

Table 5.1: Summary of MS/MS screening results at NSO 41

Table 5.2: Summary of diagnostic testing results for PKU 41
Table 5.3: Summary of diagnostic testing results for MCAD deficiency 42
Table 5.4: Clinical significance of variants that were included and excluded from ClinVar and gnomAD 44
Table 5.5: Population frequencies of each type of allelic variant for *PAH* and *ACADM* 45
Table 5.6: Estimated genotype proportions for *PAH* and *ACADM* 45
Table 5.7: Estimated genotype frequencies of infants in the NSO population 46
Table 5.8A: MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. 46
Table 5.8B: MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated true negatives with *P or LP || P or LP* genotype 47
Table 5.9: Genotype proportions of infants diagnosed with PKU 48
Table 5.8C: MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated infants who were true positives 49
Table 5.10 Genotype proportions of infants diagnosed with benign HPA 50
Table 5.8D: MS/MS results of infants in the NSO population, categorized by estimated genotype. Allocated remaining infants diagnosed with benign HPA 51
Table 5.8E: MS/MS results of infants in the NSO population, categorized by estimated genotype. Allocated remaining false positives 51
Table 5.8F: MS/MS results of infants in the NSO population, categorized by estimated genotype. Allocated infants who were true negatives 52
Table 5.11A MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. 53
Table 5.11B MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated false negatives and true negatives who were *P or LP || P or LP* 54
Table 5.12 Genotypes of patients diagnosed with MCAD deficiency through biochemical testing 54
Table 5.11C MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were true positives 55
Table 5.11D MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were MCAD deficiency carriers. 56

Table 5.11E MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were false positives. 57
Table 5.11F MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were true negatives. 57
Table 5.13A Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || P or LP* to the lower limit of the 95% confidence interval of the estimate. 58
Table 5.13B Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || P or LP* to the upper limit of the 95% confidence interval of the estimate. 59
Table 5.13C Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || VUS* to the lower limit of the 95% confidence interval of the estimate. 92
Table 5.13D Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || VUS* to the upper limit of the 95% confidence interval of the estimate. 92
Table 5.13E Sensitivity analysis for PKU. Varied proportion of infants who were *VUS ||VUS* to the lower limit of the 95% confidence interval of the estimate. 93
Table 5.13F Sensitivity analysis for PKU. Varied proportion of infants who were *VUS || VUS* to the upper limit of the 95% confidence interval of the estimate. 93
Table 5.13G Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || P or LP* genotype to the lower limit of the 95% confidence interval. 94
Table 5.13H Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || P or LP* genotype to the lower limit of the 95% confidence interval. 94
Table 5.13I Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || VUS* genotype to the lower limit of the 95% confidence interval. 95
Table 5.13J Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || VUS* genotype to the lower limit of the 95% confidence interval. 95
Table 5.14A Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || P or LP* to the lower limit of the 95% confidence interval of the estimate. 60
Table 5.14B Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || P or LP* to the upper limit of the 95% confidence interval of the estimate. 60
Table 5.14C Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || VUS* to the lower limit of the 95% confidence interval of the estimate. 97

Table 5.14D Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || VUS* to the upper limit of the 95% confidence interval of the estimate. 97
Table 5.14E Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *VUS || VUS* to the lower limit of the 95% confidence interval of the estimate. 98
Table 5.14F Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *VUS || VUS* to the upper limit of the 95% confidence interval of the estimate. 98

LIST OF FIGURES

Chapter 2
Figure 2.1: Distribution of newborn screening test measure in diseased and non-diseased populations with threshold value 9

Chapter 6
Figure 6.1: Proposed NBS model for PKU using first-tier MS/MS and second-tier T-NGS. 68

LEGEND

ACMG – American College of Medical Genetics and Genomics

CHEO – Children's Hospital of Eastern Ontario

IEM – Inborn Error of Metabolism

MCAD – Medium-Chain Acyl-CoA Dehydrogenase

MS/MS – Tandem Mass Spectrometry

NBS – Newborn Screening

NGS – Next-Generation Sequencing

NSO – Newborn Screening Ontario

Phe – Phenylalanine

PKU – Phenylketonuria

T-NGS – Targeted Next-Generation Sequencing

Tyr - Tyrosine

WES – Whole Exome Sequencing

WGS – Whole Genome Sequencing

CHAPTER 1: INTRODUCTION

1.1 Overview of Screening Programs

Screening is an essential tool of modern public health and preventive medicine.[1] The purpose of screening is to identify individuals at high risk for a particular condition within the asymptomatic stage, where intervention may change its natural history[2,3] or provide valuable risk information to the screened individual.[2]

Screening programs involve more than solely the test itself; they function in a coordinated manner and include several partners within the public health system. In 1968, Wilson and Jungner drafted a report for the World Health Organization containing ten principles, as described in **Table 1.1,** that have since become the most cited publication pertaining to screening.[2,4] Along with the technical characteristics of the screening test, these principles acknowledge the significance of political, economic, social, and health issues in shaping the development of screening programs.[1,5]

Table 1.1 Wilson and Jungner's principles of screening for disease[4]

1. The condition should be an important health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic disease.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a 'once and for all' project.

A screening program contains several elements, including the development of a suitable screening test, identification of a target population, interventions to improve screening behaviours, sampling and delivery of samples, testing and analysis, reporting and disclosure, follow-up, confirmatory

testing, and initiation of therapy.[2,6] Therefore, revisions to screening programs should consider the implications of these changes on the entire system and its individual components.

1.2 Genomics in Screening

With the emergence of new health technologies and the development of new prevention and treatment strategies, it has been predicted that many populations or specific population subgroups will be screened in the future for genetic information.[7] Proponents of the use of population-wide genomic sequencing are hopeful that the information obtained can be used to guide personalized interventions in the prevention and treatment of disease,[8] but these claims have yet to be substantiated, and may be too optimistic at this time.[9] Several factors should be considered before conducting population-based screening of genetic information. A number of frameworks have been proposed to provide guidance with the evaluation of genomic testing.[10-15] All of these frameworks suggest (i) evaluating the validity and reliability of the test in predicting the disease of interest, (ii) assessing its ability to improve clinical outcomes, (iii) weighing the risks and benefits of implementing genomic technologies, and (iv) considering ethical, legal, and social implications.[10-16] Other issues that should be addressed in an evaluation include patient and health care provider receptivity as well as public health policy and oversight.[8]

The earliest applications of genomic screening may be conducted in clinical and public health settings where testing already focuses on the detection of genetic diseases, such as newborn screening (NBS) programs.[16] Traditionally, the goal of NBS was to identify infants at risk of rare genetic disorders and enable early treatment to improve health outcomes. With the expansion of screening panels through tandem mass spectrometry (MS/MS), a new screening technology (discussed in section 2.1), panels were extended to less treatable conditions, with the framing of benefits expanded to include knowledge for parents to inform later reproductive decisions.[1,18]

Emerging genomic technology offers secondary "benefits" that have led some to suggest a re-evaluation of the goals of NBS.[1,18] For example, genomic screening may identify infants who are heterozygous carriers of mutations for genetic disorders with autosomal recessive inheritance. Although infants who are carriers do not require treatment, their parents may benefit from being notified of the carrier status information, to inform future reproductive decisions. [1,18] However, an emphasis on these secondary benefits can change how screening is perceived, as a service that offers knowledge to achieve different health care goals, rather than for the sole purpose of preventing adverse health outcomes from rare genetic conditions.[1,18] Policy discussions regarding the changing objectives of NBS and the application of genomic technology need to consider the challenges of interpreting different test results, particularly those with uncertain clinical significance.[1,18]

The application of genomic sequencing technologies in NBS may generate new categories of results that affect clinical management. The feasibility as well as the risks and benefits of using these technologies should be assessed. Several stakeholder groups can be affected by a change in screening technology, including families, clinicians, health care payers, genomic researchers, and public health decision makers.[18] Empirical data is necessary to evaluate new health technologies in NBS and inform policy decision-making regarding their use. This thesis aims to estimate the impact that introducing genomic sequencing technologies in NBS will have on the nature of screening results and their interpretation. The findings of the thesis can be used to describe how NBS programs and the wider health care system may respond to new information generated through the application of genomics in NBS.

CHAPTER 2: LITERATURE REVIEW AND THESIS OBJECTIVES

2.1 History of NBS

NBS encompasses all activities carried out within the first 28 days in the life of a newborn to identify infants at higher risk of having a range of conditions, including inborn errors of metabolism (IEMs).[5] The primary goal of NBS is to intervene at a very early stage in the course of the illness and prevent adverse health outcomes.

Throughout its history, NBS has developed in increments with advancements in technology that enabled the expansion of the number of conditions on screening panels. The origin of NBS dates back to the late 1950s, when there was a focus on improving the management of phenylketonuria (PKU), a rare metabolic disease that is not apparent at birth but leads to severe and irreversible neurologic damage if untreated.[5,6,20] At the time, evidence suggested that dietary intervention could markedly improve outcomes – children prescribed with a special diet by 1 month of age could have a normal IQ.[6,20] In 1963, Dr. Robert Guthrie developed a test for PKU by analyzing dried blood spot samples that were blotted on to filter paper and demonstrated that it was effective. Consequently, other researchers applied similar testing protocols to detect additional disorders such as sickle cell disease, congenital hypothyroidism, and congenital adrenal hyperplasia.[5] With the development of tests to adequately detect these disorders and availability of medical treatment to improve outcomes, NBS continued to grow and was eventually established as fundamental public health practice.[5,6,20,21]

Initially, conditions on NBS panels each required an individual test and only PKU and congenital hypothyroidism were universally screened for.[5,21] In the 1990s, a major shift in NBS practice began to unfold. The development of MS/MS as a screening technology led to the ability to rapidly screen for multiple IEMs.[5,6,20] In contrast to previous screening methods, MS/MS enabled

the simultaneous measurement of many analytes using a single sample and a single test.[21] Furthermore, the entire MS/MS analysis, including the generation of the final report, took approximately 2 minutes.[21] Eventually, MS/MS was implemented by most NBS programs, which resulted in the inclusion of several new IEMs within NBS panel of disorders.[21]

2.2 Inborn Errors of Metabolism

IEMs are a group of disorders that typically result from a dysfunctional enzyme in a metabolic pathway.[22,23] These are genetic disorders where pathogenic variants of a gene are inherited in an autosomal recessive manner.[22,23] Clinical signs and symptoms usually manifest from accumulation of a toxic product or deficiency of an essential nutrient, or both.[22,23] Hence, treatment often entails dietary restriction or medications to reduce a toxic substrate, or supplementation to replace a deficiency in an essential product.[22,23] The prognosis for several IEMs depends on prompt diagnosis and therapeutic intervention; and not all IEMs can be treated adequately (discussed in 2.7). Individually, IEMs are rare diseases, but when considered collectively they have an overall birth prevalence of greater than 1 in 1000.[22] More than 500 IEMs have been identified, and approximately 25% begin to show clinical manifestations in the neonatal period, usually hours or days after birth.[22] Common early signs of IEMs include lethargy, decreased activity, poor feeding, respiratory distress, or seizures. These signs, however, are non-specific and can occur due to other neonatal conditions, making it essential to maintain a high degree of suspicion.[22]

2.3 The NBS System

NBS may be viewed as a system with several components, some of which are described below:

2.31 Education

NBS programs generally include basic education about screening delivered in person by health professionals or through written literature or other tools.[5,24,25] These materials may inform parents about the nature of the screening test, the interpretation of results, implied consent, the use of dried blood spot samples, and privacy and confidentiality. In addition to parents, primary care physicians usually receive education about NBS through residency training and continuing medical education.

2.32 Screening

This component of the system includes the collection of dried blood spot samples from birthing centers, along with the newborn and mother's basic demographic information.[5,24,25] Heel-stick specimens are obtained 24-48 hours after birth and transported to a designated screening laboratory. In the laboratory, the specimens are processed and screening tests are performed, after which results are interpreted according to specified guidelines. Samples and data collected from screening may be stored for research purposes. Abnormal findings indicative of high risk for a screened condition require prompt follow-up with families.[5,24,25]

2.33 Follow-Up and Management

Parents of newborns with positive screening results are notified and the infants followed up for confirmatory diagnostic testing. In some settings, primary care providers are responsible for communicating the results with parents and, in others, this is the role of the NBS program.[26]

Primary care providers are provided with educational materials regarding the diagnostic next steps and other basic information to be communicated to parents.[5,24,25] Diagnostic testing for confirmation of the disorder requires more elaborate tests that have higher specificity than those used in NBS. After a newborn is diagnosed with a screened disorder, the NBS system arranges appropriate access to necessary health services (i.e. pediatrician, genetic counseling, special foods) for families.[5,24]

2.34 Evaluation

Laboratories are subject to external accreditation requirements.[5,24] The quality of non-laboratory components of the screening system are usually also subject to evaluation. The basis for reviewing non-laboratory NBS system components is often provided by best practice guidelines that are described by professional organizations, such as the Canadian Pediatric Society, the Clinical and Laboratory Standards Institute, or the American College of Medical Genetics and Genomics (ACMG).[5] NBS programs may also have advisory committees, composed of a multidisciplinary group of professionals and any relevant stakeholders, which identify and monitor quality assurance indicators, and provide advice regarding policy decisions.[5,24,25]

2.4 Screening Test Performance

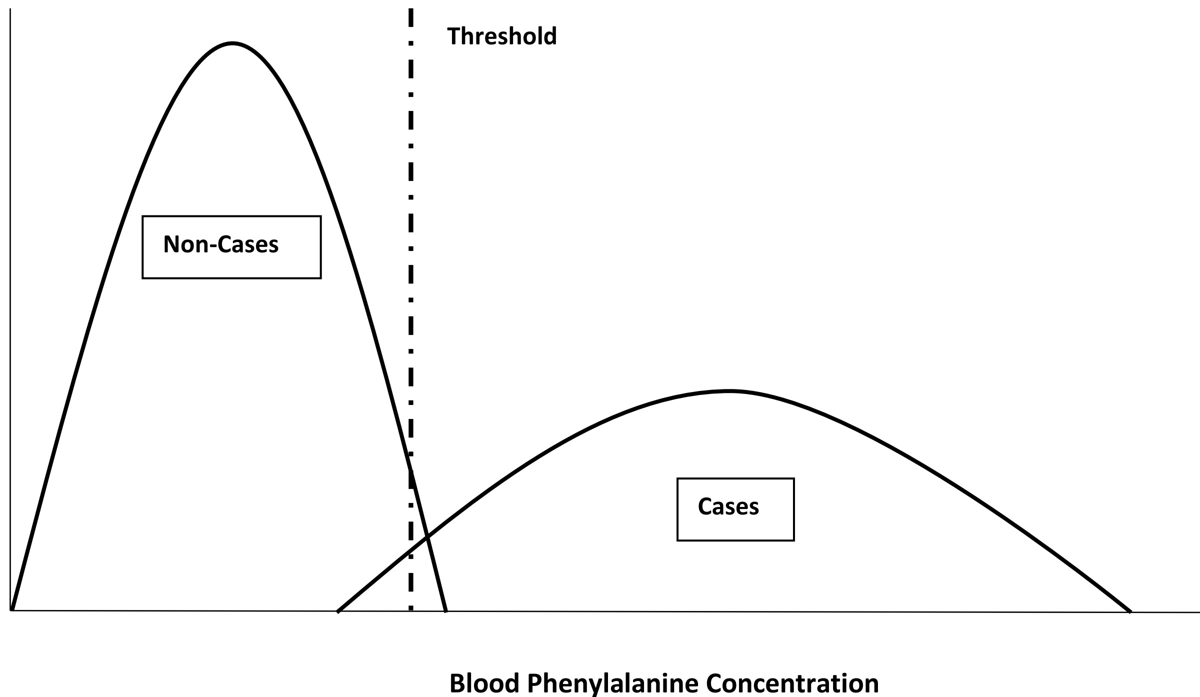
All of the IEMs screened for have an underlying mechanism of action at the molecular level that is caused by a pathogenic mutation at a particular gene locus.[22] Using the present MS/MS approach, NBS for IEMs is conducted by using biochemical tests that are designed to measure amino acid or acylcarnitine products of metabolic processes (analytes) that are regulated by these genes.[6,22] Infants who are identified with abnormal concentrations of a particular analyte may have a dysfunctional enzyme and are adjudicated to be at high risk for a screened condition. However,

following a positive NBS result, more complex diagnostic tests are still necessary to confirm that an infant has the screened condition.

Ideally, screening tests should be able to identify and distinguish infants who have a targeted condition from those who do not, making sensitivity and specificity important measures of performance in NBS. The nature of the screening test is outlined in **Figure 2.1**. Two notional population distributions, representing disease-affected and non-affected individuals, respectively, have been graphed according to the theoretical measure of a particular analyte. Most of the time, the populations will overlap on the measure of interest.[27] Hence, a threshold value that falls outside the range of most "normal" measures of the analyte must be determined. Non-diseased individuals with a measure greater than the threshold value will be false positives and diseased individuals with a measure below this point will be considered false negatives. Accordingly, an important objective of NBS programs is to establish a threshold that maximizes sensitivity while maintaining high specificity.

Traditionally, most NBS programs have emphasized the need for extremely high sensitivity to avoid false negatives (missed cases), though this depends on the severity of the disease being screened.[5,25] In most cases, the biomarkers are very accurate predictors of disease, so specificity is still very high.[5,25] However, due to the extremely low prevalence of each IEM in the population, the positive predictive value remains relatively low (i.e. a high proportion of positive screening results occur in individuals without disease and thus are eventually resolved as false positive).

Figure 2.1 Distribution of newborn screening test measure in disease-affected and non-affected populations with threshold value.



2.5 Limitations of Current Screening Technology

MS/MS-based screening is not a diagnostic test; it is intended to identify infants who are at high risk for IEMs so that they can be offered diagnostic evaluations and, if found to have disease can be offered early treatment.[21] Abnormal concentrations of screened biomarkers may be observed in some infants even if they do not have the underlying condition. For instance, infants born with low birth weight or lower gestational age are more likely to have false positive results for some of the screened conditions since their organ and endocrine function is not fully developed.[28] In addition, preterm infants are more likely to require admission in the intensive care unit and receive therapies (i.e. total parenteral nutrition, antibiotic administration, transfusion, *nulla per os* status (no breast milk or formula ingestion), and prenatal and postnatal glucocorticoid exposure) that may affect their NBS results by causing fluctuations in the concentration of screened analytes.[28] False positive results can also occur due to the nature of MS/MS technology, which is designed to test for

analytes associated with a disorder rather than detect the condition itself. A benign condition which elevates a screened analyte may result in a positive test for a screened condition that is associated with the same analyte.[21] On the other hand, in infants with milder phenotypes of a condition, screened analytes may only be elevated during specific periods, such as metabolic stress. At the time of screening, the analyte concentration may appear normal, potentially leading to missed cases.[21]

NBS results are generally available after the newborn has been discharged from the hospital. In the event of a positive screening result, families are contacted to arrange for a follow-up for confirmatory testing. However, many infants who are referred for additional testing are ultimately found not to have the disorder. Studies indicate that false positive results can have considerable impact on parents' psychosocial health.[29-31] A typical reaction to receiving a false positive result for PKU is anxiety, potentially due to unfamiliarity with the NBS process and a lack of understanding regarding the nature of the screened illnesses.[29] These effects, however, are transient and do not appear to persist a year after birth.[32] Even so, increased parental stress from false positive results has been documented to impair parent-child relationships, with parents being overprotective of their children and having an ongoing fear that their child might be developmentally delayed.[29] Distorted parental perception of their child's health may lead to more physician visits and hospitalizations, particularly within the first year of life.[33] Thus, avoiding false positive results can improve the entire NBS system by benefitting parent-child relationships and reducing unnecessary health service use.

