**Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Erin Hodson Date

# Developmental Outcomes of School-Age Children with Duarte Galactosemia: An Interim Analysis

By

Erin Hodson

Master of Public Health

Epidemiology

_____

Jennifer G. Mulle, Ph.D., M.H.S.

Committee Chair

# Developmental Outcomes of School-Age Children with Duarte Galactosemia: An Interim Analysis

By

Erin Hodson

Bachelor of Arts

College of Wooster

2015

Thesis Committee Chair: Jennifer G. Mulle, Ph.D., M.H.S.

An abstract of

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2017

# Abstract

Developmental Outcomes of School-Age Children with Duarte Galactosemia: An Interim Analysis

By Erin Hodson

**Background**: Duarte galactosemia is a genetic disorder characterized by partial impairment of galactose-1-phosphate uridylyltransferase (GALT). This enzyme is critical for metabolism of galactose, a sugar abundant in milk. Children with Duarte galactosemia are commonly detected by newborn screening programs aiming to identify infants with the more serious, potentially fatal classic galactosemia. Patients with classic galactosemia require immediate treatment with a galactose-free diet to prevent severe acute complications, but experience increased risk of long-term health effects even with dietary restriction. The proper treatment and long-term prognosis associated with Duarte galactosemia are much more controversial. No consensus exists in regards to whether newborn screening programs should seek to detect infants with Duarte galactosemia, and when cases do come to light, recommendations issued to parents differ drastically depending on state and clinician. The current study is thus centered on two primary aims: 1) to determine whether children with Duarte galactosemia are at increased risk for the developmental delays seen in classic galactosemia patients and 2) to determine whether children with Duarte galactosemia benefit from dietary restriction of galactose during infancy.

**Methods**: Relevant developmental measures were compared in 90 cases with Duarte galactosemia and 63 of their unaffected siblings, all between the ages of 6 and 12 years. Performance in a range of domains was evaluated using a combination of parental surveys and direct child assessment by trained professionals using validated instruments.

**Results**: No significant differences were detected in cases as compared with controls in any of the developmental outcomes included in the analysis. Amongst cases, there was no association between assessed development and whether or not the child experienced galactose restriction early in life.

**Discussion**: The evidence presented here suggests the prognosis for children with Duarte galactosemia is good, with or without dietary intervention. There was no indication that Duarte galactosemia patients experienced any of the same long-term outcomes known to affect those with classic galactosemia. If the cumulative results support the same conclusions, public officials would be well positioned to redirect funds away from identifying and treating this condition and towards efforts with demonstrated potential for improving health.

# Developmental Outcomes of School-Age Children with Duarte Galactosemia: An Interim Analysis

By

Erin Hodson

Bachelor of Arts

College of Wooster

2015

Thesis Committee Chair: Jennifer G. Mulle, Ph.D., M.H.S.

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2017

## Acknowledgement

I would like to express my gratitude to my thesis advisor, Dr. Jennifer Mulle for her excellent guidance throughout this process. I also wish to thank Grace Howard for her continuous support and generous availability at every stage of my efforts.

# Table of Contents

## Background

Classic galactosemia is a rare genetic metabolic disorder estimated to affect between 1 in 40,000 to 1 in 60,000 live births (1). Patients with classic galactosemia carry deleterious mutations in both copies of the gene coding for the enzyme galactose-1-phosphate uridyltransferase (GALT), which is responsible for converting ingested galactose to glucose. The near total absence of GALT activity results in the abnormal accumulation of galactose and its metabolites in various cells and tissues of the body.

Infants with classic galactosemia typically appear healthy at birth, but experience a range of serious and potentially lethal complications when exposed to breast milk or milk formula, both of which contain galactose. Within days, affected infants develop acute symptoms including poor feeding, lethargy, hypoglycemia, hepatocellular damage, bleeding diathesis, and jaundice (2). Continued exposure to galactose can rapidly lead to sepsis with *Escherichia coli*, shock, and death (3).

In the United States, these devastating consequences are effectively avoided since the introduction of population newborn screening (NBS) for galactosemia in the early 1960s. Almost all children born in the U.S. undergo state-mandated screening in which a small blood sample ("heel stick") is collected within 48 hours of birth and analyzed for a panel of genetic disorders, including classic galactosemia (4). To check for the possibility of galactosemia, NBS laboratories conduct a coupled assay quantifying the activity of the GALT enzyme and take further action to retest infants and notify parents when the measured activity falls below a specified level (5). The goal of this screening procedure is to identify infants with classic galactosemia before they become symptomatic. When lab results are suggestive of classic galactosemia, newborns are immediately switched from breast milk or milk formula to a

galactose-restricted formula (6). Once the classical galactosemia diagnosis has been confirmed, most clinicians recommend that patients adhere to a lifelong lactose and galactose-free diet (7).

The rapid identification and treatment enabled by newborn screening has been highly successful in preventing the severe acute complications described above, including neonatal death. However, even individuals who maintain a galactose-restricted diet from infancy are at increased risk for adverse outcomes as they age, including developmental delays, cognitive disability, speech problems, neurological and/or movement disorders and, in female patients, ovarian dysfunction (2, 8, 9). The pathophysiology behind these long-term complications remains largely unknown, and the most appropriate ways to prevent and manage them is a topic of continued debate amongst experts (9). Despite these lingering uncertainties, there is a high degree of consensus concerning the importance of detecting cases of classic galactosemia early and the correct treatment to maximize their likelihood of living health lives (7).

Comparatively little consensus exists in regards to a milder and much more common form of GALT deficiency known as Duarte galactosemia. Individuals with Duarte galactosemia inherit one classic galactosemia gene and 1 Duarte gene. As touched upon earlier, the classic galactosemia allele carries a mutation which results in no or barely detectable GALT enzyme activity (10). The Duarte allele, on the other hand, is characterized by mutations that reduce the biostability of the enzyme and partially impair its function (11). Because they have one allele conferring at least partial functionality, Duarte galactosemia patients typically have GALT enzyme activity levels between 14 and 25 percent of that seen in unaffected controls (10).

Duarte galactosemia has an estimated prevalence of 1 in 4,000 Caucasians, making this disorder approximately ten times as common as classic galactosemia amongst infants born in the U.S. (12, 13). Data from the state of Georgia support this conclusion: Fernhoff (2010) reported

that over the previous three years, the Georgia Newborn Screening Program had identified 8 children with classic galactosemia and 83 children with Duarte galactosemia. This example highlights the way in which infants with Duarte galactosemia come to clinical attention in the U.S. The vast majority are diagnosed following abnormal newborn screening results, as the laboratory findings for infants with Duarte galactosemia can sometimes overlap with those seen for infants with the more severe, potentially fatal classic galactosemia (13). While infants with Duarte galactosemia experience reduced GALT activity and increased bodily levels of galactose metabolites, they are not at risk for the acute disease which threatens the lives of newborns with classic galactosemia who are not rapidly identified and treated with a galactose-restricted diet. In fact, some clinicians maintain that Duarte galactosemia produces no discernable health effects at all throughout the lifetime of patients affected. Other experts dispute the claim that Duarte galactosemia is a completely benign condition and suggest that affected individuals bear some of the long-term outcomes known to impact classic galactosemia patients.

This ongoing controversy is reflected in the substantial discrepancies in how Duarte galactosemia is identified and treated across the states. It is apparent that detection rates of Duarte galactosemia differ drastically, from essentially zero in states such as New York to over 1 in 3,500 in places like Arkansas and New Jersey (13). This variation is not believed to reflect true differences in the incidence of the disorder, but rather inconsistencies in what forms of galactosemia state's newborn screening programs are tailored to detect (13). As briefly mentioned above, all states measure GALT activity to evaluate whether infants may be affected by galactosemia and all have a predetermined cutoff for the level of activity considered abnormal. Some states set this cutoff high enough that samples from infants with Duarte galactosemia are flagged and funneled into the path of further testing and parental notification.

Other states have a relatively low cutoff, meaning that only infants with classic galactosemia are identified. Those states belonging to the latter category operate under a belief that Duarte galactosemia is asymptomatic and that those affected do not require any treatment.

The issue of treatment has become controversial amongst those states that regularly detect infants with Duarte galactosemia. Depending on their state of residency and consulting clinician, parents may receive a range of recommendations regarding the most appropriate dietary intervention. Pyhtila et al. (2014) surveyed 28 state newborn screening programs and found that five recommended no intervention, seven recommended at least partial milk restriction in the first year of life, and ten reported that specialists within the state advocated different interventions. The remaining six states did not intervene if parents intended to breastfeed their child, but recommended soy options if they were to give the baby formula. Parents receiving a diagnosis of Duarte galactosemia may be understandably confused to find that their child's treatment appears to depend more on location than conclusive scientific evidence. These deep inconsistencies are unsurprising, however, considering the current dearth of epidemiological data demonstrating whether infants with Duarte galactosemia benefit from dietary restriction or if milk exposure is associated with any negative health effects as they grow older.

The unresolved questions surrounding Duarte galactosemia have resulted in considerable costs, both on the part of the individuals affected and the larger public health system. If Duarte galactosemia is truly a benign condition, identifying and diagnosing infants carrying the genotype may pointlessly subject parents to worry and confusion. Moreover, if the dietary interventions recommended by many states have no impact on development, a substantial portion of children with Duarte galactosemia may miss out on the well-known benefits of breast milk

while gaining nothing in return (14). The public health system also invests time and resources following up diagnoses, as parents require genetic and nutritional counselling at metabolism clinics and children may receive costly biochemical-monitoring for years (6). These public health funds could certainly be spent on more useful initiatives if the reality is that Duarte galactosemia patients grow up to be indistinguishable from their peers, with or without galactose restriction.

Newborn screening programs that aim to detect Duarte galactosemia cases in addition to classic galactosemia cases may also pay a price in terms of the rate of false positives. False positives are those newborns identified as needing further testing that ultimately do not have either classic or Duarte galactosemia. Some of the incorrectly flagged samples may belong to galactosemia carriers, who represent nearly 1 percent of the U.S. population (13). These individuals have one classic galactosemia gene paired with one normal gene, and are therefore not at risk for the disorder. Clinics that use a cutoff designed to capture newborns with 25 percent GALT activity have increased likelihood of also identifying carriers and other healthy infants, whose families must then endure the anxiety and inconvenience of follow-up testing. The wait period between the first test and the follow-up result could necessitate a considerable interruption to normal breastfeeding as parents wait to find if their infant's health is in danger. Lowering the GALT activity cutoff to detect only newborns with classic galactosemia could significantly curb the false positive rate, as demonstrated by the NBS program of the state of Utah. By lowering the cutoff from 3.5 U/gHb to 3.0 U/gHb, Utah officials found that the number of classic galactosemia cases detected remained consistent while the number of Duarte galactosemia cases and false positives fell dramatically. The drop in number of annual false positives was particularly striking, from between 44 and 72 before the cutoff was lowered to

between 4 and 7 afterwards (13). This change was feasible in Utah because the state does not recommend any type of dietary intervention for infants with Duarte galactosemia, and therefore does not regard their detection as a priority. As long as the long-term health outcomes and potential benefits of galactose restriction remain poorly understood, however, other states may be hesitant to follow Utah's example and allow newborns with Duarte galactosemia to go undetected.

