

**THE EPIDEMIOLOGY AND HEALTH SYSTEM
IMPACT OF MEDIUM-CHAIN ACYL-COA
DEHYDROGENASE DEFICIENCY AMONG
AFFECTED CHILDREN AND THOSE WITH
FALSE POSITIVE NEWBORN SCREENING
RESULTS IN ONTARIO, CANADA**

By

MARIA D KARACEPER

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in
partial fulfillment of the requirements for the MSc degree in Epidemiology

Department of Epidemiology and Community Medicine
Faculty of Medicine
University of Ottawa

© MARIA D KARACEPER
OTTAWA, CANADA, 2014

ABSTRACT

Objective: To describe the epidemiology and health system impact of medium-chain acyl-CoA dehydrogenase deficiency (MCADD) in Ontario.

Methods: Following a review of methods to estimating robust health event rates for small populations, this study described health services use among infants diagnosed with MCADD or received a false positive newborn screening result for MCADD from April 2006 through March 2010. Each cohort was compared with screen negative infants by linking to databases encompassing physician visits, emergency department care, and hospitalizations.

Results: Relative to comparison birth cohorts, children with MCADD (n=40) experienced significantly higher rates of all health service types, regardless of age at the time of visit; infants with false positive results for MCADD (n=43) experienced significantly higher rates of physician visits and hospitalizations in the first year of life only.

Conclusion: This study makes an important contribution to the limited existing research describing the health system impact of rare diseases.

TABLE OF CONTENTS

ABSTRACT	ii
TABLE OF CONTENTS.....	iii
LIST OF TABLES	vi
LIST OF FIGURES.....	vii
CONTRIBUTION OF AUTHORS	viii
ACKNOWLEDGMENTS.....	ix
CHAPTER ONE: INTRODUCTION	1
SUMMARY OF RATIONALE FOR THESIS PROJECT	1
OBJECTIVES OF THESIS PROJECT.....	2
REFERENCES	3
CHAPTER TWO: BACKGROUND.....	4
MEDIUM-CHAIN ACYL-COA DEHYDROGENASE DEFICIENCY.....	5
NEWBORN SCREENING FOR MCADD.....	7
THE EFFECT OF NEWBORN SCREENING PROGRAMS	9
<i>False positive results</i>	10
<i>Cost of newborn screening and MCADD</i>	12
THE IMPORTANCE OF MEASURING HEALTH SERVICES IMPACTS	13
REFERENCES	15
CHAPTER THREE: METHODS REVIEW – MANUSCRIPT #1.....	24
ABSTRACT.....	25
INTRODUCTION.....	26
EXAMPLE DISEASES.....	27

METHODS FOR ROBUST ESTIMATION OF EVENT RATES IN SMALL POPULATIONS	29
<i>Aggregation</i>	29
<i>Data Smoothing Techniques</i>	33
DISCUSSION.....	37
CONCLUSIONS AND RECOMMENDATIONS	40
REFERENCES	42
CHAPTER FOUR: HEALTH SYSTEM IMPACT OF MCADD – MANUSCRIPT #2	50
ABSTRACT.....	51
INTRODUCTION	53
PATIENTS AND METHODS	55
RESULTS	61
DISCUSSION.....	67
CONCLUSION.....	72
REFERENCES	81
CHAPTER FIVE: HEALTH SYSTEM IMPACT OF FALSE POSITIVE RESULTS –	
MANUSCRIPT #3	89
ABSTRACT.....	90
INTRODUCTION	92
METHODS.....	94
RESULTS	100
DISCUSSION.....	104
CONCLUSION.....	107
REFERENCES	113

CHAPTER SIX: FINAL DISCUSSION	117
REFERENCES	122
APPENDICES.....	123
APPENDIX A (Ethics Approval)	123
APPENDIX B (Overview of potential approaches for modeling count data and example use of the Vuong statistic to select an approach)	124
APPENDIX C (Example: Weighted moving average calculation)	133

LIST OF TABLES

Table 3.1	41
Table 4.1	74
Table 4.2	77
Table 4.3	78
Table 4.4	79
Table 4.5	80
Table 5.1	108
Table 5.2	110
Table 5.3	111
Table 5.4	112
Table B.1	132
Table B.2	132
Table B.3	132
Table C.1	135

LIST OF FIGURES

Figure 4.1	75
Figure 4.2	76
Figure 5.1	109
Figure C.2	136

CONTRIBUTION OF AUTHORS

The three manuscripts included in this thesis are co-authored by the student (MK); her thesis co-supervisors, Drs Beth Potter and Pranesh Chakraborty; and the members of her thesis advisory committee, Drs Doug Coyle and Kumanan Wilson. The student is the primary author on all manuscripts and was responsible for conducting the literature reviews, planning the statistical methodology, completing the analysis, and drafting the manuscripts. The thesis co-supervisors and thesis advisory committee members provided guidance throughout the data analysis and manuscript writing process.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisor, Dr Beth Potter, for her endless guidance and encouragement throughout this endeavour. This achievement would not have been possible without your incredible support. I would also like to thank my co-supervisor, Dr Pranesh Chakraborty, for giving me the opportunity to experience the field of rare metabolic diseases and newborn screening first hand. Further, I would like to thank my thesis advisory members, Drs Kumanan Wilson and Doug Coyle, for providing invaluable insight throughout this thesis project. I would additionally like to mention the continuous support that I received from Steve Hawken and Robin Ducharme at the Institute for Clinical Evaluative Sciences.

Finally, a very special (and large) thank you is owed to my parents, friends, and Trent – I cannot express how fortunate I was to have such supportive people surrounding me during these past two and a half years. Please stick around for the many more years to come.

CHAPTER ONE: INTRODUCTION

SUMMARY OF RATIONALE FOR THESIS PROJECT

Despite well-developed provincial health services databases, there is a lack of information describing the health system impact of rare diseases in Canada. Medium-chain acyl-CoA dehydrogenase deficiency (MCADD) is one of the most prevalent of the rare diseases included in newborn bloodspot screening programs, with an estimated one in 14,000 affected infants born in Ontario each year (1). In order to evaluate the impact of screening and clinical management of MCADD on the health care system, this thesis project investigated health care services use and estimated its associated expenditures in a cohort of Ontario newborns diagnosed with MCADD following a positive newborn screening result during the first few years of life, relative to an unaffected screen negative birth cohort. Since sociodemographic and geographic characteristics are important sources of variation in patterns of health care use among both well and ill children in general, these associations were explored in children with MCADD in order to identify potential gaps in the provision of care (2-6).

Additionally, given the inconsistent findings with respect to early health care utilization among children with false positive newborn screening results, a false positive cohort for MCADD was examined to analyze the health system impact during the first few years of life relative to an unaffected screen negative birth cohort. Finally, since methods for estimating robust event rates and associated factors are not well established in rare disease populations, one of the initial steps in this project was to critically review the relevant methodological literature.

OBJECTIVES OF THESIS PROJECT

In order to thoroughly evaluate the health system impact of MCADD in Ontario, three thesis objectives have been outlined:

1) Conduct a review to identify the most appropriate methods for estimating robust event rates and associated factors for a rare disease population.

2) Describe the epidemiology and evaluate the health system impact of MCADD during the first few years of life among a population-based cohort of infants with MCADD relative to an unaffected birth cohort, using the guidelines developed from the initial objective where appropriate and possible. Specifically:

- a) Outline the patterns of health care services use (e.g., emergency department visits, inpatient hospitalizations, and physician visits) among affected children in order to gain an understanding of the associated burden of care.
- b) Investigate the associations between patterns of health care services use and sociodemographic/geographic characteristics so as to generate hypotheses highlighting potential gaps in care or issues related to access to services.
- c) Estimate the MCADD-attributable health care services expenditures for Ontario.

3) Evaluate the health system impact of a false positive screening result by studying patterns of health care use and costs in a population-based cohort of infants with false positive results for MCADD identified by newborn screening during the first few years of life relative to an unaffected birth cohort.

REFERENCES

1. 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 Pediatr*. 2010 Nov;10:82.
2. Sturm R, Ringel JS, Andreyeva T. Geographic disparities in children's mental health care. *Pediatrics*. 2003 Oct;112(4):e308.
3. Coker TR, Elliott MN, Kataoka S, Schwebel DC, Mrug S, Grunbaum JA, et al. Racial/Ethnic disparities in the mental health care utilization of fifth grade children. *Acad Pediatr*. 2009 Mar-Apr;9(2):89-96.
4. Fulda KG, Johnson KL, Hahn K, Lykens K. Do Unmet Needs Differ Geographically for Children with Special Health Care Needs? *Matern Child Health J*. 2013 Apr;17(3):505-511.
5. Ghandour RM, Kogan MD, Blumberg SJ, Jones JR, Perrin JM. Mental health conditions among school-aged children: geographic and sociodemographic patterns in prevalence and treatment. *J Dev Behav Pediatr*. 2012 Jan;33(1):42-54.
6. Kuhlthau KA. Measures of availability of health care services for children. *Acad Pediatr*. 2011 May-Jun;11(3 Suppl):S42-8.

CHAPTER TWO: BACKGROUND

From a health policy perspective, decision-makers aim to simultaneously balance three interconnected objectives in order to generate improvements within every sector of the health care system (1). This “triple aim” concept involves improving patient and family experiences with care, ameliorating clinical health outcomes of populations, and managing health system impacts (1). For patients affected with one of an estimated 7,000 rare diseases, achieving these goals is difficult due to challenges associated with correct and timely diagnosis, availability and access to treatment, potential high opportunity costs of treatments, and a sparsity of robust empirical clinical evidence to guide long-term patient treatment (2-5).

Inborn errors of metabolism (IEM), a collection of more than 400 rare single-gene diseases, represent one class of disorders within this large domain of rare diseases (6-8). IEM are heterogeneous, with an array of potential clinical manifestations, ranging from risk of acute illness to chronic multi-system ramifications, due to inactivity or deficiency within one or more cellular metabolic pathways. While individually uncommon (birth prevalence estimated at 1:10,000 to 1:1,000,000 depending on the disease), IEM adversely influence morbidity and mortality during childhood and collectively impact affected families (6, 9, 10). A significant burden is placed on population health as a result. For several of these disorders, pre-symptomatic diagnosis along with the timely implementation of effective disease management has been shown to prevent the most serious clinical manifestations (11-14).

MEDIUM-CHAIN ACYL-COA DEHYDROGENASE DEFICIENCY

Medium-chain acyl-CoA dehydrogenase deficiency (MCADD), a mitochondrial fatty acid oxidation disorder, is one of the most common IEM with an estimated birth prevalence ranging from one in 10,000 to one in 20,000 (10, 15-17). MCADD is a single gene disorder, identified in 1982, and has an autosomal recessive inheritance pattern. It is the result of deleterious mutations on the *ACADM* gene, which is located on chromosome 1p31. The most prevalent mutation in affected individuals is an adenine to guanine transition at coding position 985 (c.985A>G); this mutation results in a substitution of glutamate for lysine at position 304 in the mature protein, i.e., the enzyme medium-chain acyl-CoA dehydrogenase (17-19). In MCADD, this enzyme is deficient or has insufficient activity to effectively catalyze the initial step in the mitochondrial beta-oxidation pathway for medium-chain fatty acids. This step is critical towards the formation of ketone bodies in the liver, which are an essential alternative energy source from the breakdown of fat in the fasting state (17). This defect in the biochemical pathway thus hinders the body's ability to adequately respond in times of physiological stress, such as intercurrent illness and prolonged fasting (17, 20). During these periods, individuals with MCADD are susceptible to experience the clinical manifestations of the disorder such as a potentially lethal acute metabolic crisis. While symptoms are variable, typical presentation can include lethargy, vomiting, hypoketotic hypoglycemia, and hypotonia (10, 17, 19). As fatty acid toxic metabolites accumulate in the body, progression to encephalopathy and seizures can occur; the most severe cases result in coma or death (17, 19).

This episodic disorder is predominantly asymptomatic at birth; clinical presentation in the form of occurrences of metabolic decompensation is most likely to develop between three months and three years of age during times of acute viral illness or lengthy fasting (18). Affected infants without a definitive diagnosis have a 50 to 75% probability of experiencing an acute metabolic crisis due to the associated delay in delivering appropriate preventative treatment (13, 17). Without a disease management plan in place, these decompensation episodes are associated with a mortality rate of up to 25% in affected yet unscreened children (18, 21, 22). Factoring in those who have and have not been diagnosed clinically, it has been estimated that approximately five to seven percent of children with MCADD die by the age of 6 years in the absence of a newborn screening program (23). Further, up to one third of young children with MCADD who survive an initial metabolic crisis episode may display some form of developmental delay as a result (10, 17, 24). However, if early diagnosis is established, prognosis is excellent since the vast majority of adverse outcomes are preventable with the immediate initiation of treatment (17, 19, 20, 25).

Standard long-term therapy for MCADD, aimed at the prevention and management of catabolic stress, includes the avoidance of fasting and close medical monitoring during intercurrent illness (25-29). Emergency protocol dictates the immediate use of aggressive intervention with intravenous glucose along with hospitalization during times of illness and severe metabolic crises (17, 29, 30). Additional treatments, such as prescription L-carnitine supplementation or dietary fatty acids restriction, may be suggested in some instances (26, 31). Specific precautions with respect to routine childhood immunizations may also be

recommended (20, 26). Despite the implementation of newborn screening and early initiation of therapy, the risk of sudden mortality persists for children with MCADD, albeit at a lesser degree (32, 33).

NEWBORN SCREENING FOR MCADD

Evidence supporting effective treatment to prevent metabolic crises with associated morbidity and mortality risk reduction has prompted the inclusion of MCADD in population-based newborn screening programs throughout the world (34-36). A newborn screening system, used to identify newborns at risk of having a rare but treatable condition, has been in place in throughout Canada for several decades. However, since the respective programs are governed at the provincial and territorial level, considerable variation exists among the disorders included on the panels across the country (34, 37). In April 2006, Newborn Screening Ontario began screening for MCADD as part of the panel of screened disorders in Ontario (38); thus, to date, approximately 980,000 babies have been screened for the disorder in the province. Screening relies on the acquisition of newborn bloodspot samples from a heel prick ideally during the first 48 to 72 hours of life. In Ontario, the bloodspot samples are sent to Newborn Screening Ontario in Ottawa, ON for analysis. The biochemical manifestation of MCADD presents as elevated medium-chain acylcarnitines in the blood, in particular octanoylcarnitine (C8), due to the defect in the fatty acid beta-oxidation pathway. To screen for MCADD, tandem mass spectrometry (MS/MS) is used to quantitatively assay the relevant acylcarnitine metabolite levels in the newborn (39). This technology is additionally relied upon to simultaneously detect a panel of other IEM through the analysis of amino acids and additional acylcarnitine levels (39).

Marked variation with respect to the selected analytes, cut-offs, and algorithms used in screening for MCADD exists among newborn screening programs internationally (15, 40). In Ontario, the primary analyte is C8 and secondary analytes include hexanoylcarnitine (C6), decanoylcarnitine (C10), decenoylcarnitine (C10:1), and relevant ratios of these analytes (38). When a sample is flagged as a potential screen positive based on these analytes and associated algorithms, a biochemical geneticist reviews the acylcarnitine profile to definitively report a positive screening result (38). Newborns with abnormal screening results are referred to one of five Newborn Screening Regional Treatment Centres in Ontario (based at tertiary children's academic health sciences centres), which carry the responsibility of organizing confirmatory diagnostic testing along with follow-up care for those ultimately diagnosed with the condition. Diagnostic results are reported back to Newborn Screening Ontario through a Diagnostic Evaluation Reporting Form and tracked to ensure timely referral, retrieval, and diagnosis of each screen positive child in the province.

The basis of clinical diagnosis following a positive newborn screen for MCADD has evolved over time and differences in diagnostic criteria persist within and among jurisdictions (17, 38). While deficient enzymatic activity is noted as the traditional gold standard for diagnosis, metabolic geneticists currently rely on a variety of strategies including the identification of the common c.985A>G mutation. While the mutation spectrum for the disorder varies, homozygotes for the common mutation account for between 30 and 80% of those afflicted with the deficiency, with many of the remaining cases being compound heterozygotes, possessing the common mutation and a second

deleterious mutation (15-17, 31, 41). The frequency of homozygosity for the c.985A>G mutation is significantly higher in cases that clinically present with the disorder as compared to those who are ascertained through screening. This observation suggests that homozygosity for the common mutation may pose a higher risk of severe clinical phenotype, since screening identifies disease asymptotically and thus may be expected to identify cases that are milder and have a lower risk of metabolic decompensation. However this finding is not consistent in the literature (15, 42, 43). Further, homozygosity for the c.985A>G mutation is primarily identified in those of White ethnic origin (44, 45). A number of additional mutations of unknown clinical importance have been described including c.199T>C and c.127G>A (17, 19, 46). Thus, homozygosity for the common mutation fulfills the diagnostic criteria for MCADD in Ontario and many other jurisdictions. Infants whose molecular results show homozygosity for mutations other than c.985A>G or compound heterozygosity for the common mutation and a second mutation typically need to have additional evidence of disease to be classified as truly affected, including an abnormal plasma acylcarnitine profile and/or abnormal levels of urine organic acids (38).

THE EFFECT OF NEWBORN SCREENING PROGRAMS

Implementation of newborn screening for MCADD has resulted in the detection of 2 to 3 fold more cases of the disease as compared to reliance on clinical diagnosis and associated family studies alone (15, 23). This discrepancy likely reflects: (i) the identification by screening of some infants who would have remained asymptomatic for the disorder in the absence of screening (and thus would never be required to come to clinical attention); and (ii) missed diagnoses of the disorder in unscreened children

who truly are at risk (including infants and young children whose deaths from metabolic decompensation are not correctly attributed to MCADD in the absence of screening) (17, 23). Overall, there is reasonable evidence that newborn screening programs result in improved outcomes for the screened infant population with fewer deaths and serious disabilities reported within the first 6 years of life, after correcting for ascertainment bias in comparing screened versus unscreened populations (27, 47). Given the prevalence of the disorder, approximately 70,000 to 80,000 infants would need to be screened in order to prevent one case of death or significant disability (17). Early diagnosis by screening and proper management may also be associated with higher verbal, communication, and socialization skills in children as compared to those diagnosed clinically (48). Despite the notably small risk of death into adulthood, adults with MCADD must continue to adequately manage the disorder and understand the challenges associated with alcohol and drug consumption, weight reduction, pregnancy and delivery (28). However, long-term outcomes for a child with MCADD diagnosed through newborn screening are not yet well described due to its recent implementation (36).

False positive results

Despite the high specificities of newborn screening, the positive predictive values of such testing are inherently low due to the small prevalence of the targeted rare diseases (49). In the case of MCADD, the positive predictive value of a screen positive in Ontario is reported to be 46.3%, which fits into the range of 35-100% as reported by other programs internationally (15, 38). While MS/MS technology has supported the expansion of newborn screening panels, it has consequently also

resulted in an increased number of newborns presenting with false positive results. Attention has been directed towards describing the short- and long-term impact of false positive results in terms of parental psychosocial experiences and child health care utilization; however, research has yet to reach a definite consensus. Some studies have indicated that false positive screening results are related to higher levels of parental stress and anxiety, additional parental care demands and dysfunction in the parent-child relationship (35, 37, 50-53). While these psychosocial consequences are short-term for the majority of families, a small proportion of parents continue to exhibit residual anxiety following diagnostic testing (54, 55). In order to minimize the harms associated with newborn screening and false positive results, parental stress and anxiety can be mitigated with improved education and identifying best practices for effective provider communication during follow-up (54, 56).

Considerable variation exists regarding the impact of these psychosocial effects on patterns of early health care service use. While some studies demonstrate an increased frequency of emergency department (ED) visits and hospitalizations among children with false positive newborn screening results during the first few months of life as compared to a screen negative cohort, this finding was not consistent (49-52). If these increased rates are demonstrated to exist, they could be related to parental anxiety, parental perceptions of the child's health, or to actual health-related characteristics of the child (49, 57). For example, because of the nature of the biochemical markers used in the newborn screening test and underlying biology associated with illness and prematurity, false positive rates are significantly higher in newborns who are ill, born preterm, or have low birth weights. These three

characteristics are respectively associated with higher rates of health services use and thus must be considered in studies focusing on early health care utilization among false positive infants, before drawing conclusions about the psychosocial impact of the experience of receiving a false positive result (57-59).

Cost of newborn screening and MCADD

The cost of newborn screening includes the actual screening test process, other direct costs (e.g., diagnostic follow-up testing, false positives, treatment as necessary, hospitalizations), indirect costs (e.g., parental absence from work), and intangible costs (e.g., parental quality of life) (54, 60). Numerous analyses demonstrate that newborn screening for IEM is cost-effective with ratios comparable to other preventative health interventions (55, 56, 60-62). According to the universal screening strategy, the cost of screening is approximately US\$4.00 per newborn screened in the United States and between C\$0.50 and C\$5.60 per newborn screened in Canada depending on the province or territory (55, 62). In the Canadian context, interventions are typically considered cost-effective if the incremental cost-effectiveness ratio (ICER) is less than C\$50,000 per quality-adjusted life-year (QALY) in general (63). When the cost for screening was set as C\$2.40 per newborn screened in a base-case scenario, the estimated ICER for MCADD screening was C\$2,676 per QALY when compared to no screening over a 77-year horizon; the ICER remained under the C\$20,000 threshold even if the cost of screening increased to C\$5.60 per newborn screened (55). Over a prospective 90-year horizon, the ICER of newborn screening for MCADD in the United States was estimated at US\$27,423 per QALY when quality of life effects following a false positive result and lifetime dietary treatment were additionally included in the

analysis (60). The projected total cost related to screening and treatment of MCADD is similar to that for phenylketonuria at an estimated \$74,200 per newborn screened detected case over a 70-year horizon (62).

THE IMPORTANCE OF MEASURING HEALTH SERVICES IMPACTS

Caring for children with a range of special health care needs, including those with IEM can be challenging and costly both for families and the health care system. For example, in the United States, these patients experience approximately 4 times the number of hospitalizations, 7 times as many hospital days, twice as many physician visits, and 1.5 times as many ED visits as children without such needs (64). This considerable increase in utilization resulted in health care expenditures that were about 3 times higher for these cases as compared to healthy controls (64). A further US study demonstrated that children with chronic illnesses accessed medical services at a rate 3 to 20 fold higher than that of healthy children, depending on the illness (65). When a child is diagnosed with a more serious disability, total health care expenditures are approximately 4 times higher than that in healthy children (66). These patients experience 4 times the number of inpatient hospitalizations, 8 times as many hospital days, twice as many physician visits, and twice as many ED visits relative to a healthy control cohort in the United States (66).

Limited research has been directed towards characterizing health service utilization and its associated expenditures in the rare disease population and particularly for IEM. Patients with rare diseases encounter notable obstacles in accessing high quality health care services due in part to inadequate knowledge among

health professionals and policy decision-makers regarding the specific health-related needs of these individuals and deficiencies in evidence to support treatment effectiveness. These difficulties faced by the rare disease population include lack of access to timely diagnosis, lack of quality information and support at the time of diagnosis, inequities and struggles relating to access to treatment, rehabilitation and care, and challenges in accessing orphan drugs (2). Similar to MCADD, sickle cell disease is a rare disease characterized by an acute crisis that relies on urgent care provided within the ED (67). As compared to other patient groups, sickle cell disease patients experience higher levels of hospitalization and ED visits (67-69). Further, children diagnosed with IEM that were admitted to the neonatal or pediatric intensive care unit due to associated clinical manifestations required more aggressive support (e.g., mechanical ventilation, extracorporeal removal therapy) and had a higher mortality rate as compared to unaffected admitted patients (70-72).

REFERENCES

1. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. *Health Aff (Millwood)*. 2008 May-Jun;27(3):759-69.
2. Posada de la Paz M, Groft SC, editors. *Rare Disease Epidemiology*. New York: Springer; 2010.
3. Making rare diseases a public-health and research priority. *Lancet*. 2008 Jun;371(9629):1972.
4. Schieppati A, Henter JI, Daina E, Aperia A. Why rare diseases are an important medical and social issue. *Lancet*. 2008 Jun;371(9629):2039-41.
5. Potter BK, Chakraborty P, Kronick JB, Wilson K, Coyle D, Feigenbaum A, et al. Achieving the "triple aim" for inborn errors of metabolism: a review of challenges to outcomes research and presentation of a new practice-based evidence framework. *Genet Med*. 2013 Jun;15(6):415-22.
6. Ellaway CJ, Wilcken B, Christodoulou J. Clinical approach to inborn errors of metabolism presenting in the newborn period. *J Paediatr Child Health*. 2002 Oct;38(5):511-7.
7. Levy HL, Albers S. Genetic screening of newborns. *Annu Rev Genomics Hum Genet*. 2000;1:139-77.
8. Garrod SAE, Harris H. *Inborn errors of metabolism*: Oxford University Press; 1963.
9. Burton BK. Inborn errors of metabolism in infancy: a guide to diagnosis. *Pediatrics*. 1998 Dec;102(6):e69-e.
10. Wilcken B, Hammond J, Silink M. Morbidity and mortality in medium chain acyl coenzyme A dehydrogenase deficiency. *Arch Dis Child*. 1994 May;70(5):410-2.
11. Seymour CA, Thomason MJ, Chalmers RA, Addison GM, Bain MD, Cockburn F, et al. Newborn screening for inborn errors of metabolism: a systematic review. *Health Technol Assess*. 1997;1(11):i-iv, 1-95.