2.6 Newborn Screening in Canada

In Canada, NBS is under provincial jurisdiction, where some provinces have their own advisory board and decision-making process.[34] Although all provinces use MS/MS as the primary screening technology for IEMs, the number of screened conditions ranges substantially, from as low

as six conditions in Newfoundland and Labrador to 28 primary and 24 secondary conditions screened in Manitoba.[35] Only three conditions are included in all NBS programs in Canada: PKU, medium chain acyl-CoA dehydrogenase (MCAD) deficiency, and congenital hypothyroidism.[34]

In Ontario, NBS is coordinated by Newborn Screening Ontario (NSO), which is based in the Children's Hospital of Eastern Ontario (CHEO) and funded by the provincial Ministry of Health and Long Term Care. Its establishment in 2006 was accompanied by a rapid expansion in the screening panel, and, at the time of writing, NSO primarily uses MS/MS to screen for 22 metabolic disorders in addition to a number of conditions that are screened using other technologies, including biotinidase deficiency, congenital adrenal hyperplasia, congenital hypothyroidism, cystic fibrosis, galactosemia, severe combined immune deficiency, and sickle cell disease.[36] NSO screens approximately 140,000 infants per year and over 1600 infants have been diagnosed with disease since its establishment.[36]

2.7 Expansion of NBS

The expansion of population-wide NBS has been described as "one of the major public health achievements" in recent times.[37] The ability to screen for numerous IEMs simultaneously led to a drastic increase in the number of conditions on NBS panels. Out of the 29 core conditions that have been recommended by the ACMG for NBS panels, only four have been assessed by experts as having existing treatments with the potential to prevent all negative consequences.[38] For the majority of the conditions on the core panel, existing treatments have been assessed as being able to prevent only some of the negative consequences of disease.[38] This indicates that the intervention treats a subset of the adverse effects, provides incremental improvement, or may not be equally effective in all individuals.[38] Therefore, in practice, in many (but not all) jurisdictions, it appears

that policy level justification of expanded NBS has been based on valuing a range of benefits beyond effectiveness of early intervention alone.

2.8 Next-Generation Sequencing

Nearly all conditions typically on NBS panels (including all IEMs) are genetic, caused by classic Mendelian patterns of inheritance. Moreover, most of the disease-causing genes for the NBS-targeted IEMs have been identified.[6,22] Coupled with the substantial reduction in the cost of whole genome, whole exome, or targeted gene sequencing methods (collectively termed "next-generation sequencing" (NGS), see below), this has created the possibility that a genomic approach to NBS could be feasible.[39] NGS enables large panels of genes to be simultaneously, rapidly, and accurately sequenced at low cost.[39-41] Previously, genetic testing needed to be performed on a gene-by-gene basis, using clinical and family histories as a guide, which cost thousands of dollars per case.[40,41] Today, NGS can be used to sequence the whole genome in a matter of days.[41] In fact, it may soon be less expensive to sequence an entire genome rather than single genes.[41] The infrastructure needed to support the use of NGS technologies is currently limited; as such, early applications of genome sequencing are likely to be restricted to settings where genetic testing is embedded in large scale program practice, such as NBS.[42] The apparent move towards a genetic approach to NBS has been driven by two key forces: first is the potential for a more rapid diagnosis to prevent adverse outcomes and second is financial savings resulting from shorter length of stay and reduced health service use associated with more rapid diagnosis.[40] Several authors have predicted that, once NGS technologies are sufficiently robust and affordable, they will be implemented across NBS programs and replace the current first-tier MS/MS.[16,18,41-43] This potential development is promoted by initiatives such as a National Institutes of Health funding program, which granted $25 million to four pilot projects in this area.[44]

Whole genome sequencing (WGS) and whole exome sequencing (WES) are laboratory processes that reveal the DNA sequence of an organism. WGS is performed to determine the sequence of the entire genome whereas WES is limited to the identification of protein-coding regions in genes.[45] A benefit of applying genomic sequencing into NBS would be the improved genetic characterization of screened conditions, which improves knowledge of the associated phenotypes of different genomic variants.[16,45] This can lead to more accurate diagnosis and better management of diseases with a strong heritable component.[16,41,43]

2.81 Targeted Next-Generation Sequencing

The technical approach used for WGS/WES can be applied in a targeted manner in NBS programs without making any fundamental changes to the present model. NGS can function strictly as a substitute technology designed to screen for exactly (and only) the same conditions currently tested for by MS/MS, essentially become an alternative screening method to the current technological approach.[40] Thus, sequencing in NBS would be restricted to a panel of target genes to identify highly penetrant disease-associated variants that are strong risk factors for currently tested conditions. In a collaborative effort, members of the Public and Professional Policy Committee of the European Society of Human Genetics, the Human Genome Organization Committee on Ethics, Law and Society, the PHG Foundation, and the P3G International Paediatric Platform have recommended that NGS should be applied to NBS in a targeted manner at this time.[16]

Nevertheless, even if sequencing is limited to only a panel of target genes, there are challenges associated with its implementation. The goal of targeted NGS is to identify rare, pathogenic mutations at target genes, but variants with other clinical manifestations may also be detected. Standards and guidelines developed by the ACMG assist in interpreting the clinical significance of different types of alleles by classifying them into five categories.

Table 2.1 describes the categories of genetic variants defined by the ACMG according to their clinical significance. It illustrates the range of associations between an allele and the probability of the disease of interest. These associations have been determined through clinical and genetic epidemiological studies, and take into account the mode of inheritance, for example autosomal dominant or recessive.

Table 2.1 Types of alleles classified by phenotypic manifestation.[46]

Type of Allele	Description
Wild-type	The most prevalent allele under normal conditions.
Pathogenic	There is very strong evidence that these variants are highly predictive of disease. They are often null variants that occur at genes where loss of regulatory function causes disease.
Likely pathogenic	These variants are moderately predictive of a disease. Conclusive evidence is still required to fully establish pathogenicity.
Benign	These are genetic mutations that do not cause damaging effects on protein function or splicing. Their allele frequency is greater than expected for a disease-associated variant.
Likely benign	Rare variants for that meet some criteria for benign impact, however the evidence is not conclusive.
Unknown significance	Variants that do not fully meet criteria for being benign or pathogenic. Conversely, they may meet criteria for both benign impact and pathogenicity, hindering the interpretation of their clinical significance.

An allele refers to an alternative form (or variant) of a gene that is located in a specific position on a particular chromosome.[47] The genotype of a person is the set of alleles that make up his or her genetic constitution.[47] The phenotype of an individual refers to the observed expression of a genotype as a morphological, clinical, cellular, or biochemical trait.[47]

A T-NGS approach would be designed to describe an individual's genotype at one or more defined gene loci . The most prevalent allele in the population for any particular genetic locus is referred to as the wild-type.[47] Apart from the wild-type allele, there can be several variant alleles,

each of which may have different phenotypic manifestations.[47] The combination of alleles (genotype), the mode of inheritance (autosomal or other, recessive or dominant), and the correlation between genotype and phenotype are all relevant to understanding the results from T-NGS. For a gene that is associated with a rare disease inherited in an autosomal recessive manner (the case for IEMs), having a genotype composed of two pathogenic alleles (strongly associated with disease) confers, by definition, very high risk for having the disease. Due the wide range of possible variants at a gene locus, a variety of results can be obtained from targeted genetic sequencing of a particular gene on the NBS panel. For instance, some infants will be identified as carriers of a pathogenic allele if their genotype is composed of one wild-type (or benign) allele and one pathogenic allele. On the other hand, the interpretation of results can be complex when variants of unknown significance (VUS) are detected; evidence regarding these variants is contradictory and their pathogenicity is yet to be established,[46] which may create a challenge for the clinical management of these newborns.

Inconsistencies in the genotype-phenotype relationship for many screened conditions can complicate the interpretation of sequencing results.[40] Even for monogenic disorders (diseases associated with a single gene), there have been reported instances where clinical manifestations of the condition are absent despite the presence of a classic disease-associated mutation.[40] In NBS, using supplementary phenotypic information can improve the interpretation of genotypic data. To facilitate the management of results, a screening test using both MS/MS and T-NGS in parallel may be more useful than applying either method alone. Combined with MS/MS, T-NGS offers the potential to enhance the performance of NBS by reducing false positives, identifying new disease-associated variants, and specifying genotypes that are associated with milder symptoms of disease.[40]

Table 2.2 describes screening results that are obtained from the present MS/MS-based approach to screening. With the addition of T-NGS, genotype may be considered along with MS/MS results to determine whether an infant should be referred for diagnostic follow-up. The use of additional genetic information prior to making a referral decision may improve the performance of NBS by enhancing sensitivity and specificity. However, the combination of these two

Table 2.2. Description of possible results using the current NBS system.

Screening Result	Description
Positive	Analyte being measured by MS/MS falls above (or below if low levels are disease-associated) the defined threshold for the test.
Negative	Analyte being measured by MS/MS is below (or above) defined threshold for the test.

technologies can increase the number and type of results that are obtained through NBS. The management of these new types of results need to be carefully considered before NGS can be implemented in NBS.

2.9 Rationale for this Study

In summary, technological advancements, particularly the emergence of MS/MS, have promoted the development and expansion of NBS. In fact, some contend that it now provides additional benefits beyond the prevention of adverse disease outcomes. There have been arguments that NGS is the next technological leap that will change the nature of NBS by improving its accuracy and providing genetic information that can be used to provide patients with more personalized care. Because of concerns about several ethical and logistical challenges associated with WGS/WES, expert committees have recommended for NGS to be applied in a targeted manner in the NBS context. As with any new health technology, the risks and benefits of T-NGS need to be evaluated and its implications on clinical practice, health services, and policymaking should be

considered. In this regard, quantifying the expected change in screening results from applying T-NGS in parallel with MS/MS can enable the assessment of some consequences that this technological shift would have on clinical management and the public health system. In addition, understanding the change in results can guide stakeholders in developing a model to implement T-NGS in NBS.

2.10 Research Question and Thesis Objectives

The overall aim of this thesis is to answer the following research question:

> ***In the context of newborn screening for IEMs, what will be the likely impact (in terms of types of screening results and their frequency) of changing the screening technology from first tier tandem mass spectrometry to an approach using targeted next-generation genomic sequencing and tandem mass spectrometry in parallel?***

This is the first study, to our knowledge, that estimates the different types of results that would be generated from using T-NGS in NBS and the frequency with which these results would occur. The knowledge generated by this study will inform our understanding of the need for associated health services and personnel required to manage the results that would be produced from the proposed approach. This study is expected to make an important contribution to research literature regarding the application of NGS technologies in the NBS context and will be useful in preparing for a possible shift in screening technologies at NSO and other NBS programs.

The <u>objectives</u> of the thesis are:

1) Evaluate the current performance of MS/MS-based NBS in Ontario through secondary analysis of data from NSO.

    i. Identify the frequency of positive and negative screening results for two case conditions, PKU and MCAD deficiency.
    ii. Identify the frequency of confirmed disease (PKU and MCAD deficiency) among newborns who were referred for follow-up diagnostic confirmatory testing following a positive screening result.
    iii. Determine measures of screening performance including sensitivity and specificity, as well as positive and negative predictive value.

2) Estimate T-NGS screening results for PKU and MCAD deficiency by querying genetic compendia for information regarding genetic variation in the genes association with the two case study conditions, *PAH* and *ACADM*, respectively.

    i. Identify known variants of *PAH* and *ACADM* and determine their clinical phenotype.
    ii. Estimate the allele frequencies of known variants of *PAH* and *ACADM*.
    iii. Use phenotypic and allele frequency data to predict possible genotypes and their prevalence in the Ontario population.

3) Estimate the results of a parallel MS/MS and T-NGS approach to NBS

    i. Estimate the genotypes of newborns who screened positive and negative for PKU using MS/MS.
    ii. Assess how adding T-NGS to NBS affects the performance of MS/MS screening.

CHAPTER 3: CASE STUDIES

Two IEMs have been selected as case studies: PKU and MCAD deficiency. These conditions have distinct characteristics that may affect NBS results and how these results are interpreted by NBS programs, making them informative for this particular study. The *ACADM* gene, associated with MCAD deficiency, has a single very common pathogenic mutation that is detected in the majority of individuals with the illness.[1] In contrast, the *PAH* gene, associated with PKU, has a broad spectrum of pathogenic variants that vary in terms of their phenotypic expression.[48] Additionally, the genotype-phenotype relationship between *PAH* and PKU is relatively well-defined. The specific mutation on *PAH* and its effect on the structure of the PAH enzyme, its gene product, can help to predict the severity of the illness.[48,49] This relationship is not consistent in the case of MCAD deficiency, where infants with two copies of a pathogenic allele may not present with symptoms of disease.[1] Both conditions are discussed in further detail below.

3.1 Phenylketonuria

PKU is an IEM with an incidence of approximately 1 in 12,000 in Canada.[50] It is a genetic disease that is inherited in an autosomal recessive manner and is caused by mutations in the *PAH* gene that affect the metabolism of phenylalanine (Phe).[49] Infants at high risk for PKU are identified by NBS programs in most developed countries.[51] Mutations in the *PAH* gene can cause a deficiency of PAH enzyme, leading to the accumulation of Phe in the blood.[49,51] If untreated, the elevated Phe affects brain function and can cause severe and irreversible intellectual disability.[49] PKU can manifest at different degrees of severity and the clinical management of the condition varies accordingly.[52] The different classes of PKU are described in **Table 3.1**.

Table 3.1 Classification of PKU according to severity.[52]

Class of PKU	Description	Untreated Blood Phe Concentration
Classic PKU	Most common and severe form of the disorder. Complete PAH enzyme deficiency.	≥ 1200 µmol/L
Moderate PKU	Less severe form of PKU.	600-1200 µmol/L
Mild PKU	Mild form of PKU.	360-600 µmol/L
Mild Hyperphenylalaninemia	Mild variant; these patients are not diagnosed with PKU. Does not require clinical intervention.	< 360 µmol/L

Elevated concentrations of Phe in the blood can be detected as early as 24 hours after birth.[52] In NBS laboratories that use MS/MS, both Phe and tyrosine (Tyr) are analyzed to improve the sensitivity and positive predictive value of the test;[51] the sensitivity of PKU has been reported to be close to 100%.[24] Newborns who screen positive for PKU must be immediately followed-up for confirmatory testing.[51]

To prevent severely adverse health outcomes in the neonate, therapy for PKU should begin within the first week of life.[52] Phe may be excluded from the diet until levels approach the treatment range.[52] The main form of treatment provided to PKU patients involves the substitution of foods containing natural protein with "medical foods" that are free of Phe.[49,51,52]

3.2 MCAD Deficiency

MCAD deficiency is a fatty acid β-oxidation disorder that has an incidence of approximately 1 in 14,000 in Ontario.[53] In unaffected individuals, β-oxidation provides an alternate, but essential, source of energy during periods of fasting or increased energy demands such as concurrent infection.[54] Under these conditions specifically, individuals with symptomatic MCAD deficiency may experience hyperammonemia, transient hypoglycemia, metabolic acidosis, cardiomyopathy and sudden death.[54] MCAD deficiency is inherited in an autosomal recessive manner and is caused

by mutations of the *ACADM* gene.[54] These mutations result in a deficiency of the MCAD protein, which catalyzes the β-oxidation of fatty acids that are six to 10 carbons long.[54]

The goal of NBS is to identify infants with MCAD deficiency before symptoms of the disease can arise. The symptomatic presentation of MCAD deficiency occurs during periods of prolonged fasting or excessive energy demand, and is associated with high morbidity and mortality.[55] The primary approach to preventing adverse outcomes is awareness of the need to avoid prolonged periods of fasting and clinical supervision during periods of illness.[54,55] Screening for MCAD deficiency is performed by analysis of the newborn's acylcarnitine profile.[55] Carnitine is an amino acid derivative that transports acyl groups from fatty acids (as acylcarnitine) into the mitochondrial matrix so they can be catabolized via β-oxidation and used for energy.[56] Newborns with suspected MCAD deficiency have elevated blood levels of medium-chain acylcarnitines (C_6-C_{10}), with increased concentrations of C_8 being the most definitive biomarker.[55] The sensitivity of NBS for MCAD deficiency is extremely high and false negative results are rare, though they are possible if an individual's overall blood carnitine concentrations are decreased due to other conditions such as secondary carnitine deficiency.[55] MCAD deficiency also has been reported to have varied clinical presentation.[57] In some cases, adverse outcomes may not develop in children with two pathogenic mutations on *ACADM;* the association with clinical symptoms is likely to depend on other genetic as well as environmental factors.[57]

Management of MCAD deficiency is relatively simple as most patients tend to be well and specific treatment is not necessary.[55] Parents are usually advised to ensure the child avoids prolonged periods of fasting or catabolic stress, with close attention during periods of infection and immunization.[54,55] High carbohydrate drinks or intravenous glucose are effective in maintaining normoglycemia in situations of fever, vomiting, and diarrhea.[55] Families are often advised to carry a

letter or information sheet to present to health professionals who may not be familiar with the management protocol for MCAD deficiency.[55]

CHAPTER 4: METHODOLOGY

4.1. Study Design

There are three components to this thesis: a secondary analysis of retrospective data collected at NSO, an analysis of genetic databases to predict T-NGS results, and an estimation of the distribution of results that would be generated using an approach that uses both MS/MS and T-NGS in NBS.

4.2. Objective 1: Retrospective Analysis of MS/MS Data

To meet the first objective of this thesis, we analyzed population-level data collected at NSO. The data set contains information regarding the screening results for PKU and MCAD deficiency for nearly all newborns who were born in Ontario between April 2006 and November 2015, and diagnostic follow-up results for those who screened positive.

In order to conduct newborn screening in Ontario, blood samples are collected from hospitals, birthing centres, and midwifery clinics across Ontario. Ideally, between 48 and 72 hours of birth, samples are collected from the newborn using a heelprick, while personal, health, and demographic information is recorded from the infant's medical chart. The samples and accompanying information are transported to the NSO laboratory (based at CHEO) for testing, which occurs in two phases. The initial phase of testing is performed on all samples, and those with abnormal analyte values undergo the confirmatory (but still screening rather than diagnostic) phase of testing, which generally involves a repeat of the initial tests. Consistently out-of-range analyte values during both phases generate a screen-positive result and the infant is referred for diagnostic follow-up, which is organized through a pediatric tertiary care centre (the specific centre responsible for follow-up depends on where the child resides), with the final diagnostic resolution

reported back to NSO. Inconsistent results between the two phases of screening may require a new sample to be collected from the infant. The criteria for a screen positive result for PKU and MCAD deficiency are described in **Table 4.1** and **Table 4.2** below.

Table 4.1 MS/MS screening criteria for PKU.

MS/MS Screening Result	Criteria
Positive	Phe ≥ 160 µmol/L **OR** Phe ≥ 130 µmol/L *and* Phe/Tyr ≥ 1.72
Negative	Does not meet criteria for positive screening result

Table 4.2 MS/MS screening criteria for MCAD deficiency.

MS/MS Screening Result	Criteria		
	Age < 7 days	Age ≥ 7 days	All Ages
Positive	C8 ≥ 0.5 **OR** C8 ≥ 0.4 *and* (C6 ≥ 0.25 *or* C10:1 ≥ 0.15 *or* C10 ≥ 0.30 *or* C8/C2 ≥ 0.01 *or* C8/C10 ≥ 2.50)	C8 ≥ 0.4	C8 ≥ 1 *and* C8/C10 ≥ 6
Negative	Does not meet criteria for positive screening result		

The process for sample collection differs for newborns who are born preterm (<33 weeks gestational age) or with very low birthweight (<1,500g). Newborns in these categories require a repeat sample either before discharge or at 3-4 weeks of age. This is repeated at 4-6 months of age if the newborn received a packed red blood cell transfusion, and an adequate pre-transfusion sample could not be collected.

The protocol for diagnostic testing varies depending on the disease under consideration. Diagnostic testing for PKU involves the measurement of additional analytes (which confer a higher degree of specificity) using more complex techniques. Diagnostic testing for MCAD deficiency involves the use of more complex biochemical tests, but infants may also undergo genetic sequencing to determine whether disease-associated mutations are present on the *ACADM* gene.

Table 4.3 List of variables in NSO data set for MS/MS screening and diagnostic follow-up.