Clearly, improved knowledge of the long-term prognosis associated with Duarte galactosemia would be highly useful in resolving the ongoing conflict in the screening and metabolic communities as to the importance of identifying and treating the disorder. Few studies have set out to address either of these issues, and the findings from the ones that have offer mixed results. Ficicioglu et al. (2008) was the first to shed light on whether children with Duarte galactosemia benefit from a galactose restricted diet by comparing biochemical and developmental outcomes in twenty-eight subjects, including seventeen that had a restricted diet in the first year of life and eleven that were on a regular diet since birth (12). The data demonstrated that subjects on a regular diet had elevated galactose metabolites relative to those on a restricted diet, but that these levels gradually decreased during the first year of life and reached near-normal levels by one year of age. All of the Duarte galactosemia children tested scored within normal IQ range and there were no significant differences between the diet treatment groups in mean IQ or in language development. Untreated Duarte galactosemics did have significantly higher scores on the measure of adaptive functioning, though this data relied entirely on parental report. The authors interpret these findings as evidence that the early differences in levels of galactose metabolites not only disappear by the end of year one, but have no negative impact on development. Clinical outcomes were also consistent across all subjects

regardless of diet: none of the children showed any signs of abnormal liver function, early cataracts, or, in females, abnormal follicle-stimulating hormone activity—all symptoms known to affect classic galactosemia patients. Taken together, the authors argue that the results indicate outcomes for Duarte galactosemics are good whether or not children receive dietary intervention during infancy. An important caveat to their findings was the young age of study participants; all children were between one and six years of age, with a mean of just 2.96 years. This limitation motivates the authors' closing note that the results could justify lowering the cutoff used by screening programs, but only after future studies establish whether older Duarte galactosemics suffer from developmental delays not yet visible in the younger cohort.

One way researchers have addressed this question is by evaluating the number of children diagnosed with Duarte galactosemia who go on to require special education services during their school-age years. Van Naarden Braun et al. (2003) provided a first look into this answer through their more general analysis which included all metabolic and endocrine disorders detected through newborn screening (15). The study uncovered nine children with some form of metabolic or endocrine dysfunction that had received special education services through the public school system. Six of these nine children were Duarte galactosemics, of whom five required services because of a speech and/or language impairment and one due to a specific learning disability. Interestingly, all six of the children had been placed on a restricted diet throughout their first year of life, causing the authors to question whether more extended treatment is necessary or if some unaccounted for factor could mediate the observed association between Duarte galactosemia and developmental delays.

The finding that Duarte galactosemics were surprisingly well-represented amongst children receiving special education services prompted a follow-up study focused on this

condition specifically. Powell et al. (2009) expanded on the earlier work by comparing the prevalence of special education services amongst children with Duarte galactosemia with the prevalence of these services in unaffected children (16). They found that of 59 children in metropolitan Atlanta with Duarte galactosemia, five, or 8.5 percent, had received special education services, compared to 4.5 percent of the general population. Because they noticed all five of the children receiving services were 8 years old, the authors performed a secondary analysis restricting the sample to just those of this age. With this adjustment, the prevalence of receiving special education services was predictably higher amongst the children with Duarte galactosemia at 15.2 percent, now compared to 5.9 percent of 8-year-old children in the general population. Consistent with the findings of Van Naarden Braun et al., the most common reason for requiring services appeared to be speech and/or language impairments, as four of the five children with Duarte galactosemia were eligible by this criteria. Speech and language deficits commonly affect classic galactosemia patients, leading the authors suggest that the delays apparent in their study's subjects could represent less severe manifestations of disorders with similar underlying etiology (17, 18). If the observed difference in prevalence of developmental issues is attributable to Duarte galactosemia, these complications resulted despite adherence to the most commonly recommended treatment, as all children in the study consumed a galactose-restricted diet up to age one. The finding that even treated children may be at increased risk for developmental issues associated with special education services challenges the idea that Duarte galactosemia is a benign condition and underscores the need for larger epidemiological studies investigating long-term effects in affected children.

Most recently, Lynch et al. (2014) performed a pilot study addressing the most obvious limitation of Powell et al.—the fact that the results relied entirely on review of secondary data

rather than direct assessment of the study subjects (19). Fifteen children participated in this preliminary exploration, ten with Duarte galactosemia and five unaffected siblings from the same group of families. Parents of seven children with Duarte galactosemia had expressed concern about some aspect of their development, while parents of the remaining three reported that they had no concerns. Researchers assessed and compared performance of all children in areas of development known to be affected in classic galactosemia cases.

Children completed assessments targeted to several aspects of cognitive development, including global intelligence, visual-motor function, memory, processing speed and sustained attention. The authors report that overall intelligence was consistent across the fifteen subjects, but that children with Duarte galactosemia showed evidence of less efficient auditory memory. Auditory processing abilities were also assessed because of the connection with language development, and cases again demonstrated slower processing of information relative to controls. On measures of communication, children with Duarte galactosemia whose parents expressed developmental concerns appeared to have weaker listening and receptive language skills than either controls or cases whose parents did not have concerns.

Physical measures included assessments of movement (balance, dexterity, and coordination) and presence of visible tremors. Once again, the authors present the results separately depending on whether cases' parents expressed concerns. They report that six of the seven children whose parents did have concerns scored below the fifth percentile on the main measure of motor development, compared to two of the five control children and zero of the three cases whose parents were not concerned. Four of the seven children in the parental concerns group also had a pronounced kinetic hand tremor while attempting to draw a smooth line, along with one child in the control group.

Findings from the final domain assessed, socio-emotional development, are highlighted by the authors as strong evidence that Duarte galactosemia can contribute to the same deficits recognized in classic galactosemia patients. Results from parent-response surveys indicated that children in the Duarte galactosemia parent concerns group were more likely than controls to demonstrate internalizing behaviors, a broad class of negative behaviors in which children direct emotions inward. Internalizing behaviors such as heightened anxiety and social withdrawal have been reported to occur with unusual frequency in individuals with classic galactosemia (20, 21, 22, 23). Scores on measures of hyperactivity, inattention, and autism spectrum disorder followed the same pattern: the Duarte galactosemia children with noted parental concerns consistently rated higher than the other two groups. This trend is of course unsurprising considering that all of the socio-emotional outcomes were classified based on parents' responses; those parents who expressed general concerns could also be expected to report specific behavioral issues. The issue of parental concern is important, however, because it relates to the primary threat to the validity of the preliminary findings presented by Lynch et al. Since the final sample size represented a subset of all those invited, it is possible that parents who had noticed delays in their children were more likely to choose to participate. The authors argue that ascertainment bias cannot fully account for their results because some of the effects applied to all Duarte galactosemia kids, including those whose parents had not discerned any developmental issues. Combined with the indirect findings of Powell et al., these data suggest children with Duarte galactosemia may be at risk for subtle delays in typical development by mid-childhood, particularly related to auditory processing, memory, and socio-emotional wellbeing.

The work carried out by Ficicioglu et al., Van Naarden Braun et al., Powell et al., and Lynch et al. has contributed to greater scientific understanding of the long-term health outcomes

associated with a diagnosis of Duarte galactosemia. However, the results are not fully concordant and the usefulness is limited by small sample sizes and lack of variability in key variables such as diet; the latter three studies were unable to assess the effectiveness of dietary intervention because all or nearly all subjects experienced galactose restriction. The critical questions of whether infants with Duarte galactosemia require detection and if they benefit from treatment therefore remain largely unanswered, to the consternation of the newborn screening community and anxious parents alike. Until larger scale epidemiological data provides insight into these unresolved issues, prevailing recommendations will continue to rely more on state of birth and the individual leanings of clinicians than hard scientific evidence. If Duarte galactosemia is in fact essentially benign, public health funds can be redirected away from identifying and treating this condition and spent on efforts with greater value. If Duarte galactosemia is associated with long-term health effects, diagnosis and access to early intervention services will become all the more crucial. Prior work has provided some indication as to which possibility appears most likely, but further research is required to finally put this matter to rest.

# Introduction

The preceding section provided a summary of the current understanding and persistent controversy surrounding the genetic condition Duarte galactosemia. As described at length above, Duarte galactosemia results in impaired ability to metabolize galactose, a sugar abundant in milk. Little consensus exists regarding the long-term prognosis for infants detected through newborn screening and diagnosed with Duarte galactosemia, with some experts maintaining that the condition is essentially benign and others pointing to data demonstrating affected children are at increased risk for developmental delays. Widespread disagreement is also attached to the issue of whether newborns with Duarte galactosemia require dietary restriction of galactose. Thus far, no adequately powered study has provided a conclusive answer to either of these critical questions, forcing clinicians to use their best judgment and parents to cope with uncertainty as to how they can best care for their child and what that they might expect as they grow older.

The study presented here has potential to provide much-needed resolution to the discussion of Duarte galactosemia. This work was conducted by the same team of researchers that collaborated on the pilot study, Mary Ellen Lynch, Nancy L. Potter, Claire D. Coles, and Judith L. Fridovich-Keil. The current study builds on the procedures described above for pilot data collection, making use of a much larger sample size to compare developmental outcomes between cases with Duarte galactosemia and unaffected controls in the domains of cognitive skills, communication processes, physical development, and socio-emotional development. Importantly, this analysis does not include data from the full set of subjects who will eventually participate in the study—data collection efforts are currently ongoing and the final sample size will be nearly twice that presented here. All of the efforts described were carried out as part of a planned preliminary analysis at the study's halfway point.

Though the sample size available for this analysis was small compared to the anticipated final subject count, it was many times larger than any of the studies explored above, granting increased power to detect subtle developmental deficits in cases relative to controls. The relatively large sample size allowed for the use of more sophisticated analysis methods than were possible with the extremely small number of participants examined by Ficicioglu et al. or in the pilot study. These substantially expanded methods were utilized to provide a more complete, reliable look into the two main study aims repeated throughout this presentation: 1) To determine whether school-age children with Duarte galactosemia are more likely to experience disorders in development relative to their unaffected peers and 2) To assess whether dietary restriction of galactose in infancy is associated with developmental outcomes amongst children with Duarte galactosemia.