12. Schulze A, Lindner M, Kohlmüller D, Olgemöller K, Mayatepek E, Hoffmann GF. Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics*. 2003 Jun;111(6):1399-406.
13. Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV, et al. Neonatal screening for inborn errors of metabolism: cost, yield and outcome. *Health Technol Assess*. 1997;1(7):i-iv, 1-202.
14. Pandor A, Eastham J, Beverley C, Chilcott J, Paisley S. Clinical effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review. *Health Technol Assess*. 2004 Mar;8(12):iii, 1-121.
15. Rhead WJ. Newborn screening for medium-chain acyl-CoA dehydrogenase deficiency: a global perspective. *J Inherit Metab Dis*. 2006 Apr-Jun;29(2-3):370-7.
16. Oerton J, Khalid JM, Besley G, Dalton RN, Downing M, Green A, et al. Newborn screening for medium chain acyl-CoA dehydrogenase deficiency in England: prevalence, predictive value and test validity based on 1.5 million screened babies. *J Med Screen*. 2011;18(4):173-81.
17. Grosse SD, Khoury MJ, Greene CL, Crider KS, Pollitt RJ. The epidemiology of medium chain acyl-CoA dehydrogenase deficiency: an update. *Genet Med*. 2006 Apr;8(4):205-12.
18. Wang SS, Fernhoff PM, Hannon WH, Khoury MJ. Medium chain acyl-CoA dehydrogenase deficiency human genome epidemiology review. *Genet Med*. 1999 Nov-Dec;1(7):332-9.
19. 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.
20. Ziadeh R, Hoffman EP, Finegold DN, Hoop RC, Brackett JC, Strauss AW, et al. Medium chain acyl-CoA dehydrogenase deficiency in Pennsylvania: neonatal screening shows high incidence and unexpected mutation frequencies. *Pediatr Res.* 1995 May;37(5):675-8.
 21. Wilcken B, Carpenter K, Wiley V. Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency. *Lancet.* 2002 Feb;359(9306):627-8.
 22. Pollitt RJ, Leonard JV. Prospective surveillance study of medium chain acyl-CoA dehydrogenase deficiency in the UK. *Arch Dis Child.* 1998 Aug;79(2):116-9.
 23. Wilcken B. Fatty acid oxidation disorders: outcome and long-term prognosis. *J Inherit Metab Dis.* 2010 Oct;33(5):501-6.
 24. Iafolla AK, Thompson RJ, Roe CR. Medium-chain acyl-coenzyme A dehydrogenase deficiency: clinical course in 120 affected children. *J Pediatr.* 1994 Mar;124(3):409-15.
 25. Leonard JV, Dezateux C. Newborn screening for medium chain acyl CoA dehydrogenase deficiency. *Arch Dis Child.* 2009 Mar;94(3):235-8.
 26. Potter BK, Little J, Chakraborty P, Kronick JB, Evans J, Frei J, et al. Variability in the clinical management of fatty acid oxidation disorders: results of a survey of Canadian metabolic physicians. *J Inherit Metab Dis.* 2012 Jan;35(1):115-23.
 27. Wilcken B, Haas M, Joy P, Wiley V, Chaplin M, Black C, et al. Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency in Australia: a cohort study. *Lancet.* 2007 Jan;369(9555):37-42.
 28. Schatz UA, Ensenauer R. The clinical manifestation of MCAD deficiency: challenges towards adulthood in the screened population. *J Inherit Metab Dis.* 2010 Oct;33(5):513-20.
 29. Dixon MA, Leonard JV. Intercurrent illness in inborn errors of intermediary metabolism. *Arch Dis Child.* 1992 Nov;67(11):1387-91.

30. Waisbren SE. Expanded newborn screening: information and resources for the family physician. *Am Fam Physician*. 2008 Apr;77(7):987-94.
31. Couce ML, Sanchez-Pintos P, Diogo L, Leao-Teles E, Martins E, Santos H, et al. Newborn screening for medium-chain acyl-CoA dehydrogenase deficiency: regional experience and high incidence of carnitine deficiency. *Orphanet J Rare Dis*. 2013 Jul;8:102.
32. Yusupov R, Finegold DN, Naylor EW, Sahai I, Waisbren S, Levy HL. Sudden death in medium chain acyl-coenzyme a dehydrogenase deficiency (MCADD) despite newborn screening. *Molecular genetics and metabolism*. 2010 Sep;101(1):33-9.
33. Nennstiel-Ratzel U, Arenz S, Maier EM, Knerr I, Baumkotter J, Roschinger W, et al. Reduced incidence of severe metabolic crisis or death in children with medium chain acyl-CoA dehydrogenase deficiency homozygous for c.985A>G identified by neonatal screening. *Mol Genet Metab*. 2005 Jun;85(2):157-9.
34. Health Quality Ontario. Neonatal screening of inborn errors of metabolism using tandem mass spectrometry: an evidence-based analysis. *Ont Health Technol Assess Ser*. 2003;3(3):1-36.
35. Wilcken B. Screening for disease in the newborn: the evidence base for blood-spot screening. *Pathology*. 2012 Feb;44(2):73-9.
36. Dezateux C. Newborn screening for medium chain acyl-CoA dehydrogenase deficiency: evaluating the effects on outcome. *Eur J Pediatr*. 2003 Dec;162 Suppl 1:S25-8.
37. Morrison DR, Clayton EW. False positive newborn screening results are not always benign. *Public Health Genomics*. 2011;14(3):173-7.
38. 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 Pediatr*. 2010 Nov;10:82.
39. Levy HL. Newborn screening by tandem mass spectrometry: a new era. *Clin Chem*. 1998 Dec;44(12):2401-2.

40. McHugh D, Cameron CA, Abdenur JE, Abdulrahman M, Adair O, Al Nuaimi SA, et al. Clinical validation of cutoff target ranges in newborn screening of metabolic disorders by tandem mass spectrometry: a worldwide collaborative project. *Genet Med.* 2011 Mar;13(3):230-54.
41. Tanaka K, Gregersen N, Ribes A, Kim J, Kolvraa S, Winter V, et al. A survey of the newborn populations in Belgium, Germany, Poland, Czech Republic, Hungary, Bulgaria, Spain, Turkey, and Japan for the G985 variant allele with haplotype analysis at the medium chain Acyl-CoA dehydrogenase gene locus: clinical and evolutionary consideration. *Pediatr Res.* 1997 Feb;41(2):201-9.
42. 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.
43. Arnold GL, Saavedra-Matiz CA, Galvin-Parton PA, Erbe R, Devincentis E, Kronn D, et al. Lack of genotype-phenotype correlations and outcome in MCAD deficiency diagnosed by newborn screening in New York State. *Mol Genet Metab.* 2010 Mar;99(3):263-8.
44. Gregersen N, Winter V, Curtis D, Deufel T, Mack M, Hendrickx J, et al. Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency: the prevalent mutation G985 (K304E) is subject to a strong founder effect from northwestern Europe. *Hum Hered.* 1993 Nov-Dec;43(6):342-50.
45. Khalid JM, Oerton J, Cortina-Borja M, Andresen BS, Besley G, Dalton RN, et al. Ethnicity of children with homozygous c.985A>G medium-chain acyl-CoA dehydrogenase deficiency: findings from screening approximately 1.1 million newborn infants. *J Medical Screen.* 2008;15(3):112-7.
46. 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;7(9):e45110.

47. Wilcken B, Haas M, Joy P, Wiley V, Bowling F, Carpenter K, et al. Expanded newborn screening: outcome in screened and unscreened patients at age 6 years. *Pediatrics*. 2009 Aug;124(2):e241-8.
48. Joy P, Black C, Rocca A, Haas M, Wilcken B. Neuropsychological functioning in children with medium chain acyl coenzyme a dehydrogenase deficiency (MCADD): the impact of early diagnosis and screening on outcome. *Child Neuropsychol*. 2009 Jan;15(1):8-20.
49. Lipstein EA, Perrin JM, Waisbren SE, Prosser LA. Impact of false-positive newborn metabolic screening results on early health care utilization. *Genet Med*. 2009 Oct;11(10):716-21.
50. 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.
51. Waisbren SE, Albers S, Amato S, Ampola M, Brewster TG, Demmer L, et al. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA*. 2003 Nov 19;290(19):2564-72.
52. Tu W-J, He J, Chen H, Shi X-D, Li Y. Psychological Effects of False-Positive Results in Expanded Newborn Screening in China. *PloS One*. 2012;7(4):e36235.
53. Tluczek A, Orland KM, Cavanagh L. Psychosocial consequences of false-positive newborn screens for cystic fibrosis. *Qual Health Res*. 2011 Feb;21(2):174-86.
54. Wildner M. Health economic issues of screening programmes. *Eur J Pediatr*. 2003 Dec;162(1):S5-S7.
55. Tran K, Banerjee S, Li H, Noorani HZ, Mensinkai S, Dooley K. Clinical efficacy and cost-effectiveness of newborn screening for medium chain acyl-CoA dehydrogenase deficiency using tandem mass spectrometry. *Clin Biochem*. 2007 Feb;40(3-4):235-41.

56. Hamers FF, Rumeau-Pichon C. Cost-effectiveness analysis of universal newborn screening for medium chain acyl-CoA dehydrogenase deficiency in France. *BMC Pediatr.* 2012 Jun;12:60.
57. 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.
58. McCandless SE, Chandrasekar R, Linard S, Kikano S, Rice L. Sequencing from dried blood spots in infants with "false positive" newborn screen for MCAD deficiency. *Mol Genet Metab.* 2013 Jan;108(1):51-5.
59. Tarini BA, Clark SJ, Pilli S, Dombkowski KJ, Korzeniewski SJ, Gebremariam A, et al. False-positive newborn screening result and future health care use in a state Medicaid cohort. *Pediatrics.* 2011 Oct;128(4):715-22.
60. Prosser LA, Kong CY, Rusinak D, Waisbren SL. Projected costs, risks, and benefits of expanded newborn screening for MCADD. *Pediatrics.* 2010 Feb;125(2):e286-94.
61. Norman R, Haas M, Chaplin M, Joy P, Wilcken B. Economic evaluation of tandem mass spectrometry newborn screening in Australia. *Pediatrics.* 2009 Feb;123(2):451-7.
62. Venditti LN, Venditti CP, Berry GT, Kaplan PB, Kaye EM, Glick H, et al. Newborn screening by tandem mass spectrometry for medium-chain Acyl-CoA dehydrogenase deficiency: a cost-effectiveness analysis. *Pediatrics.* 2003 Nov;112(5):1005-15.
63. Winquist E, Bell CM, Clarke JT, Evans G, Martin J, Sabharwal M, et al. An evaluation framework for funding drugs for rare diseases. *Value Health.* 2012 Sep-Oct;15(6):982-6.
64. Newacheck PW, Kim SE. A national profile of health care utilization and expenditures for children with special health care needs. *Arch Pediatr Adolesc Med.* 2005 Jan;159(1):10-7.

65. Shatin D, Levin R, Ireys HT, Haller V. Health care utilization by children with chronic illnesses: a comparison of medicaid and employer-insured managed care. *Pediatrics*. 1998 Oct;102(4):E44.
66. Newacheck PW, Inkelas M, Kim SE. Health services use and health care expenditures for children with disabilities. *Pediatrics*. 2004 Jul;114(1):79-85.
67. Shankar SM, Arbogast PG, Mitchel E, Cooper WO, Wang WC, Griffin MR. Medical care utilization and mortality in sickle cell disease: a population-based study. *Am J Hematol*. 2005 Dec;80(4):262-70.
68. Lanzkron S, Carroll CP, Haywood C, Jr. The burden of emergency department use for sickle-cell disease: an analysis of the national emergency department sample database. *Am J Hematol*. 2010 Oct;85(10):797-9.
69. Yusuf HR, Atrash HK, Grosse SD, Parker CS, Grant AM. Emergency department visits made by patients with sickle cell disease: a descriptive study, 1999-2007. *Am J Prev Med*. 2010 Apr;38(4 Suppl):S536-41.
70. Meng M, Zhang Y-P. Impact of inborn errors of metabolism on admission in a neonatal intensive care unit: a 4-year report. *J Pediatr Endocrinol Metab*. 2013;26(7-8):689-93.
71. Couce ML, Bana A, Boveda MD, Perez-Munuzuri A, Fernandez-Lorenzo JR, Fraga JM. Inborn errors of metabolism in a neonatology unit: Impact and long-term results. *Pediatr Int*. 2011 Feb;53(1):13-7.
72. Jouvét P, Touati G, Lesage F, Dupic L, Tucci M, Saudubray JM, et al. Impact of inborn errors of metabolism on admission and mortality in a pediatric intensive care unit. *Eur J Pediatr*. 2007 May;166(5):461-5.

INTERFACE

Analyses focused on small populations present unique challenges to researchers with regards to the reliability and stability of estimates, which may lead to a misinterpretation of the results. However, there is limited literature that describes and compares the potential statistical methods through which these issues can be addressed. The following manuscript reviews approaches for estimating robust event rates and associated factors with small population sizes, which includes examples from the field of rare diseases. The primary objective of this review was to identify the most appropriate method that could be used to estimate stable health event rates and investigate its feasibility in a subsequent portion of this project. Two categories of methods were presented for analyzing robust event rates in small populations: 1) aggregation; and 2) data smoothing techniques. Using hypothetical examples within a rare disease context, the manuscript focused on describing the methodology behind the proposed approaches and the types of health services analyses that could be completed using each method.

This manuscript was co-authored by the student (MK); her thesis co-supervisors, Drs Beth Potter and Pranesh Chakraborty; the members of her thesis advisory committee, Drs Doug Coyle and Kumanan Wilson; and a biostatistician, Steven Hawken. The student was responsible for conducting the literature review, developing examples within the rare disease context, and drafting the manuscript; the thesis co-supervisors, thesis advisory committee members, and Steven Hawken provided guidance throughout the review process and gave feedback on manuscript drafts.

CHAPTER THREE: METHODS REVIEW – MANUSCRIPT #1

The “small number” problem: a methods review for improving data reliability in health services research with small populations, using examples from the field of rare inherited metabolic diseases

Maria D Karaceper, BSc,¹; Pranesh Chakraborty, MD, FRCPC,²; Doug Coyle, MSc, PhD,¹; Steven Hawken, MSc,³; Kumanan Wilson, MD, FRCPC,⁴; Beth K Potter, MSc, PhD,¹; and in collaboration with the Canadian Inherited Metabolic Diseases Research Network

Affiliation:

1 Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada

2 Newborn Screening Ontario, Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada

3 Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada

4 Ottawa Hospital Research Institute and Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada

Keywords: health services research, rare diseases, confidentiality, epidemiologic methods, small-area analysis

Manuscript Status: To be submitted to the Journal of Clinical Epidemiology

ABSTRACT

Objective: Health services research focused on small populations presents unique statistical challenges regarding robust estimation of event rates. We present an overview of potential methodological approaches to addressing these challenges within a rare disease context.

Study Design and Setting: We identified techniques described in the literature that could be used to increase the statistical reliability of estimates based on small numbers, emphasizing their relevance to estimating rates of health services use. We critically evaluated each approach with reference to two rare inherited metabolic diseases used as examples.

Results: Two general approaches were identified: 1) aggregation, such as expanding geographic areas and using composite outcomes; and 2) data smoothing techniques, including weighted averages and Bayesian estimation methods. Empirical comparisons of these methods were limited in the literature. We summarized the advantages and disadvantages associated with each approach. Although there is no current gold standard, we suggest that investigators consider weighted moving averages and Bayesian methods as the most promising among the strategies discussed.

Conclusion: Future research aimed at estimating event rates with small populations can use this review as a basis from which to select the most appropriate approach to improve statistical reliability.

INTRODUCTION

Estimating robust rates of health events and studying the association between risk factors and health outcomes can be problematic when dealing with small populations. These small patient subgroups can be encountered in rare disease settings or in the emerging area of personalized or stratified medicine, where clinical, physiologic, or molecular markers are used to create small groups of patients for whom disease characteristics and treatment strategies are expected to be similar (1, 2). One important issue faced by researchers studying small populations is the need to protect the privacy and confidentiality of health information: when dealing with few individuals, there is a risk of inadvertently disclosing identifying information. This concern has been addressed to some extent by legislation and by privacy policies established by organizations responsible for maintaining and analyzing large health surveys and health care administrative datasets. However less attention has been directed towards a second area of concern, namely the statistical issues involved with the analysis of data based on small numbers of individuals. Specifically, estimates of event rates from these analyses, such as rates of health services use, are often unstable and may result in misinterpretation of the results.

There are, to our knowledge, only a few published guidelines that focus on addressing these reliability concerns in cases where cell sizes are small or the underlying population is small (3-5). These guidelines and further suggestions from the literature generally focus on statistical approaches that modify the existing data to improve the stability of the estimates. Proposed approaches include aggregation, weighted averages and Bayesian estimation methods. This article summarizes these

recommendations within a rare disease context with reference to hypothetical analyses of rates of health services use for example disorders. Throughout the paper, we assume that the outcome of interest is the estimate of the rate of one or more dichotomous events (e.g., hospitalizations).

As examples we have chosen two inborn errors of metabolism (IEM), Phenylketonuria (PKU) and Medium-chain acyl-CoA dehydrogenase deficiency (MCADD). IEM are a heterogeneous group of >400 rare genetic disorders which exemplify these methodological issues surrounding small populations; a research program focused on these conditions will provide an opportunity to directly investigate the approaches reviewed here (6). The discussed methods have the potential to be further applied to other situations where this “small number” problem is commonly encountered.

EXAMPLE DISEASES

IEM result in an array of detrimental clinical consequences, ranging from risk of acute illness to chronic multi-system ramifications, due to inactivity or deficiency within one or more cellular metabolic pathways. While individually uncommon (estimated birth prevalences of 1:10,000 to 1:1,000,000 depending on the specific disorder), IEM adversely influence morbidity and mortality during childhood and collectively place a significant burden on population health as a result (7). This review will rely, as examples, on two of the more prevalent IEM, PKU and MCADD, for which studies of rates of health services use are of interest.

Example disease #1: Phenylketonuria (PKU)

PKU is an autosomal recessive amino acid disorder characterized by a deficiency of the liver enzyme phenylalanine hydroxylase. It has a prevalence of about 1 in 15,000 births worldwide (8). Based on blood phenylalanine (Phe) levels, a spectrum of disorders including classic PKU, moderate PKU, mild PKU and mild hyperphenylalaninemia can be diagnosed through newborn screening (9). Early dietary intervention (low-protein diet with Phe-restriction and Phe-free medical formula) is necessary in order to prevent the major neurological sequelae associated with the disorder, such as microcephaly, seizures, and intellectual impairment (10-12). Despite lifelong monitoring and management of blood Phe levels, deficits in executive function may still develop in affected individuals (13, 14). As a result, a patient with PKU may heavily rely on certain health care services such as close physician care during pregnancy, genetic counseling, and nutritional guidance (8, 15). Hence, investigation of rates of health services use and rates of adverse health outcomes is of interest in cohort studies focused on PKU patient populations in order to evaluate the impact of screening and clinical management of PKU on the healthcare system and to identify potential gaps that could be addressed to improve delivery of care to this population.

Example disease #2: Medium-chain acyl-CoA dehydrogenase deficiency (MCADD)

MCADD, an autosomal recessive mitochondrial fatty acid oxidation disorder, is one of the most common IEM with an estimated birth prevalence between one in 10,000 to one in 20,000 (16-19). It is characterized by the inability to form ketone bodies, which are an essential alternative energy source in the fasting state. Typical

clinical manifestations include risk of acute metabolic crisis during times of increased metabolic demands or reduced dietary intake; severe complications include coma or death (19, 20). If pre-symptomatic diagnosis is established via newborn screening, prognosis is excellent given that the vast majority of adverse outcomes are preventable with early intervention (21). Standard therapy, aimed at the prevention and management of catabolic stress, includes the avoidance of fasting and close medical monitoring during intercurrent illness (21-24). As a result, a patient with MCADD relies on urgent care provided within the emergency department (ED) where aggressive intervention with intravenous glucose can be administered. By investigating rates of ED visits and of metabolic crises, delivery and quality of care for this patient population can be improved by understanding the impact of MCADD on both families (i.e., psychosocial implications) and health systems (i.e., costs of care).

METHODS FOR ROBUST ESTIMATION OF EVENT RATES IN SMALL POPULATIONS

Aggregation

Aggregation techniques result in larger cell sizes by combining multiple years of data, collapsing data categories, or expanding the geographic area (3, 4, 25). The approach seems to be a particularly useful strategy in situations where an investigator would otherwise need to suppress the data within cells to meet privacy standards. Despite improving the reliability of estimates, the application of these methods introduces a significant tradeoff: the loss of detail. As a result, there is an inherent risk of masking the differences between the characteristics of combined groups or changes over time (26, 27).

Example #1: Combining multiple years of data, with important limitations

As an example of the limitations of this approach, if one is interested in estimating the rate of ED visits over the first few years of life in a cohort of patients with MCADD born over a 10-year period, it may be tempting to aggregate data across children born in all 10 years. However, this has the potential to mask important time trends. Specifically, MCADD has been added to newborn screening panels recently (e.g., in 2006 in Ontario, Canada), so this cohort would combine two distinct populations: those who were clinically diagnosed prior to MCADD screening and those who were identified through the newborn screening program. It has been established that the clinical spectrum of MCADD is broader in screened versus unscreened populations, in part because some children diagnosed through newborn screening likely would have remained asymptomatic (and undiagnosed) in the absence of screening (21). As a result, younger, screen-diagnosed children may be expected to have lower rates of ED visits as compared to older, clinically diagnosed children. If we aggregated these birth cohorts into a 10-year group, this time trend would be masked.

Example #2: Combining multiple years of data, with more minor limitations

The above approach of aggregating across birth cohorts may be more appropriate for a disease such as PKU, for which newborn screening dates back to the 1960s. For example, given that children in the general population have the highest rate of ED visits when they are less than one year of age, an investigator may wish to investigate ED utilization patterns in the first year of life among children with PKU. In order to increase the sample size, the researcher could combine consecutive birth

cohorts to form a larger cohort that represents those diagnosed with PKU over a 15-year span. Although to the best of our knowledge there is no seminal event (such as initiation of a population-wide screening program) that would render this approach inappropriate for PKU, changes over time due to changes in care provision for PKU would be obscured (27).

Example #3: Collapsing data categories

If the number of affected cases is not sufficient to yield stable estimates for a specific disorder and/or collapsing across time is deemed too limiting, a second form of aggregation would be to group multiple disorders together, based on their pathophysiology. For example, MCADD could be grouped with other IEM characterized as fatty acid oxidation disorders, such as long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency and very long-chain acyl-CoA dehydrogenase deficiency (28). We could additionally group those with confirmed diagnoses for ‘variant’ types of disease to those with the ‘classical’ form of the targeted condition to increase the cohort size. An example of this approach would be to create a PKU study cohort that includes those with classic PKU, moderate PKU, mild PKU, and mild hyperphenylalaninemia. With this technique, the differences between or among the combined groups will be masked. The appropriateness of aggregation thus relies on whether the underlying disease characteristics that are related to the outcome of interest are stable across the categories being combined; this may be seen as analogous to the situation where an investigator must consider whether it is appropriate to combine data from individual studies into a meta-analysis. In the scenario we consider, a formal test for heterogeneity (often used to inform data combining in a meta-analysis) may be

challenging because of low statistical power inherent to analyses with small numbers of patients.

Example #4: Expanding the geographic area

Given the low prevalence estimates within an IEM context, an investigator may additionally consider expanding the scope of the analysis to include a larger geographic area in order to increase the sample size. For example, children with PKU or MCADD from various states or provinces may be grouped to form one large cohort. Important limitations of this method include loss of detail (e.g., it will be difficult to characterize the geographical variation among the combined regions) and privacy barriers. In particular, policies and legislation governing the protection of health information, including information derived from healthcare administrative data, differ across jurisdictions and present challenges to multi-jurisdictional aggregation of patient data. This is a key challenge in the field of rare disease research, where multi-centre and international collaboration would be highly beneficial.

Example #5: Creating composite definitions

For rare diseases with prevalence estimates near 1 in 1,000,000, an investigator may be unable to rely upon the above aggregation techniques to achieve the desired increase in statistical reliability. In this situation, the expansion of event definitions to create a composite outcome may be a suitable alternative (29, 30). For example, if an investigator wanted to evaluate health services use in a given rare disease cohort relative to an unaffected cohort, this resource use could be defined to include multiple services such as ED visits, repeat physician visits, hospitalizations, and diagnostic tests

within a fixed time frame. Careful consideration should be placed on what the investigator wants the data to represent (i.e., examining disease complexity vs acute events) and specific services should be selected and amalgamated accordingly. While this technique would increase the number of events for the given analysis, the specificity of the services would be lost. As a result, the method could yield misleading conclusions if the individual services were not comparable in impact, depending on the metric used (e.g., the total number of “events”, or their cost) and the specific research question.