Variable		Description
Screening	Patient ID	Patient-level identifier
	Specimen ID	Specimen-level identifier
	Year-month of collection	In the format YYYYMM
	Age at collection	In hours
	Birth weight category	
	Gestational age category	
	Unsatisfactory reason	Reason for unsatisfactory sample
	Transfusion indicator	Yes/No
	Phe	Phenylalanine concentration (µmol/L)
	Phe/Tyr	Phenylalanine to tyrosine ratio
	C8	C8 concentration
	C8/C2	C8/C2 ratio
	C8/C10	C8/C10 ratio
	C6	C6 concentration
	C10	C10 concentration
	C10:1	C10:1 concentration
	Screening outcome PKU	Positive/Negative
	Screening outcome MCAD deficiency	Positive/Negative
Diagnostics	Treatment centre	Regional treatment centre for follow-up testing
	Definitive diagnosis, PKU	Specific result of diagnostic follow-up testing for PKU (i.e. classic PKU, mild PKU, liver disease, deceased)
	Definitive diagnosis, MCAD deficiency	Specific result of diagnostic follow-up testing for MCAD deficiency (i.e. classic MCAD deficiency, glutaric aciduria type II, deceased)
	Final outcome, PKU	Final outcome of diagnostic testing for PKU (i.e. yes, no, variant, incidental)
	Final outcome, MCAD deficiency	Final outcome of diagnostic testing for MCAD deficiency (i.e. yes, no, variant, carrier)

Demographic and other brief information (e.g., birth weight, gestational age at birth, feeding status) are provided on the blood spot card that is submitted to NSO for screening by the birthing

centre or midwifery practice where the sample was taken. To ensure the quality of the NSO data set, data entry clerks identify missing data fields and contact the submitting birthing centre to obtain this information. A list of variables that were included in the data set is described in **Table 4.3.**

4.21 Outcomes of MS/MS Screening

With respect to the analysis of NBS using MS/MS, the primary outcomes were:

1. The frequencies of positive and negative screening results for both PKU and MCAD deficiency during the specified period.
2. The definitive diagnosis of infants who were referred for diagnostic follow-up after receiving a positive screening result.
3. The number of missed cases (false negatives) of PKU or MCAD deficiency.
4. Screening performance metrics including sensitivity, specificity, positive predictive value, and negative predictive value, computed by the student using record-level data regarding screening results and diagnostic follow-up information.

Outcomes 1 and 2 were obtained from the NSO data set, and used to calculate the specificity and positive predictive value. False negatives (missed cases) of PKU and MCAD deficiency, once identified, are reported back to NSO, where they are kept on record. Thus, outcome 3 was obtained by performing an inquiry of these records for reports of missed cases. This was used to calculate sensitivity and negative predictive value.

4.22 Data Cleaning

Before the NSO data set could be analyzed, several steps were taken to evaluate and clean the data. The following problems identified in the data set:

Missing Data

We examined all variables in the data set and assessed the proportion of total observations that were missing for each variable. Missing value proportions that were less than 5% were considered acceptable and would be handled by case-wise deletion. A complete list of missing variable proportions is available in **Appendix A**.

We identified a high proportion of missing values among three variables:

Birth weight category (5.2%)

Gestational age category (37.2%)

Screening outcome, MCAD deficiency (3.0%)

Upon further investigation. We found that nearly all infants whose specimen were taken prior to 2010 were missing birth weight and gestational age. Before 2010, the screening card, where patient information is recorded, only required these two fields to be entered if infants had low birth weight or were preterm. Therefore, we assumed that infants who were missing data for birth weight or gestational age were not born with low birth weight or preterm, respectively. This was acceptable since we used these two variables to create a new dichotomous variable labelled "Preterm/LBW", where infants who were preterm or low birth weight were assigned a code of '1'. Otherwise, they were assigned a code of '0'.

Although screening outcome for MCAD deficiency was missing less than 5% of observations, we investigated this further since over 40,000 values were missing for one of the most important variables in the data. We noted that nearly all of the missing values were between April 2006 and July 2006 (almost all of the infants who were screened during this period were also missing this data). This period included the first infants to be screened by NSO and there is a

possibility that there may have been issues with data entry at the time. Although the screening outcome for MCAD deficiency was missing, diagnostic follow-up data was fully available for infants who screened positive for MCAD deficiency during this period. Therefore, we assumed that if infants who were screened between April 2006 and July 2006 had diagnostic follow-up data available, then the screening outcome variable for MCAD deficiency for these infants must be "positive". If there was no diagnostic data available, then the screening outcome was "negative".

Faulty Observations

We checked all continuous variables for outliers and invalid values (i.e. year-month of collection was Jan. 1900). We also checked all categorical variables for invalid responses. We noted the following:

Age: some values were lower than 0

Date: some dates were Jan. 1900

Since a very small proportion of values for these variables were invalid, we resolved these issues by case-wise deletion.

Unsatisfactory Samples

Certain samples in the data set were labelled as unsatisfactory for a number of reasons (i.e. blood spots were scratched, missing demographic information, not enough blood in sample). All samples that were considered unsatisfactory were deleted. Patients with unsatisfactory samples were recalled for a repeat test.

Duplicate Patient IDs

The NSO data set was initially sorted by specimen. For this thesis, the analysis was performed at the population level, requiring the data set to be sorted by patient. Repeat samples may be required from infants whose initial DBS sample was classified as unsatisfactory by the laboratory. In addition, infants who are born preterm or with very low birth weight may require a transfusion. DBS samples taken immediately after a transfusion may yield invalid results. As a result, these infants needed to provide repeat samples for testing, with some requiring up to six samples. When there were multiple specimen for a single infant, we followed specific protocol to select one specimen per infant:

1. Observations that had a positive screening result either for PKU or MCAD deficiency were selected to maximize the number of positive screens in the data set.
2. Observations indicating that an infant was born preterm, with low birth weight, or required a transfusion were deleted.
3. If there were multiple but valid observations with negative screening results, then any one observation was selected at random. This was judged to be an appropriate method since we were only interested in the final screening outcome, which was the same regardless of which observation was chosen.

4.3. Objective 2: Analysis of Genetic Compendia to Estimate T-NGS Results

To meet the second objective, we queried publicly available genetic compendia (discussed below), which contain information regarding the prevalence of genetic variants in humans and the clinical significance of these variants.

4.31 Genetic Compendia

Genetic compendia contain records pertaining to human genetic variation. Various clinical and scientific laboratories report specific information on genetic variants within these archives. A report by the ACMG provides guidance in the selection of these databases. Important factors considered were: i) the frequency with which the database is updated; ii) whether the appropriate nomenclature is used and the genome build and transcript references used for naming variants; iii) the degree to which data were validated for analytical accuracy and the use of quality metrics to assess data accuracy; and iv) the source and independence of observations used.[46] We selected three genetic compendia that included a comprehensive list of variants, sampled from a large population, and verified the quality of the data.

Genome Aggregation Database (gnomAD)

gnomAD is a genetic compendium, the largest of its kind, that was developed by a team of investigators who sought to combine genome and exome sequencing data from numerous disease-specific and population sequencing projects.[58] gnomAD is the second release of the original database, the Exome Aggregation Consortium, which exclusively contained exome (coding regions) sequencing data.[58] For the two genes of interest, *PAH* and *ACADM*, the data in this particular compendium were obtained from whole genome sequencing of 15,496 individuals.

ClinVar

ClinVar is a freely accessible public archive which reports the relationship between different genetic variants, their phenotypic manifestations, and their clinical significance, with supporting evidence provided.[59] The interpretation of clinical significance is provided according to guidelines specified by the ACMG. Our classification of alleles (**Table 3.5**) corresponds to the categories used by ClinVar. A wide range of variants are included in ClinVar, including single nucleotide substitutions, small insertions and deletions, copy number changes, and cytogenetic rearrangements.[59] Interpretations of genetic variants are submitted by clinical testing laboratories, research laboratories, and locus-specific databases.[59]

dbSNP

dbSNP is a free public archive developed by the National Institute for Biotechnology Information that contains data pertaining to genetic variation. Along with single nucleotide polymorphisms, its collection includes small-scale insertions and deletions, short tandem repeats, heterozygous sequences, multinucleotide polymorphisms, and named variants.[60]

4.32 Data Extraction

We queried ClinVar for known variants of *PAH* and *ACADM*. With the exception of minor allele frequency (MAF), which was not available on ClinVar, we extracted all data fields specified in **Table 4.4**. We sorted the variants by their clinical significance prior to retrieval. If clinical significance for a variant was not provided, it was not eligible for inclusion. In addition, variants of *PAH* and *ACADM* that were not associated with PKU, hyperphenylalaninemia, or MCAD deficiency were excluded.

Variants that were identified on ClinVar were matched on gnomAD using the standard reference assembly sequence, Rs ID number, molecular consequence, and type of mutation. Once we confirmed that the variants on both databases were identical, the MAF was extracted from gnomAD. If the MAF was not available on gnomAD, we performed an additional search for the MAF on dbSNP. Variants that were unavailable on both gnomAD and dbSNP were assumed to be extremely rare, preventing MAF from being adequately estimated, and were excluded from final analysis.

Conflicting Interpretations of Clinical Significance

Some variants on ClinVar had multiple entries of clinical significance that were submitted by different genetic laboratories. Although the ACMG provided guidelines, laboratories may have varying interpretations, resulting in conflicting judgments regarding the classification of the variant. For instance, a particular variant may have been classified as pathogenic by one laboratory, and of uncertain significance by another. We managed conflicting interpretations of clinical significance based on the following criteria:

i) The submission dates were noted, as newer evidence of pathogenicity may be more pertinent.
ii) Evidence from the literature was assessed to determine whether the variant demonstrated pathogenicity in humans.
iii) Supporting evidence, if provided by the submitted laboratory, was examined.

If we could not reach a conclusion regarding the clinical significance of these conflicts, the associated variant was excluded from the analysis.

Table 4.4 Data fields extracted from genetic compendia.

Variable	Description
Standard reference assembly sequence (GRCh37)	A reference assembly is a nucleic acid sequence database intended to represent a species' set of genes and can provide information regarding the particular location of a variant
Molecular consequence	Describes the particular effect that a variant has at the nucleic acid level (i.e. change in base pair) or at the protein level (i.e. change in amino acid)
Rs ID number	A reference ID that allows single nucleotide polymorphisms (SNPs) to be linked between databases, including gnomAD, ClinVar and dbSNP
Type of Mutation	Classifies variants based on how they affect DNA (e.g. missense, nonsense, insertion, deletion, etc.)
Phenotype	The condition that is associated with the pathogenic variant of a gene (e.g., PKU, hyperphenylalaninemia, MCAD deficiency).
Clinical Significance	Identifies the degree of pathogenicity of a particular variant. Variants were classified by submitting laboratories as pathogenic, likely pathogenic, uncertain significance, benign, or likely benign, using guidelines provided by the ACMG (Richards, 2015).
Minor Allele Frequency (MAF)	Refers to the frequency at which the second most common allele for a given genotype occurs in the population.

4.33 Estimating Genotype Proportions

Deriving genotype proportions in a population can be challenging even if the allele proportions of a particular variant are known. Under certain conditions, the Hardy-Weinberg principle specifies a mathematical relationship[47] that can be used to estimate genotype proportions from allele proportions of known variants:

[1] $p + q = 1$, where:

p = proportion of dominant allele, *A*

q = proportion of recessive allele, *a*

According to Equation 1, for a given gene with two types of alleles, the sum of the allele proportions will be equal to 1. The binomial expansion of this relationship will provide the proportions of genotypes *AA*, *Aa*, and *aa*.[47]

[2] $(p + q)^2 = 1$

$p^2 + 2pq + q^2 = 1$, where:

p^2 = proportion of homozygous dominant genotype, *AA*

$2pq$ = proportion of heterozygous genotype, *Aa*

q^2 = proportion of homozygous recessive genotype, *aa*

The proportion of a given population that have the *AA*, *Aa*, and *aa* genotypes are represented by the terms p^2, $2pq$, and q^2, respectively. The Hardy-Weinberg principle states that genotype proportions for a gene remain the same as long as the allele frequencies in a given population do not change.[47] Consequently, the Hardy-Weinberg principle is applicable when the following assumptions[47] hold true:

i) There is no genetic drift
ii) There is no migration
iii) There are no mutations
iv) There is random mating
v) There is no natural selection

We combined alleles with similar clinical significance to obtain three categories: "pathogenic or likely pathogenic", "uncertain significance", and "benign, likely benign or wild-type". The equation defined by the Hardy-Weinberg principle is applicable for two alleles. Therefore, it was modified to be used for three types of alleles:

[3] $p + q + r = 1$, where:

p = proportion of benign, likely benign, or wild-type alleles

q = proportion of pathogenic or likely pathogenic alleles

r = proportion of alleles of uncertain significance

Since PKU and MCAD deficiency have an autosomal recessive mode of inheritance, for the *PAH* and *ACADM* genes, the benign, likely benign, or wild-type (p) alleles were considered the dominant allele, and pathogenic or likely pathogenic (q) alleles were considered the recessive allele. A third type of allele, uncertain significance (r), which cannot be classified into the benign or pathogenic group, was also been added to the equation. In order to obtain the genotype proportions, a trinomial expansion of the mathematical term above was necessary:

[4] $(p + q + r)^2 = 1$

$$p^2 + q^2 + r^2 + 2pq + 2pr + 2qr = 1$$

The expressions described in **Table 4.5** represent each of the possible genotypes defined in this study that can be observed through T-NGS.

Table 4.5 Hardy-Weinberg variables and their description

Expression	Description (Allele 1 || Allele 2)
p^2	Benign or likely benign or wild-type || Benign or likely benign or wild-type
q^2	Pathogenic or likely pathogenic || Pathogenic or likely pathogenic
r^2	VUS || VUS
2pq	Pathogenic or likely pathogenic || Benign or likely benign or wild-type
2pr	Benign or likely benign or wild-type || VUS
2qr	Pathogenic or likely pathogenic || VUS

Variants were stratified into three groups according to their clinical significance (pathogenic or likely pathogenic, uncertain significance, benign, likely benign, or wild-type). The MAFs of variants within each group were summed (p, q, and r in equation 3). Benign, likely benign, or wild-

type alleles may not be minor alleles because they are often not rare variants, making it inappropriate to sum the MAFs of these variants. Therefore, p was calculated as the difference between 1, q, and r. The values of p, q, and r were then substituted into equation 4 to obtain genotype frequencies (*p2, q2,* r^2*, 2pq, 2pr, 2qr* in equation 4) in the population.

4.4 Objective 3: Primary Analyses, Combining MS/MS and Estimated T-NGS Results

To meet the final objective of this thesis, we classified the MS/MS screening and diagnostic follow-up results obtained from the NSO data set into the six genotypic categories. We anticipated that an infant's genotype, as identified by T-NGS, would be associated with their MS/MS screening results, and, among infants who screened positive, their diagnostic follow-up results. We used a series of assumptions to model this particular association; these assumptions were based on data that were obtained from the literature and the etiology of the two diseases. By predicting the genotype-phenotype association, we were able to populate the empty cells in **Table 4.6** for both PKU and MCAD deficiency.

Table 4.6 Genotype proportions of infants in the NSO population.

<table>
<tr><th rowspan="3">T-NGS Result
(Allele 1 || Allele 2)</th><th colspan="4">MS/MS Result</th><th rowspan="3">T-NGS Totals</th></tr>
<tr><th colspan="2">Positive</th><th colspan="2">Negative</th></tr>
<tr><th>True</th><th>False</th><th>True</th><th>False</th></tr>
<tr><td>P or LP || P or LP</td><td></td><td></td><td></td><td></td><td>q^2 x Total Population</td></tr>
<tr><td>P or LP || VUS</td><td></td><td></td><td></td><td></td><td>2qr x Total Population</td></tr>
<tr><td>VUS || VUS</td><td></td><td></td><td></td><td></td><td>r^2 x Total Population</td></tr>
<tr><td>P or LP || B, LB, or WT</td><td></td><td></td><td></td><td></td><td>2pq x Total Population</td></tr>
<tr><td>VUS || B, LB, or WT</td><td></td><td></td><td></td><td></td><td>2pr x Total Population</td></tr>
<tr><td>B, LB, or WT || B, LB, or WT</td><td></td><td></td><td></td><td></td><td>p^2 x Total Population</td></tr>
<tr><th rowspan="2">MS/MS Totals</th><td>Total True Positives</td><td>Total False Positives</td><td>Total True Negatives</td><td>Total False Negatives</td><td rowspan="2">Total Population</td></tr>
<tr><td colspan="2">Total Screen Positives</td><td colspan="2">Total Screen Negatives</td></tr>
</table>

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

4.41 Association between Genotype and MS/MS Screening Results

For both PKU and MCAD deficiency, we expect a very strong association between T-NGS results (genotype) and MS/MS results. In particular, we expected that infants who are homozygous for a pathogenic or likely pathogenic mutation were more likely to have a positive screening result due to a deficiency in the function of the *PAH* or *ACADM* genes (Al-Hafid, 2015; Vishwanath, 2016). To account for this association, we estimated the total number of infants with this particular genotype who had a negative screening result. To obtain this estimate, we recorded the number of false-negatives in the screened population from the NSO data and allocated them into the table as "false negatives" with the "*P or LP* // *P or LP*" genotype. We would not expect any infants who are "true negatives" to have the *P or LP* // *P or LP* genotype. The clinical definition for both PKU and MCAD deficiency state that these diseases are caused by two pathogenic mutations at the *PAH* and *ACADM* gene, respectively. Therefore, it is highly unlikely that an infant with this genotype will not have PKU or one of its milder variants (or MCAD deficiency).

4.42 Association between Genotype and Diagnostic Follow-up Results

Among infants who screened positive for PKU or MCAD deficiency, we anticipated a very strong association between being a true positive and genotype (i.e. infants with a *P or LP* // *P or LP* genotype are more likely to be true positives). In order to account for this association, we conducted a literature search on PubMed for studies in the last 10 years that have assessed the genotype-phenotype association for the *PAH* and *ACADM* genes. These studies provided data that enabled us to estimate the genotype frequencies of infants diagnosed with PKU or MCAD deficiency (true positives).

4.43 PKU: Benign Hyperphenylalaninemia

After receiving a positive screening result for PKU, some newborns are diagnosed with benign hyperphenylalaninemia (HPA), a mild variant of PKU characterized by Phe levels between 120 µmol/L and 360 µmol/L that does not require any clinical intervention. Benign HPA is classified as a false-positive by NSO. According to its clinical definition, benign HPA is caused by mild, but pathogenic mutations on the *PAH* gene. Therefore, we expected an association between benign HPA and genotype. In order to account for this association, we search PubMed for studies that investigated the genotype-phenotype association between *PAH* and benign HPA. These studies provided data that enabled us to estimate the genotype of infants in the NSO population who have benign HPA.

4.44 MCAD Deficiency: Carriers

Unlike PKU carriers, infants who are carriers of a single pathogenic allele on the *ACADM* gene are more likely to have a positive screening result due to elevated acylcarnitine levels.[57,61] Since diagnostic follow-up testing for MCAD deficiency may involve mutational analysis, some carriers of a pathogenic allele may be identified. These infants were allocated as false-positives with one of the following three genotypes:

i) P or LP || VUS

ii) VUS || VUS

iii) P or LP || B, LB, or WT

4.45 False Positives and True Negatives

For the remaining false positive as well as the true negative results, we did not expect an association between MS/MS result and T-NGS-identified genotype. Therefore, assuming that these

results are independent of one another, we calculated the expected frequencies within each cell in **Table 4.6** by multiplying the marginal totals.

4.5 Sensitivity Analyses

We performed two sets of sensitivity analyses to test the robustness of assumptions made in the above-described primary analyses, for each of PKU and MCAD deficiency.

4.51 Genotype Proportions in All Newborns

First, we acknowledged that the samples in the genetic databases that were used to calculate genotype proportions may not be representative of the NSO population. We varied these proportions based on the upper and lower limits of the 95% confidence interval and assessed whether this led to changes in our conclusions.