# Methods

*Study Participants*

Volunteers in the study included 90 children with Duarte galactosemia and 63 unaffected siblings of children with Duarte galactosemia, all between the ages of 6 and 12 years. Demographic characteristics of the 153 participants are presented in Table 1. To be eligible, children had to fall into the specified age range at the time of testing, have no exclusionary medical conditions, speak English well, and have a parent who also spoke English well and had served as their primary caregiver since birth. This last criteria was important because parents needed to be familiar with the child's diet during infancy. Exclusionary medical conditions included any chronic illness or condition unrelated to Duarte galactosemia that is known to cause problems in development. None of the 153 subjects also participated in the pilot study.

Children diagnosed with Duarte galactosemia during follow-up to newborn screening (NBS) were recruited in several collaborating states. Seven of these states contributed subjects included in this analysis: Alabama, Georgia, Illinois, Iowa, Michigan, Missouri, and South Carolina. In each participating locale, study personnel partnered with the state newborn screening program and their local metabolic referral clinics to identify children with Duarte galactosemia. A liaison within the newborn screening program mailed hard copy invitation letters and "reasons to participate" summary sheets prepared by the research team to parents of children diagnosed with Duarte galactosemia in the target years. At the discretion of the NBS liaison, members of the research team also reached out to families by telephone or through texting. The initial invitation letter mailed to each family provided a description of the study and explained that the child with Duarte galactosemia whose name was listed on the envelope was potentially eligible to participate as a "case." The letter also extended the invitation to any other

children in the household between the ages of 6 and 12 years interested in volunteering either as a “case” if they also had Duarte galactosemia or as a “control” if they did not.

*Procedures*

Once a pool of potential participants was identified and informed of the study, data collection took place in two distinct phases. Part 1 consisted of a brief parent survey including questions about the family’s demographics and socioeconomic status and the child’s general health, early education and interventions, response to sensory stimuli, and diet during infancy. When parents responded to the recruitment letter, the research team sent them an online consent form and a Part 1 survey link for each eligible child in the family. The process of completing the online consent form and the Part 1 survey was estimated to take no more than twenty minutes. Families that submitted the survey within two weeks of receiving the recruitment letter were compensated for their time with a fifty-dollar gift card.

Parents’ responses to the Part 1 survey contained all of the information needed to determine their children’s eligibility for Part 2 of the study. In this second, more intensive phase, a subset of all those that completed the Part 1 survey were invited to come to a local facility for an in-person assessment of relevant developmental domains. Data obtained during Part 2 involved both direct testing of children and surveys administered to parents for aspects of development more difficult to quantify during limited appointment time, such as the child’s experience with depression and anxiety. Five members of the research team conducted Part 2 testing, including the Project Manager, in charge of walking families through the study process, and four testers (two doctoral level child psychologists and two speech/movement specialists). The only biological samples collected during the Part 2 procedures were small amounts of saliva

used to isolate DNA for GALT genotyping to confirm subjects' case versus control status. In total, the full schedule of assessments took approximately three and a half hours, at the end of which families were compensated with a two hundred dollar gift card per child.

The locations that served as sites for direct testing were determined to maximize the number of eligible Part 1 respondents within a two-hour driving radius. Each potential location was classified in terms of the number of subjects reported to live within driving distance and key characteristics of the cases in the area, including history of dietary intervention. In order to be selected, a geographic area had to cover at least 24 eligible children. Once a location was chosen, study personnel re-contacted families inside the two-hour radius and scheduled appointments for those interested in participating.

*Outcome Measures*

The trained professionals in charge of carrying out Part 2 direct assessment used validated instruments to quickly and reliably gage children's development relative to their peers. The developmental outcomes to be measured were determined based on the results of the pilot study and existing knowledge of those areas known to be affected in classic galactosemia patients. Whenever possible, testers used standardized measures adjusted for the child's age and sex so that these factors could be omitted from the model in the analysis stage. Outcomes included fell into the following domains: cognitive skills, communication processes, physical development, and socio-emotional development. Each instrument used to quantify functioning in these areas is described below and listed in Table 1.

Cognitive Skills

Several aspects of child cognition were evaluated including memory, executive function, and intelligence. The Children's Memory Scale (CMS) was one primary tool used to assess learning and memory (24). Designed for children and adolescents of ages 5 through 16 years, the CMS explores a variety of memory dimensions such as attention and working memory, verbal and visual memory, short- and long-delay memory, recall and recognition, and learning characteristics. Two index scores from this assessment are included as primary outcomes in the analysis, the Attention/Concentration index and the Delayed Recognition Index. Index scores represent summations of children's performance across multiple subtests and are standardized with a mean of 100 and a standard deviation of 15. High scores on the CMS are associated with better performance while low scores may indicate deficits in some aspect of learning and memory.

Working memory was also measured using another validated testing instrument, the Wechsler Intelligence Scale for Children–Fourth Edition Integrated (WISC–IV Integrated) (25). In this case, only one subtest of the assessment was included: Spatial Span, which provides an evaluation of spatial working memory. This test makes use of a spatial span board containing ten cubes placed in random order. During the assessment, the administrator taps cubes in a certain sequence then asks the child to replicate it, both forwards and backwards (26). Scores from the backwards version of the test were included as an outcome in this analysis. WISC-IV subtests are scored on a scaled metric with a mean of 10 and a standard deviation of 3. Children can score anywhere from 1 to 19 and are typically considered average if they fall between 8 and 12.

NEPSY-II: A Developmental Neuropsychological Assessment served as the primary measure of executive functioning, an umbrella term for mental processes required to focus

attention, remember instructions, and manage multiple tasks successfully (27, 28). The Word Generation subtest from NEPSY was included as an outcome in the analysis. This subtest comprises two related tasks: Semantic and Initial Letter word generation. In the former, children are asked to name as many animals, and then as many foods or drinks, as possible within a sixty second time limit. Then, in the latter task, children must list as many words as he or she can think of that begin with the letter S, and then the letter F, again within sixty seconds. The Initial Letter Word Generation task is administered only to children between the ages of 7 and 16. Because the study population encompassed ages 6 to 12, the youngest subjects were ineligible to take this test, meaning that this particular outcome has substantially more missing data than any of the others considered. Scores from both tasks were required to compute the outcome measurement, the Word Generation Semantic versus Initial Letter contrast scaled score. In NEPSY, contrast scores are used to compare performance across two cognitive functions, in this case ability to generate words in a specified sematic category versus this same ability when prompted with just the initial letter. NEPSY's creators designed the two tasks to measure distinct aspects of cognition; the word generation semantic score is intended to capture executive control of language production, ideation, and knowledge of vocabulary while the initial letter score theoretically requires more efficient executive functions. High contrast scores are thus interpreted to mean that the child can produce language and focus attention, but may lack an adequate search strategy for retrieving information that is not categorically organized. Low scores are less common and are indicative of undeveloped semantic association networks relative to overall word knowledge (29). Deficits in either respect could suggest problems in cognitive development.

The final instrument used to quantify performance in the cognitive domain was the Wechsler Abbreviated Scale of Intelligence-II (WASI-II), a measure of general intelligence commonly used in research settings due to its consistency and relative brevity (30, 31). Two of the overall four subtests factored into the full scale standardized IQ score used as an outcome in the analysis: Vocabulary and Matrix Reasoning. The IQ scores obtained from WASI have a mean of 100 and a standard deviation of 15. Scores between 70 and 79 are classified as borderline and those less than 70 can be considered extremely low.

Communication Processes

The second domain included in direct assessment, communication processes, also relied on measures from multiple instruments. The first of these was the Oral and Written Language Scales, Second Edition (OWLS-II), an assessment of receptive and expressive language. Study subjects participated in only two subtests: Listening Comprehension, a measure of oral language reception—the understanding of spoken language, and Oral Expression, the use of spoken language. In the Listening Comprehension section, test administrators present increasingly difficult words, phrases, and sentences and request that the child point to which of four pictures is correct. The procedure for the Oral Expression test is nearly the reverse: the administrator presents a verbal prompt along with a picture and the child must respond orally with increasingly complex language (32). Standard scores for these measures are centered on a mean of 100 with a standard deviation of 15; scores falling within one standard deviation of the mean indicate average abilities. Children with low scores may be at risk for language disorders or deficits relative to their peers.

Another critical element of communication, the accuracy of speech sound production, was assessed using the Goldman-Fristoe Test of Articulation-3 (GFTA-3). GFTA offers

systematic means of evaluating an individual's articulation of the consonant and consonant cluster sounds of Standard American English (33). It provides an informative view of subjects' true capabilities by sampling both spontaneous and imitative sound production. As with OWLS-II, standard scores have a mean of 100 and a standard deviation of 15, with low scores suggesting the possibility of problems with intelligibility and proper pronunciation.

Physical Development

The possibility of problems with movement and coordination were assessed using several instruments, one well-known and widely used tool, one more recent development, and one extremely new protocol revised for the purposes of this study. The first of these instruments was the Movement Assessment Battery for Children-2 (MABC-2), which is commonly used to identify impairments in motor performance of children and adolescents between 3 and 16 years (34). It contains eight tasks that evaluate abilities in three areas: manual dexterity, static and dynamic balance, and catching and releasing a ball. Raw scores on the eight tasks are converted to standard scores, which can then be added together to obtain the total standard test score—the outcome measurement included in this analysis. Standard MABC-2 scores have a mean of 10 and a standard deviation of 3; very low scores indicate that the child is at risk for movement difficulty.

Results from the pilot study suggested that children with Duarte galactosemia were more likely to experience visible hand tremors when attempting to draw a smooth line for one of the MABC tasks. Based on this finding, The Essential Tremor Rating Assessment Scale (TETRAS) was included as an additional instrument in the present study. Developed by the Tremor Research Group, TETRAS consists of nine items in which the degree of tremor is rating on a scale from 0 to 4, with 0 indicating no tremor and 4 indicating severe tremor (35). The test items

are as follows: head tremor, face tremor, voice tremor, upper and lower limb tremor during various maneuvers, tremor when attempting to draw Archimedes spirals, tremor when providing a sample of cursive handwriting, standing tremor, and tremor during a dot approximation task in which the subject holds the tip of a pen "as close as possible to a dot on a piece of paper without touching it." TETRAS tasks were videotaped and scored by a tester with advanced training in kinesiology. Measurement reliability of the Archimedes spirals task was additionally improved by having subjects draw on a piece of paper placed on a Wacom Intuos5 electronic tablet, which records pen movements and pressure. Steadiness and pen pressure could then be quantified using NeuroGlyphics software, improving sensitivity to small effects and inter-rater reliability. Unlike the other outcomes described thus far, scores from TETRAS could not be standardized around an accepted mean; it was therefore important to adjust for children's sex and age during the analysis. The summation of scores on the nine sub-items served as the primary outcome measurement, and since higher scores on each task corresponded with greater degree of tremor, for this outcome, high numbers are suggestive of problems.