Data Smoothing Techniques

Data smoothing is typically used to improve the reliability of data by eliminating small fluctuations, or “noise” stemming from random errors (31, 32). This smoothing process relies on integrating data from other sources, such as adjacent geographic areas and time periods, in order to stabilize estimates that are subject to random error (33, 34). For IEM and other rare diseases, these comparative data are often not readily available due to the sparse literature focusing on these populations and on disorder-related outcomes. However, as registries and cohort studies continue to develop, and given the above-mentioned challenges in directly combining data across jurisdictions, data smoothing approaches hold promise for future research within the field of rare diseases. We outline two data smoothing techniques using the example IEM: weighted averages and Bayesian estimation methods.

Example #1: Weighted averages

A weighted average can be calculated by averaging event rates from “reference” geographic regions adjacent to the area of interest, thereby stabilizing and smoothing the desired rate. Each rate is consequently weighted according to its proximity to the region of interest (33). This spatial smoothing technique, a form of “small-area analysis”, is commonly applied to analyze data and calculate event rates in regions with small populations (35, 36). In the context of MCADD, if an investigator wanted to estimate the annual ED visit rate for children with MCADD in a specific region, annual utilization data on affected children from neighboring regions could be incorporated in order to calculate a stable rate despite the small patient populations in each of these regions by “borrowing strength” from neighboring regions. Utilization rates from the closest regions would be given the greatest weights, while those that are further away would be assigned smaller weights according to the distance from the geographic area of interest.

Observations can similarly be smoothed over a period of time through the calculation of temporal exponentially smoothed moving averages. This weighted smoothing technique similarly places weights on selected data points; in this approach, the greatest weight would be assigned to the most recent data points (31, 34). Consider an investigator who wishes to estimate the annual physician visit rate for pregnant women with PKU in a given region. As opposed to relying on data from adjacent geographic areas, a stable annual event rate could be estimated by simply incorporating utilization data from multiple past and future years. The most recent year of data would be assigned the greatest weight, while earlier and later years would be given

exponentially smaller weights. In contrast to aggregation methods, area-level and time-trend differences can be ascertained with these smoothing approaches (36, 37). In addition, weighted averages are not computationally difficult to use if the neighboring rates are readily available, which makes it an appealing solution for researchers dealing with small populations. Calculating confidence intervals around descriptive rates that have been smoothed using a weighted averages approach is not straightforward, however, because the individual rates are not statistically independent. More sophisticated time series methods have been developed for modeling trend data but are beyond the scope of discussion here (38).

Example #2: Bayesian Estimation Methods

Bayesian estimation methods combine external information “prior knowledge” with the observed data, to generate a posterior distribution from which parameters of interest can be estimated for a subgroup of interest. This approach combines the available data from the target subgroup with information from external sources in order to derive the desired stabilized rates for the subgroup. Study subgroup counts are masked as a result such that the approach serves to increase the robustness of the estimate while also protecting data privacy (5, 36). Under empirical Bayesian principles, the value of using data from an external source to produce posterior distributions when limited subgroup data are available relies on the assumption that the external data provide an equal or better representation of the subgroup population as compared to the subgroup data alone (3, 36). For example, if an investigator wanted to estimate the annual physician visit rate among pregnant women with PKU in a particular geographic area, the posterior distribution could be estimated by using the

data from the target area, augmented by data from neighboring geographic regions.

Hierarchical Bayesian models can similarly be used to analyze small area rates in which prior distributions are based on hyperprior distributions and fixed a priori (40).

Both empirical and hierarchical Bayesian methods can thus be used to smooth rates of health services use. A full description of the statistical approaches involved is beyond the scope of this paper but can be found in several published overviews (36, 39, 40). Conceptually, the empirical Bayes approach is a two-step process, in which weights are applied to the sample rates as a first step. In the PKU example, the magnitude of the weight will be dependent on the size of the affected population in the city; hence, given the limited number of affected individuals, the corresponding weight may be small. However, if the investigator wanted to estimate the annual rate of physician visits on a larger scale (e.g., state- or province-wide), the corresponding weight would increase accordingly. In a second step, the physician visit rate for the city will then be stabilized by “borrowing” physician visit data from other sources on a local- (from surrounding areas) or global-scale (from the whole state or province). This would ultimately result in a weighted rate that has been pushed, or “shrunk”, towards the mean of neighboring values, with the extent of the shrinkage depending on the size of the population of interest (i.e., a city with a smaller numbers of individuals with PKU would have a smoothed visit rate that is more similar to the global or local external data, whereas the smoothed rate for a city with a larger number of affected individuals would be more similar to that city’s observed rate) (32). To ensure that the stabilized rate for the geographic area is accurate, the investigator may refer to rates from previous years as a means for comparison; this is known as a spatial time series analysis. Empirical

Bayesian estimators can additionally be spatial in nature, such that information is only imputed from local sources as opposed to oversmoothing the estimates with globally-derived source data and losing important local variation (32).

Similar to the previously described data smoothing techniques, there may be limited sources from which to impute relevant information needed to apply the Bayesian method at this point in time, for IEM and other rare diseases. However, empirical and hierarchical Bayesian methodology has been used extensively in healthcare system evaluation research and for mapping of disease rates in small geographic regions (40-44), and is thus a potentially promising approach for future rare disease research with current development of larger cohorts and registries.

DISCUSSION

Although we have reviewed a number of promising approaches for addressing the ‘small numbers’ problem in generating robust estimates of rates of health events, to date there is no gold standard, with limited comparative evidence to guide the selection of an approach. Table 3.1 summarizes the described techniques along with their respective advantages and disadvantages. It seems likely that the most suitable method will depend on the purpose of the analysis, the nature of available data, and the audience for whom the results must be interpretable.

Of the presented methods, aggregation is the simplest and most intuitive approach to improve the stability of estimates based on small numbers. While readily understandable to non-academics, its limitations with regards to loss of detail suggest that more complex methods may often be warranted for investigating rates of health

events in rare disease populations. In addition, aggregation across geographic areas may be challenging to implement due to differing privacy policies and legislation regarding the protection of health information across jurisdictions. Therefore, weighted moving averages and Bayesian estimation warrant consideration for estimating event rates in small populations such as the IEM populations that were described in the examples. While weighted moving averages are computationally simpler as compared to Bayesian estimation methods, the latter has been recognized as an important approach to small area analysis in health services research (40-44). Both approaches will become more feasible as rare disease registries and cohort studies mature and comparative data become more readily available to investigators.

Regardless of the approach selected by the investigator, often these methods will form an initial stage of descriptive analysis in preparation for research that seeks to compare event rates in two or more populations. For example, investigators may be interested in differences in health care utilization patterns among those with a rare disease as compared to the general population, or across multiple rare disease groups. Once smoothed health event rates are computed for the rare disease cohort of interest, comparisons can be achieved by calculating rate ratios and, if potential confounders are to be considered, by performing multivariable regression analysis. Exact statistical methods can then be used to compute exact confidence intervals about the rate ratios. Exact confidence intervals and p-values are based on exact probability statements as opposed to asymptotic assumptions; these estimates are more reliable for small sample sizes as a result (45, 46). For regression modeling, it would be desirable to estimate smoothed event counts for each member of the study population according to the

respective follow-up time. This may be achieved by relying on within-individual time series data, matching of individuals from populations used for smoothing, or estimation based on extrapolation from the unadjusted smoothed event rate. Once estimated, the smoothed event counts could then be modeled as dependent variables in the analysis to compare two or more distinct populations.

Health services researchers often rely on Poisson, negative binomial, zero-inflated, or Hurdle models when studying rates of health services use, to account for the distribution of health service use events and adjust for the variation in follow-up time. There are assumptions inherent to each statistical distribution, thus investigators need to ensure that inferences accurately reflect the observed data. Inappropriate selection can result in the calculation of biased standard errors, inefficient parameters, and an increase in Type I or Type II error rates due to the highly skewed nature of the resulting distribution (47-49). Some post-hoc methods of comparing and evaluating different modeling approaches have been developed, such as the Vuong statistic. Such tests may help to inform a decision about which model is the most appropriate given the nature of the data (50).

The association between risk factors and smoothed health event rates can similarly be evaluated using the above regression models if the covariates are fixed; however, time-dependent covariates often need to be considered when conducting research in a rare disease setting. The investigator may consider stratifying the selected regression model (e.g., by age or prior to vs. after initiation of therapy) or using time-dependent variables in stratified Cox regression models (51, 52). In the later modeling

technique, again smoothed event counts would need to be estimated for each member of the study population before proceeding with the regression modeling.

CONCLUSIONS AND RECOMMENDATIONS

Aside from aggregation, the methods reviewed here do not appear to be in widespread use among researchers studying healthcare use and health outcomes in small populations, particularly in the field of rare diseases such as IEM. We recommend that investigators consider approaches such as weighted moving averages and Bayesian methods to improve the statistical reliability of calculated event rates, and that they use sensitivity analyses (planned *a priori*) to compare different approaches. We suggest that, whatever approach is taken, investigators should be transparent about their chosen methods and provide as much detail about the original sample data as possible (within constraints of privacy regulations) so that those in the target audience may evaluate the suitability of the analyses and appropriateness of the conclusions. The discussed techniques may often result in more robust estimates, ensuring better value for investments in research studies focusing on rare diseases or other small populations. These methods, if more widely used, could provide better evidence to support improvements in health care and health outcomes among patients and populations.

Table 3.1 A comparison of the methods identified for analyzing robust event rates in small populations

Methods	Advantages	Disadvantages
Aggregation		
Multiple years		<ul style="list-style-type: none"> - Time trends masked - Seminal events need to be considered (i.e., when disease added to NBS panel) - Access to longitudinal data
Collapsing data categories	<ul style="list-style-type: none"> - Larger cell sizes - Computationally simple 	<ul style="list-style-type: none"> - Differences between or among the combined groups will be masked
Expanding geographic area		<ul style="list-style-type: none"> - Geographical variation masked - Multi-jurisdiction privacy barriers
Composite definitions		<ul style="list-style-type: none"> - Specificity of services masked
Data Smoothing		
Weighted averages	<ul style="list-style-type: none"> - Computationally simple - Area-level and time-trend differences not masked 	<ul style="list-style-type: none"> - Access to longitudinal and/or neighboring geographic data
Bayesian estimation methods	<ul style="list-style-type: none"> - Spatial in nature - Extensively used in health care system evaluation research - IEM mapping in the future 	<ul style="list-style-type: none"> - Computationally most difficult as compared to other methods - Access to longitudinal and/or neighboring geographic data

REFERENCES

1. Trusheim MR, Berndt ER, Douglas FL. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nat Rev Drug Discov.* 2007 Apr;6(4):287-93.
2. Trusheim MR, Burgess B, Hu SX, Long T, Averbuch SD, Flynn AA, et al. Quantifying factors for the success of stratified medicine. *Nat Rev Drug Discov.* 2011 Oct;10(11):817-33.
3. The National Association of Health Data Organizations. *Statistical Approaches for Small Numbers: Addressing Reliability and Disclosure Risk.* NAHDO-CDC Cooperative Agreement Project, CDC Assessment Initiative. The National Association of Health Data Organizations; 2004.
4. VanEenwyk J, Macdonald S. *Guidelines for Working with Small Numbers.* Washington: Washington State Department of Health; 2012.
5. Scandol J, Moore H. *Health Statistics NSW: Privacy Issues and the Reporting of Small Numbers.* Sydney: Centre for Epidemiology and Research, Population Health Division, New South Wales Department of Health; 2011.
6. Potter BK, Chakraborty P, Kronick JB, Wilson K, Coyle D, Feigenbaum A, et al. Achieving the "triple aim" for inborn errors of metabolism: a review of challenges to outcomes research and presentation of a new practice-based evidence framework. *Genet Med.* 2013 Jun;15(6):415-22.
7. Ellaway CJ, Wilcken B, Christodoulou J. Clinical approach to inborn errors of metabolism presenting in the newborn period. *J Paediatr Child Health.* 2002 Oct;38(5):511-7.
8. Mitchell JJ, Trakadis YJ, Scriver CR. Phenylalanine hydroxylase deficiency. *Genet Med.* 2011 Aug;13(8):697-707.
9. Guldberg P, Rey F, Zschocke J, Romano V, Francois B, Michiels L, et al. A European multicenter study of phenylalanine hydroxylase deficiency:

- classification of 105 mutations and a general system for genotype-based prediction of metabolic phenotype. *Am J Hum Genet.* 1998 Jul;63(1):71-9.
10. Paine RS. The variability in manifestations of untreated patients with phenylketonuria (phenylpyruvic aciduria). *Pediatrics.* 1957 Aug;20(2):290-302.
 11. Bickel H, Gerrard J, Hickmans EM. The influence of phenylalanine intake on the chemistry and behaviour of a phenyl-ketonuric child. *Acta Paediatr.* 1954 Jan;43(1):64-77.
 12. National Institutes of Health Consensus Development Panel. National Institutes of Health Consensus Development Conference Statement: phenylketonuria: screening and management, October 16-18, 2000. *Pediatrics.* 2001 Oct;108(4):972-82.
 13. Welsh MC, Pennington BF, Ozonoff S, Rouse B, McCabe ER. Neuropsychology of early-treated phenylketonuria: specific executive function deficits. *Child Dev.* 1990 Dec;61(6):1697-713.
 14. Sharman R, Sullivan K, Young R, McGill J. A preliminary investigation of the role of the phenylalanine:tyrosine ratio in children with early and continuously treated phenylketonuria: toward identification of "safe" levels. *Dev Neuropsychol.* 2010;35(1):57-65.
 15. Drogari E, Smith I, Beasley M, Lloyd JK. Timing of strict diet in relation to fetal damage in maternal phenylketonuria. An international collaborative study by the MRC/DHSS Phenylketonuria Register. *Lancet.* 1987 Oct;2(8565):927-30.
 16. Rhead WJ. Newborn screening for medium-chain acyl-CoA dehydrogenase deficiency: a global perspective. *J Inherit Metab Dis.* 2006 Apr-Jun;29(2-3):370-7.
 17. Wilcken B, Hammond J, Silink M. Morbidity and mortality in medium chain acyl coenzyme A dehydrogenase deficiency. *Arch Dis Child.* 1994 May;70(5):410-2.
 18. Oerton J, Khalid JM, Besley G, Dalton RN, Downing M, Green A, et al. Newborn screening for medium chain acyl-CoA dehydrogenase deficiency in England:

- prevalence, predictive value and test validity based on 1.5 million screened babies. *J Med Screen*. 2011;18(4):173-81.
19. Grosse SD, Khoury MJ, Greene CL, Crider KS, Pollitt RJ. The epidemiology of medium chain acyl-CoA dehydrogenase deficiency: an update. *Genet Med*. 2006 Apr;8(4):205-12.
 20. Gregersen N, Andresen BS, Corydon MJ, Corydon TJ, Olsen RK, Bolund L, et al. Mutation analysis in mitochondrial fatty acid oxidation defects: Exemplified by acyl-CoA dehydrogenase deficiencies, with special focus on genotype-phenotype relationship. *Hum Mutat*. 2001 Sep;18(3):169-89.
 21. Wilcken B, Haas M, Joy P, Wiley V, Chaplin M, Black C, et al. Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency in Australia: a cohort study. *Lancet*. 2007 Jan 6;369(9555):37-42.
 22. Potter BK, Little J, Chakraborty P, Kronick JB, Evans J, Frei J, et al. Variability in the clinical management of fatty acid oxidation disorders: results of a survey of Canadian metabolic physicians. *J Inherit Metab Dis*. 2012 Jan;35(1):115-23.
 23. Schatz UA, Ensenauer R. The clinical manifestation of MCAD deficiency: challenges towards adulthood in the screened population. *J Inherit Metab Dis*. 2010 Oct;33(5):513-20.
 24. Leonard JV, Dezateux C. Newborn screening for medium chain acyl CoA dehydrogenase deficiency. *Arch Dis Child*. 2009 Mar;94(3):235-8.
 25. Rudolph BA, Shah GH, Love D. Small numbers, disclosure risk, security, and reliability issues in Web-based data query systems. *J Public Health Manag Pract*. 2006 Mar-Apr;12(2):176-83.
 26. Stoto M. Statistical Issues in Interactive Web-based Public Health Data Dissemination Systems. *RAND Health*. 2003;WR-106.
 27. Fefferman NH, O'Neil EA, Naumova EN. Confidentiality and confidence: is data aggregation a means to achieve both? *J Public Health Policy*. 2005 Dec;26(4):430-49.

28. 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 Pediatr*. 2010 Nov;10:82.
29. Normand SL. Some old and some new statistical tools for outcomes research. *Circulation*. 2008 Aug;118(8):872-84.
30. Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials. *JAMA*. 2003 May;289(19):2554-9.
31. Simonoff J. *Springer Series in Statistics: Smoothing Methods in Statistics*. New York: Springer-Verlag; 1996.
32. Auchincloss AH, Gebreab SY, Mair C, Diez Roux AV. A review of spatial methods in epidemiology, 2000-2010. *Annu Rev Public Health*. 2012 Apr;33:107-22.
33. Shi X, Duell E, Demidenko E, Onega T, Wilson B, Hoftiezer D. A polygon-based locally-weighted-average method for smoothing disease rates of small units. *Epidemiology*. 2007 Sep;18(5):523-8.
34. Wang F. Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Ann Assoc Am Geogr*. 2012;102(5):1104-12.
35. Jia H, Muennig P, Borawski E. Comparison of small-area analysis techniques for estimating county-level outcomes. *Am J Prev Med*. 2004 Jun;26(5):453-60.
36. Ghosh M, Rao JNK. Small Area Estimation: An Appraisal. *Stat Sci*. 1994;9(1):55-76.
37. Carstairs V. Small area analysis and health service research. *Community Med*. 1981 May;3(2):131-9.
38. Zeger SL, Irizarry RA, Peng RD. On time series analysis of public health and biomedical data. *Annu Rev Public Health*. 2006;27:57-79.
39. Marshall RJ. Mapping disease and mortality rates using empirical Bayes estimators. *J R Stat Soc Ser C Appl Stat*. 1991;40(2):283-94.

40. Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res.* 2005 Feb;14(1):35-59.
41. Spiegelhalter DJ. Incorporating Bayesian ideas into health-care evaluation. *Stat Sci.* 2004:156-74.
42. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health-care evaluation. West Sussex: John Wiley & Sons Ltd; 2004.
43. O'Hagan A, Stevens JW. Bayesian methods for design and analysis of cost-effectiveness trials in the evaluation of health care technologies. *Stat Methods Med Res.* 2002 Dec;11(6):469-90.
44. Gomez-Rubio V, Best N, Richardson S, Li G. Bayesian statistics for small area estimation. London: Office for National Statistics, United Kingdom; 2010.
45. Martin DO, Austin H. Exact estimates for a rate ratio. *Epidemiology.* 1996 Jan;7(1):29-33.
46. Agresti A, Min Y. On small sample confidence intervals for parameters in discrete distributions. *Biometrics.* 2001 Sep;57(3):963-71.
47. Karazsia BT, van Dulmen MH. Regression models for count data: illustrations using longitudinal predictors of childhood injury. *J Pediatr Psychol.* 2008 Nov-Dec;33(10):1076-84.
48. Glynn RJ, Stukel TA, Sharp SM, Bubolz TA, Freeman JL, Fisher ES. Estimating the variance of standardized rates of recurrent events, with application to hospitalizations among the elderly in New England. *Am J Epidemiol.* 1993 Apr 1;137(7):776-86.
49. Baldi I, Ciccone G, Merletti F, Gregori D. Computing hospitalization rates in presence of repeated events: impact and countermeasures to avoid misinterpretation. *J Eval Clin Pract.* 2008 Apr;14(2):316-20.
50. Vuong QH. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica.* 1989 Mar;57(2):307-33.

51. Dekker FW, de Mutsert R, van Dijk PC, Zoccali C, Jager KJ. Survival analysis: time-dependent effects and time-varying risk factors. *Kidney Int.* 2008 Oct;74(8):994-7.
52. Fisher LD, Lin DY. Time-dependent covariates in the Cox proportional-hazards regression model. *Annu Rev Public Health.* 1999;20:145-57.

INTERFACE

The health system impact of rare diseases in Canada has not been well characterized to date. For those diseases that are targets of newborn screening, screening and clinical management are both associated with health services use, with important implications for both families and the health care system. Given the large population size with universal health care coverage in Ontario, there was a unique opportunity to study the impact of one of these diseases through the secure linkage of provincial newborn screening diagnostic confirmation data with population-based health care administrative datasets. This manuscript focused on describing the pattern of early health services use and estimating its associated expenditures in a cohort of Ontario infants diagnosed with medium-chain acyl-CoA dehydrogenase deficiency ascertained through newborn screening from April 1, 2006 through March 31, 2010; and comparing this pattern to a screen negative birth cohort. Sociodemographic and geographic characteristics that are known to be associated with health services use, such as gestational age and urban-rural status, were considered in the analyses. Using the information summarized in the methods review from the previous manuscript, a weighted moving average approach was selected to estimate smoothed rates of health events when examining time trends.

This manuscript was co-authored by the student (MK); her thesis co-supervisors, Drs Beth Potter and Pranesh Chakraborty; and the members of her thesis advisory committee, Drs Doug Coyle and Kumanan Wilson. The student was responsible for planning the analytical protocol, completing the analysis and drafting the manuscript. The thesis co-supervisors and thesis advisory committee members

provided guidance throughout the data analysis and manuscript writing process. In addition, Dr Pranesh Chakraborty provided clinical expertise with regards to the pathophysiology of MCADD, while Dr Beth Potter was integral towards developing the methodological approach.

CHAPTER FOUR: HEALTH SYSTEM IMPACT OF MCADD – MANUSCRIPT

#2

The epidemiology and health system impact of medium-chain acyl-CoA dehydrogenase deficiency among children ascertained through newborn screening in Ontario, Canada.

Maria D Karaceper, BSc,¹; Pranesh Chakraborty, MD, FRCPC,²; Doug Coyle, MSc, PhD,¹; Kumanan Wilson, MD, FRCPC,³; Beth K Potter, MSc, PhD,¹; and in collaboration with the Canadian Inherited Metabolic Diseases Research Network

Affiliation:

1 Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada

2 Newborn Screening Ontario, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada

3 Ottawa Hospital Research Institute and Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada

Keywords: newborn screening, metabolic diseases, health care service utilization, economic

Manuscript Status: To be submitted to the Journal of Inherited Metabolic Disease

ABSTRACT

Objective: Limited research has described the health system impact of rare childhood diseases. This observational study characterized the early health care utilization patterns and associated health expenditures for children diagnosed with medium-chain acyl-CoA dehydrogenase deficiency (MCADD) through newborn screening in Ontario, Canada.

Methods: Newborn screening diagnostic confirmation data were linked with health care administrative datasets capturing physician visits, emergency department care, and inpatient hospitalizations from April 1, 2006 through March 31, 2012. We described patterns of health services use in the MCADD cohort, stratified by the age of the child at the time of the receipt of care. We used negative binomial regression models to compare health care service visit rates and health care costs between MCADD and comparison cohorts.

Results: Among forty children diagnosed with MCADD through newborn screening between April 1, 2006 and March 31, 2010, rates of physician visits, emergency department visits, and inpatient hospitalizations were 11.7, 1.6, and 0.4 per child per year, respectively. All service visit rates were significantly higher in the MCADD cohort relative to children with negative newborn screening results, regardless of age at the time of visit. The average cost of health services for a child with MCADD during the first year of life was estimated to be \$6,258 CAD, compared to \$4,746 CAD for a child with a negative newborn screening result.

Conclusion: Children with MCADD are frequent users of health care services, highlighting the importance of a well-designed system of care to improve patient experiences and health outcomes.

INTRODUCTION

Medium-chain acyl-CoA dehydrogenase deficiency (MCADD), a mitochondrial fatty acid oxidation disorder with an autosomal recessive inheritance pattern, is one of the most common inborn errors of metabolism with an estimated birth prevalence of 1:10,000 to 1:20,000 (1-4). Mutations in the *ACADM* gene result in dysfunction of medium-chain acyl-CoA dehydrogenase and impairment of mitochondrial beta-oxidation of medium-chain fatty acids (4-6). This step is critical for hepatic ketone body production, which provides an essential alternative energy source in times of physiological stress, such as prolonged fasting and intercurrent illness (4, 7). While symptoms are variable, the typical presentation of MCADD is hypoketotic hypoglycemia causing encephalopathy during times of physiological stress; the most severe cases result in coma or death (2, 4, 6).

This episodic disorder is typically asymptomatic at birth, with metabolic decompensation being most frequent between three months and three years of age (5). Without a disease management protocol in place, these events are associated with a mortality rate of up to 25% (5, 8, 9). Further, up to one third of children with MCADD who survive an initial metabolic crisis episode may display some form of developmental delay (2, 4, 10). However, if early diagnosis is established, prognosis is excellent since the vast majority of adverse outcomes are preventable with the immediate initiation of treatment (4, 7, 11, 12). Standard long-term therapy for the disorder includes the avoidance of fasting, as well as parenteral glucose and close medical monitoring during intercurrent illness; this may involve visits to the emergency department and hospitalization (12-16). Evidence supporting effective

treatment to prevent and manage metabolic crises with associated morbidity and mortality risk reduction has prompted the inclusion of MCADD in many population-based newborn bloodspot screening programs throughout the world (17-19). However, limited research has focused on describing the impact of childhood rare diseases, including MCADD, on the health care system and families in Canada.