4.52 Genotype Proportions in True Positives

As described above, for the primary analyses, we assessed the genotype-phenotype association for PKU and MCAD deficiency by performing a literature search. Using studies identified through this search, we estimated the proportions by genotype among newborns who are "true positives" for PKU and MCAD deficiency. Acknowledging that the samples used in these studies may not be representative of the NSO population, to which the results were applied, we conducted a further sensitivity analysis by varying these proportions according to the upper and lower limits of their 95% confidence intervals.

4.6 Ethical Considerations

Newborn screening data are collected and stored for clinical, quality assurance, and research and development purposes. Parents are provided with an information package regarding NBS as

well as the collection and storage of data **(Appendix B).** Consent for NBS and the use of screening data and samples for research is implicit – parents may opt out if desired. The use of data from screening dried blood spot samples is covered by a specific policy regarding the storage and secondary use of NBS samples at NSO (**Appendix C**).

Appropriate measures were taken to protect the confidentiality of newborns. Screening data were de-identified and only non-identifying information was maintained in the version of the data set that was provided for the study. Moreover, data were stored on a secure private network at NSO and were neither transferred nor accessed outside of the institution. All analyses were performed using a laptop computer provided by NSO. Collectively, these measures substantially reduced the risk of disclosure of identifying information.

The study received approval from the Research Ethics Boards of CHEO and the Ottawa Health Sciences Network (**Appendix D**).

CHAPTER 5: RESULTS

5.1 Retrospective Analysis of NSO Data

Between April 2006 and November 2015, there were 1,380,110 infants screened in Ontario. In this population, there were 535 (0.04%) infants who screened positive for PKU and 276 (0.02%) infants who screened positive for MCAD deficiency.

Table 5.1 Summary of MS/MS screening results at NSO

MS/MS Screening Result	N (%)	
	PKU	MCAD Deficiency
Positive	535 (0.04%)	276 (0.02%)
Negative	1,379,575 (99.96%)	1,379,834 (99.98%)
Total	**1,380,110**	**1,380,110**

5.11 PKU Screening and Diagnosis

Table 5.2 describes the results of diagnostic testing for PKU. Of the 535 infants who screened positive for PKU, 86 were diagnosed with the condition during follow-up testing. This results in an overall prevalence of 1 case per 16,048 infants. From the 449 newborns who were false positives, 79 had persistently elevated levels of Phe and were diagnosed with benign HPA; these infants do not require any further clinical intervention and were classified as false positives by NSO.

Table 5.2 Summary of diagnostic testing results for PKU

PKU Outcome	Frequency
Yes	86
No	326
Benign Hyperphenylalaninemia (Benign HPA)*	79
Incidental**	11
Other***	33
Total	**535**

*Benign HPA is considered a false positive by the NBS system. These infants do not require clinical intervention.
**Includes liver disease, 6-pyruvoyl tetrahydropterin synthase deficiency, and persistent lab abnormalities as explanations for the positive screening result.
***Includes infants who were lost to follow-up or deceased.

Eleven infants were diagnosed with secondary, non-targeted conditions including liver disease and 6-pyruvoyl tetrahydropterin synthase deficiency. There were 33 infants in the "Other" category, of whom 24 had died prior to diagnostic testing and 9 were lost to follow-up. No false negatives for PKU were reported to NSO in the screened population.

Based on these results, we calculated test performance for MS/MS-based screening of PKU. Sensitivity and specificity were 100% and 99.97%, respectively, while positive and negative predictive value were 16.1% and 100%, respectively.

5.12 MCAD Deficiency Screening and Diagnosis

Table 5.3 describes the results of diagnostic testing for MCAD deficiency. Of the 276 infants who screened positive for MCAD deficiency, 94 received a confirmed diagnosis with the condition, 16 of whom had a variant of MCAD deficiency that was predicted to be clinically mild. The prevalence of MCAD deficiency in this population was 1 in 14,682 infants. Of the 182 infants who were false positives, 37 were carriers of a single pathogenic mutation, and six had a secondary condition called glutaric acidemia II (GA II). There were 15 false positives categorized as "Other", of which 12 died and three were lost to follow-up. One false negative in the screened population was reported to NSO during the study period.

Table 5.3 Summary of diagnostic testing results for MCAD deficiency

MCAD Deficiency Outcome	Frequency
Yes – Classic MCAD Deficiency	78
Yes – Variant MCAD Deficiency	16
No	124
Incidental	43 (37 carriers and 6 GA II)
Other*	15
Total	**276**

*Includes infants who were lost to follow-up or deceased.
GA II – Glutaric acidemia II

Test performance for MS/MS-based screening of MCAD deficiency was calculated according to the observed results. Sensitivity and specificity were 98.95% and 99.99%, respectively, and the positive and negative predictive values were 34.1% and ≈100% (99.99992%), respectively.

5.2 Analysis of Genetic Databases

5.21 Query for Genetic Variants

A search of ClinVar identified 309 entries for *PAH* and 233 entries for *ACADM*.

Some variants in ClinVar had multiple entries, with one entry described as a pathogenic variant and another described as likely pathogenic (or one as benign and the other as likely benign). Since we combined these pathogenic and likely pathogenic variants (and benign and likely benign) for this study, we treated the second entry as a duplicate, and it was excluded.

There were four variants of *PAH* and 11 variants of *ACADM* with conflicting interpretations. All conflicts for *PAH* were resolved while a resolution could not be reached for five of the conflicts for *ACADM* (**Appendix E**). After removing duplicates and addressing conflicting interpretations, 231 *PAH* variants and 184 *ACADM* variants remained.

Variants were queried on gnomAD and dbSNP for MAF data; if data were not available, then the variants were assumed to be extremely rare and excluded from the final dataset. After exclusion of these rare variants, 147 *PAH* variants and 114 *ACADM* variants were retained for final analysis.

5.22 Characteristics of Genetic Variants

Table 5.4 shows the distribution of variants by clinical significance cateogies retained for final analysis or excluded due to unresolved conflicts or unavailable MAFs. For both *PAH* and *ACADM*, a chi-squared test indicated that there was no significant difference between included variants and excluded variants with respect to their clinical significance category (p=0.25 for *PAH*; p=0.58 for *ACADM*).

Table 5.4. Clinical significance of variants that were included and excluded from ClinVar and gnomAD.

Clinical Significance	PAH	ACADM
Included Variants	**N, [% (95% CI)]**	**N, [% (95% CI)]**
Pathogenic or likely pathogenic	111, [75.5 (68.6-82.5)]	50, [43.9 (34.8-53.0)]
Uncertain significance	15, [10.2 (5.3-15.1)]	44, [38.6 (29.7-47.5)]
Benign or likely benign*	21, [14.3 (8.6-19.9)]	20, [17.5 (10.6-24.5)]
Total	**147**	**114**
Excluded Variants		
Pathogenic or likely pathogenic	64, [76.2 (67.1-85.3)]	35, [50.0 (38.3-61.7)]
Uncertain significance	13, [15.5 (7.7-23.2)]	26, [37.5 (26.2-48.8)]
Benign or likely benign*	7, [8.3 (2.4-14.2)]	9, [12.9 (5.0-20.8)]
Total	**84**	**70**

*This category does not include the wild-type allele

5.23 Allele Frequencies

We added together the MAFs of individual variants in the pathogenic, likely pathogenic category, and uncertain significance categories. The MAF of benign, likely benign, and wild-type variants was determined by subtracting the MAF of the other variants from 1. **Table 4.5** reports the MAFs of variants according by clinical significance category. The confidence intervals were calculated based on a sample of 15,496 whole-genome sequences included in gnomAD.

Table 5.5 Population frequencies of each type of allelic variant for *PAH* and *ACADM*.

Type of Allele	PAH		ACADM	
	MAF	95% CI	MAF	95% CI
Pathogenic or likely pathogenic	9.0×10^{-3}	7.5×10^{-3} - 1.0×10^{-2}	7.7×10^{-3}	6.4×10^{-3} - 9.0×10^{-3}
Uncertain significance	7.0×10^{-3}	5.7×10^{-3} - 8.3×10^{-3}	3.8×10^{-3}	2.8×10^{-3} - 4.8×10^{-3}
Benign, likely benign, or wild-type	9.8×10^{-1}	Very close to estimate	9.8×10^{-1}	Very close to estimate

*Estimates rounded to 1 decimal place in table

5.24 Hardy-Weinberg Genotype Proportions

We estimated genotype proportions, as described in **Table 5.6,** using the modified Hardy-Weinberg equation described previously **(Equation 3)**. The vast majority of infants would be homozygous for a benign, likely benign or a wild-type allele (97% for *PAH*; 98% for *ACADM*). Approximately 1 out of 12,345 (proportion = 8.1×10^{-5}; 95% CI: 5.6-11.0 x 10^{-5}) and 1 out of 16,949 (proportion = 5.9×10^{-5}; 95% CI: 4.1-8.1 x 10^{-5}) infants in the screened population would have pathogenic or likely pathogenic mutations on both alleles of *PAH* and *ACADM*, respectively.

Table 5.6 Estimated genotype proportion for *PAH* and *ACADM*.

Genotype (Allele 1 || Allele 2)	Variable	PAH		ACADM	
		Proportion	95% CI	Proportion	95% CI
P or LP || P or LP	q^2	8.1×10^{-5}	5.6×10^{-5} - 1.1×10^{-4}	5.9×10^{-5}	4.1×10^{-5} - 8.1×10^{-5}
P or LP || VUS	2qr	1.3×10^{-4}	8.5×10^{-5} - 1.7×10^{-4}	5.9×10^{-5}	3.6×10^{-5} - 8.6×10^{-5}
VUS || VUS	r^2	4.9×10^{-5}	3.2×10^{-5} - 6.9×10^{-5}	1.4×10^{-5}	7.8×10^{-6} - 2.3×10^{-5}
P or LP || B, LB, or WT	2pq	1.8×10^{-2}	1.5×10^{-2} - 2.1×10^{-2}	1.5×10^{-2}	1.3×10^{-2} - 1.8×10^{-5}
VUS || B, LB, or WT	2pr	1.4×10^{-2}	1.1×10^{-2} – 1.6×10^{-2}	7.4×10^{-3}	5.5×10^{-3} - 9.4×10^{-3}
B, LB, or WT || B, LB, or WT	p^2	9.7×10^{-1}	Very close to estimate	9.8×10^{-1}	Very close to estimate

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

*Estimates rounded to 1 decimal place in table

We applied Hardy-Weinberg genotype proportions to the NSO population to estimate the number of newborns expected with each genotype for the *PAH* and *ACADM* genes, as described in **Table 5.7**. We estimated that 112 (95% CI: 77-152) and 82 (95% CI: 57-112) babies in the

population as homozygous for a pathogenic or likely pathogenic allele for *PAH* and *ACADM*, respectively.

Table 5.7 Estimated genotype frequencies of infants in the NSO population.

Genotype (Allele 1 || Allele 2)	Variable	PAH	ACADM
		N (95% CI)	N (95% CI)
P or LP || P or LP	q^2	112 (77-152)	82 (57-112)
P or LP || VUS	2qr	173 (117-235)	82 (50-119)
VUS || VUS	r^2	67 (44-95)	20 (11-32)
P or LP || B, LB, or WT	2pq	24,435 (20,702-28,982)	21,079 (16,561-24,842)
VUS || B, LB, or WT	2pr	18,964 (15,181-22,082)	10,437 (7,590-12,973)
B, LB, or WT || B, LB, or WT	p^2	1,336,359	1,348,410

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

5.3 Primary Analyses: Estimating Results from a Parallel MS/MS and T-NGS Approach

5.31 Combined PKU Results

The outcomes obtained from the retrospective analysis of PKU screening at NSO as well as the analysis of *PAH* variants on ClinVar and gnomAD were brought together in **Table 5.8A.** Through a series of steps, we populated the table in order to estimate the *PAH* genotype categories of infants screened by first-tier MS/MS at NSO. An updated table is presented at the end of each step to reflect the changes that were made.

Table 5.8A MS/MS screening results for PKU in the NSO population, categorized by estimated genotype.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP				0	112
P or LP || VUS				0	173
VUS || VUS				0	67
P or LP || B, LB, or WT				0	24,435
VUS || B, LB, or WT				0	18,964
B, LB, or WT || B, LB, or WT				0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Step 1: Association between Genotype and MS/MS Screening Result

We expected a very strong association between an infant's *PAH* genotype and MS/MS screening result, particularly that infants with a *P or LP || P or LP* genotype were more likely to screen positive on MS/MS. We expected that infants with this genotype had a deficiency in *PAH* function that led to elevated Phe levels which exceeded the threshold for a positive screening result (Al-Hafid, 2015). Infants with this genotype who screened negative were considered false negatives. Since there were no missed cases (false negatives) of PKU reported to NSO, all infants with a *P or LP || P or LP* genotype were assumed to have a positive screening result. This change is highlighted in grey in **Table 5.8B**.

Table 5.8B MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated true negatives with *P or LP || P or LP* genotype.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP			**0**	0	112
P or LP || VUS				0	173
VUS || VUS				0	67
P or LP || B, LB, or WT				0	24,435
VUS || B, LB, or WT				0	18,964
B, LB, or WT || B, LB, or WT				0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Step 2: Association between Genotype and Diagnostic Follow-up Result

In neonates who screened positive for PKU, we anticipated a very strong association between their genotype and diagnostic testing result. We expected that those who were diagnosed with PKU were more likely to have a *P or LP || P or LP* genotype. To estimate the strength of this

association, we searched the literature for studies that examined the genotype-phenotype association between *PAH* and PKU, the results of which are presented in **Table 5.9**.

Table 5.9 Genotype proportions of infants diagnosed with PKU.[62-66]

Genotype (Allele 1 || Allele 2)	Study					Total	% (95% CI)
	Karam et al, 2013	Murad et al, 2013	Bueno et al, 2013	Vela-Amieva et al, 2015	Couce et al, 2013		
P or LP || P or LP	10	20	46	39	42	157	77.7 (72.0-83.4)
P or LP || VUS	1	2	5	5	5	18	8.9 (5.0-12.8)
VUS || VUS	1	4	0	2	0	7	3.5 (1.0-6.0)
P or LP || B, LB, or WT	4	4	0	0	3	11	5.4 (2.3-8.5)
VUS || B, LB, or WT	1	0	0	0	0	1	0.5 (0-1.4)
B, LB, or WT || B, LB, or WT	3	5	0	0	0	8	4.0 (1.3-6.7)
Total	**20**	**35**	**51**	**46**	**50**	**202**	**100**

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

We identified five studies that assessed the genotype of individuals with a confirmed diagnosis of PKU. For each, the case definition for PKU was identical to that used by NSO. We extracted data regarding the proportion of PKU patients with each of the specified genotypes. Data from each individual study was combined by adding their genotype frequencies together to produce a more precise estimate of the genotype proportions of PKU patients **(Table 5.9)**.

The proportions that were determined were applied to infants diagnosed with PKU (true positives) in the NSO population; this enabled us to estimate the number of infants with each of the specified genotypes. The corresponding cells have been populated in **Table 5.8C** and highlighted in grey.

Table 5.8C MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated infants who were true positives.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	**67**		0	0	112
P or LP || VUS	**8**			0	173
VUS || VUS	**3**			0	67
P or LP || B, LB, or WT	**5**			0	24,435
VUS || B, LB, or WT	**0**			0	18,964
B, LB, or WT || B, LB, or WT	**3**			0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Step 3: Benign HPA

Benign HPA is a variant of PKU that is characterized by a mild, but detectable elevation of Phe levels. Infants are diagnosed with this condition after receiving a positive NBS result and undergoing follow-up tests. Benign HPA is classified by NSO as a false positive and does not require any clinical intervention. Like classic PKU, benign HPA is caused by two pathogenic mutations on the *PAH* gene. Therefore, it is likely that there is an association between a benign HPA diagnosis and genotype.

In **Table 5.8B**, we noted 45 infants with the *P or LP || P or LP* genotype who were yet to be allocated, but must have been false positives. Since we assumed that all infants predicted to have *P or LP || P or LP* genotype had some functional deficiency in *PAH*, we concluded that all 45 false positive infants with this genotype were diagnosed with benign HPA.

We determined that there were a total of 79 infants with benign HPA. Following the allocation of 45 false positive infants (above), 34 infants remained who needed to be classified

according to genotype. We conducted a literature search for articles reporting an association between genotype and the benign HPA phenotype, the results of which are reported in **Table 5.10**.

Table 5.10 Genotype proportions of infants diagnosed with benign HPA.[62,65,67]

Genotype (Allele 1 || Allele 2)	Study			Total	% (95% CI)
	Mallolas et al, 1999	Karam et al, 2013	Vela-Amieva et al, 2015		
P or LP || P or LP	19	0	2	21	61.8 (45.5-78.1)
P or LP || VUS	9	2	0	11	32.4 (16.7-48.1)
VUS || VUS	0	0	1	1	2.9 (0-8.5)
P or LP || B, LB, or WT	0	1	0	1	2.9 (0-8.5)
VUS || B, LB, or WT	0	0	0	0	0
B, LB, or WT || B, LB, or WT	0	0	0	0	0
Total	28	3	3	34	100

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

We identified three studies that investigated this association with a total sample size of 34 patients with benign HPA. For each of the studies, the case definition of benign HPA was identical to that specified at NSO. The proportion of infants with benign HPA with the *P or LP || P or LP* genotype (61.8%; 95% CI: 45.5-78.1%) in these studies was not substantially different from our estimate of this proportion in the NSO population (57.0%). We therefore calculated the genotype proportions for benign HPA patients and applied these to the NSO population. However, we did not consider the *P or LP || P or LP* genotype in these calculations as we did not seek to estimate the proportion of benign HPA patients with this genotype category. The infants with benign HPA have been indicated in **Table 5.8D** in grey.

Table 5.8D MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated infants diagnosed with benign HPA.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	**45**	0	0	112
P or LP || VUS	8	**28**		0	173
VUS || VUS	3	**3**		0	67
P or LP || B, LB, or WT	5	**3**		0	24,435
VUS || B, LB, or WT	0			0	18,964
B, LB, or WT || B, LB, or WT	3			0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Step 4: Remaining False Positives and True Negatives

There were 370 false positives remaining to be allocated. For these results, we assumed no association with the five possible genotypes where they could have been allocated (all except *P or LP || P or LP*). The expected distribution across categories is highlighted in grey in **Table 5.8E**. We noted that the results indicated none were expected to have *P or LP || VUS* or *VUS || VUS* genotypes.

Table 5.8E MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated remaining false positives.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	45	0	0	112
P or LP || VUS	8	28		0	173
VUS || VUS	3	3		0	67
P or LP || B, LB, or WT	5	**10**		0	24,435
VUS || B, LB, or WT	0	**5**		0	18,964
B, LB, or WT || B, LB, or WT	3	**358**		0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Lastly, to calculate the number of infants in each genotype category who were true negatives, we subtracted the total number of estimated true positives, false positives, and false negatives from the total number of infants estimated in the corresponding genotype. Cells populated by the true negatives have been highlighted in grey in **Table 5.8F**, which also presents the full results of the calculations that have been described.

Table 5.8F MS/MS screening results for PKU in the NSO population, categorized by estimated genotype. Allocated infants who were true negatives.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	45	0	0	112
P or LP || VUS	8	28	**137**	0	173
VUS || VUS	3	3	**61**	0	67
P or LP || B, LB, or WT	5	10	**24,420**	0	24,435
VUS || B, LB, or WT	0	5	**18,959**	0	18,964
B, LB, or WT || B, LB, or WT	3	358	**1,335,998**	0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

5.32 Combined MCAD Deficiency Results

The outcomes of from the retrospective analysis of MCAD deficiency screening at NSO as well as the analysis of *ACADM* variants on ClinVar and gnomAD are presented together in **Table 5.11A.** As with PKU, we populated the table in a series of steps to estimate the *ACADM* genotype of infants who were screened by first-tier MS/MS at NSO. An updated table is presented at the end of each step to reflect the changes that were made.