The final instrument used to assess subjects' physical functioning was the structure, function, praxis (SFP) exam. Because no gold standard protocol exists, the methods used were first developed in collaboration between this study's experts and colleagues at Mayo Clinic and later revised to be better suited to the purposes of this analysis. The final rubric includes a three-pronged examination of the child's oral motor development. In the Structure and Tone section, testers assess the anatomical structure of the face and mouth and check for the presence of drooling or open mouth posture. The Function portion of the protocol unsurprisingly involves examining the function of the tongue, lips, jaw, and soft palate. One element of this exam also focused on the phonatory quality of the subject's voice in a task during which they made the

"ah" sound as long as possible. The last segment of the exam measured coordination during non-speech movement of the mouth and tongue. Children were asked to smile, puff out their cheeks, smack their lips together, clear their throat, bite their lower lip, and imitate blowing out a candle while the examiner rated their level of success on each task. Every element of SFP was scored as 0 if no observable issues were present and as a higher number if the child struggled to perform a requested action or if the examiner noted some structural abnormality. The total score included as an outcome represents the summation from all of the SFP subsections. Like for TETRAS, standardization was not a possibility due to the fact that the instrument is very recently developed and lacks broad application.

Socio-emotional Development

The instruments used to assess the subjects' socio-emotional development are distinct from those described thus far in that they relied on parental report of their child's behavior rather than direct assessment by experts who were blinded to case-control status. Parents were first asked to complete the Child Behavior Checklist (CBCL), a 118 item questionnaire which requests information about a wide range of behavioral and emotional problems (36). The CBCL provides summary scores which aim to capture the degree to which the child exhibits Internalizing and Externalizing behavior, two outcomes included in this study. Internalizing behaviors reflect mood disturbance, including depression, anxiety and social withdrawal while externalizing behaviors are characterized by conflict with others and violation of social norms (37). Raw scores can be converted to standardized T scores with a mean of 50 and a standard deviation of 10. For the Internalizing and Externalizing behavior summary scores, scores less than 60 are considered normal, scores between 60 and 63 are borderline, and scores greater than 63 are in clinical range.

Problems with internalizing behavior were also evaluated more directly using the short forms of the Children's Depression Inventory-2 (CDI-2) and the Revised Children's Manifest Anxiety Scale: Second Edition (RCMAS-2). The short form of the CDI-2 is a twelve-item tool used to quickly screen children for depressive symptoms (38). Like the CBCL, raw scores from the CDI-2 can be standardized to T scores that have a mean of 50 and a standard deviation of 10, with high scores posing cause for concern. RCMAS in its short form consists of ten yes/no items designed to measure the level and nature of anxiety (39). Results from the Social Anxiety subscale were included as an outcome in this analysis, again in the standardized T score form. Interpretation follows the same pattern seen with the CBCL and CDI-2, with scores falling close to the mean of 50 considered normal and those at least one standard deviation above the mean generally thought to be of clinical interest.

Response to Sensory Stimuli

The variables described above are the primary outcomes the study was designed to measure and compare in cases and controls. A supplementary analysis included an additional set of outcomes derived not from direct testing but from the Part 1 surveys completed by parents. These surveys included a short series of yes/no questions addressing whether the child experienced heightened sensitivity to touch or texture, smell or taste, visual signals, or sound. Though not the emphasis of the present work or marked by same level of rigor as the primary outcomes, results regarding unusual sensitivity to each of these four sensory channels are also presented here.

*Analyses of Data*

Data analysis was performed using SAS 9.4 software. Two key variables were generated for the purpose of the analysis. The case control variable was created based on information from two separate fields of the original dataset, one containing the GALT genotyping results and the other containing the child's diagnostic status as reported by parents. GALT genotyping results were not available for all subjects at the point of this analysis. The case control variable is therefore coded as equal to 1 ("case") if GALT genotyping results confirmed a diagnosis of Duarte galactosemia or, for subjects whose results were unavailable, if parents reported that the child had the condition. The remaining subjects are classified as controls ("case=0"). In rare cases in which the GALT genotyping information contradicted parental report, the conclusion supported by the genotype data took precedence.

The variable used to represent the child's exposure to dietary galactose also drew information from multiple sources, this time several items of the Part 1 parent survey. Parents responded to a series of questions asking about their child's consumption of breast milk or other dairy products in specific windows of time. The galactose exposure variable took into account parents' answers as to the child's breastfeeding between 2 and 6 months, breastfeeding between 7 and 12 months, and consumption of other dairy products from birth to 1 year. Only children meeting criteria for the lowest level of exposure in each of these windows were classified as unexposed ("milk exposure=0"). All other children with some degree of galactose consumption reported in one or more of the questions were classified as exposed ("milk exposure=1").

Because controls were recruited from the families of cases, it was necessary to use an analysis method capable of adjusting for the expected correlation between siblings. To this end, the GENMOD procedure in SAS, which fits models to correlated responses by the generalized

estimating equations (GEE) method, was implemented for all analyses. An exchangeable correlation structure was determined to be most appropriate, as the association between any given pair of siblings within a family can be assumed to be equal.

A variety of covariates were considered for inclusion in the final model: state of birth, race, total combined annual household income, presence of neonatal complications, comorbid health conditions, parent/caregiver stress, and highest level of parental education. Collinearity was first assessed using the GENMOD-specific SAS macro (SAS Macro, Department of Epidemiology, Rollins School of Public Health at Emory University). The criterion for identifying collinearity concerns was a condition index greater than 30 accompanied by at least two variance decomposition proportions over 0.5.

Presence of confounding was next evaluated with the all possible subsets approach, in which results obtained from every subset of covariates are compared to determine whether some variables can be dropped without changing the estimate of the main exposure. The issue of confounding is a complex one in this analysis, as few of the covariates mentioned above would be expected to have a causal relationship with a genetic disorder like Duarte galactosemia, and would thus fail to meet the definition of a true confounder. In fact, the only variable included on this list with the capacity to influence whether a child carries the Duarte galactosemia genotype is race, and the confounding potential of this factor is reduced based on the fact that participants were overwhelmingly white. Because the diagnostic status of some subjects relied on parental report rather than genotyping results, state of birth could also qualify as a confounder, as likelihood of detection differs dramatically depending on state of birth (13). Other variables described such as socioeconomic status or parental stress have clear causal ties to developmental outcomes but questionable potential to have any impact on Duarte galactosemia diagnostic

status. Variables like presence of neonatal complications and diagnosis with comorbid health conditions are also worthy of careful consideration because they could feasibly exist as intermediates on the hypothesized causal pathway between Duarte galactosemia and developmental difficulties. For example, Duarte galactosemia could cause an infant to suffer from prenatal complications such as failure to thrive, and challenges during this early window could lead to delays later on in the child's development. Because of the nuances highlighted here, care was taken to approach adding variables to the model with great deliberation, though some covariates specifically identified by the primary investigators for this role were retained.

For the most part, interaction was not relevant to the main study questions and was thus not considered. One interaction term, between case control status and the milk exposure variable, was included in the model because of its critical importance to the second primary aim. This variable was used to identify if scores on developmental measures differed amongst cases depending on whether the child consumed dietary galactose during the first year of life. Significant interactions terms would suggest the effect of being a case was not consistent across the entire pool of children with Duarte galactosemia but variable contingent on exposure to milk.

The final models vary somewhat for the different outcome variables examined, as most are standardized scores already adjusted for age and sex but some had no such in-built correction and thus needed to include these extra terms. In addition, the dichotomous models for the sensory sensitivity outcomes are reduced relative to the continuous models due to challenges with model fit. Models with many covariates did not converge properly for this portion of the analysis, making it necessary to simplify.

One final point to address under the umbrella of data analysis relates to the high number of tests performed—fifteen for the scores from direct assessment and four more for the sensory

sensitivity results. Conducting nineteen tests on the same dataset invites an enormous potential for Type I error, that is, the probability of rejecting a null hypothesis when it is actually true. For *M* tests, the probability of committing at least one Type I error is $1 - (1-\alpha)^M$ where $\alpha$ represents the desired significance level, conventionally set at 0.05 for single comparisons. Adopting the standard alpha in this analysis while performing nineteen tests would therefore result in a 0.62 probability of observing at least one false positive. To ensure an overall Type I error rate of less than 0.05, the Bonferroni correction was applied to obtain an adjusted alpha level of 0.05/M, or 0.003. Terms in the final model were not considered statistically significant unless they met this conservative criterion.

# Results

*Demographics and Family Information*

Demographic characteristics of the 153 study subjects are presented in Table 1. Cases and controls were highly similar in regards to mean age, gender, racial background, level of parental education and socioeconomic status, measured by total combined annual household income. As seen in the pilot study, a strong majority of families were European-American. This trend likely reflects true differences in the prevalence of Duarte galactosemia, as previous research has found that the condition is much more common amongst Caucasians than in African Americans, and is rarer still amongst those of Asian descent (13). While the overall breakdown of socioeconomic status was consistent for cases and controls, some differences existed in the number of people supported by the stated income, with controls tending to come from larger families. Upon reflection, this observation is unsurprising considering that cases were eligible for inclusion whether or not they had siblings, but controls could participate only if they had a brother or sister with Duarte galactosemia. Despite this subtle variability introduced by the method of case selection, similar percentages of parents reported that financial concerns posed major limitations to their child's care and opportunities.

The vast majority of parents judged their child's health to be "good," "very good" or "great" at the time of the study; two parents of children with Duarte galactosemia children chose "fair." Some discrepancy existed in the percentage of children reported to suffer from conditions other than Duarte galactosemia: 30 percent of cases' parents reported that their child had at least one additional diagnosis while the number for controls was closer to 18 percent. Very few parents stated that developmental concerns motivated their decision to participate in the study, with only two parents of children with Duarte galactosemia choosing this option. Far more

claimed that they were inspired by desire to help future families or willingness to learn about Duarte galactosemia.

Seven states appeared on the lists of birthplace and current residence for the 153 subjects: Alabama, Georgia, Illinois, Iowa, Michigan, Missouri, and South Carolina. Georgia was the most well represented in both respects, serving as the state of birth and the state of residence for 21 percent of cases and 27 percent of controls (Tables 3 and 4). Illinois, on the other extreme, was the birthplace of just two subjects, one case and one control from the same family. The distribution of subjects across the seven states remained fairly consistent from time of birth to the point of the study, with a small number of shifts in between.