Newborn Screening Ontario is responsible for coordinating the provincial newborn bloodspot program for the approximately 140,000 babies born each year in Ontario, Canada. If a child screens positive for any of the 29 screened diseases on the panel, a referral is made to one of five Newborn Screening Regional Treatment Centres in Ontario for diagnostic testing and follow-up care. MCADD was added to the panel in April 2006 (20). All citizens of Ontario are eligible for universal government funded health insurance, which includes a wide range of health care services. These services are captured within provincial population-based health administrative databases housed at the Institute for Clinical Evaluative Sciences (ICES), which can be securely linked to provincial newborn screening diagnostic confirmation data. As a result, there is a unique opportunity in Ontario to characterize health service utilization among children with MCADD. As part of a larger program of observational research designed to improve health outcomes and health care services for children with inborn errors of metabolism in Canada, we investigated the epidemiology and health system impact of MCADD among affected children in Ontario, Canada.

PATIENTS AND METHODS

Study Population and Data Sources

The source population for this study was all children born in Ontario who received bloodspot screening through Newborn Screening Ontario between April 1, 2006 and March 31, 2010. The MCADD study cohort included children with a confirmed diagnosis of MCADD identified through newborn screening and with confirmatory follow-up testing (20). A primary comparison cohort included all children with negative newborn screening results for all screened conditions in Ontario during the same time period. To achieve better control over potential confounding variables, a secondary matched comparison cohort was created using factors that are known to be strong predictors of health services use; thus, ten controls were randomly matched to each diagnosed child based on sex, calendar year of birth, a measure of the urban-rural status of the child's home residence, and neighbourhood income quintile. Individuals were excluded from the analysis if they were ineligible for Ontario Health Insurance Plan (OHIP) coverage at the time of birth, received a positive screening result for another disorder on the newborn screening panel, or were deceased within 24 hours following birth; this latter group was excluded because newborn screening bloodspot samples must be collected after 24 hours of age in Ontario for the result to be considered satisfactory.

Newborn screening data used to identify children in each cohort were securely linked to Ontario's Registered Persons Database, the province's healthcare patient registry, at ICES by encrypted health card numbers. Screening and confirmatory testing

results were subsequently linked at the individual level with population-based healthcare administrative databases encompassing health services use from April 1, 2006 through March 31, 2012. Specifically, hospitalization data were obtained from the Canadian Institute for Health Information Discharge Abstract Database; this database includes administrative and clinical information from all acute inpatient facilities throughout Ontario (21). Emergency department (ED) care data were obtained from the Canadian Institute for Health Information National Ambulatory Case Reporting System, which reports information for all hospital- and community-based ambulatory care services as well as day surgeries and outpatient clinics; close to 100% of ED claims in Ontario are captured within this database (22). Primary care and specialist physician billing data for all ambulatory visits were identified using the OHIP database; while some providers still follow an enhanced fee-for-service model, there has been a significant shift towards capitated payment models in the past decade (23).

Independent Variables

To account for factors that were associated with having MCADD and were also strong predictors of health service utilization, the following variables were included in the analysis: sex, season of birth, birth weight, gestational age, an area-based proxy measure of relative socioeconomic status (neighbourhood income quintile) and a measure of residential urban-rural status. The measured concentration of the primary analyte used in the screening algorithm for MCADD (octanoylcarnitine, C8) was also available for the true positive cohort and thus was included in descriptive analysis for that cohort only. These variables were ascertained from the newborn screening record (C8 concentration), hospital records at the time of birth (sex, birth weight, gestational

age), and Canadian census data linked to the other datasets by child's postal code of residence at the time of birth (urban-rural status, socioeconomic status).

Season of birth: The following three categories of season of birth were considered, in order to account for seasonal variation in care: January-April, May-August, and September-December.

Birth weight and gestational age: Participants were separated into two birth weight categories for the analysis: low birth weight (<2,500 grams) and normal or high birth weight (≥ 2500 g). Gestational age was dichotomized to term/post-term (≥ 37 weeks of gestation) or pre-term (<37 weeks of gestation).

Socioeconomic status: Income quintile at the neighborhood-level was selected as a proxy measure of socioeconomic status. We derived this variable using children's postal codes at the time of birth from hospitalization data linked to small area-level socioeconomic information from the 2006 Canadian Census using Statistic Canada's Postal Code Conversion File. Specifically, we used the household-size-adjusted mean neighborhood income value per single person at the dissemination area level (dissemination areas in Canada have a population size of approximately 400 to 700 persons) (24). Each dissemination area within a larger census metropolitan area was then divided into five income quintile groups according to the ranked average neighborhood household income values (25). The two lowest quintiles [1 and 2] were combined to form a "lower" relative income group, while the other three quintiles [3, 4, and 5] were merged to create a "higher" relative income group.

Urban-rural status

We defined urban-rural status according to the revised Rurality Index for Ontario, RIO-2008, developed by the Ontario Medical Association and the Ontario Ministry of Health and Long-Term Care in 2008 (26), which is designed specifically to characterize urban versus rural areas with respect to access to health services. It was derived using three geographical components: the community population and population density, the travel time to the nearest tertiary care centre, and the travel time to the nearest Level 2 centre. Rural physician eligibility requirements in Ontario define a rural community as one whose RIO2008 score is 40 or greater, thus we established this value as a cut-off for rural status in our analysis (27).

Utilization Outcomes

We included every original health service encounter made by a child within the study period (hospitalizations, ED visits, and primary care/specialist physician visits) in the analysis. If a child had multiple OHIP billed procedures within a single physician visit, these were considered together as one visit. However, if a child saw multiple physicians on a single day, these were considered as separate encounters. Similarly, if a child visited the ED more than once on the same date, each visit was considered an original encounter; children affected with MCADD may access these services more than once per day during episodic events. A child could only have a single inpatient hospitalization on any given day.

Statistical Analysis

We accessed de-identified, study-specific datasets at ICES; cell sizes less than six were not reported in accordance with stringent confidentiality and privacy policies with respect to small numbers. Descriptive statistics were used to characterize the independent variables and utilization outcomes of interest. Counts and percentages were presented for categorical data, while medians along with ranges were reported for continuous variables. Chi-square tests were used to investigate bivariate associations between disease status (children with MCADD versus comparison cohorts) and geographic and sociodemographic characteristics.

The number of hospitalizations, ED visits, and physician visits during the study period were respectively summed for each child. The length of follow-up for each individual was subsequently calculated as the time elapsed in years between the date of birth and one of the following three time points: the date of death, the date of OHIP eligibility loss, or the last date of follow-up for the study. The earliest of these three possible dates was selected as the end point for the calculation. Unadjusted visit rates and incidence rate ratios (IRR) to compare cohorts for each type of visit category along with 95% confidence intervals were calculated. Stratified unadjusted visit rates and IRR were subsequently computed with strata based on age at the time of visit (<1 year of age and ≥ 1 year of age) as suggested by the literature (28, 29). To examine time trends based on year of birth while accounting for instability in rates due to small numbers, unadjusted smoothed visit rates for each service type were calculated using a weighted moving average approach with Hanning weights. This calculation was completed for

children with MCADD born in 2007, 2008, and 2009 respectively, stratified by age at the time of visit (<1 year of age and 1–2 years of age) (30).

We used the Vuong statistic, a likelihood ratio-based statistical test, to compare regression models for calculating adjusted visit IRR for each health service type (31). Based on the results of the Vuong test, we selected a negative binomial regression model to investigate the difference in visit rates for health care services in the MCADD versus comparison cohorts while adjusting for potential confounding factors. Influential observations in the regression models were identified using standardized Pearson residuals, Cook's D statistic, and DFBETA statistic; an observation was considered influential if the following three conditions were met: the standardized Pearson residual was greater than the absolute value of 2, the value of Cook's D statistic was larger than $4/n$, and the value of the DFBETA statistic was greater than $2/n$, where n is equal to the number of observations (32, 33). Distinct 99th percentile values were calculated for the three service types in each of the three cohorts, and then applied to the appropriate influential observations for truncation. In analyses with the matched comparison cohort, we adjusted for season of birth, birth weight, and gestational age; sex, socioeconomic status, and urban-rural status were additionally included in unmatched analyses. All models were stratified by age at the time of the visit (<1 year of age and ≥ 1 year of age). As a sensitivity analysis to account for a potential residual confounding effect of premature birth, we re-ran our final models restricted to children with term births (≥ 37 weeks of gestation).

We relied on the Ontario Case Costing Initiative to compute individual-level health care costs for children diagnosed with MCADD; costs were adjusted for inflation

and expressed in 2012 Canadian dollars (34). These costs encompassed OHIP billings, inpatient hospitalizations, same day surgeries, ED visits, home care services, rehabilitation, and complex and continuing care. Costs were calculated for the first year of life for each individual; means and medians along with ranges were reported for the MCADD and the primary comparison group separately. All statistical analyses were performed using SAS® software version 9.3 (SAS Institute Inc., Cary, North Carolina, USA). The study protocol was approved by the Ottawa Health Science Network Research Ethics Board.

RESULTS

Study Population

Forty children were diagnosed with MCADD through Newborn Screening Ontario during the specified study period. The primary and secondary matched comparison cohorts consisted of 545,355 and 400 children respectively. We were unable to report the number of children with MCADD with a birth weight of <2500g or a gestational age of <37 weeks due to the previously described privacy guidelines. Children with MCADD were more likely to live in rural communities as compared to the primary comparison group (Table 4.1). No other statistically significant differences between the cohorts were observed across geographic and sociodemographic characteristics, however we did not evaluate the significance of differences in birth weight and gestational age due to the privacy suppression of those data. Those with MCADD had a median C8 screening level of 8.6 umol/L (range: 0.75-30.4). While all children in the MCADD cohort were followed for at least 2 years (and up to

approximately 6 years), the larger comparison cohorts included children who died before the end of follow-up and thus follow-up time ranged from less than one month to approximately 6 years.

Patterns of Care in MCADD and Comparison Cohorts

We calculated unadjusted visit rates for children for each type of health service visit (Figure 4.1). Children with MCADD had a total of 1956 recorded physician visits over the entire follow-up period, with an average rate of 11.7 physician visits per child per year; children in the primary and secondary comparison groups had an average rate of 8.4 and 8.0 physician visits per child per year respectively. Those with MCADD had a total of 259 ED visits over the follow-up period with an average rate of 1.6 ED visits per child per year. Those in the primary comparison group had an average of 0.7 ED visits per child per year, while the secondary comparison group experienced an ED visit rate of 0.8 visits per child per year. Finally, children with MCADD had a total of 70 inpatient hospitalizations over the follow-up period with an average rate of 0.4 hospitalizations per child per year. Those in the primary comparison group had an inpatient hospitalization rate of 0.06 stays per child per year, while the secondary comparison group had a hospitalization rate of 0.05 stays per child per year.

Unadjusted visit rates for each type of health service visit were then stratified by age at the time of visit. Children in all 3 cohorts had the highest physician use, ED visits, and hospitalization rates during their first year of life (Figure 4.1). In the first year of life, children with MCADD had on average 20.0 physician visits, 2.0 ED visits, and 0.5 hospital admissions per child. The rates decreased after age one to, on average, 9.1

physician visits, 1.4 ED visits, and 0.4 hospitalizations per child per year. In the first year of life, those in the primary comparison group had on average 14.7 physician visits, 0.8 ED visits, and 0.1 hospitalizations per child; these rates similarly decreased in the subsequent years of age to an average of 6.3 physician visits, 0.6 ED visits, and 0.04 hospitalizations per child per year. Finally, in the first year of life, children in the matched comparison group had on average 14.4 physician visits, 1.0 ED visits, and 0.1 hospitalizations per child over one year; these rates also decreased in the subsequent years of age to an average of 5.9 physician visits, 0.7 ED visits, and 0.03 hospitalizations per child per year.

In the first year of life, the rate differences between children with MCADD and those in the primary comparison cohort were 5.3 physician visits, 1.2 ED visits, and 0.4 hospitalizations per child per year; when compared to those in the secondary comparison group, the rate differences were 5.6 physician visits, 1.0 ED visits, 0.4 hospitalizations per child per year. In the subsequent years of life, the rate differences were 2.8 physician visits, 0.8 ED visits, 0.36 hospitalizations per child per year between those with MCADD and the primary comparison cohort; finally, the rate differences were 3.2 physician visits, 0.7 ED visits, and 0.37 hospitalizations per child per year when those with MCADD were compared to the secondary comparison group.

Smoothed Visit Rates for MCADD

We calculated unadjusted smoothed visit rates for each service type stratified by age at the time of visit (<1 year of age and 1–2 years of age) for children with MCADD born in 2007, 2008, and 2009 respectively (Figure 2). At less than one year of age,

children with MCADD had average smoothed rates of 20.7, 21.7, and 19.0 physician visits per child per year, for those born in 2007, 2008, and 2009, respectively. At less than one year of age, the cohort additionally experienced average smoothed rates of 2.1, 2.0, and 1.7 ED visits per child per year, again for those born in 2007, 2008, and 2009, respectively. Finally, again at less than one year of age, children with MCADD had average smoothed rates of 0.4, 0.7, and 0.5 inpatient hospital admissions per child per year, for those born in 2007, 2008, and 2009 respectively. Similar smoothed trends were observed in the older age category for each service type. For those born in 2007, 2008, and 2009 respectively, children with MCADD between 1 and 2 years of age experienced average smoothed rates of 13.7, 12.0, and 11.6 physician visits per child per year. In the same age range, the cohort also had average smoothed rates of 2.6, 2.0, and 1.5 ED visits per child per year for those born in 2007, 2008, and 2009, respectively. Additionally, for those born in 2007, 2008, and 2009 respectively, children with MCADD had average smoothed rates of 0.6, 0.5, and 0.5 inpatient hospital admissions per child per year between 1 and 2 years of age.

Incidence Rate Ratios for Individual Health Care Service Types

In unadjusted and adjusted analyses, children with MCADD displayed a statistically significant higher frequency of all health services types we investigated, compared to both comparison groups, regardless of the age of the child at the time of visit (Table 4.2). In unadjusted analyses, there was a statistically significant higher rate of physician visits (IRR: 1.38 [95% CI: 1.32-1.45]), ED visits (IRR: 2.33 [95% CI: 2.06-2.63]) and hospitalizations (IRR: 6.99 [95%CI: 5.53-8.84]) for children with MCADD when compared to the primary comparison group. The results were not appreciably

different with adjustment for sex, season of birth, birth weight, gestational age, socioeconomic status, and urban-rural status: children with MCADD had higher rates of physician visits (IRR: 1.42 [95%CI: 1.20-1.71]), ED visits (IRR: 2.02 [95%CI: 1.47-2.87]), and hospitalizations (IRR: 8.39 [95%CI: 5.18-14.22]). The IRRs were very similar when the matched cohort was used as the comparison (Table 4.2). When the analysis was stratified by the age of the child at the time of visit, the IRRs were similar to the overall unstratified results with the exception of the rate of hospitalizations (Table 4.2). The difference between the MCADD and comparison groups in the rate of hospitalizations was much larger for children ≥ 1 year of age compared with < 1 year of age; the adjusted IRR for hospitalizations in the MCADD cohort compared to the primary comparison group was 12.97 (95%CI: 6.96-26.52) for children ≥ 1 year of age and 2.87 (95%CI: 1.41-5.56) for children < 1 year of age. The comparable adjusted IRRs for hospitalizations in the MCADD cohort versus the matched comparison group were 8.70 (6.48-11.54) for children ≥ 1 year of age and 2.31 (95%CI: 1.43-3.55) for children < 1 year of age.

With respect to other predictors in the full, age-stratified models that included the MCADD primary comparison cohorts, children in the population who were males, had lower birth weights, or were born pre-term generally had higher rates of health services use across service types and age categories (Table 4.3). Those with lower socioeconomic status had higher rates of ED visits and hospitalizations in both age categories as compared to those with a higher socioeconomic status, whereas those with a higher socioeconomic status had a higher rate of physician visits among children ≥ 1 year of age. Similarly, while children living in rural areas had higher rates of ED

visits and hospitalizations than those living in urban areas for both age categories, children living in urban areas had higher rates of physician visits. Season of birth was marginally and variably associated with health services use.

The stratified models with the primary comparison cohort were subsequently restricted to children born at term (≥ 37 weeks gestational age) as a sensitivity analysis, to account for the potential residual confounding effect of premature births (Table 4.4); the results were similar to those reported in the unrestricted full model (Table 4.3). Term children with MCADD had statistically significant higher rates of physician visits, ED visits, and hospitalizations as compared to term children in the primary comparison group regardless of the age of the child at the time of visit. However, the IRRs comparing hospitalization rates in children with MCADD versus controls were slightly increased in the analysis restricted to children born at term, both for hospital visits at < 1 year of age (IRR: 4.97) and those at ≥ 1 year of age (IRR: 14.78) (Table 4.4); as compared to the IRRs for hospitalization in those with MCADD versus controls in the full model that included children born both term and pre-term (< 1 year of age, IRR: 2.87; ≥ 1 year of age, IRR: 12.97) (Table 4.3).

Costs of Care for MCADD Cohort in Ontario in the First Year of Life

We calculated the costs of care for children with MCADD in the first year of life, which were adjusted for inflation and expressed in 2012 Canadian dollars (Table 4.5). The average costs of care for a child affected with MCADD were \$1,458 (median: \$1,112; interquartile range (IQR): \$843 - \$1,703), \$423 (median: \$251; IQR: \$0 - \$549), and \$4,116 (median: \$1,412; IQR: \$829 - \$4,470) for physician visits, ED care, and

hospitalizations respectively during this one-year period. For a child in the screen negative primary comparison group, the average costs of care in the first year of life were \$983 (median: \$662, IQR: \$457 - \$998), \$149 (median: \$0, IQR: \$0 - \$163), and \$3,605 (median: \$853; IQR: \$829 - \$1,581) for physician visits, ED care, and hospitalizations respectively. The average total cost of care for a child with MCADD during the first year of life was \$6,258 (median: \$3,335; IQR: \$2,158 - \$6,108); for a child in the primary comparison cohort, the average total cost of care was \$4,746 (median: \$1,876, IQR: \$1,373 - \$2,756).

DISCUSSION

Caring for children with special health care needs can be challenging and costly both for families and the health care system. Studies have shown that these patient populations access health care services at a significantly higher rate as compared to healthy children (35-39). However, limited research has been directed towards characterizing healthcare service utilization for those diagnosed with an inborn error of metabolism. To our knowledge, this study is the first population-based analysis to date that evaluates the healthcare utilization patterns and associated health system expenditures for a cohort of children diagnosed with MCADD through a newborn screening program.

In this study, we found that children diagnosed with MCADD accessed physician services, ED care, and were hospitalized at statistically significant higher rates as compared to those with negative newborn screening results during the study period. The higher rate of physician visits among children with MCADD relative to the screen

negative population may be related to diagnostic follow-up and metabolic monitoring in this cohort. Similarly, children with MCADD had approximately twice as many ED visits and 7 times as many inpatient hospitalizations as compared to the comparison cohorts; increased access to these services would be necessary during times of intercurrent illness or for the treatment of episodic metabolic decompensation events. Our results with respect to increased use of health services across these service types are similar to previous studies of sickle cell disease, which has in common with MCADD its rarity and being characterized by acute crises that rely on urgent care (37-39). For example, in one US study, children under the age of five with sickle cell disease had approximately twice as many ED visits and 7.8 times as many hospitalizations as compared to same-aged children without the disease (38).

While not unexpected, our findings quantify the frequency and patterns of use in this patient population, which is relevant for understanding the burden of disease from the perspective of families as well as the health care system. Specifically, on average, children with MCADD experienced approximately 20 physician visits and 2 ED visits over the first year of life; further, about half of affected children would be expected to be hospitalized during that year based on our findings. These rates, particularly for physician visits, were lower after children reached 1 year of age. The resultant cost of care for a child with MCADD in the first year of life, on average, was estimated to be approximately 32% higher at \$6,258 as compared to \$4,746 for a child with a negative newborn screening result. In addition to the increased rates of physician visits and ED care that contribute to these elevated costs for the healthcare system and associated impacts for family life, previous research has identified that children diagnosed with

inborn errors of metabolism who were admitted to the neonatal or pediatric intensive care unit due to disorder-associated clinical manifestations required more aggressive and costly support relative to other illnesses (although not all of the supports documented in these studies would apply to most patients with MCADD) (40-42).

Thus, children with MCADD are frequent users of health care services, with associated costs to the health care system and impacts on families with respect to travel time and the stress of requiring care. These findings underscore the need to ensure that services are organized effectively for this group of children. A multidisciplinary family-centered care team approach, which involves the coordinated collaboration of primary care and subspecialist providers along with other medical professionals (e.g., nurses, dietitians, genetic counselors), has been identified as a beneficial model of care for children with rare, inherited disorders (43-46). Regional metabolic treatment centres across Canada were established to provide a multidisciplinary approach to clinical care for these pediatric populations as a result, however variation in disease management persists among these centres (13, 20). Future research should be aimed towards characterizing different models of care in relation to health outcomes and costs in order to select the most effective and efficient care pathways.

A limitation of this study that is common to all rare disease research was the challenge in interpreting differences in health services use given small patient populations. As a result, we were limited in the number of stratified analyses that we could perform due to privacy standards for small numbers. For example, we were unable to distinguish between extremely low birth weight (<1,000 grams), very low birth weight (<1,500 grams), and high birth weight (\geq 4,500 grams) as suggested in the

literature (47, 48). In order to address this challenge, a review on methods for estimating robust event rates with small populations was completed. One of the identified approaches was the calculation of weighted moving averages in order to stabilize and smooth the desired visit rates using health care utilization data from consecutive MCADD birth cohorts (49). In this approach, health care utilization rates from the most recent years are assigned the greatest weights, while those that are further away are assigned smaller weights according to the length of time from the year of interest (50, 51). When we calculated smoothed health care visit rates for each type of service using this methodology, trends over calendar time (for children born in 2007, 2008, and 2009) were relatively stable, although there was some evidence of a decrease in health care use over time for physician visits and particularly rates of ED use. Weighted averages are not computationally difficult to use if the neighboring rates are readily available, which makes it an appealing solution for researchers dealing with small populations. Further sociodemographic and geographic stratification will become possible as the cohort size grows and follow-up period increases.

While our small sample size did not allow us to examine rates of health care use by sociodemographic characteristics such as rural/urban status and socioeconomic position within the MCADD population separately, these factors were associated with health care use in the entire study population (those with MCADD together with the comparison group). Based on an adjusted negative binomial model, children living in a rural area were twice as likely to access services in the ED as compared to those living in an urban area during the study period, yet those living in an urban area had higher rates of physician visits. This finding suggests a potential need to improve patient

access to physician services in rural areas to relieve the burden currently being placed on emergency services in these geographic regions. It also suggests that it would be worthwhile to explore potential geographic inequities in access to care among children with MCADD and other complex diseases in Ontario, once further data are accrued from the newborn screening program. Multiple Canadian provinces including Ontario presently have rural recruitment and retention initiatives to attract physicians to establish full-time practices in eligible communities (52).

Clinical and psychosocial details are limited in health care administrative datasets, thus we were unable to account for several potential factors that may affect health services use in affected children, such as comorbidities and parental perception of need (53, 54). Integration of clinical and patient- and family-reported information with administrative data would provide richer information to further characterize the epidemiology of MCADD and other inborn errors of metabolism as well as to investigate factors that may be associated with differences in outcomes and have an impact on the health care system.

An additional limitation to the current study is the OHIP eligibility criterion, which shortened the length of follow-up for some children in the affected and screen negative cohorts respectively. If a child emigrated from Ontario, we were unable to access any prospective health care utilization data beyond the date that OHIP eligibility was lost. While less than six children lost OHIP eligibility during the study period in the MCADD cohort, this loss of information still represents a sizable proportion of the potentially available data due to the small population size. In addition to increasing the

overall affected cohort size, future expansion of these methods to include other provinces would allow for the continued follow-up of children who move residence.

Lastly, the reported total cost values in this study include costs associated with all OHIP billings, inpatient hospitalization, emergency department care, same day surgeries, Ontario Drug Benefit claims, rehabilitation, complex and continuing care, home care services, and assisted devices. However, only children that meet stringent eligibility criteria are covered under the Ontario Drug Benefit Program, such as those receiving social assistance; the costs associated with prescription drugs are not fully captured for the study cohort as a result. Furthermore, we were unable to account for disease-related costs of a few additional services and treatments that may be pertinent to children with MCADD, such as L-carnitine supplementation, specialized formulas, and dietetic services.

CONCLUSION

To our knowledge, this study represents the first population-based study to date that describes early health care utilization patterns for a cohort of children affected with MCADD ascertained through expanded newborn screening. The findings demonstrate that children with MCADD access physician services, ED care, and are hospitalized at significantly higher rates as compared to those with negative newborn screening results regardless of age at the time of visit, with associated costs to the health care system and impacts on families. As a result, there is a need to ensure that services are organized effectively for this rare disease subgroup in order to promote better health outcomes and patient experiences with care. There is strong potential to

adapt and expand these methods to investigate other rare diseases, taking advantage of unique opportunities in Ontario and other Canadian provinces. Future research should concentrate on describing current models of care in place across regional metabolic treatment centres, and on comparing their effectiveness towards improving outcomes.