Table 5.11A MS/MS screening results for MCAD deficiency in the NSO population, categorized by genotype.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP					82
P or LP || VUS					82
VUS || VUS					20
P or LP || B, LB, or WT					21,079
VUS || B, LB, or WT					10,437
B, LB, or WT || B, LB, or WT					1,348,410
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Step 1: Association between Genotype and MS/MS Screening Results

We expected an association between *ACADM* genotype and MS/MS screening result, specifically that infants with a *P or LP || P or LP* genotype would be more likely to receive a positive result. According to the case definition of MCAD deficiency, infants homozygous for the pathogenic mutation will be classified as affected. Hence, infants with a *P or LP || P or LP* genotype could not be “true negatives”. There was one false negative reported in the screened population at NSO we could not ascertain the genotype of this individual with certaint, but judged that this infant was most likely to have the *P or LP || P or LP* genotype. The cells affected by this assumption have been indicated in grey in **Table 5.11B**.

Table 5.11B MS/MS screening results for MCAD deficiency in the NSO population, categorized by genotype. Allocated false negatives and true negatives who were *P or LP || P or LP*.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP			**0**	**1**	82
P or LP || VUS				**0**	82
VUS || VUS				**0**	20
P or LP || B, LB, or WT				**0**	21,079
VUS || B, LB, or WT				**0**	10,437
B, LB, or WT || B, LB, or WT				**0**	1,348,410
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Step 2: Association between Genotype and Diagnostic Testing Results

We also expected an association between *ACADM* genotype and diagnostic testing results. As previously mentioned, infants with two pathogenic mutations on *ACADM* are diagnosed with MCAD deficiency. Therefore, all 81 babies who received a positive MS/MS screening result and were predicted to have a *P or LP || P or LP* genotype were assumed to be true positives.

To allocate the 13 remaining true positives, we searched PubMed for studies that reported genotype proportions in infants who received an MCAD deficiency diagnosis. Since diagnostic testing for MCAD deficiency may involve mutational analysis, we restricted our search to include only studies that diagnosed MCAD deficiency using biochemical tests.

Table 5.12 Genotypes of patients diagnosed with MCAD deficiency through biochemical testing.[68]

Genotype (Allele 1 || Allele 2)	Frequency	% (95% CI)
P or LP || P or LP	47	82.5 (73.8-90.2)
P or LP || VUS	9	15.8 (6.3-25.3)
VUS || VUS	1	1.8 (0-5.3)
Total	57	100

P or LP – pathogenic or likely pathogenic; VUS – variant of uncertain significance

We identified only one study that assessed the genotype-phenotype relationship between *ACADM* and MCAD deficiency and also used biochemical testing for MCAD deficiency diagnosis.[68] The criteria used to make a biochemically confirmed diagnosis in this study was similar to the criteria at NSO. Following a positive NBS result by MS/MS, confirmatory investigations included an assessment of concentrations of acylcarnitines of repeat dried blood spot samples, plasma acylcarnitines, and urinary organic acids.[68] Of 57 patients in this study, 47 had the *P or LP || P or LP genotype*. The proportion of MCAD deficiency patients in this study with this genotype (82.5%; 95% CI: 73.8-90.2%) was not substantially different from the proportion estimated for the NSO population (86.2%). We calculated the other genotype proportions for MCAD deficiency patients and applied these to the NSO population to estimate the frequency of infants with each genotype. We excluded the *P or LP || P or LP* genotype from these calculations as we did not seek to estimate the proportion of true positive patients with these genotype. Based on these calculations, we populated the corresponding cells in **Table 5.11C** and indicated the changes in grey.

Table 5.11C MS/MS screening results for MCAD deficiency in the NSO population, categorized by genotype. Allocated infants who were true positives.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	**81**	**0**	0	1	82
P or LP || VUS	**12**			0	82
VUS || VUS	**1**			0	20
P or LP || B, LB, or WT	**0**			0	21,079
VUS || B, LB, or WT	**0**			0	10,437
B, LB, or WT || B, LB, or WT	**0**			0	1,348,410
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Step 3: MCAD Deficiency Carriers

The protocol for diagnostic testing for MCAD deficiency may require mutational analysis, which enables the identification of some infants who are carriers of a pathogenic allele (37 infants in the NSO population). These infants may have the *P or LP/VUS*, *VUS/VUS*, or *P or LP/B, LB, or WT* genotypes. We assumed no association between being an MCAD deficiency carrier and having one of these three genotypes and calculated expected distribution proportionately, resulting in all 37 allocated to the *P or LP || B, LB, or WT* genotype (**Table 5.11D**, highlighted in grey).

Table 5.11D MS/MS screening results for MCAD deficiency in the NSO population, categorized by genotype. Allocated infants who were MCAD deficiency carriers.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	81	0	0	1	82
P or LP || VUS	12	**0**		0	82
VUS || VUS	1	**0**		0	20
P or LP || B, LB, or WT	0	**37**		0	21,079
VUS || B, LB, or WT	0			0	10,437
B, LB, or WT || B, LB, or WT	0			0	1,348,410
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Step 4: Remaining False Positives and True Negatives

There were 145 false positives remaining to be classified. We assumed no association between false positive and genotype, so calculated the distribution proportionately. (**Table 5.11E**, highlighted in grey).

Table 5.11E MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were false positives.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	81	0	0	1	82
P or LP || VUS	12	0		0	82
VUS || VUS	1	0		0	20
P or LP || B, LB, or WT	0	**39**		0	21,079
VUS || B, LB, or WT	0	**1**		0	10,437
B, LB, or WT || B, LB, or WT	0	**142**		0	1,348,410
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

To calculate the number of infants in each genotype category who were true negatives, we subtracted the total number of true positives, false positives, and false negatives estimated to be in each genotype category, from the total number of infants estimated to have that particular genotype. We allocated true negatives in **Table 5.11F** and highlighted the corresponding cells in grey. This completed the table.

Table 5.11F MS/MS screening results for MCAD deficiency in the NSO population, categorized by estimated genotype. Allocated infants who were true negatives.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	81	0	0	1	82
P or LP || VUS	12	0	**70**	0	82
VUS || VUS	1	0	**19**	0	20
P or LP || B, LB, or WT	0	39	**21,040**	0	21,079
VUS || B, LB, or WT	0	1	**10,436**	0	10,437
B, LB, or WT || B, LB, or WT	0	142	**1,348,268**	0	1,348,410
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

4.4 Sensitivity Analyses

Selected results from sensitivity analyses are presented below. **Table 5.13A** and **Table 5.13B** describe the predicted PKU screening results of the parallel T-NGS and MS/MS approach, after modifying the proportion of infants with the *P or LP || P or LP* genotype. We estimated a minimum of 77 and a maximum of 152 infants with this genotype in the NSO population. In a population where this number is 77 **(Table 5.13A)**, the number of false positives with benign HPA who have the *P or LP || P or LP* genotype decreases from 45 to 10, compared to the primary analyses. Correspondingly, the number of infants who are false positives with benign HPA and have the *P or LP || VUS* genotype increases from 28 to 59, compared to the primary analyses.

Table 5.13A Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || P or LP* to the lower limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	10	0	0	77
P or LP || VUS	8	59	106	0	173
VUS || VUS	3	5	59	0	67
P or LP || B, LB, or WT	5	12	24,418	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	358	1,336,998	0	1,336,394
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

In a population where the number of infants with the *P or LP || P or LP* genotype is 152 **(Table 5.13B)**, we predict an increase from 45 to 85 false positive infants with this particular genotype. Out of these 85 false positives, 79 would have benign HPA, and 6 would not have any form of PKU. In addition, we predict no instances of benign HPA among babies with the *P or LP/VUS* or *VUS/VUS* genotypes in this scenario.

Table 5.13B. Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || P or LP* to the upper limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	85	0	0	152
P or LP || VUS	8	0	165	0	173
VUS || VUS	3	0	64	0	67
P or LP || B, LB, or WT	5	7	24,423	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	352	1,335,998	0	1,336,319
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.14A and **Table 5.14B** describe the predicted MCAD deficiency screening results of the parallel T-NGS and MS/MS approach, after modifying the proportion of infants with the *P or LP || P or LP* genotype. We predicted that there was a minimum of 57 and a maximum of 112 infants with this genotype in the NSO population. In a population where the number of infants with this genotype was 57 **(Table 5.14A)**, the number of true-positives who have the *P or LP || P or LP* genotype decreased from 81 to 56, compared to the primary analyses. Correspondingly, the number of infants who were true positives and had the *P or LP || VUS* genotype increased from 12 to 34, compared to the primary analyses.

Table 5.14A. Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || P or LP* to the lower limit of the 95% confidence interval of the estimate.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	56	0	0	1	57
P or LP || VUS	34	0	48	0	82
VUS || VUS	4	0	16	0	20
P or LP || B, LB, or WT	0	39	21,040	0	21,079
VUS || B, LB, or WT	0	1	10,436	0	10,437
B, LB, or WT || B, LB, or WT	0	142	1,348,293	0	1,348,435
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

In a population where the number of infants with the *P or LP/P or LP* genotype was 112 **(Table 5.14B)**, we predict that all 94 true-positives had this particular genotype. In addition, there were 18 predicted false negatives, compared to 1 in the primary analyses.

Table 5.14B. Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || P or LP* to the upper limit of the 95% confidence interval of the estimate.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	94	0	0	18	112
P or LP || VUS	0	0	82	0	82
VUS || VUS	0	0	20	0	20
P or LP || B, LB, or WT	0	39	21,040	0	21,079
VUS || B, LB, or WT	0	1	10,436	0	10,437
B, LB, or WT || B, LB, or WT	0	142	1,348,238	0	1,348,380
MS/MS Totals	**94**	**182**	**1,379,816**	**18**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

CHAPTER 6: DISCUSSION

Some observers have suggested that NGS technologies have the potential to improve the performance of NBS. Before NGS can be extensively applied in NBS, there is a need to assess its feasibility and associated risks and benefits. We sought to quantitatively estimate how the interpretation of NBS results for PKU and MCAD deficiency might change if T-NGS was applied to the current NBS model that uses first-tier MS/MS.

6.1 Performance of MS/MS Screening

We evaluated the performance of first-tier MS/MS screening for PKU and MCAD deficiency from April 2006 to November 2015. The sensitivity of MS/MS for PKU was excellent (100%), with no false negatives reported in the screened population during this period. The specificity was also considerably high (99.97%), but due to the extremely low prevalence of PKU in the population, the positive predictive value was relatively low (16.1%). Furthermore, without timely detection and treatment, PKU is a condition that leads to severely adverse health outcomes. As a result, the threshold for a positive screening result is kept relatively low to ensure that infants with the condition are not missed. Thus, the specificity (and positive predictive value) of the screening test is reduced in order to maximize sensitivity.

Because one false negative was reported, the sensitivity of MS/MS screening for MCAD deficiency was below ideal (98.95%) for the study period, considering that one of the critical objectives of NBS is to avoid missed cases.[25] On the other hand, the specificity (99.99%) and positive predictive value (34.1%) of MS/MS screening for MCAD deficiency are substantially higher than MS/MS screening for PKU.

6.2 A Combined T-NGS and MS/MS Approach

We used MS/MS screening data from NSO and data pertaining to human genetic variation to predict the results of a screening model that uses both T-NGS and MS/MS in parallel. In MS/MS, blood spot samples whose concentrations of a measured biomarker surpass a threshold constitute a positive screening result and are considered at high-risk for a screened condition. In T-NGS, three genotypes (*P or LP || P or LP*; *P or LP || VUS*; *VUS || VUS*) may be considered higher risk than the others. Individuals with one of these genotypes have or may have two pathogenic alleles, which would constitute a genotype that forms the genetic basis for PKU or MCAD deficiency.

<u>6.21 Primary Analyses: PKU</u>

In our primary analyses for PKU, we reported that approximately 40% of infants in the study population who were estimated to have the genotype of highest risk (*P or LP || P or LP*) would ultimately be diagnosed with benign HPA. Infants with this diagnosis require no further follow-up or treatment and are classified as false positives by NSO. This emphasizes the idea that, even for monogenic disorders, some variants that are classified as pathogenic or likely pathogenic may not necessarily cause classic symptoms of the disorder if an individual has two of these variants.

We also estimated that eight out of 86 infants with a confirmed diagnosis of PKU had one of three lower-risk genotypes (*P or LP || B, LB, or WT; VUS || B, LB, or WT; B, LB, or WT || B, LB, or WT*). This is unexpected considering that the clinical definition of PKU states that a case must have two pathogenic mutations on the *PAH* gene, which does not appear to be possible with one of those three genotypes. These estimates were formed from the results of a literature review for studies that investigated the genotype-phenotype association between the *PAH* gene and PKU. Two of the five

studies[62,63] that were retrieved had classified a considerable proportion of PKU patients with one of the three lower-risk genotypes. In these studies, the authors reported that existing pathogenic mutations in PKU patients may have been missed during mutational analysis. These mutations may have been present in regions of the *PAH* gene that were not screened since they were non-coding regions (introns).[62,63] Furthermore, these studies used Sanger sequencing to conduct mutational analysis as opposed to NGS. Sanger sequencing is slower and more costly than NGS,[41,69] which may explain why non-coding regions were not sequenced in the two studies above. The remaining three studies identified in the review used NGS to conduct mutational analysis. One of these three studies, conducted in France,[66] categorized a small proportion of PKU patients as *P or LP || B, LB, or WT*. One explanation is that the existing mutations were missed by NGS due to problems with the analytic validity, which refers to the ability of a genetic test to accurately and reliably measure the genotype of interest. It may also be possible that these patients had new, unknown mutations on the *PAH* gene with unclassified clinical significance. It appears that NGS for the *PAH* gene has some limitations with respect to analytic validity, which can reduce its sensitivity if employed as a first-tier test in NBS for PKU.

We estimated that over 80% of infants who were false positives had a low-risk genotype (i.e. *B, LB, or WT || B, LB, or WT*). If genotypic information was available in addition to the results of MS/MS, this may have informed decision-making with regards to the interpretation of the screening (i.e. whether to consider the screening result to be "positive" and thus pursue more invasive diagnostic follow-up testing). In particular, genotypic information may guide clinical decision-making of MS/MS results that are difficult to interpret (i.e. MS/MS is positive but close to the threshold value). Consequently, the number of false positives generated by NBS may be reduced with the addition of genotypic information. Nevertheless, there is still a need for the NBS program

to consider whether information generated by NGS is robust enough to change the screening decision that would have be made based on MS/MS results. This may require consideration of the analytic validity of NGS, which refers to the technical performance of the test and the clinical validity of NGS, which refers to the test's ability to accurately and reliably predict the disorder of interest[12].

6.22 Primary Analyses: MCAD Deficiency

In our primary analyses for MCAD deficiency, we judged that there were no false positives or true negatives among infants who had the *P or LP || P or LP* genotype. Infants who have this particular genotype meet the case definition of MCAD deficiency and, from a clinical perspective, are diagnosed with the disease.

One false negative was reported back to NSO in this particular population. This infant may have had one of the three higher risk genotypes, but we assumed that it was *P or LP || P or LP* since this genotype is most likely to cause MCAD deficiency. There is potential for sensitivity of NBS for MCAD deficiency to be improved with the addition of genotypic information. Specifically, by using T-NGS to supplement MS/MS results, a more well-informed assessment of an infant's risk might be made. A high index of suspicion would be maintained for infants with one of the higher risk genotypes. For example, infants with MS/MS results that are negative, but close to the threshold, may have a change in the interpretation of their screening results if they were identified to have a genotype which places them at high risk for the screened condition.

We also estimated that over 80% of infants who were false positives had a low-risk genotype. As with NBS for PKU, using T-NGS may help reduce these false positives by improving the specificity of NBS for MCAD deficiency by providing genotypic information.

6.3 Sensitivity Analyses

6.31 Sensitivity Analyses: PKU

When we varied the estimated genotype proportion in the NSO population for the *P or LP || P or LP* genotype, our results were different in some notable ways. For PKU, varying this genotype proportion to the lower limit of the 95% confidence interval led to the finding that there would be fewer infants with this particular genotype who also have benign HPA, compared to the findings in the primary analyses. Consequently, a greater proportion of infants with the *P or LP || P or LP* genotype on *PAH* were estimated to receive a confirmed diagnosis for PKU. Therefore, in a population where the genotype proportion for *P or LP || P or LP* is lower than estimated in the primary analyses, infants who have this particular genotype are more likely to ultimately be diagnosed with PKU.

Varying the genotype proportion of *P or LP || P or LP* to the upper limit of the 95% confidence interval led us to estimate that there were 85 infants with this genotype who would have a positive MS/MS screening result but will not be diagnosed with PKU at follow-up. We assumed that infants with this genotype who are not diagnosed with PKU will have benign HPA. However, results from the NSO data indicated that there were 79 infants with benign HPA in the population. Therefore, there were six infants in this scenario who have a *P or LP || P or LP* genotype but do not have PKU or benign HPA. Although some mutations of *PAH* are associated with milder phenotypes, two copies of a pathogenic mutation on the *PAH* gene are still expected to, at the very least, cause in elevation in blood Phe levels that would result in benign HPA.[70] The scenario presented here may not reflect any real-world population because we expect that there would be a greater number of infants with a confirmed diagnosis of PKU in a population with such a high proportion of infants who have the *P or LP || P or LP* genotype.

6.32 Sensitivity Analysis: MCAD Deficiency

We also varied the genotype proportion of *P or LP || P or LP* for *ACADM* that was predicted in the primary analyses. With the genotype proportion at the lower limit of the 95% confidence interval, we noted that, compared to the primary analyses, a greater number of infants who were true positives fell into the *P or LP || VUS* genotype.

When we varied the genotype proportion to the upper limit of the 95% confidence interval, we predicted that all true positives would fall into this particular genotype category. Additionally, we estimated 18 false negatives, of which, if present, 17 were not reported to NSO. MCAD deficiency is a condition where infants may have two pathogenic mutations on the *ACADM* gene but may not present with any phenotypic manifestations of the illness.[57] Thus, infants with the *P or LP || P or LP* genotype who are false negatives may have mutations that are less severe and do not immediately present symptoms of MCAD deficiency. Nonetheless, it is necessary to identify these infants and be able to make decisions about preventive interventions to avoid clinical manifestation of the disease in the future.

6.4 Models for Implementing T-NGS in NBS

In our evaluation of first-tier MS/MS-based NBS for PKU, we reported that the present screening test has extremely high sensitivity and negative predictive value. Specificity is also extremely high, but the positive predictive value remains relatively low because PKU is such a rare disease. In this regard, it appears that there is potential for NBS for PKU to be improved by new technologies that increase specificity and positive predictive value. It may be effective to implement T-NGS into NBS as a second-tier test for PKU, where the samples of infants who have been identified as high-risk by MS/MS can be sequenced for genetic information regarding their risk for

developing PKU. For these infants, the decision for diagnostic referral can be made by considering both genotypic and phenotypic information.

In our evaluation of MS/MS-based NBS for MCAD deficiency, we reported a higher specificity than MS/MS screening for PKU. However, due to the false negative that had been reported during the study period, the sensitivity was comparatively lower than for PKU and can be improved. Furthermore, sensitivity analyses highlighted a scenario where infants with two pathogenic mutations on *ACADM* may not be detected by MS/MS and this may not come to the attention of the screening program (i.e., there could be additional false negatives that are unreported) since MCAD deficiency symptoms may not immediately (or ever) arise. This may not pose an immediate risk to infants' health, but these infants may need to be identified and monitored by the NBS program to ensure that the appropriate health services and interventions are readily available if symptoms do manifest in the future.

6.41 Second-tier NGS

The model in **Figure 6.1** describes a scenario, using PKU as the case condition, where first-tier MS/MS screening would be performed to rule out all individuals whose Phe level is below the present threshold for a positive screening result. Samples of infants whose Phe levels exceed this threshold would then undergo T-NGS to identify their genotype at the *PAH* gene. The MS/MS screening result would then be interpreted differently, depending on the genotype of the infant. For example, the MS/MS threshold for referral would not change for infants with a *P or LP || P or LP* genotype – all infants with this genotype would be referred for diagnostic follow-up. On the other hand, infants with a *B, LB, or WT || B, LB, or WT* genotype would require "Very High" Phe levels in order to be referred for diagnostic follow-up. This model has to potential to maintain the sensitivity of NBS while improving its specificity and reducing the number of false positives. In

addition, since T-NGS would only be applied as a second-tier test, it is likely to be more feasible than universal T-NGS.