*Exposure to Prenatal and Neonatal Risk Factors*

Prenatal risk factors were infrequently reported by parents of the study subjects and, for the most part, were evenly distributed amongst cases and controls (Table 5). The most common exposure, maternal health problems, was also the only one to show any perceivable difference across the two groups: 12.2 percent of cases' mothers said they experienced general health issues during their pregnancy compared to 6.4 percent of controls' mothers. All chemical exposures were extremely rare based on parents' interview responses: no parents reported using alcohol, seven (three case parents and four control parents) reported smoking cigarettes, one control parent reported using prescription drugs, and two (one case parent and one control parent) reported using recreational drugs. Cases were slightly more likely than controls to suffer from neonatal complications such as jaundice and failure to thrive. According to their parents, approximately 24 percent of cases experienced these types of health issues early in life compared to 16 percent of controls (Table 6).

*Experience with Special Education Services*

Results from the parent questionnaire demonstrated that cases were more likely to receive special education services during the younger age frames, but this trend disappeared as the children grew older, resulting in no detectable difference by the point of the study. Parents of 8 of the 90 cases (9 percent) reported that their child had received early intervention or special education services before the age of three. No control children were reported to receive services during this period (Table 7). Six of the eight cases required services due to a broadly defined developmental delay while the remaining two qualified because they were considered to be at risk for delays. The apparent gap in need for special education services remained present, though to a slightly smaller extent, during the interval from age 3 to 5. Over these three years, 5 cases (6 percent of this group) received special services or accommodations while the analogous percentage for controls stayed at zero. Parents supplied a variety of reasons for why these services were necessary, including significant developmental delay, emotional disturbance, speech or language disorder, and specific learning disability. Speech and language disorders appeared to be the most common cause for eligibility, as parents of four of the five cases chose this option as at least one factor driving their child's need for extra support.

The questionnaire designed by this work's researchers also asked about the period between age six and participation in the study. However, the wording of this section resulted in an unfortunate conflation between special education services and gifted programs, making it difficult to parcel out which subjects may have been experiencing delays in this age range. The breakdown of subjects within each category specified in this section were similar for cases and controls (Table 7). At the point of the study, identical percentages of cases and controls were reported by parents to currently attend a classroom with some special education services.

A final segment of the questionnaire assessed the possibility of delays more directly by asking whether the child had ever been diagnosed with a developmental or attention/emotional disorder by a licensed professional. Contingent on the reliability of parental report, similar percentages of cases and controls had received a diagnosis, though cases were slightly more likely than controls to have a developmental disorder and controls included a relatively high percentage of children with attention or emotional disorders (Table 8).

*Exposure to Dietary Galactose*

The majority of cases experienced total restriction of galactose throughout the first year of life (Table 9). Of the ninety children with Duarte galactosemia, 56 (62 percent) did not drink breast milk or consume other dairy products from the age of two months (when parents can be expected to have received their child's diagnosis and been counselled as to dietary recommendations) to twelve months. The remaining 34 cases experienced varying degrees of milk exposure, with most falling on the lower end of the spectrum. Only 3 children with Duarte galactosemia were reported by their parents to have drank breast milk exclusively between two and twelve months and to have also regularly consumed other dairy products. Unsurprisingly, the proportion of controls exposed to milk in this time frame was comparatively high. Relatively few children in this category, 6 of the total 63, consumed neither breast milk nor other dairy products on a regular basis during their infant stage.

*Cognitive Skills*

Scores on the five tests used to assess abilities in the cognitive domain followed a consistent pattern: all were similar comparing across cases and controls and means for both groups tended to fall slightly above the average (Table 10). As described in the Methods section above, high scores on these measures are associated with better functioning, suggesting that subjects performed well overall. No significant differences were detected between cases and controls in any of the areas examined. Memory, working memory, executive functioning and intelligence did not appear to be compromised in cases relative to controls.

*Communication Processes*

Subjects also demonstrated adequate ability to understand and use spoken language, with scores from each OWLS-II assessment clustering almost exactly at the mean for both cases and controls (Table 10). Performance on the test of articulation was slightly lower overall than might be expected, but average scores for cases and controls still fell well within normal limits and were indistinguishable from one another. None of the three evaluations of communication differed significantly depending on case-control status, suggesting that at the point of study, cases did not experience deficits in language skills or in speech production compared to their unaffected siblings.

*Physical Development*

Cases and controls received highly comparable average scores on the three measures of movement and coordination (Table 10). MABC-2 scores were slightly lower than the standardized test average for cases, but fell inside one standard deviation of this figure and differed from the control subjects' mean score by an even lesser degree. Results from the

TETRAS and SFP assessments were virtually equivalent. Overall, the three tests of physical functioning offered no evidence to suggest that cases were impaired in any aspect of motor skills relative to controls.

*Socio-emotional Development*

The three instruments that measured functioning in the socio-emotional realm were mainly targeted towards capturing differences in internalizing behaviors. Scores on the CBCL, CDI-2, and RCMAS-2 all indicated no difference in internalizing behaviors including depression and social anxiety between cases and controls (Table 10). Average T scores for each group on the three assessments consistently fell extremely close to the standardized mean, far from clinical range. The CBCL evaluation of externalizing behavior also resulted in no perceivable difference between cases and controls. Across-the-board insignificance of findings in this domain suggests that cases are no more likely than controls to experience poor socio-emotional outcomes during the age frame relevant to this study.

*Response to Sensory Stimuli*

Based on parental report, cases appeared to be somewhat more likely to experience unusual sensitivity to sensory stimuli, most particularly sound and smell or taste. The adjusted odds ratios comparing reports of heightened sensitivity to these channels in cases versus controls were 2.03 and 2.27, respectively (Table 11). Both confidence intervals included the null, however, forbidding any conclusion about individuals with Duarte galactosemia being more likely to encounter sensory difficulties as they develop. The mild but consistent trend across categories of sensory sensitivity is suggestive of a potential difference, but the findings could

easily result from chance alone and are undermined by the fact that they stem entirely from parental report and not direct assessment.

*Impact of Dietary Galactose Exposure on Developmental Outcomes*

Discerning the potential effect of exposure to galactose was difficult because such a large proportion of cases were treated with a galactose-free diet over the first year of life. The mean scores on each of the fifteen primary outcomes are listed in Table 12 for the 56 cases whose parents reported adhering to a galactose restricted diet and the 34 cases whose parents reported some level of galactose exposure between 2 and 12 months. The fully adjusted model for each outcome included an interaction term for the effect of case status by galactose exposure level, intended to reveal differences in cases' developmental outcomes that were related to diet early in life. This term was insignificant for every one of the fifteen outcome variables, as indicated in the final column of Table 12. Reviewing the respective means of the two exposure groups, this is unsurprising, as most differ by negligible amounts. When slightly more substantial variation did exist, it was in the opposite of the hypothesized direction; for example, children that were exposed to galactose as infants had WASI-2 IQ scores that were on average eight points higher than their galactose-restricted peers. Cases treated with a galactose-free diet also had marginally higher CBCL Internalizing Behavior and RCMAS-2 Social Anxiety T Scores (indicating greater degree of difficulties), and lower scores on the OWLS-II measure of listening comprehension and the CMS standard score for attention and concentration (indicating poorer performance). These trends certainly do not demonstrate any type of harmful impact of galactose restriction, but they do undermine the belief that this intervention meaningfully improves long-term health outcomes for individuals with Duarte galactosemia.

## Discussion

The present analysis was centered on two primary goals. The first of these was to determine whether school-age children with Duarte galactosemia were at increased risk for disorders in the domains of cognitive skills, communication, physical development, and socio-emotional functioning. If delays were evident, the second central aim was to explore whether dietary restriction during infancy was associated with developmental outcomes amongst children with Duarte galactosemia.

To the extent these questions are reliably addressed by the methods presented here, the answer to both appears to be no. Direct testing of cases and controls offered no evidence that children with Duarte galactosemia lagged behind their unaffected siblings in any of the relevant developmental measures. Data obtained from parental report, theoretically more vulnerable to bias as parents are not blinded to their child's diagnostic status, pointed to the same conclusion. For the most part, mean scores for all fifteen outcome variables were extremely similar for cases and controls and differences between groups fell far short of statistical significance by even the conventional 0.05 alpha level let alone the more conservative Bonferroni corrected value. The absence of any meaningful difference suggests cases fared no worse than controls performing tasks that demanded skills in developmental areas known to be affected in classic galactosemia patients. Results from the supplementary analysis of sensory sensitivity data were consistent with the trends seen for the main outcomes, with cases once again no more likely to experience difficulties than controls.

In regards to the second prong of the analysis, cases who were exposed to dietary galactose early in life experienced no observable difference in long-term outcomes compared to cases whose diets were restricted during this stage. Average scores on the fifteen outcome

variables were by and large indistinguishable regardless of whether cases were classified as galactose exposed or galactose restricted. When weak trends did exist, they were in the opposite direction than might be expected, with exposed cases faring better than those treated with the dietary intervention. Overall, the results of the galactose exposure analysis offered no evidence supporting the value of restricting diet early in life relative to allowing for some degree of breastfeeding or consumption of dairy products.

An important limitation to the galactose exposure analysis was the heavy skew amongst cases towards total galactose restriction. The majority (56 percent) of cases did not drink any milk throughout their first year of life. This trend towards adherence to a completely galactose restricted diet is likely not representative of the target population, the entire population of children with Duarte galactosemia living in the United States, and likely resulted from sampling subjects primarily from states that recommend this intervention. Such a large proportion of cases who were galactose restricted left a relatively small number of galactose exposed individuals on which to base key conclusions regarding comparative risk of disorders in development.

A second, similar issue relates to the exposure classification scheme imposed by this skewed distribution. Because so few cases had a high reported exposure level and most had none at all, it was necessary to consider children with any degree of galactose consumption as "exposed" in the analysis. The result of this was that some portion of the "exposed" cases drank milk only during one window of time in limited quantities. To some extent, the comparison therefore reduced to cases with no exposure to dietary galactose in the first year of life versus cases with low to moderate exposure to galactose in the first year of life. It is possible that if the comparison group had instead involved a substantial number of children with Duarte

galactosemia who had breast fed and consumed dairy without any restriction since birth, the conclusions would have been different.

As the findings summarized here represent an interim analysis and data collection is still ongoing at the point of this submission, it is possible that the final results will yield different conclusions. The total sample size will be at least 288 subjects, substantially larger than the 153 included at this midway point. Adding this number of subjects could increase the study's power to detect subtle differences showing promising trends in the smaller dataset but falling short of statistical significance. However, as explained above, none of the primary outcome variables analyzed here produced results that were even suggestive of a difference between cases and controls. Because scores were so comparable across the two groups, it would be unusual and somewhat concerning to suddenly see sharp differences after doubling the number of participants. Most likely, results from the completed study will lead to similar inferences regarding the effect of Duarte galactosemia on development within the assessed domains.