Table 4.1. Geographic and sociodemographic characteristics of the study population

	Study Cohort		
	MCADD, n (%) (n=40)	Primary Comparison, n (%) (n=545,355)	Secondary Comparison, n (%) (n=400)
Sex^a			
Male	20 (50.0)	279 638 (51.3)	-
Female	20 (50.0)	265 717 (48.7)	-
Season of Birth			
January – April	12 (30.0)	177 918 (32.6)	115 (28.8)
May – August	16 (40.0)	192 896 (35.4)	144 (36.0)
Sept. – Dec.	12 (30.0)	174 541 (32.0)	141 (35.2)
Birth weight^b			
<2500g	<6 (≤12.5)	33 027 (6.1)	<60 (≤12.5)
≥2500g	35–40 (≥87.5)	508 466 (93.2)	350–400 (≥87.5)
Gestational age^b			
<37 weeks	<6 (≤12.5)	42 235 (7.7)	<60 (≤12.5)
≥37 weeks	35–40 (≥87.5)	489 232 (89.7)	300–400 (≥87.5)
Socioeconomic status^a			
‘Lower’	16 (40.0)	232 269 (42.6)	-
‘Higher’	24 (60.0)	310 003 (56.8)	-
Urban-rural status^{a, c}			
Rural	6 (15.0)	34 111 (6.3)	-
Urban	34 (85.0)	505 236 (92.6)	-

^a Counts are not presented for secondary matched comparison group on matched variables

^b Results are repressed for cell sizes <6

^c $P < 0.05$ for difference in proportion in the MCADD cohort versus the comparison cohort

Figure 4.1. Unadjusted health service visit rates stratified by age at the time of visit (<1 year of age and ≥1 year of age) for children in all three cohorts

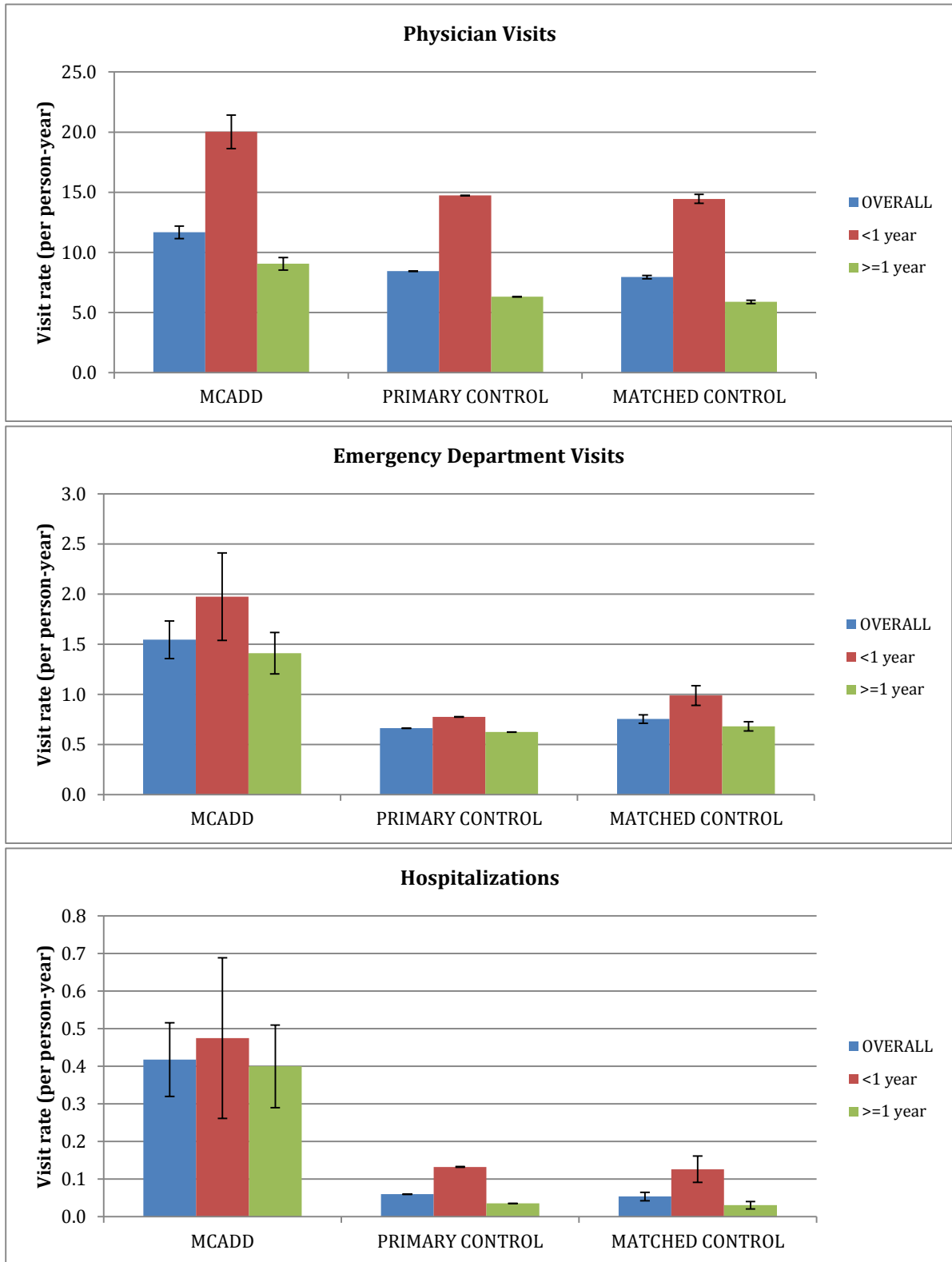


Figure 4.2. Smoothed health service visit rates for children (<1 year of age and 1-2 years of age) with MCADD using weighted moving averages by year of birth

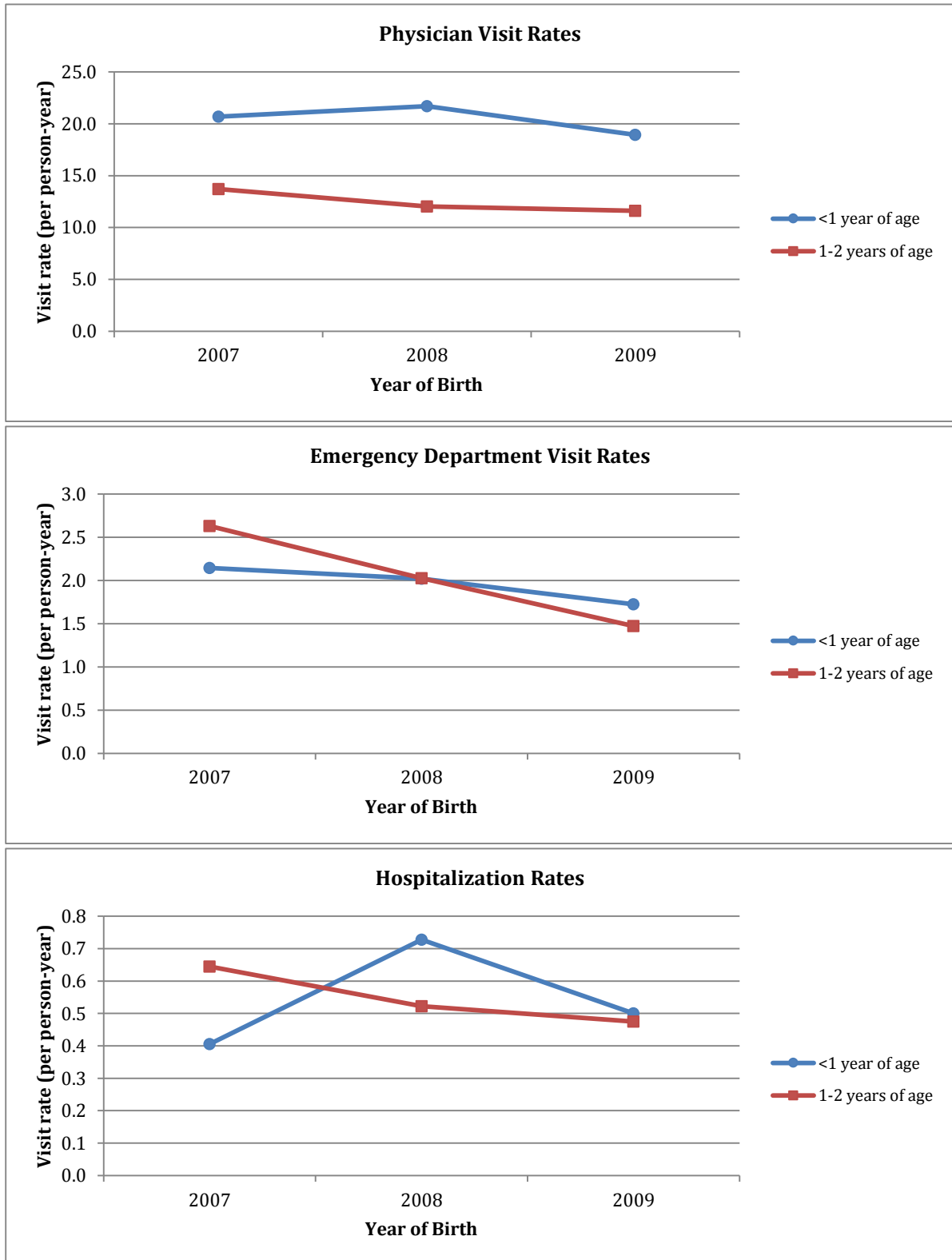


Table 4.2. Unadjusted and adjusted incidence rate ratios stratified by age at the time of visit (<1 year of age and ≥1 years of age), MCADD cohort versus primary (unmatched) comparison cohort or secondary matched comparison cohort

	MCADD vs Primary Comparison Group		MCADD vs Matched Comparison Group	
	IRR _{unadj} (95% CI)	IRR _{adj} (95% CI) ^a	IRR _{unadj} (95% CI)	IRR _{adj} (95% CI) ^b
Physician visits				
Overall	1.38 (1.32–1.45)	1.42 (1.20–1.71)	1.47 (1.40–1.68)	1.43 (1.20–1.73)
<1 year of age	1.36 (1.27–1.46)	1.39 (1.18–1.65)	1.39 (1.29–1.49)	1.38 (1.16–1.64)
≥1 year of age	1.44 (1.35–1.52)	1.51 (1.22–1.88)	1.54 (1.44–1.64)	1.50 (1.20–1.89)
ED visits				
Overall	2.33 (2.06–2.63)	2.02 (1.47–2.87)	2.05 (1.79–2.34)	2.08 (1.49–2.96)
<1 year of age	2.54 (2.04–3.17)	2.41 (1.59–3.73)	2.00 (1.57–2.54)	2.22 (1.53–3.31)
≥1 year of age	2.26 (1.95–2.62)	1.98 (1.39–2.89)	2.07 (1.76–2.43)	2.24 (1.55–3.32)
Hospitalizations				
Overall	6.99 (5.53–8.84)	8.39 (5.18–14.22)	7.81 (5.70–10.69)	7.91 (4.76–13.71)
<1 year of age	3.60 (2.29–5.64)	2.87 (1.41–5.56)	3.76 (2.22–6.38)	2.31 (1.43–3.55)
≥1 year of age	11.35 (8.62–14.93)	12.97 (6.96–26.52)	13.14 (8.64–20.00)	8.70 (6.48–11.54)

^a Incidence rate ratios adjusted for sex, season of birth, gestational age, birth weight, socioeconomic status, and urban-rural status

^b Incidence rate ratios adjusted for season of birth, gestational age, and birth weight

Table 4.3. Stratified adjusted models (<1 year of age and ≥1 years of age) showing incidence rate ratios for the three service types, MCADD cohort versus primary (unmatched) comparison cohort

	Adjusted rate ratio (95% CI)					
	<1 year of age			≥1 year of age		
	Physician visits	ED visits	Hospitalizations	Physician visits	ED visits	Hospitalizations
MCADD	1.39 (1.18-1.65)	2.41 (1.59-3.73)	2.87 (1.41-5.56)	1.51 (1.22-1.88)	1.98 (1.39-2.89)	12.97 (6.96-26.52)
Sex						
Female	0.92 (0.91-0.92)	0.86 (0.85-0.86)	0.76 (0.75-0.77)	0.91 (0.91-0.92)	0.86 (0.85-0.87)	0.78 (0.77-0.80)
Male	Reference	Reference	Reference	Reference	Reference	Reference
Season of birth						
Jan. – Apr.	Reference	Reference	Reference	Reference	Reference	Reference
May – Aug.	1.01 (1.00-1.01)	1.04 (1.03-1.05)	0.98 (0.96-1.00)	0.96 (0.95-0.96)	0.95 (0.95-0.96)	0.94 (0.92-0.97)
Sept. – Dec.	1.00 (0.99-1.00)	1.02 (1.00-1.03)	1.08 (1.06-1.11)	1.00 (0.99-1.00)	1.00 (0.99-1.01)	1.03 (1.00-1.05)
Birth weight						
<2500g	1.75 (1.74-1.77)	1.04 (1.02-1.06)	2.14 (2.08-2.21)	1.13 (1.12-1.14)	1.03 (1.01-1.05)	1.59 (1.51-1.66)
≥2500g	Reference	Reference	Reference	Reference	Reference	Reference
Gestational age						
<37 weeks	1.57 (1.56-1.58)	1.23 (1.21-1.26)	2.65 (2.57-2.73)	1.09 (1.09-1.10)	1.17 (1.15-1.19)	1.42 (1.36-1.48)
≥37 weeks	Reference	Reference	Reference	Reference	Reference	Reference
Socioeconomic status						
‘Lower’	1.00 (1.00-1.00)	1.23 (1.22-1.24)	1.09 (1.07-1.11)	0.96 (0.96-0.97)	1.12 (1.11-1.13)	1.10 (1.08-1.13)
‘Higher’	Reference	Reference	Reference	Reference	Reference	Reference
Urban-rural status						
Rural	0.79 (0.78-0.79)	2.34 (2.30-2.37)	1.16 (1.13-1.20)	0.73 (0.72-0.73)	2.28 (2.25-2.31)	1.15 (1.11-1.20)
Urban	Reference	Reference	Reference	Reference	Reference	Reference

Table 4.4. Stratified adjusted models (<1 year of age and ≥1 years of age) showing incidence rate ratios for the three service types for term infants only, MCADD cohort versus primary (unmatched) comparison cohort

	Adjusted rate ratio (95% CI)					
	<1 year of age			≥1 year of age		
	Physician visits	ED visits	Hospitalizations	Physician visits	ED visits	Hospitalizations
MCADD	1.44 (1.22-1.72)	2.61 (1.68-4.20)	4.97 (2.50-10.01)	1.53 (1.22-1.95)	2.24 (1.54-3.37)	14.78 (7.63-32.00)
Sex						
Female	0.92 (0.91-0.92)	0.86 (0.85-0.87)	0.75 (0.74-0.77)	0.86 (0.86-0.87)	0.86 (0.85-0.87)	0.78 (0.76-0.80)
Male	Reference	Reference	Reference	Reference	Reference	Reference
Season of birth						
Jan. – Apr.	Reference	Reference	Reference	Reference	Reference	Reference
May – Aug.	1.01 (1.00-1.01)	1.04 (1.03-1.05)	0.98 (0.96-1.01)	0.96 (0.95-0.96)	0.95 (0.95-0.96)	0.94 (0.91-0.96)
Sept. – Dec.	1.00 (1.00-1.00)	1.02 (1.01-1.03)	1.09 (1.07-1.12)	1.00 (0.99-1.00)	1.00 (0.99-1.01)	1.02 (0.99-1.05)
Birth weight						
<2500g	1.39 (1.38-1.41)	1.08 (1.04-1.11)	1.88 (1.78-1.99)	1.12 (1.10-1.13)	1.02 (1.00-1.05)	1.60 (1.50-1.71)
≥2500g	Reference	Reference	Reference	Reference	Reference	Reference
Socioeconomic status						
‘Lower’	1.00 (1.00-1.00)	1.22 (1.21-1.24)	1.10 (1.07-1.12)	0.97 (0.96-0.97)	1.12 (1.11-1.13)	1.10 (1.08-1.13)
‘Higher’	Reference	Reference	Reference	Reference	Reference	Reference
Urban-rural status						
Rural	0.78 (0.78-0.79)	2.37 (2.33-2.41)	1.21 (1.16-1.25)	0.73 (0.72-0.73)	2.30 (2.27-2.33)	1.16 (1.11-1.21)
Urban	Reference	Reference	Reference	Reference	Reference	Reference

Table 4.5. Individual-level costs of care in the first year of life for a cohort of Ontario infants diagnosed with MCADD through newborn screening and a primary (unmatched) comparison cohort

	Costs of Care^a		
	Mean	Median	Interquartile Range
Physician visits			
MCADD	\$1,458	\$1,112	\$843 - \$1,703
Primary Comparison	\$983	\$662	\$457 - \$998
ED visits			
MCADD	\$423	\$251	\$0 - \$549
Primary Comparison	\$149	\$0	\$0 - \$163
Hospitalizations			
MCADD	\$4,116	\$1,412	\$829 - \$4,470
Primary Comparison	\$3,605	\$853	\$829 - \$1,581
Total costs of care^b			
MCADD	\$6,258	\$3,335	\$2,158 - \$6,108
Primary Comparison	\$4,746	\$1,876	\$1,373 - \$2,756

^a Costs of care are adjusted for inflation and expressed in 2012 CAD dollars.

^b Total costs of care includes all OHIP billings, inpatient hospitalization, same day surgery, ED visits, Ontario Drug Benefit, rehabilitation, complex & continuing care, home care services, and assisted device costs.

REFERENCES

1. Rhead WJ. Newborn screening for medium-chain acyl-CoA dehydrogenase deficiency: a global perspective. *J Inher Metab Dis.* 2006 Apr-Jun;29(2-3):370-7.
2. Wilcken B, Hammond J, Silink M. Morbidity and mortality in medium chain acyl coenzyme A dehydrogenase deficiency. *Arch Dis Child.* 1994 May;70(5):410-2.
3. Oerton J, Khalid JM, Besley G, Dalton RN, Downing M, Green A, et al. Newborn screening for medium chain acyl-CoA dehydrogenase deficiency in England: prevalence, predictive value and test validity based on 1.5 million screened babies. *J Med Screen.* 2011;18(4):173-81.
4. Grosse SD, Khoury MJ, Greene CL, Crider KS, Pollitt RJ. The epidemiology of medium chain acyl-CoA dehydrogenase deficiency: an update. *Genet Med.* 2006 Apr;8(4):205-12.
5. Wang SS, Fernhoff PM, Hannon WH, Khoury MJ. Medium chain acyl-CoA dehydrogenase deficiency human genome epidemiology review. *Genet Med.* 1999 Nov-Dec;1(7):332-9.
6. Gregersen N, Andresen BS, Corydon MJ, Corydon TJ, Olsen RK, Bolund L, et al. Mutation analysis in mitochondrial fatty acid oxidation defects: Exemplified by acyl-CoA dehydrogenase deficiencies, with special focus on genotype-phenotype relationship. *Hum Mutat.* 2001 Sep;18(3):169-89.
7. Ziadeh R, Hoffman EP, Finegold DN, Hoop RC, Brackett JC, Strauss AW, et al. Medium chain acyl-CoA dehydrogenase deficiency in Pennsylvania: neonatal screening shows high incidence and unexpected mutation frequencies. *Pediatr Res.* 1995 May;37(5):675-8.
8. Wilcken B, Carpenter K, Wiley V. Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency. *Lancet.* 2002 Feb;359(9306):627-8.
9. Pollitt RJ, Leonard JV. Prospective surveillance study of medium chain acyl-CoA dehydrogenase deficiency in the UK. *Arch Dis Child.* 1998 Aug;79(2):116-9.

10. Iafolla AK, Thompson RJ, Jr., Roe CR. Medium-chain acyl-coenzyme A dehydrogenase deficiency: clinical course in 120 affected children. *J Pediatr.* 1994 Mar;124(3):409-15.
11. 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.
12. Leonard JV, Dezateux C. Newborn screening for medium chain acyl CoA dehydrogenase deficiency. *Arch Dis Child.* 2009 Mar;94(3):235-8.
13. Potter BK, Little J, Chakraborty P, Kronick JB, Evans J, Frei J, et al. Variability in the clinical management of fatty acid oxidation disorders: results of a survey of Canadian metabolic physicians. *J Inherit Metab Dis.* 2012 Jan;35(1):115-23.
14. Wilcken B, Haas M, Joy P, Wiley V, Chaplin M, Black C, et al. Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency in Australia: a cohort study. *Lancet.* 2007 Jan;369(9555):37-42.
15. Schatz UA, Ensenauer R. The clinical manifestation of MCAD deficiency: challenges towards adulthood in the screened population. *J Inherit Metab Dis.* 2010 Oct;33(5):513-20.
16. Dixon MA, Leonard JV. Intercurrent illness in inborn errors of intermediary metabolism. *Arch Dis Child.* 1992 Nov;67(11):1387-91.
17. Health Quality Ontario. Neonatal screening of inborn errors of metabolism using tandem mass spectrometry: an evidence-based analysis. *Ont Health Technol Assess Ser.* 2003;3(3):1-36.
18. Wilcken B. Screening for disease in the newborn: the evidence base for blood-spot screening. *Pathology.* 2012 Feb;44(2):73-9.

19. Dezateux C. Newborn screening for medium chain acyl-CoA dehydrogenase deficiency: evaluating the effects on outcome. *Eur J Pediatr*. 2003 Dec;162 Suppl 1:S25-8.
20. 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 Pediatr*. 2010 Nov;10:82.
21. Discharge Abstract Database Data Quality Documentation - Current-Year Information, 2011-2012. Toronto: Canadian Institute for Health Information; 2012.
22. National Ambulatory Care Reporting System Data Quality Documentation- Current-Year Information, 2011-2012. Toronto: Canadian Institute for Health Information; 2012.
23. Collier R. Shift toward capitation in Ontario. *CMAJ*. 2009 Nov;181:668-9.
24. Statistics Canada. 2006 Census Dictionary. Ottawa: Statistics Canada; 2007.
25. Wang C, Guttman A, To T, Dick PT. Neighborhood income and health outcomes in infants: how do those with complex chronic conditions fare? *Arch Pediatr Adolesc Med*. 2009 Jul;163(7):608-15.
26. Kralj B. Measuring Rurality - RIO2008 BASIC: Methodology and Results. Ottawa: Ontario Medical Association; 2010.
27. Rural Physician Eligibility Requirements. Toronto: Ontario Ministry of Health and Long-Term Care; 2013 [cited 2014 April 1]. Available from: http://www.healthforceontario.ca/en/Home/Physicians/Training_%7C_Practising_in_Ontario/Locum/Rural_Family_Medicine_Locum_Program/Rural_Physicians/Eligibility.
28. Emergency Departments and Children in Ontario. Toronto: Canadian Institute for Health Information; 2008.

29. Cohen E, Berry JG, Camacho X, Anderson G, Wodchis W, Guttman A. Patterns and costs of health care use of children with medical complexity. *Pediatrics*. 2012 Dec;130(6):e1463-70.
30. Selvin S. *Statistical Analysis of Epidemiological Data*. 3rd ed. New York: Oxford University Press; 2004.
31. Vuong QH. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*. 1989;57(2):307-33.
32. Bollen KA, Jackman RW. Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. *Sociol Methods Res*. 1985;13(4):510-42.
33. Belsley DA, Kuh E, Welsch RE. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Hoboken: Wiley-Interscience; 2005.
34. Wodchis W, Bushmeneva K, Nikitovic M, McKillop I. *Guidelines on Person-Level Costing Using Administrative Databases in Ontario*. Toronto: Health System Performance Research Network (HSPRN); 2013.
35. Newacheck PW, Kim SE. A national profile of health care utilization and expenditures for children with special health care needs. *Arch Pediatr Adolesc Med*. 2005 Jan;159(1):10-7.
36. Shatin D, Levin R, Ireys HT, Haller V. Health care utilization by children with chronic illnesses: a comparison of medicaid and employer-insured managed care. *Pediatrics*. 1998 Oct;102(4):E44.
37. Lanzkron S, Carroll CP, Haywood C, Jr. The burden of emergency department use for sickle-cell disease: an analysis of the national emergency department sample database. *Am J Hematol*. 2010 Oct;85(10):797-9.
38. Shankar SM, Arbogast PG, Mitchel E, Cooper WO, Wang WC, Griffin MR. Medical care utilization and mortality in sickle cell disease: a population-based study. *Am J Hematol*. 2005 Dec;80(4):262-70.