Figure 6.1 Proposed NBS model for PKU using first-tier MS/MS and second-tier T-NGS. The thicker line between "low" and "very low" represents the MS/MS threshold for a positive result.

T-NGS Result	**MS/MS Result (Phe level)**					
(Allele 1 || Allele 2)	**Very High**	**High**	**Intermediate**	**Intermediate**	**Low**	**Very Low**
P or LP || P or LP						
P or LP || VUS						
VUS || VUS						
P or LP || B, LB, or WT						
VUS || B, LB, or WT						
B, LB, or WT || B, LB, or WT						

P or LP – pathogenic or likely pathogenic; VUS – variant of unknown significance; B, LB, or WT – benign, likely benign or wild-type

- Referred for follow-up

- Not referred

<u>6.42 Universal T-NGS</u>

The model in **Figure 6.1** has been suggested based on the observation that no false negatives have occurred among infants screened for PKU at NSO. However, we acknowledge that it is possible for false negatives to occur in NBS programs, even for PKU. In some cases, biological variation between infants may cause some cases of PKU to be missed by NBS programs.

For example, an infant may not demonstrate elevated Phe levels for a limited period of time despite having the genetic basis of the disease. This is more common among infants whose blood sample may have been obtained before 72 hours of age.[71] Recognizing the possibility of false negatives for PKU in NBS programs, we conservatively estimate that one false negative may occur during a 10-year period. In this scenario, sensitivity may be improved using T-NGS in NBS; this would require all infants to be screened using both T-NGS and MS/MS, even infants in the "very low" MS/MS category would be screened using T-NGS.

To improve sensitivity of NBS, T-NGS may be applied in parallel with MS/MS to all infants, including those with MS/MS analyte readings below the threshold. This would enable screening laboratories to make decisions by considering genotypic and phenotypic information. Before T-NGS can be applied in a universal context, NBS programs should consider whether this would be feasible, given the potential costs to the health care system and the benefits gained by using this technology. The technical performance of the technology should also be considered (i.e., is it able to accurately detect genotypes of infants who are screened).

6.5 Issues to Consider for Universal T-NGS

Universal T-NGS has the potential to improve the quality of NBS, but the benefits that can be gained from its use must be weighed against the costs associated with its use. Furthermore, there are other factors that should also be considered.

6.51 Storage

An important issue that surfaces from the implementation of newborn genomic sequencing is the storage of large amounts of data[16] as several questions are raised regarding their privacy and accessibility. The long-term storage of NBS samples and sequencing data has become a controversial topic.[16,39,42,43] The additional sequencing data that would be generated from NGS would provide opportunities for future biomedical research that could improve knowledge of disease pathophysiology and positively impact health.[16,39,43] Studies investigating public attitudes towards the storage of genetic information have found that participants expressed concerns about stigmatization, breaches in privacy, and misuse of information.[72] Stored sequencing data will become part of the medical record, which may potentially affect health and disability insurance as well as employability.[43,72] Therefore, genomic sequencing data needs to be treated like all personal

health information; appropriate measures need to be taken to protect privacy and confidentiality.[16,39] The costs and infrastructure required to securely store and maintain these data over the long term should be thoroughly considered.[16,39]

6.52 Consent

NBS is presently conducted without explicit consent from parents since it is considered as a health service that clearly provides medical benefit for the child. To communicate important information related to NBS, in many NBS programs, parents are provided with some education (verbally or through pamphlets, etc) prior to or at the time of screening and may choose to opt-out. The application of T-NGS will lead NBS programs to yield unsolicited findings that may be communicated to parents. Furthermore, the genetic information produced from T-NGS will be stored over the long term and parents need to be informed of these changes. Parents have stated that they require more information about screening and the management of dried blood spot samples and NBS data.[73] Hence, there may be a need to revisit the consent process for NBS. Introducing an informed consent process into NBS is a logistical challenge since additional health care professionals, genetic counselors in particular, would be required to adequately obtain consent.[16,39] Moreover, concerns have been raised that an explicit consent process could reduce the uptake of screening by families, though this notion is not supported by concrete evidence.[16,74] If the process for consent needs to be re-evaluated, then the added time and resources necessary to make any modifications should be thoroughly considered.

6.53 Evidence Generation and Feasibility

It is still not clear whether the implementation of NGS technologies in the NBS context would be an effective strategy, from a public health perspective.[16] More conclusive evidence

pertaining to the validity of information generated by NGS may be necessary before it can be used to influence screening decisions.[16] According to the results of this study, T-NGS still has limitations as a standalone screening technology for some conditions (i.e. PKU) for NBS due to sensitivity that may be lower than ideal. Thus, it appears that MS/MS might need to be retained for some conditions on the NBS panel; this undermines the case to switch to T-NGS as the sole modality to screen any condition. Furthermore, not all variants of genes associated with screened conditions are known,[16] hence the detection of new, unknown variants may pose a dilemma when interpreting results and may potentially reduce the sensitivity of NBS, especially if T-NGS is used as the sole or first-tier screening technology.

A formal cost-effectiveness analysis can help inform whether NGS is feasible, given the possible benefits of its use. In addition to the cost of the NGS screening test itself, subsidiary costs will also need to be considered. This includes the cost of storing, managing, or analyzing the data and of health professionals' time to provide genetic counseling or interpretation of results.[16]

6.54 Unsolicited Findings

In screening, unsolicited findings refer to results that are not directly related to the primary purpose of NBS, such as detection of carriers of a pathogenic allele. According to our results, PKU and MCAD deficiency carriers (i.e., infants who do not have the disease but have one mutation) represent considerable proportion of infants (3-4%) in the population. Although there are no immediate health concerns for the child, these results could potentially inform parents' reproductive decisions (e.g., parents may wish to undergo testing to determine whether they are both carriers, in which case, future children would be at risk of having the disease) and may be of interest to other family members. Evidence from quantitative and qualitative survey data suggests that parents would prefer to be informed of their child's carrier status, even though they understand that their child is

not at risk for the associated disease.[75-78] In addition, recommendations from the ACMG Working Group on Incidental Findings in Clinical Exome and Genome Sequencing state that variants identified through NGS which fall into the defined categories of “pathogenic” and “likely pathogenic” should be reported.[79] Commentators in favour of disclosing such findings argue that withholding information constitutes paternalism, where the public is judged to be unable to manage important health data – this violates their right to knowledge about themselves.[75] Some contend that parents should be required to possess carrier status information about their infant as it enables them to make well-informed reproductive decisions.[75] Those with opposing views support the notion that individuals have a right to ignorance based on a preference to not know medical information related to their own person, particularly genetic information.[75] Since the preferences of the screened individual (or their parents) are unknown, there must be “compelling reasons” to reveal such information to families.[75] At the moment, discussion about the management of carrier results is still ongoing and there does not appear to be international consensus regarding their disclosure to parents.[80] Annually, we predicted that T-NGS will identify over a combined 4,000 carriers for both PKU and MCAD deficiency. If this information is disclosed to parents, it may require an expansion in specific health infrastructure and personnel.

6.6 Strengths

There are notable strengths in this study that merit discussion. First, the sample size was extremely large and representative of the population of Ontario. Virtually all infants who were born in Ontario since the establishment of NSO in 2006 (with the exception of infants whose parents opted out from NBS) were included in this study. This enabled us to generate extremely precise estimate (essentially population parameters) during our analysis of the NSO data. Second, the ascertainment of disease outcomes in this study was facilitated by excellent data linkage. We were

able to use screening data that was stored at NSO and link it with data pertaining to diagnostic follow-up that was conducted at treatment centres in Ontario. Third, we conducted a comprehensive search of genetic compendia to consider all alleles (within the limits of current knowledge) that have been reported by various clinical laboratories. However, these compendia do have some limitations, which are discussed in further detail below.

6.7 Limitations

Several assumptions were made in estimating the results of a combined NBS approach using T-NGS and MS/MS in parallel.

First, while analyzing genetic compendia, we were unable to find complete data for a substantial proportion of variants of *PAH* and *ACADM* that were identified in ClinVar. In some cases, there were conflicts between submitters' interpretations of the clinical significance of variants. Although we attempted to resolve these conflicts, for five variants of *ACADM* we were unable to do so. Ideally, we would have preferred to include as many variants as possible to achieve accurate estimates of allele and genotype frequencies. Nevertheless, due to the rarity of each variant, excluding five variants out of 114 is unlikely to have a substantial effect. We were also unable to identify MAFs of many variants that were identified in ClinVar. We assumed that these variants were so rare that it was not yet possible to make an estimate of their prevalence and excluded them from further analyses. We compared the clinical significance of variants that were included for final analyses to those that were excluded and did not find a significant difference between the two. This suggests that perhaps excluding these variants may not have had a substantial effect on Hardy-Weinberg genotype proportions.

Second, the validity of the Hardy-Weinberg principle rests on the notion that the genetic makeup of a population remains stable over time. Some assumptions may not hold completely true in real-world populations. The first assumption is that there is no genetic drift. Genetic drift refers to a phenomenon where certain alleles are passed on more frequently from one generation to the next as a result of chance (i.e. in a given population, all carriers of a rare allele pass on this allele to their offspring by chance).[47] In small populations, this phenomenon can change the genetic makeup of the next generation.[47] However, in a population as large as Ontario, even if genetic drift occurs, it is unlikely to have an effect on the entire population, because random effects will average out.[47] The second assumption is that there is no natural selection in the population (i.e. there is no survival disadvantage to allele carriers or homozygotes). For both the *PAH* and *ACADM* genes, this assumption is relatively robust. Children with PKU and MCAD deficiency are identified through newborn screening and are given treatment or monitored from a very early age, which prevents adverse outcomes and allows them to live to the age of reproduction. Therefore, having recessive allele(s) either *PAH* or *ACADM* should not limit an individual's reproductive ability. The third assumption is that there are no new mutations in the population. While mutations continually occur in all populations, the rate at which new mutations arise is relatively slow and it is unlikely that they will have a significant impact on the genetic makeup of the Ontario population during the time period that is being considered for this study. The remaining two assumptions of the Hardy-Weinberg principle are that there is completely random mating in the population and that there is no migration in or out of the population. These assumptions are less likely to be robust in the Ontario population. Mating is not likely to be random as many ethnic groups within the overall population tend to marry within their ethnic communities. If the recessive allele for *PAH* or *ACADM* is more prevalent within certain ethnic groups, then this could lead to an increase in the prevalence of that

particular allele. Migration in and out of Ontario is also a continuous process, and may affect the genetic makeup of the population. This effect is further emphasized when a large group of individuals migrates in or out of the population (i.e. the recent resettlement of Syrian refugees in Canada). The violation of the assumptions of the Hardy-Weinberg principle may affect our estimates of genotype proportions. Hence, we conducted sensitivity analyses by varying these proportions according to their 95% confidence intervals.

Third, we applied genotype proportions to the NSO population using information that was abstracted from genetic compendia. The genetic makeup of the population that was used to determine the MAFs for each variant of *PAH* and *ACADM* is unclear and may be different from that of the NSO population. In order to address this limitation, we calculated 95% confidence intervals for the predicted genotype proportions based the size of the sample that was used to determine the MAFs in gnomAD. We then tested our assumption by performing sensitivity analyses where these genotype proportions were varied according to the limits of the 95% confidence interval.

Fourth, we based our genotype-phenotype correlation assumptions on five studies for PKU and one study for MCAD deficiency. We cannot be certain whether the study populations in the retrieved articles were representative of the NSO population with respect to their genetic makeup. We did not note any notable changes in our conclusions in any of the scenarios from our sensitivity analyses. A similar issue pertains to our assessment of the genotype-phenotype association between *PAH* and benign HPA. We determined that the proportion of infants in the retrieved studies with the *P or LP || P or LP* genotype was not significantly different from the proportion of infants in the NSO population predicted to have this genotype, so suggest that these populations have a similar genetic makeup to the NSO population.

Fifth, certain clinical characteristics (i.e. preterm birth, low birth weight, underlying illness) may be associated with increased false positive results.[28] We were unable to account for this possible association in our analysis, which may have led us to underestimate the number of false positives with a low-risk genotype.

6.8 Conclusion

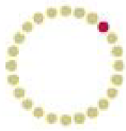
Using population-level screening data from NSO and evidence regarding the variation in the *PAH* and *ACADM* genes in humans, we have estimated the genotype of infants who were screened for PKU and MCAD deficiency at NSO between April 2006 and November 2015. In our evaluation screening for PKU, we noted 100% sensitivity for MS/MS, which could not be achieved by T-NGS alone; however, we found potential for improvement in the specificity achieved by MS/MS, and using T-NGS as a second-tier test may be a feasible and effective way to achieve this goal. On the other hand, we reported MS/MS for screening MCAD deficiency to have sensitivity that was below ideal, given the nature of NBS and the conditions being screened. In this regard, screening all infants for T-NGS may help improve both sensitivity and specificity of NBS for MCAD deficiency. Several factors should be considered, however, when contemplating the use of T-NGS on all screened infants. This includes feasibility, storage, consent, and unsolicited findings. Future studies should evaluate the costs and benefits of using T-NGS to propose a cost-effective model for implementing T-NGS in NBS. In addition, the methodology used in this study to estimate the NBS results from the application of T-NGS may be extended to a wider range of IEMs to provide a comprehensive picture of the implications across the screening panel.

APPENDICES

APPENDIX A: Proportion of Missing For Each Variable

Variable		Proportion Missing (%)
Screening	Patient ID	0
	Specimen ID	0
	Year-month of collection	0
	Age at collection	0
	Birth weight category	5.2
	Gestational age category	37.0
	Transfusion indicator	0
	Phe	0.004
	Phe/Tyr	0.004
	C8	0.004
	C8/C2	0.004
	C8/C10	0.55
	C6	0.004
	C10	0.004
	C10:1	0.004
	Screening outcome PKU	0.03
	Screening outcome MCAD deficiency	3.0
Diagnostics	Treatment centre	~ 0
	Definitive diagnosis, PKU	~ 0
	Definitive diagnosis, MCAD deficiency	~ 0
	Final outcome, PKU	~ 0
	Final outcome, MCAD deficiency	~ 0

APPENDIX B: Parent Information Form



NEWBORN SCREENING **ONTARIO**
DÉPISTAGE NÉONATAL **ONTARIO**
Information For Parents

My Baby's Blood Sample

What happens to the blood sample after testing is finished?
Why are the blood samples stored?
What are the stored blood samples used for?
Could the use of stored samples change in the future?
Will my baby's sample be used for research?
What happens when a researcher wants to access stored samples for research?
Can I have my baby's sample returned to me and/or destroyed?

What happens to the blood sample after testing is finished?

NSO began receiving blood samples on April 3, 2006. Dried blood spot samples are stored in a secure facility for 19 years, as they are a part of a child's medical record. After 19 years, the samples are destroyed. The samples for babies born before April 2006 are stored by the Public Health Laboratory in Toronto, Ontario.

Why are the blood samples stored?

The main reason is to ensure quality screening for all babies born in Ontario – to make sure that the newborn screening system and laboratory tests are working properly. NSO strives to ensure that every baby is offered the newborn screening test, and that every baby with one of the diseases tested for is found in the newborn period so that treatment can begin. Because the test is only a screening test and is not diagnostic, the detection rate may not be 100% for every disease. The stored samples help us improve our testing to try to make sure we don't miss any babies who need early treatment.

To try and find as many affected babies as possible, NSO regularly checks the screening cutoffs and decides what the normal ranges are for the chemicals we measure in the blood– the stored samples assist NSO in performing this task.

If a baby with a negative ("normal") newborn screen is diagnosed with one of the diseases we screen for, the baby's stored sample can be tested again. This helps NSO try and figure out why the baby was missed in the newborn period, and potentially stops the same thing from happening again in the future.

What are the stored blood samples used for?

Occasionally, the dried blood spot samples may be used for other purposes after testing is finished. These include:

-Quality control and quality assurance within the NBS laboratory (i.e. making sure that the laboratory and its testing are working properly);
-Retesting the sample to help make a diagnosis at the request of the baby's health care provider(s);
-Usage after a legal warrant or court order (e.g. by the Coroner's office if the baby has died unexpectedly);
-Release of part of the sample to another laboratory for other testing at the parent or guardian's written request;
-Samples may be used for research approved by a research ethics board if all identifying information has been removed so it is impossible to link an individual with the research results.

This is in compliance with the provisions in the Ontario Personal Health Information Privacy Act (PHIPA) 2004. Identifiable samples can only be used for research with the written consent of the individual or their surrogate decision maker (parent or guardian).

Could the use of stored samples change in the future?

Samples may be used for other purposed in the future, but only as authorized by PHIPA or any other applicable law, and following review by Ministry of Health and Long Term Care with the advice of the Maternal Child Screening Committee. We will communicate any changes in sample storage and use via this website, and, as always, we will continue to protect your baby's privacy.

Will my baby's sample be used for research?

There are currently only two ways that your baby's sample could be used for research:

1. Research that needs your baby's sample linked with your baby's identity. This could only happen after obtaining written consent from your child (if they were old enough to give consent) or from their surrogate decision maker (a parent or guardian). The study would have to be approved by a research ethics board.

· You/your child would be fully informed of the purpose of the research as well as the pros and cons of participating in the research.

· You/your child would have the ability to choose to participate or decline to participate in such a research study.

· If you/your child declined – your child's sample would NOT be used in the research.

· A research ethics board is a group of people representing the fields of medicine, science, ethics and the general public. A lay person sits on this board to ensure that the research being done is in the best interest of society. The remainder of the committee make sure the research is scientifically sound and does not violate ethical principles.

2. Research that requires your baby's sample may be allowed without obtaining your child's (or their surrogate decision maker's) consent ONLY IF:

· Your baby's sample is de-identified. This means that the portion of your child's sample used for such a study can NOT be traced back to your child in any way (e.g. all identifying information is removed from the sample: name, date of birth, health card number, address, postal code, mother's name, birth hospital, etc.). This means **no one**, including the researcher, the government or insurance companies, will be able to link the research results to the baby or their family.

· AND the study has been approved by a research ethics board.

What happens when a researcher wants to access stored samples for research?

A researcher who wants to access to stored samples for a study must submit a written request to NSO and gain approval from CHEO and their own institutional research ethics board (REB). An application for a research study's approval through a REB is a lengthy and rigorous process designed to guarantee the privacy and best interests of the research subjects (the babies whose samples will be used for the study). Details on the REB application process at the Children's Hospital of Eastern Ontario can be found at: www.cheori.org.

Can I have my baby's sample returned to me and/or destroyed?

Yes. To request destruction or release of sample from NSO, the parents/legal guardian/or child must complete a notarized request form or must attend NSO offices with originals of the required identifying documents to complete the forms in person. An original copy of the notarized form is required to complete the request.

A task force of the Ontario Advisory Committee on Newborn and Childhood Screening was created in 2008 to consider issues related to blood spot storage and use. A minimum storage length of 5 years was recommended to provide effective screening testing and quality assurance

for the program. This length of time was recommended for a number of reasons, most importantly because the diseases targeted by the screening program would be expected to cause health problems for an affected child by the time they were 5 years of age. Therefore storing samples for this length of time would allow investigation and possible re-testing if a child was diagnosed with one of the conditions on our panel following a negative newborn screen. It would also allow confirmation of whether or not a screening sample was obtained on the child. For this reason, the task force also recommended that an individual (or their parents or guardians) be able to request return or destruction of the sample after this five year period.

If you still wish destruction or release of the sample and your child is under five years of age, this request will be honoured. If you wish to have the sample destroyed after the child's fifth birthday, you will be asked to make this request at that time. To review NSO's policy on storage and secondary uses of the dried blood spot samples, please click here.

Parents/guardians who have further questions or concerns about the potential use of their child's stored sample are encouraged to contact NSO directly at 1-613-738-3222 or 1-877-627-8330.