Conclusions about the impact of galactose exposure have more potential to meaningfully shift by the point of the final analysis. To counteract the current problem of having too few cases exposed to milk during infancy, the primary investigators are making every effort to increase enrollment of children that did not experience dietary restriction. If they are successful in recruiting a substantial number of children with high exposure, new trends could emerge that are indiscernible as of now because of the lack in variability in reported milk consumption.

One obvious question raised by the results reported here is why this study produced such different findings compared to the earlier pilot work performed by the same team of researchers. Lynch et al. (2014) reported a range of deficits experienced by children with Duarte galactosemia relative to their unaffected siblings. Several of the exact testing instruments were

used once again in this larger-scale investigation, but this time demonstrated no notable difference between groups. Why were trends that appeared promising in the pilot data completely absent in the results of the formal study? Part of the reason for this discrepancy may stem from the fact that the pilot study was extremely small in nature, involving just fifteen subjects. Because the number of participants was so limited, data analyses primarily took a descriptive form. It is possible that trends that appeared important at this small scale were appropriately muted when the sample size was multiplied many times over and analysis relied on more robust statistical techniques.

The authors also address the potential for bias introduced by the fact that seven of the ten cases had parents who had concerns about their development. They acknowledge that some results could be explained by parents who were already noticing delays being more likely to want to join a study investigating long term outcomes in children with Duarte galactosemia. Most of the deficits noted by Lynch et al. (2014) occurred exclusively in this subset while the other three cases were more similar to controls, making it seem more plausible that selection bias may have contributed to some of the observed effects.

This same type of bias could also impact the results of this analysis. Parents who have no worries about their child's development are likely to be less inclined to respond to requests for participation compared to parents who have come to believe Duarte galactosemia is the reason for some negative health outcome in their child. The risk of this bias having a strong influence on the results presented here is mitigated by two factors, however: first, that parents of only two of the 90 cases reported that they were primarily motivated by suspicion that their child had a developmental problem and second the overarching conclusion that no substantial differences existed. Cases consistently mapped close to the population means on the instruments used in

direct assessment, lessening the possibility that they represent a pool of poorly functioning children volunteered by their parents because of preexisting concerns. These indications point to the conclusion that the subjects of this study are representative of the group investigators wanted to measure, while the original fifteen children's participation may have been tied to the fact that they were experiencing delays.

Data collection for this study will come to a close over the next year and the final results will be published sometime within 2018. If the findings are consistent with those presented here, it could have major implications for the metabolic and screening communities, currently so divided on the issue of Duarte galactosemia. Given evidence that children with Duarte galactosemia develop normally regardless of dietary intervention, it could be reasonably argued that there is no point investing time and resources detecting infants with this condition. More newborn screening programs would be justified in following the example of states like Utah that intentionally adjusted protocol so that infants with classic galactosemia would be identified while those with Duarte galactosemia would go largely undetected. Public health dollars expended uncovering cases of a disorder that requires no treatment and may be all but benign would be better spent on countless more effective causes. In this sense, the study's results have the potential to be incredibly useful and could bring about an end to decades of diagnosing children and assigning interventions unnecessarily. Finally, policymakers and clinicians will be able to cite reliable scientific evidence in making recommendations, and parents of infants carrying the Duarte galactosemia genotype will have the comfort of knowing that in all likelihood their child will experience completely normal development.

The remaining year of data collection of this study will seek to resolve the most pressing limitation of this analysis: the skew towards low galactose exposure which hindered ability to

address the second central aim. If the expanded range of exposure does reveal differences, future research should explore the fascinating question of how consumption of milk products and the corresponding buildup of galactose metabolites affects children in a way that makes them more likely to encounter developmental difficulties later in life. On the other hand, if the complete results support the conclusions described here, it may be more valuable to invest time and energy redesigning screening programs and updating recommendations than executing further studies searching for differences between cases and controls. Classic galactosemia is known to cause a wide variety of health outcomes in patients, however, so future work could focus on the physiological symptoms not explored here, perhaps in an expanded age range of study participants. Further research may uncover yet unexplored differences distinguishing cases of Duarte galactosemia, but the results of this analysis are clear: children with Duarte galactosemia were not at increased risk for disorders in development for any of the domains assessed. All evidence presented here suggests the prognosis for children diagnosed with Duarte galactosemia is good, with or without dietary intervention.

# References


1. Fridovich-Keil JL, Walter JH (2008) Galactosemia. In: Valle D, Beaudet A, Vogelstein B, Kinzler K, Antonarakis S, Ballabio A (eds) The online metabolic & molecular bases of inherited disease. McGraw Hill, Available from: http://www.ommbid.com/
2. Berry GT (2000) Classic Galactosemia and Clinical Variant Galactosemia. In: Pagon RA, Adam MP, Ardinger HH, et al., (eds). GeneReviews [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2015. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1518/
3. Levy HL, Sepe SJ, Shih VE, Vawter GF, Klein JO (1977) Sepsis due to Escherichia coli in neonates with galactosemia. N Engl J Med. 297:823–825.
4. Genetic Alliance; District of Columbia Department of Health (2010) Understanding Genetics: A District of Columbia Guide for Patients and Health Professionals. In: Washington (DC): Genetic Alliance. Chapter 4, Newborn Screening. Available from: https://www.ncbi.nlm.nih.gov/books/NBK132148/
5. Beutler E, Baluda M (1966) A simple spot screening test for galactosemia. J Lab Clin Med 68:137-141.
6. Fernhoff PM (2010) Duarte galactosemia: how sweet is it? Clin Chem 56:1045-1046.
7. Welling L, Bernstein LE, Berry GT, Burlina AB, Eyskens F, Gautschi M, Grünewald S, Gubbels CS, Knerr I, Labrune P, van der Lee JH, MacDonald A, Murphy E, Portnoi PA, Õunap K, Potter NL, Rubio-Gozalbo ME, Spencer JB, Timmers I, Treacy EP, Van Calcar SC, Waisbren SE, Bosch AM. Galactosemia Network (GalNet). (2017) International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. J Inherit Metab Dis 40: 171–176.
8. Jumbo-Lucioni, P. P., Garber, K., Kiel, J., Baric, I., Berry, G. T., Bosch, A., Fridovich-Keil, J. L., et al. (2012). Diversity of approaches to classic galactosemia around the world: a comparison of diagnosis, intervention, and outcomes. J Inherit Metab Dis 35: 1037–1049.
9. Ridel, KR, ND Leslie, DL Gilbert (2005). An updated review of the long-term neurological effects of galactosemia. Pediatric Neurology 33: 153-161.
10. Fridovich-Keil JL, Gambello MJ, Singh RH, et al. (2014) Duarte Variant Galactosemia. In: Pagon RA, Adam MP, Ardinger HH, et al., (eds) GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2017. Available from: https://www.ncbi.nlm.nih.gov/books/NBK258640/
11. Lai, K., SD Langley, PP Dembure, LN Hjelm, LJ Elsas (1998). The Duarte allele impairs biostability of galactose-1-phosphate uridyltransferase in human lymphoblasts. Hum. Mutat. 11: 28–38.
12. Ficicioglu C, Thomas N, Yager C, Gallagher PR, Hussa C, Mattie A, Day-Salvatore DL, Forbes BJ (2008). Duarte (DG) galactosemia: a pilot study of biochemical and neurodevelopmental assessment in children detected by newborn screening. Mol Genet Metab. 95: 206–12.
13. Pyhtila B, Shaw K, Neumann S, Fridovich-Keil J (2015). Newborn screening for galactosemia in the United States: looking back, looking around, and looking ahead. JIMD Rep. 15: 79–93.

14. American Academy of Pediatrics (2012). Breastfeeding and the use of human milk. Pediatrics 129: e827–e841.
15. Van Naarden Braun K, M Yeargin-Allsopp, D Schendel, P Fernhoff (2003). Long-term developmental outcomes of children identified through a newborn screening program with a metabolic or endocrine disorder: a population-based approach. J Pediatr 143: 236-242.
16. Powell KK, Van Naarden BK, RH Singh, SK Shapira, RS Olney, M Yeargin-Allsopp (2009). Long-term speech and language developmental issues among children with Duarte galactosemia. Genet Med 11: 874-879.
17. Waisbren SE, Norman TR, Schnell RR, Levy HL (1983). Speech and language deficits in early treated children with galactosemia. J Pediatr 102: 75–77.
18. Nelson CD, DD Waggoner, GN Donnell, JM Tuerck, NR Buist (1991). Verbal dyspraxia in treated galactosemia. Pediatrics 88: 346 –350.
19. Lynch ME, NL Potter, CD Coles, JL Fridovich-Keil (2015). Developmental Outcomes of School-Age Children with Duarte Galactosemia: A Pilot Study. JIMD Reports 19: 75–84.
20. Antshel K, I Epstein I, S Waisbren (2004). Cognitive strengths and weaknesses in children and adolescents homozygous for the galactosemia Q188R mutation: a descriptive study. Neuropsychology 18:658–664.
21. Bosch AM, MA Grootenhuis, HD Bakker, HS Heijmans, FA Wijburg, BF Last (2004). Living with classical galactosemia: health-related quality of life consequences. Pediatrics 113:e423–e428.
22. Ryan E, M Lynch, E Taddeo, T Gleason, M Epstein, J Fridovich-Keil (2013). Cryptic residual GALT activity is a potential modifier of scholastic outcome in school age children with classic galactosemia. J Inher Metab Dis 36:1049–1061.
23. Waisbren S, N Potter, C Gordon, et al (2012). The adult galactosemic phenotype. J Inherit Metab Dis 35:279–286.
24. Cohen, M. (1997). Children's Memory Scale. San Antonio, TX, The Psychological Corporation.
25. Kaplan, E., D. Fein, J. Kramer, D. Delis and R. Morris (2004). Wechsler Intelligence Scale for Children, Fourth Edition Integrated (WISC-IV Integrated). San Antonio, TX, PsychCorp, Pearson. http://psychcorp.pearsonassessments.com.
26. Locascio, G, EM Mahone, SH Eason, LE Cutting (2010). Executive dysfunction among children with reading comprehension deficits. J Learn Disabil 43: 441–454.
27. Korkman, M., U. Kirk and S. Kemp (1998). NEPSY: A developmental neuropsychological assessment. San Antonio, TX, Psychological Corporation.
28. Korkman, M., U. Kirk and S. Kemp (2007). NEPSY II, Second Edition: Clinical and Interpretive Manual. San Antonio, TX, NCS Pearson, Inc. http://www.pearsonassessments.com.
29. Miller, DC. "How To Administer And Interpret The NEPSY-II". 2015. Presentation.
30. Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence (WASI). San Antonio, TX, The Psychological Corporation, Harcourt Brace & Co. Wechsler, D. (2011).
31. Wechsler Abbreviated Scale of Intelligence- Second Edition (WASI-II) San Antonio, TX, The Psychological Corporation, Harcourt Brace & Co.
32. Carrow-Woolfolk, E. (2011). Oral and Written Language Scales, 2nd Edition. Minneapolis, MN, Pearson Assessments.