39. Yusuf HR, Atrash HK, Grosse SD, Parker CS, Grant AM. Emergency department visits made by patients with sickle cell disease: a descriptive study, 1999-2007. *Am J Prev Med.* 2010 Apr;38(4 Suppl):S536-41.
40. Meng M, Zhang Y-P. Impact of inborn errors of metabolism on admission in a neonatal intensive care unit: a 4-year report. *J Pediatr Endocrinol Metab.* 2013;26(7-8):689-93.
41. Couce ML, Bana A, Boveda MD, Perez-Munuzuri A, Fernandez-Lorenzo JR, Fraga JM. Inborn errors of metabolism in a neonatology unit: Impact and long-term results. *Pediatr Int.* 2011 Feb;53(1):13-7.
42. Jouvett P, Touati G, Lesage F, Dupic L, Tucci M, Saudubray JM, et al. Impact of inborn errors of metabolism on admission and mortality in a pediatric intensive care unit. *Eur J Pediatr.* 2007 May;166(5):461-5.
43. Johnson C, Butler SM, Konstan MW, Morgan W, Wohl ME. Factors influencing outcomes in cystic fibrosis: a center-based analysis. *Chest.* 2003 Jan;123(1):20-7.
44. Grosse SD, Schechter MS, Kulkarni R, Lloyd-Puryear MA, Strickland B, Trevathan E. Models of comprehensive multidisciplinary care for individuals in the United States with genetic disorders. *Pediatrics.* 2009;123(1):407-12.
45. Baker JR, Crudder SO, Riske B, Bias V, Forsberg A. A model for a regional system of care to promote the health and well-being of people with rare chronic genetic disorders. *Am J Public Health.* 2005 Nov;95(11):1910-6.
46. Padman R, McColley SA, Miller DP, Konstan MW, Morgan WJ, Schechter MS, et al. Infant care patterns at epidemiologic study of cystic fibrosis sites that achieve superior childhood lung function. *Pediatrics.* 2007 Mar;119(3):e531-7.
47. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10). Geneva: World Health Organization (WHO); 1992.
48. Statistics Canada. Data quality, concepts and methodology: Definitions. Ottawa: Statistics Canada; 2013.

49. Shi X, Duell E, Demidenko E, Onega T, Wilson B, Hoftiezer D. A polygon-based locally-weighted-average method for smoothing disease rates of small units. *Epidemiology*. 2007 Sep;18(5):523-8.
50. Simonoff J. *Springer Series in Statistics: Smoothing Methods in Statistics*. New York: Springer-Verlag; 1996.
51. Wang F. Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Ann Assoc Am Geogr*. 2012;102(5):1104-12.
52. HealthForceOntario Northern and Rural Recruitment and Retention Initiative Guidelines. Sudbury: Ontario Ministry of Health and Long-Term Care; 2013 [cited 2014 April 1]. Available from: <http://www.health.gov.on.ca/en/pro/programs/northernhealth/nrrr.aspx>.
53. Janicke DM, Finney JW, Riley AW. Children's health care use: a prospective investigation of factors related to care-seeking. *Med Care*. 2001 Sep;39(9):990-1001.
54. Riley AW, Finney JW, Mellits ED, Starfield B, Kidwell S, Quaskey S, et al. Determinants of children's health care use: an investigation of psychosocial factors. *Med Care*. 1993 Sep;31(9):767-83.

INTERFACE

While the implementation of newborn screening programs has allowed newborns at risk for rare, treatable diseases to be identified, the recent expansion of these public health initiatives has led to an increase in the number of newborns presenting with false positive newborn screening results. Findings in the literature with respect to early health care services utilization among children with false positive newborn screening results have been inconsistent to date. This manuscript focused on describing the patterns of health services use in early life in a cohort of Ontario infants who received false positive newborn screening results for medium-chain acyl-CoA dehydrogenase deficiency from April 1, 2006 through March 31, 2010; and comparing these patterns to health services use in a screen negative birth cohort. Provincial newborn screening diagnostic confirmation data was securely linked at the individual level with population-based health care administrative datasets at the Institute for Clinical Evaluative Sciences. Characteristics known or postulated to be related to false positive newborn screening results and patterns of early health care services use, such as gestational age and birth weight, were included in the analyses as potential confounding variables.

This manuscript was co-authored by the student (MK); her thesis co-supervisors, Drs Beth Potter and Pranesh Chakraborty; and the members of her thesis advisory committee, Drs Doug Coyle and Kumanan Wilson. The student was responsible for planning the analytical protocol, completing the analysis and drafting the manuscript. The thesis co-supervisors and thesis advisory committee members provided guidance throughout the data analysis and manuscript writing process. In

addition, Dr Pranesh Chakraborty provided clinical expertise with regards to false positive newborn screening results, while Dr Beth Potter was integral towards developing the methodological protocol.

CHAPTER FIVE: HEALTH SYSTEM IMPACT OF FALSE POSITIVE

RESULTS – MANUSCRIPT #3

The health system impact of false positive newborn screening results for medium-chain acyl-CoA dehydrogenase deficiency in Ontario, Canada

Maria D Karaceper, BSc,¹; Pranesh Chakraborty, MD, FRCPC,²; Doug Coyle, MSc, PhD,¹; Kumanan Wilson, MD, FRCPC,³; Beth K Potter, MSc, PhD,¹; and in collaboration with the Canadian Inherited Metabolic Diseases Research Network

Affiliation:

1 Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada

2 Newborn Screening Ontario, Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada

3 Ottawa Hospital Research Institute and Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada

Keywords: newborn screening, false positive results, metabolic diseases, health care service utilization

Manuscript Status: To be submitted to Pediatrics

ABSTRACT

Objective: There is no consensus in the published literature to date regarding the impact of false positive newborn screening results on early health care utilization patterns. Using medium-chain acyl-CoA dehydrogenase deficiency (MCADD) as an example, we aimed to evaluate the health system impact of false positive newborn screening results in a cohort of Ontario infants.

Methods: Health care administrative datasets (physician visits, emergency department care, and inpatient hospitalizations) from April 1, 2006 through March 31, 2012 were linked to provincial newborn screening diagnostic confirmation data. Unadjusted and adjusted incidence rate ratios were used to compare those with false positive results for MCADD to a birth cohort with negative newborn screening results, overall and stratified by age at the time of receiving care.

Results: We identified 43 infants with a false positive newborn screening result for MCADD between April 1, 2006 and March 31, 2010 in Ontario. These infants experienced statistically significant higher rates of physician visits and hospitalizations in the first year of life relative to screen negative birth cohorts in adjusted analyses. When restricted to term infants, this population continued to have statistically significant higher rates of physician visits. Differences in health services use were not observed after the first year of life.

Conclusion: Explanations for the higher use of some health services among false positive infants during the first year of life include residual confounding, visits related

to follow-up confirmatory diagnostic testing, or the psychosocial impact of false positive screening results on parental perceptions of infant health.

INTRODUCTION

Population-based newborn screening programs are used to pre-symptomatically identify newborns with rare but treatable conditions. While in place since the early 1960s to detect the disorder phenylketonuria, many programs throughout the world including those in North America have expanded within the last decade, in part in response to the availability of tandem mass spectrometry. This analytical technique allows for the simultaneous detection of risk for over 20 inherited metabolic diseases (1, 2). Despite the high specificities associated with newborn screening tests, the positive predictive values of these tests are often relatively low due to the low birth prevalence of the screened diseases in the population (3). The expansion of newborn screening panels has consequently resulted in an increased number of newborns presenting with false positive results.

Several studies have aimed to describe the short- and long-term impact of false positive newborn screening results in terms of parental psychosocial experiences, however there is no consensus on the presence and magnitude of any potential harms (4-9). Variation in study findings similarly exists regarding the impact of any psychosocial effect of a false positive result on patterns of early healthcare service use. While some studies demonstrate an increased frequency of emergency department visits and hospitalizations among children with false positive newborn screening results during the first few months of life as compared to a screen negative cohort (4, 5, 7), others have not found such increases (3, 10). An inherent challenge in this literature lies in ascribing any observed increases in health services use to the psychosocial impact of receiving a false positive result. Infants born preterm or with a low birth

weight are more likely to receive positive newborn screening tests results in the absence of disease, due to the underlying biology associated with prematurity and the nature of biochemical markers used in the screening tests (10-12). Thus, increased health services use in the first years of life may be attributed to health needs associated with prematurity or fetal growth restriction rather than to the screening result.

Medium-chain acyl-CoA dehydrogenase deficiency (MCADD), an inherited mitochondrial fatty acid oxidation disorder, is one of the most common inborn errors of metabolism and is included on newborn screening panels in most North American jurisdictions (13, 14). Deleterious mutations on the *ACADM* gene result in the formation of an enzyme that is unable catalyze the initial step in the mitochondrial beta-oxidation pathway for medium-chain fatty acids. This step is critical towards the formation of ketone bodies in the liver, an essential alternative energy source; the body's ability to adequately respond in times of physiological stress, such as intercurrent illness and prolonged fasting, is thus hindered in children with MCADD (14, 15). During these periods, individuals with MCADD are susceptible to experience potentially lethal acute metabolic crises (16). However, if early diagnosis is established, prognosis is excellent since the vast majority of adverse outcomes are preventable with the immediate initiation of treatment (14, 15, 17, 18).

Newborn Screening Ontario, a provincial public health program, is responsible for coordinating the newborn bloodspot program for approximately 140,000 babies born in Ontario, Canada each year; MCADD was added to the panel in April 2006 (19). The vast majority of Ontario residents are eligible for universal health coverage through the Ontario Health Insurance Plan (OHIP), which encompasses a wide range of

medically necessary health services. Billed services are captured within population-based health administrative databases housed at the Institute for Clinical Evaluative Sciences, which can be securely linked to provincial newborn screening data. These provincial databases are near complete and also include characteristics such as gestational age and birth weight (20-22). This presents a unique opportunity to investigate the impact of a false positive newborn screening result while adjusting for these potentially important confounders. Using MCADD as an example disorder, the objective of this study was to evaluate if children who have received false positive newborn screening results experienced higher rates of health services use in the first few year of life relative to a screen negative cohort in Ontario, Canada.

METHODS

This study was reviewed and approved by the Ottawa Health Science Network Research Ethics Board.

Study Population and Data Sources

The source population for this study was all children born in Ontario and screened by Newborn Screening Ontario between April 1, 2006 and March 31, 2010. The false positive study cohort included children who received a positive newborn screening result for MCADD but were determined to be unaffected based on confirmatory testing. A primary comparison cohort included all children with negative newborn screening results for all disorders during the same time period. To achieve better control over potential confounders, a secondary matched comparison cohort was formed using factors that are known to be strongly associated with health services

utilization; specifically, ten controls were randomly matched to each child with a false positive result for MCADD, based on sex, calendar year of birth, a measure of urban-rural status of the child's residence, and a neighbourhood-level proxy indicator of socioeconomic status. Due to incomplete data on matching criteria, controls could not be matched to one false positive child; this child was thus excluded from analyses involving the secondary comparison cohort. Individuals were excluded from the study if they were ineligible for universal health care coverage at the time of birth through OHIP (this is extremely rare and mainly affects recent immigrants to Ontario), were diagnosed with MCADD through newborn screening, received a positive newborn screening result for another disorder on the panel, or were deceased within 24 hours following birth; the latter group was excluded since a bloodspot sample must be collected at greater than 24 hours of age in Ontario to be considered satisfactory.

Newborn screening diagnostic confirmation data for the relevant cohorts were securely linked to Ontario's Registered Persons Database, the provincial health care patient registry, at the Institute for Clinical Evaluative Sciences via encrypted health card numbers. Subsequent linkage with population-based healthcare administrative databases was completed, encompassing health service visits from April 1, 2006 through March 31, 2012. The OHIP database encompasses all primary care and specialist physician billing data; in the last decade, there has been a shift away from a fee-for-service model towards capitated payment approaches (20). Emergency department (ED) care data were retrieved from the Canadian Institute for Health Information's National Ambulatory Case Reporting System, which reports information for all hospital- and community-based ambulatory care services; almost 100% of ED

claims in Ontario are captured within this electronic system (21). Finally, inpatient hospitalization data were obtained from the Canadian Institute for Health Information's Discharge Abstract Database, which consists of administrative and clinical information from all acute inpatient facilities throughout Ontario (22).

Independent Variables

It was important to control for variables that were associated with having a false positive screening result and were also strong predictors of health service utilization; independent variables of interest included in the analysis were sex, season of birth, birth weight, gestational age, an area-based proxy measure of relative socioeconomic status, and a measure of urban-rural status. The measured concentration of the primary analyte used in the screening algorithm for MCADD (octanoylcarnitine, C8) was also available for the false positive cohort and thus was included in descriptive analysis for that cohort only. Independent variables were ascertained from hospital records at the time of birth (sex, birth weight, gestational age), Canadian census data linked to other datasets by the child's residential postal code at the time of birth (socioeconomic status, urban-rural status) and the newborn screening records (concentration of C8). In order to account for potential seasonal variation in care, three categories for season of birth were considered: January – April, May – August, and September – December. For birth weight, children were separated into two groups for the analysis: low birth weight (<2,500 grams) and normal or high birth weight (≥ 2500 g). Similarly, we dichotomized gestational age to pre-term (<37 weeks) and term/post-term (≥ 37 weeks) births.

A proxy measure of socioeconomic status was defined as the neighborhood-level income quintile. This variable was derived using small area-level socioeconomic information from the 2006 Canadian Census linked to a child's residential postal code at the time of birth (postal code was determined through hospitalization data); this linkage was accomplished using Statistic Canada's Postal Code Conversion File. We relied on the household-size-adjusted mean neighborhood income value per single person at the dissemination area-level (23). Dissemination areas have a population size of approximately 400 to 700 persons in Canada. Each dissemination area within a larger census area is divided into five income quintile categories according to ranked average neighborhood household income values (24). For our study, the two lowest quintiles [1 and 2] were merged to form a group with a "lower" relative socioeconomic status, while the upper three quintiles [3, 4, and 5] were combined to create a "higher" relative socioeconomic status group.

Urban-rural status was defined based on the revised Rurality Index for Ontario, RIO-2008, developed by the Ontario Medical Association and the Ontario Ministry of Health and Long-Term Care in 2008 (25). This measure was designed in response to the need for a continuous measure of rurality with respect to access to health care services in order to inform incentives aimed at physician recruitment and retention in Ontario. It is based on three geographical components: the community population and population density, the travel time to the nearest tertiary care centre, and the travel time to the nearest Level 2 centre (25). Rural physician eligibility requirements in Ontario defined a rural community as one whose RIO2008 score was 40 or greater, thus we established this value as a cut-off for rural status in our study cohorts (26).

Utilization Outcomes

We included every original health care service encounter made by a child within the study period (primary care/specialist physician visits, ED visits, and hospitalizations) in the analysis. If a child had multiple OHIP billed procedures within a single physician visit, these were collectively considered as one original encounter. However, if a child saw multiple physicians on any given day, these were considered as separate visits. Similarly, if a child visited the ED more than once on the same date, each visit was considered an original encounter. A child could only have a single inpatient hospitalization on any given day.

Statistical Analysis

De-identified, study-specific datasets were accessed at the Institute for Clinical Evaluative Sciences; cells sizes less than six were not reported in accordance with stringent confidentiality and privacy policies. Descriptive statistics were used to characterize the independent variables and health care utilization outcomes of interest. Counts and percentages were calculated for categorical data, while medians and ranges were presented for continuous variables. Chi-square tests were used to examine the bivariate associations between geographic and sociodemographic characteristics and cohort membership (false positive versus screen negative).

The number of physician visits, ED visits, and hospitalizations during the study period were respectively summed for each child in the study population. The length of follow-up for each individual was subsequently calculated as the time elapsed between the date of birth and one of the following three times points in years: the date of OHIP

eligibility loss (a rare occurrence, mainly related to emigration from Ontario), the date of death, or the last date of follow-up for the study. The earliest of these three possible dates was selected as the end point for the calculation of follow-up time. Unadjusted visit rates and incidence rate ratios (IRR) for each type of visit category along with 95% confidence intervals were calculated to compare the false positive and screen negative cohorts. These analyses were carried out on the full study sample and with stratification based on the child's age at the time of visit (<1 year of age versus ≥ 1 year of age).

We relied on the Vuong statistic, a likelihood ratio-based statistical test, to compare four regression models for calculating adjusted visit IRR for each health care service type: Poisson, negative binomial, zero-inflated Poisson, and zero-inflated negative binomial (27). Based on the results of the Vuong test, we selected a negative binomial regression model to investigate the IRR for health care services in the false positive versus comparison cohorts while adjusting for potential confounding factors. Influential observations were identified using standardized Pearson residuals, Cook's D statistics, and DFBETA statistics. An observation was considered influential if the following three conditions were met: the standardized Pearson residual was greater than the absolute value of 2, the value of Cook's D statistic was larger than $4/n$, and the value of the DFBETA statistic was greater than $2/n$, where n is equal to the number of observations (28, 29). Distinct 99th percentile values were calculated for the three service types in each of the three cohorts, and then applied to the appropriate influential observations for truncation. In analyses with the matched comparison cohort, we adjusted for season of birth, birth weight, and gestational age; sex,

socioeconomic status, and urban-rural status were additionally included in unmatched analyses. All models were stratified by age at the time of the visit (<1 year of age and ≥1 year of age). In order to account for a potential residual effect of premature birth, we re-ran our final models restricted to children with term births (≥37 weeks of gestation) as a sensitivity analysis. All statistical analyses were performed using SAS® software version 9.3 (SAS Institute Inc., Cary, North Carolina, USA).

RESULTS

Study Population

Forty-three children had false positive newborn screening test results for MCADD through Newborn Screening Ontario during the specified study period. The primary and secondary matched comparison cohorts consisted of 545,355 and 420 children respectively. We were unable to report the number of children with false positive results for MCADD that lived in a rural setting due to confidentiality and privacy policies. Children with false positive results for MCADD were more likely to be male, have a birth weight <2500g and have a gestational age <37 weeks as compared to the primary comparison group; no other statistically significant differences between the cohorts were found (Table 5.1). Those with false positive screening results for MCADD had a median C8 screening level of 0.5 umol/L (range: 0.3-1.2). All cohorts included children who died before the end of the follow-up period, thus follow-up time in the entire study population ranged from less than one month to approximately 6 years of age.

Health Care Visit Rates

Children with false positive newborn screening results for MCADD had a total of 1939 recorded physician visits over the entire follow-up period, with an average unadjusted visit rate of 10.98 physician visits per child per year (Figure 5.1). Children in the primary and secondary comparison groups had average rates of 8.44 and 9.00 physician visits per child per year respectively. Those with false positive results for MCADD had a total of 105 ED visits over the follow-up period with an average rate of 0.60 ED visits per child per year; those in the comparison cohorts had averages of 0.66 and 0.72 ED visits per child per year respectively. Finally, children with false positive newborn screening results for MCADD had a total of 27 inpatient hospitalizations over the follow-up period with an average rate of 0.15 hospitalizations per child per year. Those in the primary comparison group had an inpatient hospitalization rate of 0.06 stays per child per year, while the secondary comparison group had a hospitalization rate of 0.08 stays per child per year.

Unadjusted visit rates for each type of health service visit were then stratified by age at the time of visit. Children in all 3 cohorts had the highest physician use, ED visits, and hospitalization rates during their first year of life (Figure 5.1). In the first year of life, children with false positive newborn screening results for MCADD had an average of 26.1 physician visits, 1.0 ED visits, and 0.5 hospital admissions per child. These rates decreased to, on average, 6.3 physician visits, 0.5 ED visits, and 0.05 hospitalizations per child per year in the subsequent years of age. In the first year of life, those in the primary comparison group had on average 14.7 physician visits, 0.8 ED visits, and 0.1 hospitalizations per child; these rates similarly decreased in the subsequent years of

age to an average of 6.3 physician visits, 0.6 ED visits, and 0.04 hospitalizations per child per year. Finally, in the first year of life, children in the matched comparison group had on average 18.0 physician visits, 0.9 ED visits, and 0.2 hospitalizations per child; these rates also decreased in the subsequent years of age to an average of 6.6 physician visits, 0.7 ED visits, and 0.05 hospitalizations per child per year.

Incidence Rate Ratios for Individual Health Care Service Types

In adjusted analysis, children with false positive newborn screening results for MCADD had statistically significant higher rates of physician visits (IRR: 1.22 [95% CI: 1.03–1.46]) as compared to the primary comparison group (results for all ages, not shown in tables). No significant differences were noted for the false positive versus primary comparison cohorts for ED visits (IRR: 0.96 [95% CI: 0.67–1.40]) and hospitalizations (IRR: 1.57 [95% CI: 0.82–3.01]). The IRRs were very similar when the matched cohort was used as the comparison group.

These models were subsequently stratified by the age of the child at the time of visit (Table 5.2). In the first year of life, after adjusting for a range of potential confounding factors, children with false positive results for MCADD had a statistically significant higher rate of physician visits (IRR: 1.42 [95%CI: 1.21–1.67]) and inpatient hospitalizations (IRR: 2.32 [95%CI: 1.22–4.34]) as compared to the primary comparison group; however, there was no significant difference in the frequency of ED visits (IRR: 1.27 [95%CI: 0.80–2.03]). There were no statistically significant differences for any of the three service types in the false positive versus primary comparison cohort among children older than 1 year of age, however the magnitude of the effect

size for hospitalizations was relatively large and reflective of a lower rate of hospitalization in the false positive cohort (IRR: 0.68 [95%CI: 0.15–2.25]). These results were similar in the matched comparison cohort, however an exception was noted in ED visits made during the first year of life (IRR: 0.81 [95%CI: 0.52–1.27]) (Table 5.3). This IRR suggested that children with false positive newborn screening results with MCADD had a lower frequency of ED visits as compared to the matched comparison cohort in the first year of life, however the difference was not statistically significant.

As a sensitivity analysis, the stratified models with the primary comparison cohort were subsequently restricted to children with term births, to account for the potential residual confounding effect of premature births (Table 5.4). The results were similar to those reported in the full model (Table 5.2). In the first year of life, term children with false positive screening results for MCADD had a statistically significant higher rate of physician visits (IRR: 1.35 [95%CI: 1.13–1.63]) as compared to term children with negative screening results. While there was no statistically significant difference in hospitalization rates in the two cohorts among those less than 1 year of age, the estimated IRR (IRR: 2.32 [95%CI: 0.87–5.65]) was virtually identical to the IRR from the full model (IRR: 2.32 [95%: 1.22–4.34]). Similar to the full model, no statistically significant differences were found for each of the three service types for children aged one year and older although the magnitude of the effect size for inpatient hospitalizations remained relatively large and reflected a lower rate in the false positive cohort relative to the screen negative cohort (IRR: 0.65 [95%CI: 0.10–2.60]).

DISCUSSION

While the psychosocial consequences of a false positive newborn screening test are short-term for the majority of families, the available evidence suggests that a small proportion of parents continue to experience residual anxiety following diagnostic testing (30, 31). Some studies have shown that this anxiety can alter parental perceptions of a child's health and vulnerability to illness, which can result in those families seeking health care services at an increased frequency for their child; nevertheless, conflicting results on health system impact of false positive results are found in the literature (3-5, 7).

In this study, we found that children with false positive newborn screening results for MCADD accessed health services at statistically significant higher rates as compared to those with negative newborn screening results over the span of the study. These increased rates could be explained by multiple factors including increased parental anxiety, parental perceptions of the child's health, or actual health-related characteristics of the child (3, 12). With respect to the latter explanation, false positive rates in newborn screening tests are typically higher in newborns who are born preterm or have low birth weights (10-12) and, indeed, children with false positive results for MCADD in this study were more likely to have a gestational age <37 weeks and have a birth weight <2500g as compared to the primary comparison group. Since both of these characteristics are respectively associated with higher rates of health services use, analysis focusing on early healthcare utilization patterns among false positive children must control for these variables (10-12).

After adjustment for birth weight, gestational age, and additional potential confounders, we found that children with false positive newborn screening results for MCADD still experienced a statistically significant higher frequency of physician visits and hospitalizations, but not ED visits, as compared to both screen negative comparison groups in the first year of life. One potential explanation for this result is that it reflects supplementary health service visits related to the confirmatory diagnostic testing for the screen positive newborn screening result. This explanation is supported by the finding that these differences in health care use were no longer apparent among children older than one year of age. However, while this reasoning could help to explain the higher rate of physician visits among infants with false positive results, it seems unlikely that hospitalizations would be required to rule out MCADD. Further, while non-significant, the magnitude of the effect size for inpatient hospitalizations remained large in each stratified model for children one year of age or older and reflected a lower rather than higher hospitalization rate in the false positive cohort. Since we did not exclude those with severe chronic illnesses that could require additional hospitalizations, it is possible that an increased frequency of hospitalizations experienced by a subset of ill children from the comparison cohort contributed to the effect size.

Another possible explanation for the higher rate of health services use in the first year of life in children with false positive results is a potential residual confounding effect of premature birth. To explore this possibility, we restricted the models to children with gestational age ≥ 37 weeks. While the higher frequency of hospitalizations in the false positive cohort during the first year of life was no longer

statistically significant in this population of term infants, the IRR (2.32) was virtually identical to the IRR in the original analysis. Thus, while we cannot rule out chance as an explanation for this doubling in frequency of hospitalizations, the findings do not support residual confounding by gestational age as the explanation for the main results. There also continued to be a statistically significant higher frequency of physician visits in the false positive group in the sensitivity analysis restricted to term infants.