APPENDIX C: Policy Regarding Storage and Secondary Use of NBS Samples

Storage and Secondary Use of the Newborn Screening Samples

STORAGE

For the purpose of newborn screening, a small sample of a baby's blood is collected on filter paper and sent to Newborn Screening Ontario (NSO). The dried blood spot sample (DBS) is used to screen for the diseases on the NSO panel. For approximately 6 months following the use of the DBS for screening, any residual sample is stored in NSO at the Children's Hospital of Eastern Ontario (CHEO). After this time, the DBS is stored by NSO in a secure facility off-site.

Dried blood spot samples are stored for 19 years, as they are a part of a child's medical record. After 19 years, the samples are destroyed. The samples for babies born before April 2006 are stored by the Public Health Laboratory in Toronto, Ontario, also for 19 years.

PURPOSE OF STORAGE

The primary reason to store the residual samples is to use them for screening and assuring the quality of screening provided by NSO. NSO strives to ensure that every baby is offered the newborn screening test, and to maximize the identification of babies affected with one of the diseases targeted on the screening panel to allow early treatment and prevention of morbidity or mortality. Conversely, the number of babies with a false positive screening result must be minimized. The stored samples can be used in a number of ways to help improve screening services. For example:

- If a baby has a positive screen but diagnostic testing proves the baby is not affected, it is sometimes necessary to re-run the tests on the original sample to make sure the original results are accurate. Sometimes a second sample is needed from the baby to complete the screening.
- NSO regularly checks the screening cutoffs and the normal ranges for the chemicals measured in the blood; the stored samples assist NSO in performing this task.
- If a baby with a negative newborn screen is later diagnosed with a targeted disease, the baby's stored sample (if one was received) can be tested again. This helps NSO and the baby's other health care providers determine why the baby was missed in the newborn period. If a preventable root cause is discovered, this allows corrective action to be taken to prevent a recurrence of the problem.
- NSO is constantly working to improve existing testing and develop new tests to provide better screening for the targeted diseases. Testing stored samples is the only way to validate these new or improved tests for use in the screening system.

SECONDARY USES

The residual DBS may be also be used for the following purposes:

1. Provision of Health Care. Examples include:
   i) Retesting the sample to help establish a diagnosis
   ii) Release of part of the sample to another laboratory for other testing
2. Sharing anonymized samples with other Canadian screening laboratories to provide external quality assurance for all Canadian newborn screening labs, including Ontario's.
3. Use under a Legal Warrant or Court Order. Examples include:
   i) Use by the Coroner's office if the baby has died unexpectedly
   ii) Use in a forensic investigation.
4. Research (see below)

NEWBORN SCREENING ONTARIO
DÉPISTAGE NÉONATAL ONTARIO

401 Smyth Road
Ottawa Ontario K1H 8L1
(613)738-3222
www.newbornscreening.on.ca

Version 1.0 – July 4, 2011

POLICIES FOR RESEARCH USE OF SAMPLES

There are two ways that a baby's sample can be used for research:

1. Research that needs the baby's sample linked with the baby's identity.

- This is only permitted with written consent from the child (if they are old enough to give consent) or from their surrogate decision maker (a parent or guardian).
- The study must be approved by a research ethics board.
- The parent/ guardian/ child must be fully informed of the purpose of the research as well as the pros and cons of participating in the research.
- The parent/ guardian/ child must have the ability to choose to participate or decline to participate in such a research study.
- If the parent/ guardian/ child declines – the child's sample will NOT be used in the research.

2. Research that requires a baby's sample may be allowed without obtaining the parent's/ guardian's/ child's consent ONLY IF:

- The baby's sample is fully de-identified, meaning that the portion of the child's sample used for such a study can NOT be traced back to the child in any way (e.g. ALL identifying information is removed from the sample: name, date of birth, health card number, address, postal code, mother's name, birth hospital, etc.).
- The study has been approved by a research ethics board.
- The study has been scientifically reviewed and approved by NSO.

DESTRUCTION OR RELEASE OF SAMPLE

To request destruction or release of sample from NSO, the parents/ legal guardian/ or child must complete a notarized request form or must attend NSO offices with originals of the required identifying documents to complete the forms in person.

- An original copy of the notarized form is required to complete the request.
- The parents will be notified of the recommendation to store samples until a child is five years of age. If the parents still wish destruction or release of the sample, this request will be honored. If they wish to have the sample destroyed after the child's fifth birthday, they will be asked to make this request at that time.
- The preferred method for releasing the sample is destruction versus returning the sample to the family however, requests to send the sample to the family will be honored.
- If the request form is not completed appropriately, a letter will be sent back to the family informing them of what further action they need to take to process their request.
- If the request arrives when there is a Legal Warrant or Court Order already in place, the request will not be processed until the Legal Warrant or Court Order are complete. The parents will be informed of the Legal Warrant or Court Order, and the relevant authorities will be informed of the parent's request.

APPENDIX D: REB Approval Letters

CHEO Research Ethics Board
Approval - Delegated Review

Principal Investigator: Dr. Pranesh Chakraborty
REB Protocol No: 17/16X
Romeo File No: 20170010
Project Title: CHEOREB# 17/16X - A simulation study to assess the impact on screening results of introducing next-generation sequencing (NGS) technology into newborn screening
Primary Affiliation: Biomedicine\Newborn Screening
Protocol Status: Active
Approval Date*: February 24, 2017
Valid Until:** February 15, 2018
Annual Renewal Submission Deadline: January 15, 2018

Documents Reviewed & Approved:

Document Name	Comments	Version Date
Case Report Form	Case Report Form	2017/01/08
Protocol	Study protocol	2017/02/15

This is to notify you that the Children's Hospital of Eastern Ontario Research Ethics Board has granted approval to the above named research study on the date noted above. Your project was reviewed under the delegated review stream, which is reserved for projects that involve no more than minimal risk to human subjects.

Final approval is granted for the above noted study, with the understanding that the investigator agrees to comply with the following requirements:

1. The investigator must conduct the study in compliance with the protocol and any additional conditions set out by the Board.
2. Investigators must submit an annual renewal report to the REB 30 days prior to the expiration date stated above.
3. The investigator must not implement any deviation from, or changes to, the protocol, consents or assents without the approval of the REB.

4. The investigator must, prior to use, submit to the Board changes to the study documentation, e.g., changes to the informed consent letters, recruitment materials.
5. Investigators must provide the Board with French versions of the consent form, unless a waiver has been granted. An interpreter should be offered to participants as required or at the request of the participant throughout the course of research.
6. The investigator must promptly report to the REB all unexpected and untoward occurrences (including the loss or theft of study data and other such privacy breaches).
7. Investigators must notify the REB of any study closures (closed to accrual, temporary, premature or permanent).
8. Investigators must submit a final report at the conclusion of the study.

Should you have any questions or concerns, please do not hesitate to contact the Research Ethics Board Office at 613-737-7600 ext. 3350 or 2128.

Regards,

Richard Carpentier, PhD
Chair, Research Ethics Board
Président, Comité d'éthique de la recherche

*The final approval date for initial delegated study applications approved with or without modifications will be the date the REB has determined that the conditions of approval have been satisfied.

**The expiry date of REB approval for initial study application that required no modifications will be as follows:

- If the date of review and approval was **on or before** the 15th of the month, the expiry date will be the 15th of the month prior to the date of review and approval by the Chair and/or delegate *in the following year*;
- If the date of review and approval was **after** the 15th the expiry date will be the 15th of the month in which the date of review and approval by the REB *in the following year*.

The expiry date of REB approval for initial study applications that **require modifications** will be as follows:

- If the initial feedback was sent **on or before** the 15th of the month, the expiry date will be the 15th of the month prior to the date the letter of REB feedback is issued to the investigator(s) *in the following year*;
- If the initial feedback was sent **after** the 15th the expiry date will be the 15th of the month in which the feedback was sent *in the following year*.

Ottawa Health Science Network Research Ethics Board/ Conseil d'éthique de la recherche du Réseau de science de la santé d'Ottawa

Civic Box 411 725 Parkdale Avenue, Ottawa, Ontario K1Y 4E9 613-798-5555 ext. 14902 Fax : 613-761-4311
http://www.ohri.ca/ohsn-reb

March 1, 2017

Dr. Brenda Wilson
University of Ottawa
Epidemiology & Community Medicine
451 Smyth Road
Ottawa, ON
K1H 8M5

Dear Dr. Wilson:

Re: Protocol #	**20170118-01H**	**A simulation study to assess the impact on screening results of introducing next-generation sequencing technology into newborn screening**
Protocol approval valid until -		**February 28, 2018**

I am pleased to inform you that your Application for Chart Review underwent expedited review by the Ottawa Health Science Network Research Ethics Board (OHSN-REB), and is approved. No changes, amendments or addenda may be made to the protocol without the OHSN-REB's review and approval.

Approval is conditional upon receipt of the Children's Hospital of Eastern Ontario Research Ethics Board approval letter.

Approval is for the following:
- Protocol (version 1) dated February 15, 2017

If the study is to continue beyond the expiry date noted above, a Renewal Form should be submitted to the OHSN-REB approximately six weeks prior to the current expiry date. If the study has been completed by this date, a Termination Report should be submitted.

The OHSN-REB complies with the membership requirements and operates in compliance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans; the International Conference on Harmonization - Good Clinical Practice: Consolidated Guideline, the provisions of the Personal Health Information Protection Act 2004.

Yours sincerely,

Raphael Saginur, M.D.
Chairperson
Ottawa Health Science Network Research Ethics Board

/kd

APPENDIX E: Variants with Conflicting Interpretations

PKU Variants with Conflicting Phenotype[81-87]

Variant	Conflict	Final Decision	Rationale	Submitters (Year of Submission)
12:103238111	Pathogenic vs. VUS	**VUS**	1. The Emory Genetics Lab provided more recent findings and classified the variant as VUS. 2. A study by Guldberg et al (1994) classified this variant as pathogenic and associated with hyperphenylalaninemia, rather than PKU. It was stated in the article that "the association with non-PKU HPA cannot be established unambiguously". This indicates that the variant is likely to be VUS.	Guldberg (1994); Emory Genetics Lab (2015)
12:103246701	Pathogenic vs. VUS	**Pathogenic**	1. Four of the submitting laboratories had classified the variant as pathogenic in 2016. One laboratory classified the variant as VUS in 2014. 2. In addition, the study by Guldberg et al (1994) concluded that the variant was pathogenic. Overall, the evidence appears to indicate that this is a pathogenic variant.	Guldberg (1994); Emory Genetics Lab (2015); Division of Human Genetics, Children's Hospital of Philadelphia (2016); University of Washington Medical Center (2014); GeneDx (2016); Counsyl (2016)
12:103271326	Likely pathogenic vs. VUS	**Likely pathogenic**	1. The Emory Genetics Lab classified this variant as VUS in 2013, and as likely pathogenic in 2014. The latest submission was retained.	Emory Genetics Lab (2013); Emory Genetics Lab (2014)
12:103260377	Likely pathogenic vs. VUS	**VUS**	VUS 1. Both submissions were made in the same year. 2. There is more evidence suggesting that this is a VUS. GeneDx did not provide any supporting evidence with their submission.	GeneDx (Aug. 2016); Illumina Clinical Services Laboratory (Jun. 2016); (Reblova, 2013; Aulehla-Scholz, 2003; Bercovich, 2008; Desviat, 1999)

MCADD Variants with Conflicting Phenotype[88-93]

Variant	Conflict	Final Decision	Rationale	Submitters (Year of Submission)
1:76198337	Pathogenic vs. VUS	**Pathogenic**	1. The submissions are all relatively recent. 2. Out of four submissions from laboratories, three reported the variant to be pathogenic, including the most recent submission.	Emory Genetics Lab (Aug. 2012); Center for Pediatric Genomic Medicine, Children's Mercy Hospital (Mar. 2015); ARUP Laboratories (Feb. 2015)
1:76200531	Pathogenic/Likely pathogenic vs. VUS	**Pathogenic/likely pathogenic**	1. The Emory Genetics Lab had initially classified it as VUS, but this was later changed to Pathogenic.	Emory Genetics Lab (May 2015; Oct. 2015); GeneDx (Aug. 2016); ARUP Laboratories (Feb. 2015); Counsyl (Apr. 2016)
1:76211507	Pathogenic/Likely pathogenic vs. VUS	**Pathogenic/Likely pathogenic**	1. Submissions for this variant are all relatively recent. 2. All evidence from the literature indicates that this mutation is associated with at least symptoms of mild MCAD deficiency, and may later develop into a more severe form. 3. Four submissions out of 5 have reported this to be either a pathogenic or likely pathogenic variant. Only one submission was reported as VUS.	Maier (2005); Waddell (2006); Maier (2009); Smith (2010); Sturm (2012); Jank (2014); Emory Genetics Lab (Aug. 2012); GeneDx (Dec. 2015); ARUP Laboratories (Feb. 2015); Counsyl (Nov. 2015)
1:76211589	Likely pathogenic vs. VUS	**No resolution**	There was insufficient evidence provided by submitters to arrive at a resolution for this conflict.	ARUP Laboratories (Feb. 2015); Emory Genetics Lab (Oct. 2015)
1:76216214	Pathogenic vs. VUS	**Pathogenic**	1. Submissions for this variant are both relatively recent. 2. Evidence from the literature indicates that this variant, when	Andresen (2001); Touw (2013); Emory Genetic Laboratories (Oct. 2014); ARUP Laboratories (Feb. 2015)

			identified in combination with the classic 985A>G mutation, results in elevated C8 levels beyond the threshold for MCAD deficiency.	
1:76198574	Likely pathogenic vs. VUS	**Likely pathogenic**	1. In 2013, it was classified as uncertain significance by Emory Genetics Lab, but this was updated to likely pathogenic in 2015.	Emory Genetics Lab (2013; 2015)
1:76199314	Likely pathogenic vs. VUS	**VUS**	1. There were two submissions, both of which are relatively recent. 2. The literature indicates that there has been no evidence in humans that his variant is associated with MCAD deficiency. Assertions of pathogenicity are based on experimental evidence.	Maier (2005); Sturm (2012); Emory Genetics Lab (Oct. 2015); Counsyl (Mar. 2014)
1:76200518-76200520	Likely pathogenic vs. VUS	**No resolution**	There was insufficient evidence provided by submitters to arrive at a resolution for this conflict.	ARUP Laboratories (Feb. 2015); Emory Genetics Laboratories (Oct. 2015)
1:76211574	Likely pathogenic vs. VUS	**No resolution**	There was insufficient evidence provided by submitters to arrive at a resolution for this conflict.	GeneDx (Jun. 2016); ARUP Laboratories (Feb. 2015)
1:76211569	Benign vs. VUS	**No resolution**	There was insufficient evidence provided by submitters to arrive at a resolution for this conflict.	ARUP Laboratories (Feb. 2015); GeneDx (Feb. 2015); Illumina Clinical Services (Jun. 2016)
1:76228525	Benign vs. VUS	**No resolution**	There was insufficient evidence provided by submitters to arrive at a resolution for this conflict.	ARUP Laboratories (Feb. 2015); Illumina Clinical Services (Jun. 2016)

APPENDIX F: Complete Sensitivity Analyses

Table 5.13A Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || P or LP* to the lower limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	10	0	0	77
P or LP || VUS	8	59	106	0	173
VUS || VUS	3	5	59	0	67
P or LP || B, LB, or WT	5	12	24,418	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	358	1,336,998	0	1,336,394
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13B. Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || P or LP* to the upper limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	85	0	0	152
P or LP || VUS	8	0	171	0	179
VUS || VUS	3	0	64	0	67
P or LP || B, LB, or WT	5	7	24,423	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	352	1,335,998	0	1,336,319
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13C Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || VUS* to the lower limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	45	0	0	112
P or LP || VUS	8	28	89	0	117
VUS || VUS	3	3	61	0	67
P or LP || B, LB, or WT	5	10	24,420	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	358	1,336,060	0	1,336,421
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13D Sensitivity analysis for PKU. Varied proportion of infants who were *P or LP || VUS* to the upper limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	45	0	0	112
P or LP || VUS	8	28	199	0	235
VUS || VUS	3	3	61	0	67
P or LP || B, LB, or WT	5	10	24,420	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	358	1,335,942	0	1,336,303
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13E Sensitivity analysis for PKU. Varied proportion of infants who were *VUS || VUS* to the lower limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	45	0	0	112
P or LP || VUS	8	28	137	0	173
VUS || VUS	3	3	38	0	44
P or LP || B, LB, or WT	5	10	24,420	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	358	1,335,942	0	1,336,303
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13F Sensitivity analysis for PKU. Varied proportion of infants who were *VUS || VUS* to the upper limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	67	45	0	0	112
P or LP || VUS	8	28	137	0	173
VUS || VUS	3	3	89	0	95
P or LP || B, LB, or WT	5	10	24,420	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	358	1,335,970	0	1,336,331
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13G Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || P or LP* genotype to the lower limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	62	50	0	0	112
P or LP || VUS	10	25	137	0	173
VUS || VUS	4	2	61	0	67
P or LP || B, LB, or WT	6	9	24,420	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	4	358	1,335,997	0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13H Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || P or LP* genotype to the upper limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	72	40	0	0	112
P or LP || VUS	6	33	134	0	173
VUS || VUS	2	3	62	0	67
P or LP || B, LB, or WT	4	10	24,421	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	2	358	1,335,999	0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13I Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || VUS* genotype to the lower limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	70	42	0	0	112
P or LP || VUS	4	28	141	0	173
VUS || VUS	3	3	61	0	67
P or LP || B, LB, or WT	6	10	24,419	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	358	1,335,998	0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.13J Sensitivity analysis for PKU. Varied proportion of infants who were true positives and had *P or LP || VUS* genotype to the upper limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	64	48	0	0	112
P or LP || VUS	11	28	134	0	173
VUS || VUS	3	3	61	0	67
P or LP || B, LB, or WT	5	10	24,420	0	24,435
VUS || B, LB, or WT	0	5	18,959	0	18,964
B, LB, or WT || B, LB, or WT	3	358	1,335,998	0	1,336,359
MS/MS Totals	**86**	**449**	**1,379,575**	**0**	**1,380,110**
	535		**1,379,575**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.14A. Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || P or LP* to the lower limit of the 95% confidence interval of the estimate.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	56	0	0	1	57
P or LP || VUS	34	0	48	0	82
VUS || VUS	4	0	16	0	20
P or LP || B, LB, or WT	0	39	21,040	0	21,079
VUS || B, LB, or WT	0	1	10,436	0	10,437
B, LB, or WT || B, LB, or WT	0	142	1,348,293	0	1,348,435
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.14B. Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || P or LP* to the upper limit of the 95% confidence interval of the estimate.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	94	0	0	18	112
P or LP || VUS	0	0	82	0	82
VUS || VUS	0	0	20	0	20
P or LP || B, LB, or WT	0	39	21,040	0	21,079
VUS || B, LB, or WT	0	1	10,436	0	10,437
B, LB, or WT || B, LB, or WT	0	142	1,348,238	0	1,348,380
MS/MS Totals	**94**	**182**	**1,379,816**	**18**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.14C Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || VUS* to the lower limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	81	0	0	1	82
P or LP || VUS	12	0	38	0	50
VUS || VUS	1	0	19	0	20
P or LP || B, LB, or WT	0	39	21,040	0	21,079
VUS || B, LB, or WT	0	1	10,436	0	10,437
B, LB, or WT || B, LB, or WT	0	142	1,348,236	0	1,348,378
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.14D Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *P or LP || VUS* to the upper limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	81	0	0	1	82
P or LP || VUS	12	0	107	0	119
VUS || VUS	1	0	19	0	20
P or LP || B, LB, or WT	0	39	21,040	0	21,079
VUS || B, LB, or WT	0	1	10,436	0	10,437
B, LB, or WT || B, LB, or WT	0	142	1,348,305	0	1,348,447
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.14E Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *VUS || VUS* to the lower limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	81	0	0	1	82
P or LP || VUS	12	0	70	0	82
VUS || VUS	1	0	10	0	11
P or LP || B, LB, or WT	0	39	21,040	0	21,079
VUS || B, LB, or WT	0	1	10,436	0	10,437
B, LB, or WT || B, LB, or WT	0	142	1,348,277	0	1,348,419
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

Table 5.14F Sensitivity analysis for MCAD deficiency. Varied proportion of infants who were *VUS || VUS* to the upper limit of the 95% confidence interval.