33. Goldman, R., and Fristoe, M. (2015). Goldman-Fristoe Test of Articulation, Third Edition (GFTA-3). San Antonio, TX, PsychCorp, Pearson. http://psychcorp.pearsonassessments.com.
34. Henderson, S., D. Sugden and A. Barnett (2007). Movement Assessment Battery for Children-2 second edition (Movement ABC-2). London, UK, The Psychological Corporation.
35. Elble, R., C. Comella, S. Fahn, M. Hallett, J. Jankovic, J. L. Juncos, P. Lewitt, K. Lyons, W. Ondo, R. Pahwa, K. Sethi, N. Stover, D. Tarsy, C. Testa, R. Tintner, R. Watts and T. Zesiewicz (2012). Reliability of a new scale for essential tremor. Mov Disord 27: 1567–1569.
36. Achenbach, T. and L. Rescorla (2001). Manual for ASEBA School-Age Forms & Profiles. Burlington, VT, University of Vermont, Research Center for Children, Youth, & Families.
37. Lande, MB, H Adams, B Falkner, SR Waldstein, GJ Schwartz, PG Szilagyi, H Wang, D Palumbo (2009). Parental assessments of internalizing and externalizing behavior and executive function in children with primary hypertension. J Pediatr 154: 207–212.
38. Kovacs, M. (2010). Children's Depression Inventory-2. Pearson Assessments. http:// pearsonclinical.com.
39. Reynolds, C. and B. Richmond (2008). Revised Children's Manifest Anxiety Scale: Second Edition (RCMAS- 2). Los Angeles, CA, Western Psychological Services. www.wpspublish.com.
40. SAS. Cary: SAS Institute Inc.

# Tables

**Table 1**. Demographic characteristics for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| Characteristic | Duarte galactosemia (*n*=90) | Control (*n*=63) |
|---|---|---|
| Mean Age at Part 2 Testing ± SD (years) | 9.44 ± 2.09 | 9.35 ± 1.97 |
| Male—no. (%) | 53 (58.9) | 35 (55.6) |
| Race/ethnicity—no. (%) | | |
| White/not of Hispanic origin | 82 (91.1) | 58 (92.1) |
| White/of Hispanic origin | 0 (0) | 1 (1.6) |
| Black/not of Hispanic origin | 4 (4.4) | 3 (4.8) |
| Mixed | 4 (4.4) | 1 (1.6) |
| Main reason for participation—no. (%) | | |
| I wanted to help future families | 48 (65.8) | 35 (62.5) |
| I wanted to learn more about Duarte galactosemia | 18 (24.7) | 13 (23.2) |
| I liked the idea of earning a gift card while helping with research | 5 (6.9) | 8 (14.3) |
| I suspected my child might have a developmental problem | 2 (2.7) | 0 (0) |
| Other diagnosed conditions—no. (%) | | |
| Yes | 27 (30.0) | 11 (17.5) |
| No | 63 (70.0) | 52 (82.5) |
| Overall health rating | | |
| Great | 53 (58.9) | 30 (47.7) |
| Very good | 24 (26.7) | 26 (41.3) |
| Good | 11 (12.2) | 7 (11.1) |
| Fair | 2 (2.2) | 0 (0) |
| Poor | 0 (0) | 0 (0) |
| Highest level of education obtained by either parent—no. (%)[a] | | |
| Did not complete high school | 6 (6.9) | 2 (3.4) |
| High school degree or equivalent | 4 (4.6) | 2 (3.4) |
| Some college | 27 (31.0) | 22 (37.3) |
| Bachelor's degree | 22 (25.3) | 15 (25.4) |
| Graduate degree | 28 (32.2) | 18 (30.5) |
| Total combined household annual income—no. (%)[a] | | |
| less than $25,000 | 8 (9.2) | 2 (3.4) |
| $25,000-$49,999 | 14 (16.1) | 11 (18.6) |
| $50,000-$74,999 | 17 (19.5) | 12 (20.3) |
| $75,000-$99,999 | 15 (17.2) | 10 (17.0) |
| $100,000-$149,999 | 25 (28.7) | 15 (25.4) |
| $150,000 or more | 8 (9.2) | 9 (15.3) |

Table 1, continued

| Characteristic | Duarte galactosemia (*n*=90) | Control (*n*=63) |
|---|---|---|
| Number of people supported by income—no. (%)[a] | | |
| 2 | 5 (5.8) | 0 (0) |
| 3 | 11 (12.6) | 5 (8.5) |
| 4 | 36 (41.4) | 20 (33.9) |
| 5 | 19 (21.8) | 18 (30.5) |
| 6 | 16 (18.4) | 16 (27.1) |
| Financial concerns pose major limitations—no. (%)[a] | | |
| Yes | 6 (6.9) | 3 (5.1) |
| No | 81 (93.1) | 56 (94.9) |

[a] For families that provided more than one child as study subjects, the same family-level information is included in the table multiple times, once for each child who participated.

**Table 2**. Measures used in child evaluation.

| Variable | Measure |
|---|---|
| *Cognitive Skills* | |
| Memory | Children's Memory Scale |
| Working memory | Wechsler Intelligence Scale for Children IV-Integrated (WISC-IV-Integrated)<br>• Spatial Span subtest |
| Executive Function | NEPSY-II<br>• Word Generation subtest |
| Intelligence | Wechsler Abbreviated Scales of Intelligence-II (WASI-II)<br>• Vocabulary and Matrix Reasoning subtests |
| *Language/communication* | |
| Receptive and expressive language | Oral and Written Language Scales, Second Edition (OWLS-II)<br>• Listening Comprehension (receptive) and Oral Expression (expressive) subtests only |
| Articulation | Goldman Fristoe Test of Articulation-3 (GFTA-3) |
| *Physical Development* | |
| Balance, coordination, manual dexterity | Movement Assessment Battery for Children-2 (MABC-2) |
| Tremor | The Essential Tremor Rating Assessment Scale (TETRAS) |
| *Social skills, behavior problems, and socio-emotional development* | |
| Behavior problems | Child Behavior Checklist (CBCL)[a] |
| Internalizing problems | Children's Depression Inventory-2 (CDI-2) (Short Form)[a]<br>Revised Children's Manifest Anxiety Scale Second Edition (RCMAS-2) (Short Form)[a] |

[a] The CBCL, CDI-2 and RCMAS-2 scores were determined based on parents' survey responses. All other measures described in this table were obtained during direct child assessments administered by trained professionals who were blinded to the children's case-control status.

**Table 3**. State of birth for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| State of Birth | Duarte galactosemia (*n*=90) *n* (%) | Control (*n*=63) *n* (%) |
|---|---|---|
| Alabama | 6 (6.67) | 2 (3.17) |
| Georgia | 20 (22.22) | 17 (26.98) |
| Illinois | 1 (1.11) | 1 (1.59) |
| Iowa | 16 (17.78) | 13 (20.63) |
| Michigan | 16 (17.78) | 12 (19.05) |
| Missouri | 19 (21.11) | 14 (22.22) |
| South Carolina | 12 (13.33) | 4 (6.35) |

**Table 4**. State of residence for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| State of Residence | Duarte galactosemia (*n*=90) *n* (%) | Control (*n*=63) *n* (%) |
|---|---|---|
| Alabama | 6 (6.67) | 2 (3.17) |
| Georgia | 19 (21.11) | 17 (26.98) |
| Illinois | 3 (3.33) | 3 (4.76) |
| Iowa | 15 (16.67) | 12 (19.05) |
| Michigan | 16 (17.78) | 12 (19.05) |
| Missouri | 18 (20.00) | 13 (20.63) |
| South Carolina | 13 (14.44) | 4 (6.35) |

**Table 5**. Prenatal conditions for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| Prenatal Condition | Duarte galactosemia (*n*=90) *n* (%) | Control (*n*=63) *n* (%) |
|---|---|---|
| Twin pregnancy | | |
| Yes | 4 (4.4) | 3 (4.76) |
| No | 86 (95.6) | 60 (95.2) |
| Preterm birth | | |
| Yes | 5 (5.6) | 2 (3.2) |
| No | 85 (94.4) | 61 (96.8) |
| Very low birthweight | | |
| Yes | 1 (1.1) | 0 (0) |
| No | 89 (98.9) | 63 (100) |
| Maternal health problems | | |
| Yes | 11 (12.2) | 4 (6.4) |
| No | 79 (87.8) | 59 (93.7) |
| Maternal alcohol use during pregnancy | | |
| Yes | 0 (0) | 0 (0) |
| No | 90 (100) | 63 (100) |
| Maternal cigarette smoking during pregnancy | | |
| Yes | 3 (3.3) | 4 (6.4) |
| No | 87 (96.7) | 59 (93.7) |
| Maternal prescription drug use during pregnancy | | |
| Yes | 0 (0) | 1 (1.6) |
| No | 90 (100) | 62 (98.4) |
| Maternal recreational drug use during pregnancy | | |
| Yes | 1 (1.1) | 1 (1.6) |
| No | 89 (98.9) | 62 (98.4) |
| Other | | |
| Yes | 3 (3.3) | 1 (1.6) |
| No | 87 (96.7) | 62 (98.4) |

**Table 6**. Neonatal complications for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| Presence of Neonatal Complications | Duarte galactosemia (*n*=90) no. (%) | Control (*n*=63) no. (%) |
|---|---|---|
| Yes | 22 (24.4) | 10 (15.9) |
| No | 68 (75.6) | 53 (84.1) |

**Table 7**. Experience with special education services for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| | Duarte galactosemia (*n*=90) no. (%) | Control (*n*=63) no. (%) |
|---|---|---|
| Received early intervention or special education services before age 3 | | |
| Yes | 8 (8.9) | 0 (0) |
| No | 82 (91.1) | 63 (100) |
| Received special education services or accommodations between ages 3 and 5 | | |
| Yes | 5 (5.6) | 0 (0) |
| No | 85 (94.4) | 63 (100) |
| Received special education services since age 6 | | |
| Yes–special services or a combination of gifted programs and special services | 22 (24.4) | 16 (25.4) |
| Yes–gifted program only | 9 (10.0) | 7 (11.1) |
| No special services | 59 (65.6) | 40 (63.5) |
| Current educational setting | | |
| Regular classroom | 72 (80.0) | 47 (74.6) |
| Regular classroom with some special education services | 10 (11.1) | 7 (11.1) |
| Home school setting | 8 (8.9) | 9 (14.3) |