Collectively, the results suggest that there may be an additional variable influencing the healthcare service patterns observed during the first year of life among those infants with false positive newborn screening results for MCADD, such as the psychosocial impact of the false positive screening result and its possible influence on parental perceptions of infant health. However, a limitation with using health care administrative datasets is the lack of clinical and patient- and family-reported information; as a result, we are unable to explore the role of psychosocial factors. Should additional research support a potential psychosocial effect leading to higher health service rates, the findings support the use of strategies that have been shown to mitigate parental stress and anxiety in the face of false positive findings, including improved education and identifying best practices for effective provider communication during confirmatory follow-up (30, 32).

An additional limitation of this study that is common to rare disease research was the challenge in interpreting differences in health services use given the small patient populations. As a result, the number of stratified analyses that we could perform was limited due to privacy standards for small numbers. For example, we were unable to distinguish between extremely low birth weight (<1,000 grams), very low

birth weight (<1,500 grams), and high birth weight ($\geq 4,500$ grams) as suggested in the literature (33, 34). As a result, we cannot completely rule out residual confounding as an explanation for our findings. Further sociodemographic and geographic stratification will become possible as the cohort size grows and follow-up period increases.

CONCLUSION

Newborns who are born preterm or have low birth weights are more likely to receive a positive newborn screening result in the absence of disease. With this in mind, after adjustment for a range of potential confounders including birth weight and gestational age, we found that children with false positive newborn screening results for MCADD experience statistically significantly higher rates of physician visits and hospitalizations in the first year of life, relative to children with negative screening results. However, there were no significant differences between the cohorts in health services use in the subsequent years of life. Potential explanations for this finding include the effect of health care use related to confirmatory diagnostic testing, residual confounding, or the potential psychosocial impact of false positive results on parental perceptions of infant health. By linking additional clinical and family-reported information to the health care administrative data, future studies will be able to investigate whether parental psychosocial effects are partially responsible for our study findings.

Table 5.1. Geographic and sociodemographic characteristics of the study population

	Study Cohort		
	False Positive, n (%) (n=43)	Primary Comparison, n (%) (n=545 355)	Secondary Comparison, n (%) (n=420)
Sex^{a,c}			
Male	28 (65.1)	279 638 (51.3)	-
Female	15 (34.9)	265 717 (48.7)	-
Month of Birth			
January – April	19 (44.2)	177 918 (32.6)	142 (33.8)
May – August	8 (18.6)	192 896 (35.4)	134 (31.9)
September – December	16 (37.2)	174 541 (32.0)	144 (34.3)
Birth weight^d			
<2500g	6 (14.0)	33 027 (6.1)	28 (6.7)
≥2500g	37 (86.0)	508 466 (93.2)	390 (92.9)
Gestational age^e			
<37 weeks	11 (25.6)	42 235 (7.7)	34 (8.1)
≥37 weeks	32 (74.4)	489 232 (89.7)	373 (88.8)
Relative income^a			
‘Lower’	21 (48.8)	232 269 (42.6)	-
‘Higher’	21 (48.8)	310 003 (56.8)	-
Urban-rural status^{a,b}			
Rural	<6 (≤11.6)	34 111 (6.3)	-
Urban	38–43 (≥88.4)	505 236 (92.6)	-

^a Counts are not presented for secondary matched comparison group on matched variables

^b Results are repressed for cell sizes <6

^c $P < 0.10$ for difference in proportion in the false positive cohort versus the comparison cohort

^d $P < 0.05$ for difference in proportion in the false positive cohort versus the comparison cohort

^e $P < 0.01$ for difference in proportion in the false positive cohort versus the comparison cohort

Figure 5.1. Unadjusted health service visit rates stratified by age at the time of visit

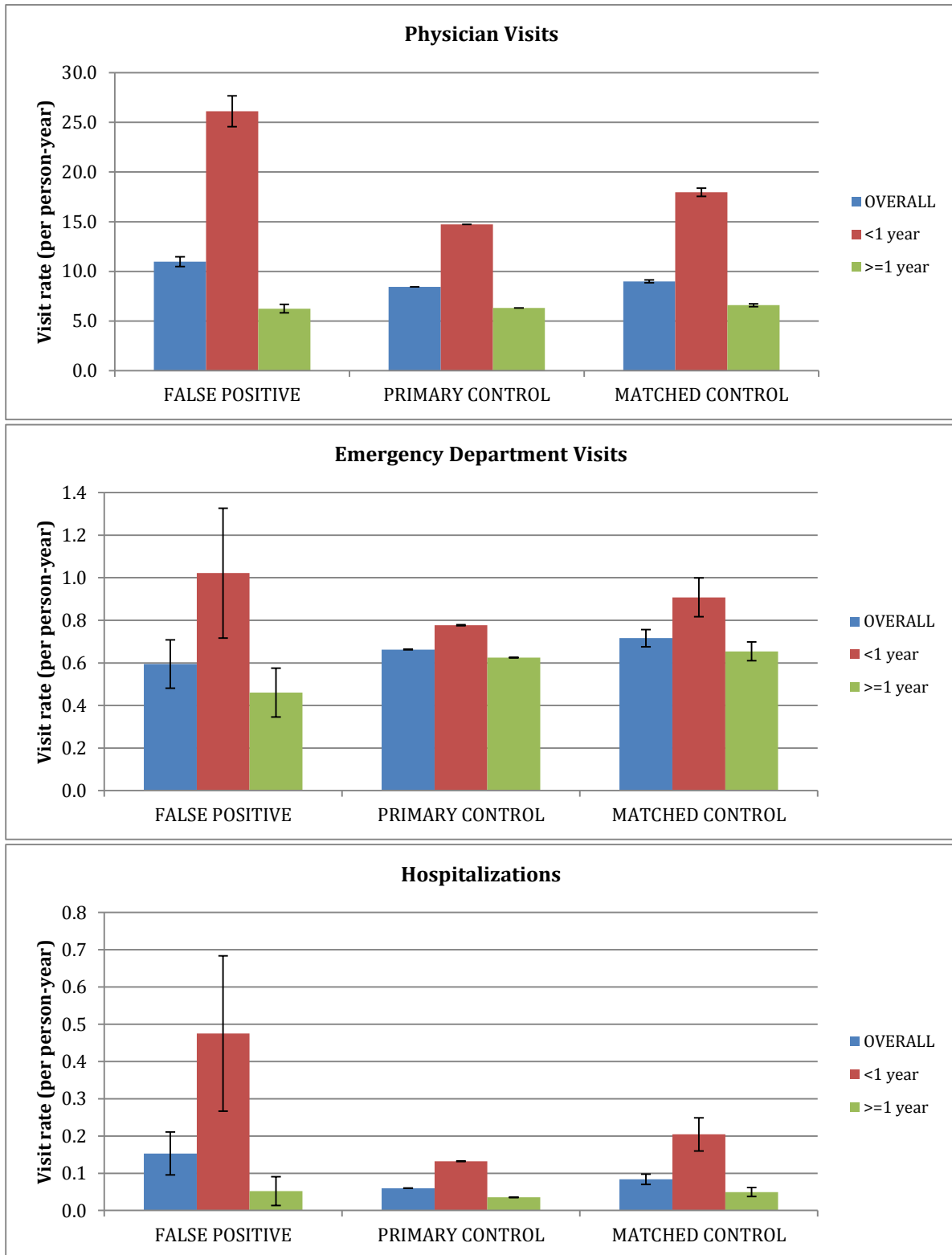


Table 5.2. Stratified adjusted models (<1 yr old and ≥1 yr old) showing incidence rate ratios for the three service types, comparing the false positive cohort with the primary (unmatched) comparison cohort

	Adjusted incidence rate ratio (95% CI)					
	<1 year of age			≥1 year of age		
	Physician visits	ED visits	Hospitalizations	Physician visits	ED visits	Hospitalizations
False Positive	1.42 (1.21-1.67)	1.27 (0.80-2.03)	2.32 (1.22-4.34)	0.96 (0.77-1.20)	0.82 (0.54-1.26)	0.68 (0.15-2.25)
Sex						
Female	0.92 (0.91-0.92)	0.85 (0.85-0.86)	0.76 (0.75-0.77)	0.91 (0.91-0.92)	0.86 (0.85-0.87)	0.78 (0.77-0.80)
Male	Reference	Reference	Reference	Reference	Reference	Reference
Season of birth						
Jan. – Apr.	Reference	Reference	Reference	Reference	Reference	Reference
May – Aug.	1.01 (1.00-1.01)	1.04 (1.03-1.05)	0.98 (0.96-1.00)	0.96 (0.95-0.96)	0.95 (0.95-0.96)	0.94 (0.92-0.97)
Sept. – Dec.	1.00 (1.00-1.01)	1.02 (1.00-1.03)	1.08 (1.06-1.11)	1.00 (0.99-1.00)	1.00 (0.99-1.00)	1.03 (1.00-1.05)
Birth weight						
<2500g	1.75 (1.74-1.77)	1.04 (1.01-1.06)	2.14 (2.08-2.21)	1.13 (1.12-1.14)	1.03 (1.01-1.05)	1.59 (1.51-1.66)
≥2500g	Reference	Reference	Reference	Reference	Reference	Reference
Gestational age						
<37 weeks	1.57 (1.56-1.58)	1.23 (1.21-1.26)	2.65 (2.57-2.73)	1.09 (1.09-1.10)	1.17 (1.15-1.19)	1.42 (1.36-1.48)
≥37 weeks	Reference	Reference	Reference	Reference	Reference	Reference
Socioeconomic status						
‘Lower’	1.00 (1.00-1.00)	1.23 (1.22-1.24)	1.09 (1.07-1.11)	0.96 (0.96-0.97)	1.12 (1.11-1.13)	1.10 (1.08-1.13)
‘Higher’	Reference	Reference	Reference	Reference	Reference	Reference
Urban-rural status						
Rural	0.79 (0.78-0.79)	2.34 (2.30-2.37)	1.16 (1.13-1.20)	0.73 (0.72-0.73)	2.28 (2.25-2.31)	1.15 (1.11-1.20)
Urban	Reference	Reference	Reference	Reference	Reference	Reference

Table 5.3. Stratified adjusted models (<1 yr old and ≥1 yr old) showing incidence rate ratios for the three service types (matched comparison group)

	Adjusted incidence rate ratio (95% CI)					
	<1 year of age			≥1 year of age		
	Physician visits	ED visits	Hospitalizations	Physician visits	ED visits	Hospitalizations
False Positive	1.45 (1.23-1.72)	0.81 (0.52-1.27)	2.20 (1.45-3.22)	1.00 (0.80-1.26)	0.81 (0.53-1.27)	0.60 (0.22-1.28)
Season of birth						
Jan. – Apr.	Reference	Reference	Reference	Reference	Reference	Reference
May – Aug.	1.03 (0.94-1.13)	1.07 (0.86-1.32)	0.73 (0.53-1.02)	0.97 (0.86-1.09)	1.07 (0.87-1.33)	1.67 (1.20-2.34)
Sept. – Dec.	1.04 (0.95-1.13)	1.13 (0.91-1.41)	1.03 (0.77-1.39)	1.04 (0.92-1.16)	1.11 (0.90-1.38)	1.29 (0.90-1.86)
Birth weight						
<2500g	1.64 (1.38-1.96)	1.26 (0.85-1.90)	1.46 (0.96-2.21)	1.18 (0.95-1.49)	1.27 (0.86-1.91)	1.60 (1.00-2.52)
≥2500g	Reference	Reference	Reference	Reference	Reference	Reference
Gestational age						
<37 weeks	1.72 (1.46-2.02)	1.12 (0.77-1.65)	3.07 (2.08-4.43)	1.21 (0.99-1.50)	1.13 (0.78-1.65)	2.18 (1.38-3.34)
≥37 weeks	Reference	Reference	Reference	Reference	Reference	Reference

Table 5.4. Stratified adjusted models (<1 yr old and ≥1 yr old) showing incidence rate ratios for the three service types for term infants only, false positive cohort versus primary (unmatched) comparison cohort

	Adjusted incidence rate ratio (95% CI)					
	<1 year of age			≥1 year of age		
	Physician visits	ED visits	Hospitalizations	Physician visits	ED visits	Hospitalizations
False Positive	1.35 (1.13-1.63)	1.46 (0.87-2.51)	2.32 (0.87-5.65)	0.92 (0.72-1.19)	0.97 (0.60-1.55)	0.65 (0.10-2.60)
Sex						
Female	0.92 (0.91-0.92)	0.86 (0.85-0.87)	0.75 (0.74-0.77)	0.86 (0.86-0.87)	0.86 (0.85-0.87)	0.78 (0.76-0.80)
Male	Reference	Reference	Reference	Reference	Reference	Reference
Season of birth						
Jan. – Apr.	Reference	Reference	Reference	Reference	Reference	Reference
May – Aug.	1.01 (1.00-1.01)	1.04 (1.03-1.05)	0.98 (0.96-1.01)	0.96 (0.95-0.96)	0.95 (0.95-0.96)	0.94 (0.91-0.96)
Sept. – Dec.	1.00 (1.00-1.00)	1.02 (1.01-1.03)	1.09 (1.07-1.12)	1.00 (0.99-1.00)	1.00 (0.99-1.01)	1.02 (0.99-1.05)
Birth weight						
<2500g	1.39 (1.38-1.41)	1.08 (1.04-1.11)	1.88 (1.78-1.99)	1.12 (1.10-1.13)	1.02 (1.00-1.05)	1.60 (1.50-1.71)
≥2500g	Reference	Reference	Reference	Reference	Reference	Reference
Socioeconomic status						
‘Lower’	1.00 (1.00-1.00)	1.22 (1.21-1.24)	1.10 (1.07-1.12)	0.97 (0.96-0.97)	1.12 (1.11-1.13)	1.10 (1.08-1.13)
‘Higher’	Reference	Reference	Reference	Reference	Reference	Reference
Urban-rural status						
Rural	0.78 (0.78-0.79)	2.37 (2.33-2.41)	1.21 (1.16-1.25)	0.73 (0.72-0.73)	2.30 (2.27-2.33)	1.16 (1.11-1.21)
Urban	Reference	Reference	Reference	Reference	Reference	Reference

REFERENCES

1. Moyer VA, Calonge N, Teutsch SM, Botkin JR. Expanding newborn screening: process, policy, and priorities. *Hastings Cent Rep.* 2008 May-Jun;38(3):32-9.
2. Therrell BL, Adams J. Newborn screening in North America. *J Inherit Metab Dis.* 2007 Aug;30(4):447-65.
3. Lipstein EA, Perrin JM, Waisbren SE, Prosser LA. Impact of false-positive newborn metabolic screening results on early health care utilization. *Genet Med.* 2009 Oct;11(10):716-21.
4. 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.
5. Waisbren SE, Albers S, Amato S, Ampola M, Brewster TG, Demmer L, et al. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA.* 2003 Nov 19;290(19):2564-72.
6. Wilcken B. Screening for disease in the newborn: the evidence base for blood-spot screening. *Pathology.* 2012 Feb;44(2):73-9.
7. Tu W-J, He J, Chen H, Shi X-D, Li Y. Psychological Effects of False-Positive Results in Expanded Newborn Screening in China. *PloS One.* 2012;7(4):e36235.
8. Tluczek A, Orland KM, Cavanagh L. Psychosocial consequences of false-positive newborn screens for cystic fibrosis. *Qual Health Res.* 2011 Feb;21(2):174-86.
9. Morrison DR, Clayton EW. False positive newborn screening results are not always benign. *Public Health Genomics.* 2011;14(3):173-7.
10. Tarini BA, Clark SJ, Pilli S, Dombkowski KJ, Korzeniewski SJ, Gebremariam A, et al. False-positive newborn screening result and future health care use in a state Medicaid cohort. *Pediatrics.* 2011 Oct;128(4):715-22.

11. McCandless SE, Chandrasekar R, Linard S, Kikano S, Rice L. Sequencing from dried blood spots in infants with "false positive" newborn screen for MCAD deficiency. *Mol Genet Metab.* 2013 Jan; 108(1):51-5.
12. 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.
13. Rhead WJ. Newborn screening for medium-chain acyl-CoA dehydrogenase deficiency: a global perspective. *J Inherit Metab Dis.* 2006 Apr-Jun;29(2-3):370-7.
14. Grosse SD, Khoury MJ, Greene CL, Crider KS, Pollitt RJ. The epidemiology of medium chain acyl-CoA dehydrogenase deficiency: an update. *Genet Med.* 2006 Apr;8(4):205-12.
15. Ziadeh R, Hoffman EP, Finegold DN, Hoop RC, Brackett JC, Strauss AW, et al. Medium chain acyl-CoA dehydrogenase deficiency in Pennsylvania: neonatal screening shows high incidence and unexpected mutation frequencies. *Pediatr Res.* 1995 May;37(5):675-8.
16. Wang SS, Fernhoff PM, Hannon WH, Khoury MJ. Medium chain acyl-CoA dehydrogenase deficiency human genome epidemiology review. *Genet Med.* 1999 Nov-Dec;1(7):332-9.
17. 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.
18. Leonard JV, Dezateux C. Newborn screening for medium chain acyl CoA dehydrogenase deficiency. *Arch Dis Child.* 2009 Mar;94(3):235-8.

19. 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 Pediatr.* 2010 Nov;10:82.
20. Collier R. Shift toward capitation in Ontario. *CMAJ.* 2009 Nov;181:668-9.
21. National Ambulatory Care Reporting System Data Quality Documentation - Current-Year Information, 2011-2012. Toronto: Canadian Institute for Health Information; 2012.
22. Discharge Abstract Database Data Quality Documentation - Current-Year Information, 2011-2012. Toronto: Canadian Institute for Health Information; 2012.
23. Statistics Canada. 2006 Census Dictionary. Ottawa: Statistics Canada; 2007.
24. Wang C, Guttman A, To T, Dick PT. Neighborhood income and health outcomes in infants: how do those with complex chronic conditions fare? *Arch Pediatr Adolesc Med.* 2009 Jul;163(7):608-15.
25. Kralj B. Measuring Rurality - RIO2008 BASIC: Methodology and Results. Ottawa: Ontario Medical Association; 2010.
26. Rural Physician Eligibility Requirements. Toronto: Ontario Ministry of Health and Long-Term Care; 2013 [cited 2014 April 1]. Available from: http://www.healthforceontario.ca/en/Home/Physicians/Training_%7C_Practising_in_Ontario/Locum/Rural_Family_Medicine_Locum_Program/Rural_Physicians/Eligibility.
27. Vuong QH. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica.* 1989;57(2):307-33.
28. Bollen KA, Jackman RW. Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. *Sociol Methods Res.* 1985;13(4):510-42.
29. Belsley DA, Kuh E, Welsch RE. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. Hoboken: Wiley-Interscience; 2005.

30. Wildner M. Health economic issues of screening programmes. *Eur J Pediatr.* 2003 Dec;162(1):S5-S7.
31. Tran K, Banerjee S, Li H, Noorani HZ, Mensinkai S, Dooley K. Clinical efficacy and cost-effectiveness of newborn screening for medium chain acyl-CoA dehydrogenase deficiency using tandem mass spectrometry. *Clin Biochem.* 2007 Feb;40(3-4):235-41.
32. Hamers FF, Rumeau-Pichon C. Cost-effectiveness analysis of universal newborn screening for medium chain acyl-CoA dehydrogenase deficiency in France. *BMC Pediatr.* 2012;12:60.
33. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10). Geneva: World Health Organization (WHO); 1992.
34. Statistics Canada. Data quality, concepts and methodology: Definitions. Ottawa: Statistics Canada; 2013.

CHAPTER SIX: FINAL DISCUSSION

Limited research has been focused on describing the health system impact of rare diseases in Canada. There is a need to study both the impact of being diagnosed and, for those diseases that are targets of newborn screening, the impact of false positive screening results. Due to the large population size with universal health care coverage in Ontario, there is a unique opportunity here to study the impact of rare disorders identified by newborn screening through the secure linkage of provincial newborn screening diagnostic confirmation data with population-based health care administrative datasets. Using MCADD as an example disorder, early health care utilization patterns among children that received a positive newborn screening result for MCADD were analyzed relative to an unaffected, screen negative birth cohort. However, since methods for estimating robust event rates and associated factors are not well established for a rare disease population, a methodological review of the literature was conducted as a first step.

Two general approaches were identified to address the ‘small numbers’ problem in generating robust estimates of rates for health events: 1) aggregation – such as, collapsing data categories or expanding geographic areas; and 2) data smoothing – for example, weighted averages to smooth ‘noisy’ data if neighboring rates are readily available and Bayesian estimation methods (1-3). Despite the respective advantages and disadvantages of each method, no gold standard exists and empirical comparisons of these techniques are limited in the literature. Selection of the most suitable method should thus be guided by the purpose of the analysis, the nature of available data, and the audience for whom the results are presented. Future studies should focus on

examining the differences in the proposed methods with respect to their effectiveness of smoothing rates and ease of implementation, in general and specific to the rare disease context.

The next step of the project involved an analysis of health services data encompassing physician visits, emergency department care, and hospitalizations for forty children diagnosed with MCADD through Newborn Screening Ontario following a positive newborn screening result between April 1, 2006 and March 31, 2010. Newborn screening diagnostic confirmation data were linked with health care administrative datasets at the Institute for Clinical Evaluative Sciences, including data representing health services visits and associated costs. To our knowledge, this is the first population-based study to date that analyzes early health care utilization patterns and estimates associated health care expenditures for a cohort of children affected with a rare disease ascertained through newborn screening. To demonstrate the feasibility of the methods discussed in the review on robust estimation techniques with small population sizes, a weighted moving average approach was selected and successfully implemented as part of the analysis to highlight its ability to estimate smoothed health event rates using small numbers (4). Children diagnosed with MCADD accessed physician services, emergency department care, and were hospitalized at statistically significant higher rates relative to those with negative newborn screening results after adjustment for potential confounders, regardless of age at the time of visit. The total cost of care for a child with MCADD during the first year of life was estimated to be, on average, approximately 32% higher as compared to a child with a negative newborn screening result. While an increased frequency of health care services use was not

unexpected in this affected cohort, this study was able to quantify the patterns of care and associated costs of services, which is essential towards effectively characterizing the burden of disease from the perspective of the health care system as well as families. Given the necessarily high use of health care services in children with MCADD, these findings emphasize the need to establish and maintain an organized system of care for this patient population in order to promote favorable health outcomes and patient experiences with care. Future research should concentrate on characterizing current models of care in place across regional metabolic treatment centres in Canada and internationally, and on comparing their effectiveness towards improving health outcomes in this pediatric population.

In part due to the advent of tandem mass spectrometry, the recent expansion of newborn screening panels has resulted in an increased number of newborns experiencing false positive results. To date, there is no consistency in the empirical literature regarding the impact of false positive newborn screening results on health care utilization patterns in early life (5-7). After adjustment for gestational age and birth weight along with other potential confounders, we found that 43 children with false positive newborn screening results for MCADD experienced statistically significant higher rates of physician visits and hospitalizations in the first year of life relative to a screen negative birth cohort. However, the false positive cohort did not experience a higher frequency of emergency department visits relative to the screen negative comparison group in the first year of life, and there were no significant differences between the cohorts in health services use after one year of age. Possible explanations for the higher rates of early physician visits and hospitalizations in the

false positive cohort include: residual confounding, either by a variable other than those we considered, or by a variable such as gestational age that we included but that could not be completely accounted for given the inability to use finer categories of exposure in our small sample; health care use related to follow-up confirmatory diagnostic testing; or the potential psychosocial impact of false positive screening results on parental perceptions of infant health. An interesting result was an apparent but statistically non-significant lower rate of hospitalization among those in the false positive cohort relative to controls at over one year of age. This may well be due to chance given the non-significance of the finding; or it may be due to differences in underlying health conditions between the groups that we did not consider. Collectively, the findings from this analysis demonstrate that false positive newborn screening results are associated with increased health care service rates in the first year of life. Linking additional clinical and family-reported data to the health care administrative data would allow for future studies to investigate whether parental psychosocial effects are partially responsible for this observation.

In addition to the limited clinical and psychosocial details available in health care administrative datasets, a further key limitation in this project was the small population size. While approaches for estimating robust event rates with small numbers were identified in the review of methodological literature, a challenge in interpreting differences in health services use between the cohort of interest and the screen negative comparison groups remained. Only a limited number of stratified analyses could be performed in order to respect stringent privacy standards for small numbers and to generate results that were relatively robust statistically, thus residual

confounding could not be entirely ruled out in the models. Additional sociodemographic and geographic stratification will become possible as the cohort size grows and follow-up period increases.

This project represents one of the objectives within the practice-based research framework of a broader multidisciplinary research program, the Canadian Inherited Metabolic Diseases Research Network, which aims to provide evidence to improve health outcomes and inform effective health care for children with inborn errors of metabolism (8). Future plans involve the adaptation and expansion of these methods to characterize the health system impact of other rare diseases by taking advantage of the unique opportunities available for linkage to health administrative datasets in Ontario and other Canadian provinces. Further, the network will collect clinical and patient- and family-reported outcomes through chart review, parent surveys and interviews. By integrating this newly collected information with administrative data, future health services research focused on these rare disease populations will be able to incorporate richer information.