T-NGS Result (Allele 1 || Allele 2)	MS/MS Result				T-NGS Totals
	Positive		Negative		
	True	False	True	False	
P or LP || P or LP	81	0	0	1	82
P or LP || VUS	12	0	70	0	82
VUS || VUS	1	0	31	0	32
P or LP || B, LB, or WT	0	39	21,040	0	21,079
VUS || B, LB, or WT	0	1	10,436	0	10,437
B, LB, or WT || B, LB, or WT	0	142	1,348,256	0	1,348,398
MS/MS Totals	**94**	**182**	**1,379,833**	**1**	**1,380,110**
	276		**1,379,834**		

P or LP – pathogenic or likely pathogenic; B, LB, or WT – benign, likely benign, or wild-type; VUS – variant of uncertain significance

REFERENCES

1. 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 Nov;50(6):142–62.
2. Raffle A, Gray M. *Screening: evidence and practice*. New York: Oxford University Press; 2007.
3. Wilson BJ, Nicholls SG. The Human Genome Project, and recent advances in personalized genomics. Risk Manag Healthc Policy. 2015 Feb 16;8:9–20.
4. Wilson J, Jungner M. *Principles and practice for screening of disease*. Geneva: World Health Organization; 1968.
5. Driscoll CJ, McPherson B. *Newborn screening systems: the complete perspective*. San Diego: Plural Publishing Inc; 2010.
6. Berry SA. Newborn screening. Clin Perinatol. 2015 Jun;42(2):441–453, x.
7. Khoury MJ, McCabe LL, McCabe ERB. Population Screening in the Age of Genomic Medicine. New England Journal of Medicine. 2003 Jan 2;348(1):50–8.
8. Khoury MJ, Coates RJ, Fennell ML, Glasgow RE, Scheuner MT, Schully SD, et al. Multilevel Research and the Challenges of Implementing Genomic Medicine. J Natl Cancer Inst Monogr. 2012 May;2012(44):112–20.
9. van El CG, Cornel MC. Genetic testing and common disorders in a public health framework. Eur J Hum Genet. 2011 Apr;19(4):377–81.
10. Haddow JE, Palomaki GE. ACCE: A model process for evaluating data on emerging genetic tests. In: *Human genome epidemiology: A scientific foundation for using genetic information to improve health and prevent disease.* Khoury M, Little J, Burke W. Oxford University Press; 2003.
11. Potter BK, Avard D, Graham ID, Entwistle VA, Caulfield TA, Chakraborty P, et al. Guidance for considering ethical, legal, and social issues in health technology assessment: application to genetic screening. Int J Technol Assess Health Care. 2008;24(4):412–22.
12. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genet Med. 2009 Jan;11(1):3–14.

13. Mattocks CJ, Morris MA, Matthijs G, Swinnen E, Corveleyn A, Dequeker E, et al. A standardized framework for the validation and verification of clinical molecular genetic tests. Eur J Hum Genet. 2010 Dec;18(12):1276–88.
14. Veenstra DL, Roth JA, Garrison LP, Ramsey SD, Burke W. A Formal Risk-Benefit Framework for Genomic Tests: Facilitating the Appropriate Translation of Genomics Into Clinical Practice. Genet Med. 2010 Nov;12(11):686–93.
15. Manolio TA, Chisholm RL, Ozenberger B, Roden DM, Williams MS, Wilson R, et al. Implementing genomic medicine in the clinic: the future is here. Genet Med. 2013 Apr;15(4):258–67.
16. Howard HC, Knoppers BM, Cornel MC, Wright Clayton E, Sénécal K, Borry P, et al. Whole-genome sequencing in newborn screening? A statement on the continued importance of targeted approaches in newborn screening programmes. Eur J Hum Genet. 2015 Dec;23(12):1593–600.
17. Brand A, Brand H, Schulte in den Bäumen T. The impact of genetics and genomics on public health. Eur J Hum Genet. 2007 Oct 24;16(1):5–13.
18. Burke W, Tarini B, Press NA, Evans JP. Genetic screening. Epidemiol Rev. 2011;33:148–64.
19. Francescatto L, Katsanis N. Newborn screening and the era of medical genomics. Semin Perinatol. 2015 Dec;39(8):617–22.
20. Kuehn BM. After 50 years, newborn screening continues to yield public health gains. JAMA. 2013 Mar 27;309(12):1215–7.
21. Frazier DM. Tandem Mass Spectrometry Newborn Screening and Its Impact on Inborn Errors of Metabolism. Topics in clinical nutrition [Internet]. 2009 [cited 2017 Aug 7]; Available from: http://agris.fao.org/agris-search/search.do?recordID=US201301796136
22. El-Hattab AW. Inborn errors of metabolism. Clin Perinatol. 2015 Jun;42(2):413–439, x.
23. Ezgu F. Inborn Errors of Metabolism. Adv Clin Chem. 2016;73:195–250.
24. Wilcken B, Wiley V. Newborn screening. Pathology. 2008 Feb;40(2):104–15.
25. Ozben T. Expanded newborn screening and confirmatory follow-up testing for inborn errors of metabolism detected by tandem mass spectrometry. Clin Chem Lab Med. 2013 Jan;51(1):157–76.
26. Hayeems RZ, Miller FA, Carroll JC, Little J, Allanson J, Bytautas JP, et al. Primary care role in expanded newborn screening: After the heel prick test. Can Fam Physician. 2013 Aug;59(8):861–8.

27. Wassertheil-Smoller, Smoller J. *Biostatistics and epidemiology: A primer for health and biomedical professionals*. New York: Springer-Verlag; 2015.
28. Slaughter JL, Meinzen-Derr J, Rose SR, Leslie ND, Chandrasekar R, Linard SM, et al. The effects of gestational age and birth weight on false-positive newborn-screening rates. Pediatrics. 2010 Nov;126(5):910–6.
29. Gurian EA, Kinnamon DD, Henry JJ, Waisbren SE. Expanded newborn screening for biochemical disorders: the effect of a false-positive result. Pediatrics. 2006 Jun;117(6):1915–21.
30. Tluczek A, Orland KM, Nick SW, Brown RL. Newborn Screening: An Appeal for Improved Parent Education. J Perinat Neonatal Nurs. 2009;23(4):326–34.
31. Waisbren SE, Weipert CM, Walsh RC, Petty CR, Green RC. Psychosocial Factors Influencing Parental Interest in Genomic Sequencing of Newborns. Pediatrics. 2016 Jan;137 Suppl 1:S30-35.
32. Hayeems RZ, Miller FA, Barg CJ, Bombard Y, Kerr E, Tam K, et al. Parent Experience With False-Positive Newborn Screening Results for Cystic Fibrosis. Pediatrics. 2016 Sep;138(3).
33. Karaceper MD, Chakraborty P, Coyle D, Wilson K, Kronick JB, Hawken S, et al. The health system impact of false positive newborn screening results for medium-chain acyl-CoA dehydrogenase deficiency: a cohort study. Orphanet Journal of Rare Diseases. 2016 Feb 3;11:12.
34. Therrell BL, Padilla CD, Loeber JG, Kneisser I, Saadallah A, Borrajo GJC, et al. Current status of newborn screening worldwide: 2015. Semin Perinatol. 2015 Apr;39(3):171–87.
35. Canadian Organization for Rare Disorders. *Newborn screening in Canada status report.* 2015. https://www.raredisorders.ca/content/uploads/Canada-NBS-status-updated-Sept.-3-2015.pdf (accessed 10 July 2016)
36. Newborn Screening Ontario. *Disease information*. https://www.newbornscreening.on.ca/en/disease. (accessed 11 July 2016)
37. Centers for Disease Control and Prevention. Ten great public health achievements – United States, 2001-2010. JAMA. 2011; 306(1): 36-8.
38. Bailey DB Jr, Beskow LM, Davis AM, Skinner D. Changing perspectives on the benefits of newborn screening. *Ment Retard Dev Disabil Res Rev*. 2006; 12(4):270-9.

39. Friedman E. Next generation sequencing for newborn screening: are we there yet? Genet Res (Camb). 2015 Sep 22;97:e17.

40. Bhattacharjee A, Sokolsky T, Wyman SK, Reese MG, Puffenberger E, Strauss K, et al. Development of dna confirmatory and high-risk diagnostic testing for newborns using targeted next-generation dna sequencing. Genet Med. 2015; 17(5): 337-47.

41. PHG Foundation. Next steps in the sequence: The implications of whole genome sequencing for health in the UK. Cambridge: PHG Foundation; 2011.

42. Goldenberg AJ, Sharp RR. The ethical hazards and programmatic challenges of genomic newborn screening. JAMA. 2012 Feb 1;307(5):461–2.

43. Knoppers BM, Sénécal K, Borry P, Avard D. Whole-genome sequencing in newborn screening programs. Sci Transl Med. 2014 Mar 26;6(229):229cm2.

44. National Institutes of Health. *NIH program explores the use of genomic sequencing in newborn healthcare.* 2013. https://www.nih.gov.news-events/news-releases/nih-program-explores-use-genomic-sequencing-newborn-healthcare *(accessed 14 July 2016)*

45. Botkin JR, Rothwell E. Whole Genome Sequencing and Newborn Screening. Curr Genet Med Rep. 2016 Mar 1;4(1):1–6.

46. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015 May;17(5):405–24.

47. Nussbaum RL, McInnes RR, Willard HF. *Genetics in medicine*, 7th ed. Philadelphia: Elsevier; 2007.

48. Blau N. Genetics of phenylketonuria: Then and now. *Human Mutation*. 2016; 37(6): 508-15.

49. Al Hafid N, Christodoulou J. Phenylketonuria: A review of current and future treatments. *Transl Pediatr*. 2015; 4(4): 304-17.

50. Alberta Health Services. *Phenylketonuria (PKU)*. 2013. http://www.albertahealthservices.ca/assets/info/hp/nms/if-hp-nms-pku.pdf. (accessed 14 July 2016).

51. Blau N, Shen N, Carducci C. Molecular genetics and diagnosis of phenylketonuria: State of the art. *Expert Rev. Mol. Diagn*. 2014; 14(6): 655-71.

52. Vockley J, Andersson HC, Antshel KM, Braverman NE, Burton BK, Frazier DM, et al. Phenylalanine hydroxylase deficiency: diagnosis and management guideline. Genet Med. 2014 Feb;16(2):188–200.

53. Kennedy S, Potter BK, Wilson K, Fisher L, Geraghty M, Milburn J, et al. The first three years of screening for medium chain acyl-CoA dehydrogenase deficiency (MCADD) by newborn screening ontario. BMC Pediatrics. 2010 Nov 17;10:82.

54. Vishwanath VA. Fatty Acid Beta-Oxidation Disorders: A Brief Review. Ann Neurosci. 2016 Mar;23(1):51–5.

55. Leonard JV, Dezateux C. Newborn screening for medium chain acyl CoA dehydrogenase deficiency. Arch Dis Child. 2009 Mar;94(3):235–8.

56. Steiber A, Kerner J, Hoppel CL. Carnitine: a nutritional, biosynthetic, and functional perspective. Mol Aspects Med. 2004 Dec;25(5–6):455–73.

57. Sturm M, Herebian D, Mueller M, Laryea MD, Spiekerkoetter U. Functional Effects of Different Medium-Chain Acyl-CoA Dehydrogenase Genotypes and Identification of Asymptomatic Variants. PLOS ONE. 2012 Sep 17;7(9):e45110.

58. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016 18;536(7616):285–91.

59. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016 Jan 4;44(D1):D862-868.

60. Kitts A, Phan L, Ward M, Holmes JB. The database of short genetic variation (dbSNP). National Center for Biotechnology Information. *The ncbi handbook*. Bethesda: National Center for Biotechnology Information; 2013.

61. Woolf LI, Cranston WI, Goodwin BL. Genetics of phenylketonuria. *Nature*. 1967; 213: 882-3.

62. Karam PE, Alhamra RS, Nemer G, Usta J. Spectrum of mutations in Lebanese patients with phenylalanine hydroxylase deficiency. Gene. 2013 Feb 15;515(1):117–22.

63. Murad H, Dabboul A, Moassas F, Alasmar D, Al-Achkar W. Mutation spectrum of phenylketonuria in Syrian population: genotype-phenotype correlation. Gene. 2013 Oct 10;528(2):241–7.

64. Bueno MA, González-Lamuño D, Delgado-Pecellín C, Aldámiz-Echevarría L, Pérez B, Desviat LR, et al. Molecular epidemiology and genotype-phenotype correlation in phenylketonuria patients from South

Spain. J Hum Genet. 2013 May;58(5):279–84.

65. Vela-Amieva M, Abreu-González M, González-del Angel A, Ibarra-González I, Fernández-Lainez C, Barrientos-Ríos R, et al. Phenylalanine hydroxylase deficiency in Mexico: genotype–phenotype correlations, BH4 responsiveness and evidence of a founder effect. Clin Genet. 2015 Jul 1;88(1):62–7.
66. Couce ML, Bóveda MD, Fernández-Marmiesse A, Mirás A, Pérez B, Desviat LR, et al. Molecular epidemiology and BH4-responsiveness in patients with phenylalanine hydroxylase deficiency from Galicia region of Spain. Gene. 2013 May 25;521(1):100–4.
67. Mallolas J, Vilaseca MA, Campistol J, Lambruschini N, Cambra FJ, Estivill X, et al. Mutational spectrum of phenylalanine hydroxylase deficiency in the population resident in Catalonia: genotype-phenotype correlation. Hum Genet. 1999 Nov;105(5):468–73.
68. Maier EM, Liebl B, Röschinger W, Nennstiel-Ratzel U, Fingerhut R, Olgemöller B, et al. Population spectrum of ACADM genotypes correlated to biochemical phenotypes in newborn screening for medium-chain acyl-CoA dehydrogenase deficiency. Hum Mutat. 2005 May;25(5):443–52.
69. Schuster SC. Next-generation sequencing transforms today's biology. Nat Methods. 2008 Jan;5(1):16–8.
70. Blau N, Spronsen FJ van, Levy HL. Phenylketonuria. The Lancet. 2010 Oct 23;376(9750):1417–27.
71. Centres for Disease Control and Prevention. New issues in newborn screening for phenylketonuria and congenital hypothyroidism: A commentary from the committee on genetics of the american academy of pediatrics. *Morbidity and Mortality Weekly Report*. 1982; 31(15): 185-89.
72. Cunningham S, O'Doherty KC, Sénécal K, Secko D, Avard D. Public concerns regarding the storage and secondary uses of residual newborn bloodspots: an analysis of print media, legal cases, and public engagement activities. J Community Genet. 2015 Apr;6(2):117–28.
73. Rothwell E, Anderson R, Goldenberg A, Lewis MH, Stark L, Burbank M, et al. Assessing public attitudes on the retention and use of residual newborn screening blood samples: a focus group study. Soc Sci Med. 2012 Apr;74(8):1305–9.
74. Jepson RG, Forbes CA, Sowden AJ, Lewis RA. Increasing informed uptake and non-uptake of screening: Evidence from a systematic review. *Health Expect*. 2001; 4(2):116-26
75. Miller FA, Robert JS, Hayeems RZ. Questioning the Consensus: Managing Carrier Status Results Generated by Newborn Screening. Am J Public Health. 2009 Feb;99(2):210–5.

76. Hayeems RZ, Bytautas JP, Miller FA. A systematic review of the effects of disclosing carrier results generated through newborn screening. J Genet Couns. 2008 Dec;17(6):538–49.
77. Ulph F, Cullinan T, Qureshi N, Kai J. Parents' responses to receiving sickle cell or cystic fibrosis carrier results for their child following newborn screening. Eur J Hum Genet. 2015 Apr;23(4):459–65.
78. Noke M, Wearden A, Peters S, Ulph F. Disparities in current and future childhood and newborn carrier identification. J Genet Couns. 2014 Oct;23(5):701–7.
79. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing. Genet Med. 2013 Jul;15(7):565–74.
80. Rigter T, Henneman L, Kristoffersson U, Hall A, Yntema HG, Borry P, et al. Reflecting on Earlier Experiences with Unsolicited Findings: Points to Consider for Next-Generation Sequencing and Informed Consent in Diagnostics. Hum Mutat. 2013 Oct;34(10):1322–8.
81. Guldberg P, Henriksen KF, Thöny B, Blau N, Güttler F. Molecular heterogeneity of nonphenylketonuria hyperphenylalaninemia in 25 Danish patients. Genomics. 1994 May 15;21(2):453–5.
82. Aulehla-Scholz C, Heilbronner H. Mutational spectrum in German patients with phenylalanine hydroxylase deficiency. Hum Mutat. 2003 Apr;21(4):399–400.
83. Bercovich D, Elimelech A, Yardeni T, Korem S, Zlotogora J, Gal N, et al. A mutation analysis of the phenylalanine hydroxylase (PAH) gene in the Israeli population. Ann Hum Genet. 2008 May;72(Pt 3):305–9.
84. Desviat LR, Pérez B, Gámez A, Sánchez A, García MJ, Martínez-Pardo M, et al. Genetic and phenotypic aspects of phenylalanine hydroxylase deficiency in Spain: molecular survey by regions. Eur J Hum Genet. 1999 Apr;7(3):386–92.
85. Smith EH, Thomas C, McHugh D, Gavrilov D, Raymond K, Rinaldo P, et al. Allelic diversity in MCAD deficiency: the biochemical classification of 54 variants identified during 5 years of ACADM sequencing. Mol Genet Metab. 2010 Jul;100(3):241–50.
86. Bercovich D, Elimelech A, Zlotogora J, Korem S, Yardeni T, Gal N, et al. Genotype-phenotype correlations analysis of mutations in the phenylalanine hydroxylase (PAH) gene. J Hum Genet. 2008;53(5):407–18.

87. Réblová K, Hrubá Z, Procházková D, Pazdírková R, Pouchlá S, Zeman J, et al. Hyperphenylalaninemia in the Czech Republic: genotype-phenotype correlations and in silico analysis of novel missense mutations. Clin Chim Acta. 2013 Apr 18;419:1–10.
88. Touw CML, Smit GPA, Niezen-Koning KE, Bosgraaf-de Boer C, Gerding A, Reijngoud D-J, et al. In vitro and in vivo consequences of variant medium-chain acyl-CoA dehydrogenase genotypes. Orphanet J Rare Dis. 2013 Mar 20;8:43.
89. Waddell L, Wiley V, Carpenter K, Bennetts B, Angel L, Andresen BS, et al. Medium-chain acyl-CoA dehydrogenase deficiency: genotype-biochemical phenotype correlations. Mol Genet Metab. 2006 Jan;87(1):32–9.
90. Andresen BS, Dobrowolski SF, O'Reilly L, Muenzer J, McCandless SE, Frazier DM, et al. Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by MS/MS-based prospective screening of newborns differ from those observed in patients with clinical symptoms: identification and characterization of a new, prevalent mutation that results in mild MCAD deficiency. Am J Hum Genet. 2001 Jun;68(6):1408–18.
91. Maier EM, Liebl B, Röschinger W, Nennstiel-Ratzel U, Fingerhut R, Olgemöller B, et al. Population spectrum of ACADM genotypes correlated to biochemical phenotypes in newborn screening for medium-chain acyl-CoA dehydrogenase deficiency. Hum Mutat. 2005 May;25(5):443–52.
92. Maier EM, Gersting SW, Kemter KF, Jank JM, Reindl M, Messing DD, et al. Protein misfolding is the molecular mechanism underlying MCADD identified in newborn screening. Hum Mol Genet. 2009 May 1;18(9):1612–23.
93. Jank JM, Maier EM, Reiβ DD, Haslbeck M, Kemter KF, Truger MS, et al. The domain-specific and temperature-dependent protein misfolding phenotype of variant medium-chain acyl-CoA dehydrogenase. PLoS ONE. 2014;9(4):e93852.