**Table 8**. Diagnosis with a developmental disorder or an attention/emotional disorder by a licensed professional for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| Diagnosis Type | Duarte galactosemia (*n*=90) no. (%) | Control (*n*=63) no. (%) |
|---|---|---|
| Developmental disorder | | |
| Yes | 6 (6.7) | 2 (3.2) |
| No | 84 (93.3) | 61 (96.8) |
| Attention or emotional disorder | | |
| Yes | 11 (12.2) | 11 (17.5) |
| No | 79 (87.8) | 52 (82.5) |

**Table 9**. Exposure to dietary galactose in the first year of life for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| Type and Time Period of Exposure | Duarte galactosemia (*n*=90) no. (%) | Control (*n*=63) no. (%) |
|---|---|---|
| **Exposure to breast milk** | | |
| 2 months to 6 months | | |
| Drank only breast milk | 10 (11.1) | 17 (27.0) |
| Drank some breast milk | 17 (18.9) | 17 (27.0) |
| Did not drink breast milk | 63 (70.0) | 29 (46.0) |
| 7 months to 12 months | | |
| Drank only breast milk | 5 (5.6) | 9 (14.3) |
| Drank some breast milk | 12 (13.3) | 10 (15.9) |
| Did not drink breast milk | 73 (81.1) | 44 (69.8) |
| **Consumption of other forms of dairy milk or other dairy products on a regular basis** | | |
| Birth to 12 months | | |
| Yes | 15 (16.7) | 47 (74.6) |
| No | 75 (83.3) | 16 (25.4) |

**Table 10**. Comparison of fifteen direct testing variables for 90 cases and 63 controls.

| | Number of Observations | Mean | Standard Deviation | Crude p-value | Adjusted p-value[a] |
|---|---|---|---|---|---|
| **Internalizing Behavior T Score** | | | | | |
| Case | 88 | 52.06 | 11.68 | 0.74 | 0.66 |
| Control | 63 | 50.83 | 11.50 | | |
| **Externalizing Behavior T Score** | | | | | |
| Case | 88 | 50.22 | 8.93 | 0.24 | 0.77 |
| Control | 63 | 48.19 | 10.40 | | |
| **Articulation Standard Score** | | | | | |
| Case | 90 | 91.32 | 22.56 | 0.55 | 0.72 |
| Control | 63 | 93.32 | 21.47 | | |
| **OWLS II-Listening Comprehension Standard Score (receptive language)** | | | | | |
| Case | 90 | 101.50 | 11.86 | 0.82 | 0.30 |
| Control | 63 | 101.03 | 12.09 | | |
| **OWLS II-Expressive Communication Standard Score (receptive language)** | | | | | |
| Case | 90 | 100.12 | 11.83 | 0.41 | 0.29 |
| Control | 63 | 98.67 | 14.58 | | |
| **MABC-2—Total Standard Score** | | | | | |
| Case | 90 | 8.81 | 3.20 | 0.14 | 0.73 |
| Control | 62 | 9.44 | 3.17 | | |
| **CMS Attention and Concentration Index Standard Score** | | | | | |
| Case | 90 | 101.18 | 14.82 | 0.22 | 0.58 |
| Control | 63 | 104.86 | 16.46 | | |
| **CMS Delayed Recognition Index Standard Score** | | | | | |
| Case | 90 | 102.18 | 15.15 | 0.89 | 0.95 |
| Control | 63 | 102.62 | 13.97 | | |
| **WASI II Full Scale IQ Standard Score** | | | | | |
| Case | 90 | 105.23 | 13.92 | 0.87 | 0.91 |
| Control | 62 | 105.55 | 14.22 | | |

Table 10, continued

| | Number of Observations | Mean | Standard Deviation | Crude p-value | Adjusted p-value[a] |
|---|---|---|---|---|---|
| **NEPSY II Word Generation Semantic Versus Initial Letter Score** | | | | | |
| Case | 78 | 6.92 | 2.80 | 0.081 | 0.63 |
| Control | 52 | 7.90 | 3.39 | | |
| **WISC IV Integrated Spatial Span Backward Standard Scaled Score** | | | | | |
| Case | 90 | 10.83 | 2.80 | 0.74 | 0.15 |
| Control | 63 | 10.73 | 3.10 | | |
| **CDI 2 Depression T Score** | | | | | |
| Case | 90 | 54.90 | 13.02 | 0.54 | 0.75 |
| Control | 63 | 53.63 | 11.29 | | |
| **RCMAS 2 Social Anxiety T Score** | | | | | |
| Case | 90 | 47.23 | 9.59 | 0.33 | 0.69 |
| Control | 63 | 45.56 | 11.87 | | |
| **Total Structure Function Praxis Score** | | | | | |
| Case | 90 | 8.97 | 6.24 | 0.93 | 0.68[b] |
| Control | 63 | 8.95 | 6.07 | | |
| **TETRAS—The Essential Tremor Rating Scale** | | | | | |
| Case | 90 | 15.22 | 8.52 | 0.64 | 0.28[b] |
| Control | 63 | 14.69 | 7.70 | | |

[a] Adjusted for state of birth, total combined annual household income level, and whether or not child experienced full galactose restriction throughout first year of life

[b] Adjusted for state of birth, total combined annual household income level, sex, age at time of direct testing, and whether or not child experienced full galactose restriction throughout first year of life

**Table 11**. Report of sensory sensitivity for 63 Controls without Duarte galactosemia and 90 Cases with Duarte galactosemia.

| | Duarte galactosemia (*n*=90) no. (%) | Control (*n*=63) no. (%) | Crude p-value | Adjusted p-value | Adjusted OR |
|---|---|---|---|---|---|
| Unusual sensitivity to sound | | | | | |
| Yes | 21 (23.3) | 8 (12.7) | 0.086 | 0.13 | 2.03[a] (0.81, 5.12) |
| No | 69 (76.7) | 55 (87.3) | | | |
| Unusual sensitivity to visual signals | | | | | |
| Yes | 5 (5.6) | 2 (3.2) | 0.48 | 0.59 | 1.62[b] (0.27, 9.71) |
| No | 85 (94.4) | 61 (96.8) | | | |
| Unusual sensitivity to touch or texture | | | | | |
| Yes | 20 (22.2) | 9 (14.3) | 0.20 | 0.26 | 1.61[b] (0.68, 3.82) |
| No | 70 (77.8) | 54 (85.7) | | | |
| Unusual sensitivity to smell or taste | | | | | |
| Yes | 16 (17.8) | 6 (9.5) | 0.13 | 0.12 | 2.27[a] (0.77, 6.74) |
| No | 74 (82.2) | 57 (90.5) | | | |

[a] Adjusted for age, sex, and total combined annual household income level

[b] Adjusted for age and sex

**Table 12**. Comparison of fifteen direct testing variables for 56 cases that experienced dietary galactose restriction during the first year of life versus 34 cases that consumed breast milk or other dairy products throughout this period.

| | Number of Observations | Mean | Standard Deviation | Adjusted p-value of interaction term[a] |
|---|---|---|---|---|
| **Internalizing Behavior T Score** | | | | |
| Galactose Restricted | 55 | 53.53 | 11.22 | 0.56 |
| Galactose Exposed | 33 | 49.61 | 12.19 | |
| **Externalizing Behavior T Score** | | | | |
| Galactose Restricted | 55 | 50.36 | 8.97 | 0.64 |
| Galactose Exposed | 33 | 49.97 | 8.98 | |
| **Articulation Standard Score** | | | | |
| Galactose Restricted | 56 | 91.04 | 23.88 | 0.59 |
| Galactose Exposed | 34 | 91.79 | 20.53 | |
| **OWLS II-Listening Comprehension Standard Score (receptive language)** | | | | |
| Galactose Restricted | 56 | 99.77 | 11.28 | 0.84 |
| Galactose Exposed | 34 | 104.35 | 12.39 | |
| **OWLS II-Expressive Communication Standard Score (receptive language)** | | | | |
| Galactose Restricted | 56 | 100.55 | 10.47 | 0.48 |
| Galactose Exposed | 34 | 99.41 | 13.93 | |
| **MABC2—Total Standard Score** | | | | |
| Galactose Restricted | 56 | 8.63 | 3.19 | 0.90 |
| Galactose Exposed | 34 | 9.12 | 3.23 | |
| **CMS Attention and Concentration Index Standard Score** | | | | |
| Galactose Restricted | 56 | 98.46 | 14.78 | 0.52 |
| Galactose Exposed | 34 | 105.65 | 13.98 | |
| **CMS Delayed Recognition Index Standard Score** | | | | |
| Galactose Restricted | 56 | 102.20 | 14.50 | 0.85 |
| Galactose Exposed | 34 | 102.15 | 16.39 | |

Table 12, continued

| | Number of Observations | Mean | Standard Deviation | Adjusted p-value of interaction term |
|---|---|---|---|---|
| **WASI II Full Scale IQ Standard Score** | | | | |
| Galactose Restricted | 56 | 102.21 | 13.07 | 0.35 |
| Galactose Exposed | 34 | 110.21 | 14.03 | |
| **NEPSY II Word Generation Semantic Versus Initial Letter Score** | | | | |
| Galactose Restricted | 49 | 6.92 | 2.61 | 0.71 |
| Galactose Exposed | 29 | 6.93 | 3.13 | |
| **WISC IV Integrated Spatial Span Backward Standard Scaled Score** | | | | |
| Galactose Restricted | 56 | 10.71 | 2.78 | 0.24 |
| Galactose Exposed | 34 | 11.03 | 2.85 | |
| **CDI 2 Depression T Score** | | | | |
| Galactose Restricted | 56 | 55.84 | 13.36 | 0.95 |
| Galactose Exposed | 34 | 53.35 | 12.47 | |
| **RCMAS 2 Social Anxiety T Score** | | | | |
| Galactose Restricted | 56 | 49.14 | 9.50 | 0.52 |
| Galactose Exposed | 34 | 44.09 | 9.00 | |
| **Total Structure Function Praxis Score** | | | | |
| Galactose Restricted | 56 | 9.13 | 6.50 | 0.69 |
| Galactose Exposed | 34 | 8.71 | 5.88 | |
| **TETRAS—The Essential Tremor Rating Scale** | | | | |
| Galactose Restricted | 56 | 15.63 | 8.84 | 0.34 |
| Galactose Exposed | 34 | 14.54 | 8.05 | |

[a] Provided p-value is that of the interaction term between case status and milk exposure in the full model also adjusted for state of birth, total combined annual household income level, and, for the final two outcomes listed, age and sex