Prior to the completion of this project, there was limited attention directed towards characterizing the health system impact of rare diseases in Canada. Health services research focused on this population ultimately aims to inform effective and appropriate care for children with MCADD and those with false positive newborn screening results.

REFERENCES

1. Fefferman NH, O'Neil EA, Naumova EN. Confidentiality and confidence: is data aggregation a means to achieve both? *J Public Health Policy*. 2005 Dec;26(4):430-49.
2. Simonoff J. *Springer Series in Statistics: Smoothing Methods in Statistics*. New York: Springer-Verlag; 1996.
3. Ghosh M, Rao JNK. Small Area Estimation: An Appraisal. *Stat Sci*. 1994;9(1):55-76.
4. Selvin S. *Statistical Analysis of Epidemiological Data*. 3rd ed. New York: Oxford University Press; 2004.
5. Lipstein EA, Perrin JM, Waisbren SE, Prosser LA. Impact of false-positive newborn metabolic screening results on early health care utilization. *Genet Med*. 2009 Oct;11(10):716-21.
6. 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.
7. Waisbren SE, Albers S, Amato S, Ampola M, Brewster TG, Demmer L, et al. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA*. 2003 Nov 19;290(19):2564-72.
8. Potter BK, Chakraborty P, Kronick JB, Wilson K, Coyle D, Feigenbaum A, et al. Achieving the "triple aim" for inborn errors of metabolism: a review of challenges to outcomes research and presentation of a new practice-based evidence framework. *Genet Med*. 2013 Jun;15(6):415-22.

APPENDICES

APPENDIX A (Ethics Approval)



**Ottawa Health Science Network Research Ethics Board/Réseau des sciences de la santé
d'Ottawa Conseil d'éthique de la recherche**

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>

Wednesday, April 02, 2014

Dr. Elizabeth Potter
University of Ottawa
Department of Epidemiology and Community Medicine

Dear Dr. Potter:

RE: Protocol# - Emerging Team in Rare Diseases: Part 1, Descriptive Epidemiology and Health Services Impact of Inborn Errors of Metabolism in an Ontario Newborn Cohort

Renewal Expiry Date - Thursday, April 30, 2015

I am pleased to inform you that your Annual Renewal Request was reviewed by the Ottawa Health Science Network Research Ethics Board (OHSN-REB) and is approved. No changes, amendments or addenda may be made in the protocol without the OHSN-REB's review and approval.

Renewal is valid for a period of one year. Approximately one month prior to that time, a single renewal form should be sent to the REB office.

The Ottawa Health Science Network Research Ethics Board (OHSN-REB) was created by the merger of both the Ottawa Hospital Research Ethics Board (OHREB) and the Human Research Ethics Board (HREB) for meetings held at the University of Ottawa Heart Institute.

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; and the provisions of the Personal Health Information Protection Act 2004.

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

APPENDIX B (Overview of potential approaches for modeling count data and example use of the Vuong statistic to select an approach)

OVERVIEW OF APPROACHES TO MODELING COUNT DATA

A count is defined as the number of given events that occur over a specified period of time, such as health services event data. Special consideration needs to be given to modeling counts since distributions of such data often violate basic assumptions of commonly used multivariate regression techniques. In public health, researchers often rely on either Poisson regression or negative binomial regression to estimate the rates of health events or in studying the associations between risk factors and health outcomes. However, when count outcomes are relatively rare, there is typically a preponderance of zeros within the distribution that needs to be considered prior to selecting an approach. In these cases, Poisson or negative binomial regression may not be the most appropriate statistical technique. Inappropriate selection can result in the calculation of biased standard errors, inefficient parameters, and an increase in Type I or Type II error rates due to the highly skewed nature of the resulting distribution (1-3). Numerous models have been developed specifically for count data, with model-specific adaptations designed for those with zero-inflated variables. There are assumptions inherent to each statistical distribution and inferences need to accurately reflect the observed data. Different methods that can be used to model health care service count data are summarized in this short review.

Poisson regression

Poisson regression methods are commonly used to predict event counts, which may sometimes be rare, given a group of predictors. This form of regression assumes that the observed counts are generated from a *Poisson* distribution, which is most commonly referred to as the default statistic for discrete count models (4). Due to its similarities with aspects of linear regression, it is relatively easy to interpret which makes it an attractive option for modeling count data. The probabilities are described as:

$$P(Y = y|\mu) = \frac{e^{-\mu} \mu^y}{y!},$$

where y is the observed count of events ($y \geq 0$) and μ is the expected count of the same event in a specified period of time (4, 5). The Poisson random variable Y is restricted to being a non-negative integer. However, this distribution holds two major assumptions (6, 7). The first assumption is equal dispersion, which implies that the mean of the count variable is equal to its variance. Unlike the normal distribution that has separate parameters for the mean and variance, this assumption does not account for unobserved heterogeneity. However, in reality, data can be overdispersed in some populations (i.e. the variance is greater than predicted from a simple Poisson distribution) and analysis must be adjusted in order to account for this limitation as a result. In the case of health services count data, there may be an excess of observed zeros in the dataset that cannot be predicted by the Poisson distribution. The second assumption is independence of occurrences of events in time, which implies that occurrences of the same event by the same person in a given period of time are

independent of each other (6). Due to the restrictive Poisson assumptions, health service researchers proposing to analyze health service outcomes often elect to model their data with a negative binomial distribution.

Negative binomial regression

The negative binomial, a derivation of the Poisson distribution, assumes that the mean of the distribution λ is gamma distributed. The conditional variance and conditional mean of the outcome variable are not assumed to be equal in this distribution; in addition, random error is incorporated into the conditional mean. The probabilities for the negative binomial are defined as:

$$P(Y = y|r, p) = \binom{y+r-1}{y} p^r (1-p)^y = \frac{r(y+r-1)!}{y! (r-1)!} p^r (1-p)^y,$$

where y is the observed count of events, r is the count of non-events (also called the dispersion parameter), and p is the probability of no events (5, 7). The symbol $\Gamma(\cdot)$ represents the gamma function. While the standard Poisson regression approach can be extended to account for variability between individuals by adding an overdispersion parameter, the negative binomial regression model can model both observed and unobserved heterogeneity. As a result, overdispersion is explicitly accounted for by the negative binomial approach (7). The limitation of the independence of occurrences of events in time is partially addressed when this derived model is applied; as such, this model can additionally account for a certain degree of inter-person heterogeneity (8). Further, comparisons between Poisson and negative binomial regression models demonstrate that the latter better fits count data as compared to the former in most

instances; however, negative binomial models may still not be the most appropriate when a dataset displays a preponderance of zeros. In these instances, health service researchers should focus their attention towards either zero-inflated Poisson or negative binomial models.

Zero-inflated regression models

There are certain situations in which the Poisson or negative binomial distribution may have to be adjusted in order to model the count data appropriately. When overdispersion is due to a preponderance of zero counts (i.e. a large proportion of individuals fail to experience the outcome of interest), a zero-inflated Poisson or negative binomial regression model should be considered (9, 10). In addition to taking the excessive zero counts into account, these methods can simultaneously model the distribution of the positive counts. However, as mentioned previously, the negative binomial version will additionally model the overdispersion that arises from unobserved heterogeneity arising from sources other than the excess zeros. In the zero-inflated models, the underlying assumption is that individuals can belong to one of two categories: the first in which he/she never experiences the outcome of interest and the second in which he/she records at least one encounter. The probability of an excess zero is modeled using logistic regression, whereas positive outcome counts are modeled by the Poisson or negative binomial probability density function (9, 10). However, the zero-inflated approach assumes that an unknown proportion of the zero counts were never at risk for experiencing the outcome of interest, which is likely not realistic with a health services outcome; this results in a form of discrete unobserved heterogeneity since we cannot differentiate between the two types of zeros (5). Hurdle

models approach modeling excess zeros differently and should be considered as an alternative to the zero-inflated models if the data presents with a preponderance of zeros.

Hurdle regression models

There are two components to a hurdle model, which can either be Poisson or negative binomial in nature similar to the zero-inflated models (11). The first part is based on a Bernoulli or right censored count density, which is used to estimate the probability of an event occurring (a zero versus a non-zero count) through a probit or logistic regression process. In the case of modeling a health service event, this first portion would model whether an individual does or does not decide to pursue the outcome of interest over the study period. The second part is based on a left truncated Poisson or negative binomial count density, which is used to estimate the positive integer count for each non-zero count obtained from the first part of the model. Thus, given that the individual decided to seek the outcome of interest, this second part will model the frequency at which they pursue the health service (12-14). However, due to the two-part complexity of the Hurdle model, literature suggests that sample populations be adequately large in order to estimate the second part of the hurdle model (15). As a result, Hurdle models should not be considered as a potential regression modeling approach when analyzing small populations.

Conclusion

There is no single gold standard analytic approach for count data describing health services events. Factors such as excess “zero” counts, high inter-person

heterogeneity, and lack of independence of repeat events within individuals need to be considered. Since no single approach can be universally recommended, health services researchers often rely on a statistical test that empirically compares the suitability of the different methods for a specific dataset. I used the Vuong statistic for this purpose in my analyses for manuscripts 2 and 3. Below is an example of its use for ED data from manuscripts 2 and 3.

EXAMPLE: VUONG STATISTIC

Background

Greater than expected inter-person heterogeneity in the number of health service events experienced was expected among children with MCADD and those with false positive newborn screening results for MCADD. For example, an asymptomatic MCADD patient may have no emergency department encounters within a specified period of time, while a symptomatic patient may have 30 or more within the same time frame. Careful consideration must be placed on selecting the most appropriate regression technique for modeling health services count data for these populations as a result.

Objective

To compare regression models using the Vuong statistic for each respective service type in three separate cohorts.

Methods

Four models were selected for comparison using the Vuong statistic: 1) Poisson; 2) negative binomial (NB); 3) zero-inflated Poisson (ZIP); and, 4) zero-inflated

negative binomial (ZINB). The Vuong statistic, a likelihood-based ratio testing framework, was designed to test the null hypothesis that the two models being compared are equally close to the true model; large Z-scores (i.e., small p -values, $p < 0.05$) would suggest that one of the two models is closer to the actual model, thus rejecting the null hypothesis (16). A Vuong macro available on the SAS® software website was used to complete the count regression model comparison and selection process (17). For each cohort, the four possible regression models were compared for physician visits, ED events, and inpatient hospitalizations respectively. An offset variable equal to the logarithm of the follow-up time was created in order to account for subjects with varying lengths of follow-up. Selection of the preferred model was made through consensus on a case-by-case basis.

Results

We calculated Vuong statistics for the cohorts of children affected with MCADD ascertained through newborn screening and those with false positive newborn screening results for MCADD, for each type of health care service. We also used the Vuong statistics with the screen negative cohorts. Here we present results for ED visits, as an example.

For affected children, the results for ED events suggested that the NB model had a significantly better fit as compared to the Poisson model ($p=0.009$) and ZIP model ($p=0.003$) respectively (Table B.1). Since the ZINB model did not converge for this service type with this cohort, it was not considered in the analysis. For the cohort of children with false positive newborn screening results, the NB model had a better fit as

compared to the Poisson model ($p=0.067$), the ZIP model ($p=0.004$), and the ZINB model ($p<0.001$) (Table B.2). Lastly, Vuong statistics for ED visits in the primary screen negative comparison group suggested that the NB model had a significantly better fit as compared to the Poisson model ($p<0.001$) and ZIP model ($p<0.001$) respectively (Table B.3). The ZINB model did not converge for this service type with this comparison population, thus it was not considered in the analysis.

Discussion

It can be challenging to determine if a particular distribution is zero-inflated since no pre-determined frequency of zero counts or ratio of zero to non-zero counts is widely accepted. As a result, post hoc analyses are the sole methods through which to determine if zero-inflated models are the most appropriate for the given count data. Based on these respective Vuong statistics, a negative binomial regression model was selected to model ED visits in all three cohorts. Results for physician visits and inpatient hospitalization are not shown, however negative binomial models were similarly found to be the most appropriate models.

Conclusion

The Vuong statistic can be used to select the most appropriate regression technique for modeling health services count data. A negative binomial regression model was selected to model physician visits, ED events, and inpatient hospitalizations respectively in all three cohorts.

Table B.1. Vuong statistics (Z-values) comparing models on emergency department use in children diagnosed with MCADD population

	Poisson	NB	ZIP	ZINB
Poisson	---	---	---	---
NB	2.61 ^a	---	---	---
ZIP	-0.74	-2.96 ^b	---	---
ZINB^c	---	---	---	---

^a Indicates that the row model had a significantly better fit as compared to the column model (p<0.05)

^b Indicates that the column model had a significantly better fit as compared to the row model (p<0.05)

^c Indicates a model that was unable to converge

Table B.2. Vuong statistics (Z-values) comparing models on emergency department use in children with false positive newborn screening results for MCADD

	Poisson	NB	ZIP	ZINB
Poisson	---	---	---	---
NB	1.83	---	---	---
ZIP	-0.80	-2.89 ^a	---	---
ZINB	-0.53	-3.39 ^a	0.86	---

^a Indicates that the column model had a significantly better fit as compared to the row model (p<0.05)

Table B.3. Vuong statistics (Z-values) comparing models on emergency department use in children with negative newborn screening results (primary comparison group)

	Poisson	NB	ZIP	ZINB
Poisson	---	---	---	---
NB	169.80 ^a	---	---	---
ZIP	167.07 ^a	-136.13 ^b	---	---
ZINB^c	---	---	---	---

^a Indicates that the row model had a significantly better fit as compared to the column model (p<0.05)

^b Indicates that the column model had a significantly better fit as compared to the row model (p<0.05)

^c Indicates a model that was unable to converge

APPENDIX C (Example: Weighted moving average calculation)

Objective

To examine time trends in health services use based on year of birth using a weighted moving average approach in a cohort of infants diagnosed with MCADD ascertained through newborn screening.

Methods

Unadjusted smoothed visit rates for each type of health care service were calculated using a weighted moving average approach with Hanning weights (18). This calculation was completed for children with MCADD born in 2007, 2008, and 2009 respectively, stratified by age at the time of visit (<1 year of age and 1-2 years of age). For each respective birth year from 2006 through 2010, the total number of visits for each service type was summed and the total length of follow-up was calculated for children based on the age groupings; unadjusted raw visit rates were then computed for each service type according to year of birth stratified by age at the time of visit. Stratified unadjusted smoothed visit rates for children in each birth year were then calculated by applying Hanning weights.

Results

The unadjusted raw rates for each of the three service types stratified by age at the time of visit for children diagnosed with MCADD were computed for each respective year of birth (Table C.1). Using these rates, the unadjusted smoothed visit rates stratified by age at the time of visit were calculated according to year of birth (Figure

C.1). For example, to calculate the smoothed physician visit rate for children born in 2008 at less than one year of age, the following equation was applied:

$$\begin{aligned} WMA &= 0.25(2007\text{ rate}) + 0.50(2008\text{ rate}) + 0.25(2009\text{ rate}) \\ &= 0.25(18.33) + 0.50(26.40) + 0.25(15.70) \\ &= 21.71 \end{aligned}$$

Discussion

The weighted moving average approach allows researchers to calculate smoothed event rates by incorporating neighboring data, which is particularly useful when dealing with small sample sizes. While comparative data for rare disease populations is currently limited, this technique will become increasingly relevant as cohort studies progress and rare disease registries develop.

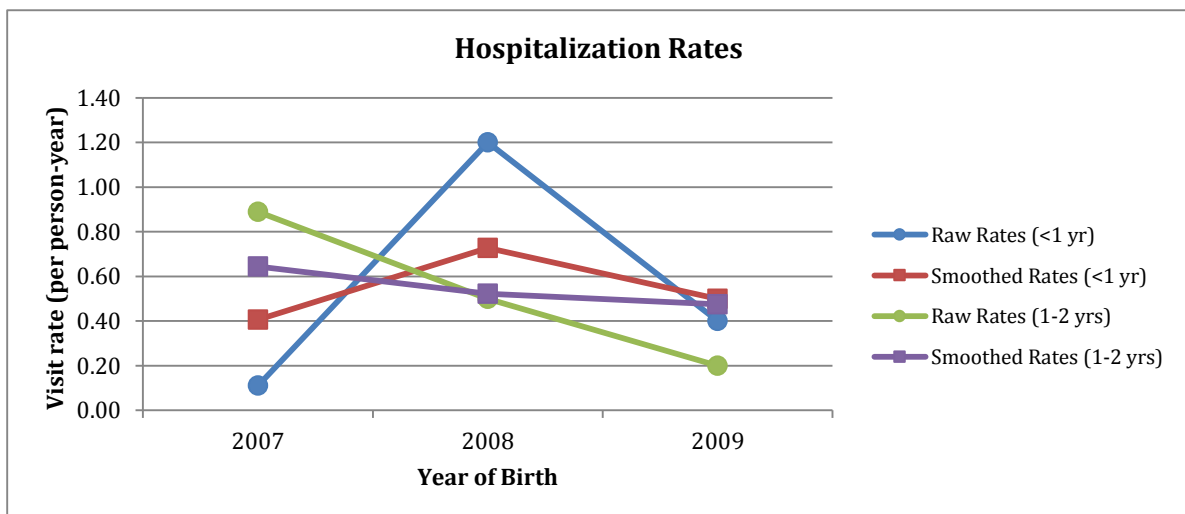
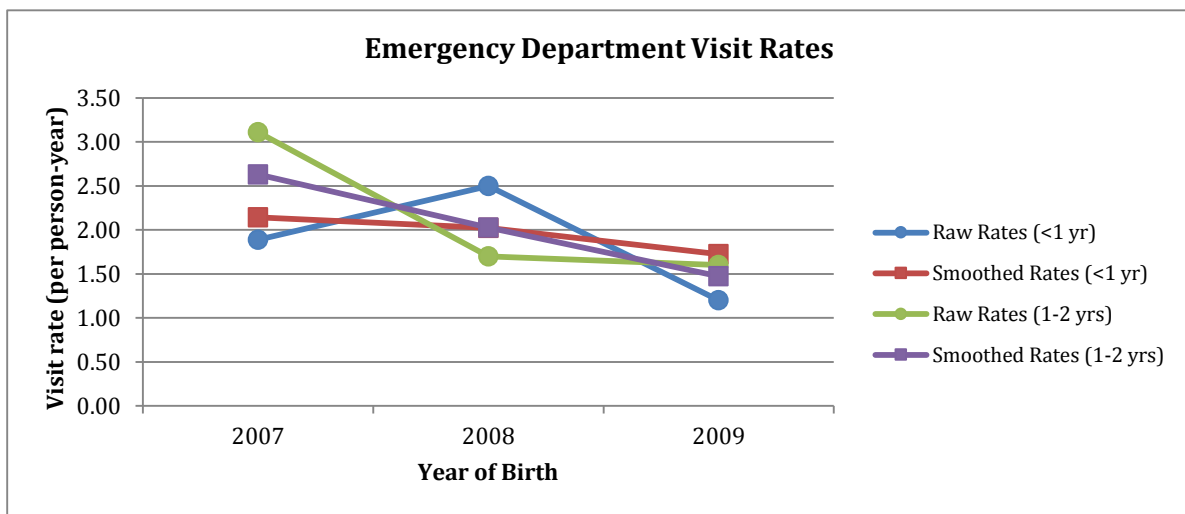
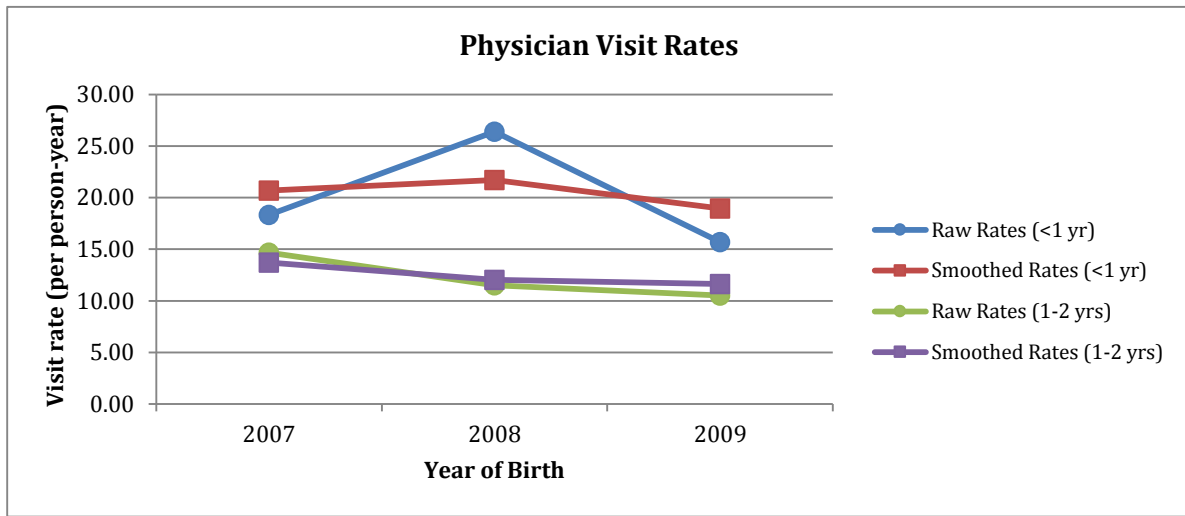
Conclusion

Weighted moving averages can be used to improve the statistical reliability of calculated event rates.

Table C.1. Unadjusted raw rates by age at the time of visits (<1 year of age and 1-2 years of age) in a MCADD birth cohort according to year of birth

Year of Birth	Unadjusted visit rates (95% CI)					
	<1 year of age			1-2 years of age		
	Physician visits	ED visits	Hospitalizations	Physician visits	ED visits	Hospitalizations
2007	18.33 (15.54-21.13)	1.89 (0.99-2.79)	0.11 (-0.11-0.33)	14.67 (12.17-17.17)	3.11 (1.96-4.26)	0.89 (0.27-1.51)
2008	26.40 (23.22-29.59)	2.50 (1.52-3.48)	1.20 (0.52-1.88)	11.50 (9.40-13.60)	1.70 (0.89-2.51)	0.50 (0.06-0.94)
2009	15.70 (13.24-18.16)	1.20 (0.52-1.88)	0.40 (0.01-0.79)	10.50 (8.49-12.51)	1.60 (0.82-2.38)	0.20 (-0.08-0.48)

Figure C.2. Unadjusted raw and smoothed health service visit rates for children stratified by age at the time of visit (<1 year of age and 1-2 years of age) for children with MCADD using weighted moving averages by year of birth



REFERENCES

1. Karazsia BT, van Dulmen MH. Regression models for count data: illustrations using longitudinal predictors of childhood injury. *J Pediatr Psychol.* 2008 Nov-Dec;33(10):1076-84.
2. Glynn RJ, Stukel TA, Sharp SM, Bubolz TA, Freeman JL, Fisher ES. Estimating the variance of standardized rates of recurrent events, with application to hospitalizations among the elderly in New England. *Am J Epidemiol.* 1993 Apr 1;137(7):776-86.
3. Baldi I, Ciccone G, Merletti F, Gregori D. Computing hospitalization rates in presence of repeated events: impact and countermeasures to avoid misinterpretation. *J Eval Clin Pract.* 2008 Apr;14(2):316-20.
4. Cameron AC, Trivedi PK. *Regression Analysis of Count Data.* New York: Cambridge University Press; 1998.
5. Long JS. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks: SAGE Publications Inc; 1997.
6. Pohlmeier W, Ulrich V. An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care. *J Hum Resour.* 1995;30(2):339-61.
7. Cameron AC, Trivedi PK, Milne F, Piggott J. A Microeconomic Model of the Demand for Health Care and Health Insurance in Australia. *Rev Econ Stud.* 1988 Jan;55(1):85-106.
8. Cameron AC, Trivedi PK. Econometric models based on count data. Comparisons and applications of some estimators and tests. *J Appl Econ.* 1986;1(1):29-53.
9. Lambert D. Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics.* 1992;34(1):1-14.
10. Greene WH. *Accounting for excess zeros and sample selection in the Poisson and negative binomial regression models.* New York: New York University, Leonard N. Stern School of Business, Department of Economics; 1994.

11. Mullahy J. Specification and testing of some modified count data models. *J Econom.* 1986 Dec;33(3):341-65.
12. Moineddin R, Meaney C, Agha M, Zagorski B, Glazier RH. Modeling factors influencing the demand for emergency department services in Ontario: a comparison of methods. *BMC Emerg Med.* 2011 Aug;11:13.
13. Hu MC, Pavlicova M, Nunes EV. Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial. *Am J Drug Alcohol Abuse.* 2011 Sep;37(5):367-75.
14. Rose CE, Martin SW, Wannemuehler KA, Plikaytis BD. On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. *J Biopharm Stat.* 2006;16(4):463-81.
15. Gurmu S. Semi-parametric estimation of Hurdle regression models with an application to Medicaid utilization. *J Appl Econ.* 1997;12(3):225-42.
16. Vuong QH. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica.* 1989;57(2):307-33.
17. Sample 42514: Tests for comparing nested and nonnested models. North Carolina: SAS Institute Inc; 2013 [cited 2014 April 1]; Available from: <http://support.sas.com/kb/42/514.html#ref>.
18. Selvin S. *Statistical Analysis of Epidemiological Data.* 3rd ed. New York: Oxford University Press; 2004.