Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a 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 in whole or in part in all forms of media, now or hereafter now, 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. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Danielle Orloff April 8, 2023

Investigating Whether or Not Long-Term Outcomes in Classic Galactosemia are Progressive

by

Danielle Orloff

Dr. Judith Fridovich-Keil
Adviser

Biology

Dr. Judith Fridovich-Keil

Adviser

Dr. Andreas Fritz

Committee Member

Dr. William Kelly

Committee Member

2023

Investigating Whether or Not Long-Term Outcomes in Classic Galactosemia are Progressive

By

Danielle Orloff

Dr. Judith Fridovich-Keil

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Biology

2023

Abstract

Investigating Whether or Not Long-Term Outcomes in Classic Galactosemia are Progressive
By Danielle Orloff

Classic galactosemia (CG) is a rare autosomal recessive disorder that arises from a profound deficiency of galactose-1-phosphate uridylyltransferase (GALT), an enzyme necessary for galactose metabolism. Despite early detection of the disease and immediate restriction of dietary galactose, individuals with CG can still encounter long-term complications. A majority of the existing cross-sectional and longitudinal studies of speech, cognitive, motor, and reproductive outcomes in CG suggest that the disorder is not progressive. However, this point remains controversial. The objective of this study was to utilize a subset of participants enrolled in an ongoing case-control observational study to explore whether or not long-term outcomes in CG are progressive. CG cases and control participants were invited to complete a series of surveys that inquired about known long-term outcomes in CG. Based on survey responses, a subset of participants who completed all initial surveys (82 CG cases and 50 controls) were assigned scores to represent the severity of each health outcome experienced for cross-sectional analyses. Additionally, participants were invited to designate someone to complete the Vineland Adaptive Behavior Scales. All participants whose assessment was completed on their behalf (102 CG cases and 67 controls) were included in this investigation. Results indicated that a majority of participants perceived no changes or improvements to their adverse outcomes over time. In addition, when comparing the severity of health outcomes by CG status, scores for the severity of outcomes experienced by CG cases decreased or remained stagnant over time. Results from the Vineland-3 assessment indicated a general increase in subdomain raw scores of both CG cases and controls over time, though CG cases did have significantly lower raw scores in younger age groups. This suggests that while younger CG cases may struggle to reach milestones at the same pace as their peers, CG cases in adulthood are not losing adaptive behaviors they had previously achieved. The results from this thesis will allow for more accurate prognostic information that will help families with CG in planning for the future as well as inform how future interventions would best be used to treat patients with the disorder.

Investigating Whether or Not Long-Term Outcomes in Classic Galactosemia are Progressive

By

Danielle Orloff

Dr. Judith Fridovich-Keil

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Biology

2023

Acknowledgements

I would like to thank Dr. Judith Fridovich-Keil, my adviser, for the opportunity to work as an undergraduate research assistant in her lab and the tremendous support she provided while undertaking this honors thesis. Her mentorship throughout the process largely contributed to my successful completion of this project. In addition, I would like to thank Dr. Andreas Fritz and Dr. William Kelly for taking the time to serve on my committee and provide me with valuable feedback. Moreover, I would like to thank all the members of the Fridovich-Keil Lab for their support while working on this project, specifically Nikki Smith for monitoring participant completion of Vineland-3 as well as Nathan Paull for enrolling participants in the observational study and assisting with REDCap. I would also like to thank Nichole Stettner for helping score the severity of health outcomes for child participants in the study and Dr. Shauna Rasmussen for her guidance when preparing my honors thesis defense. Lastly, I would like to thank all of the individuals who participated in this observational study. Without their contributions, this study would not have been possible.

Table of Contents

Tables

Figures

CHAPTER 1: INTRODUCTION

Genetics of Galactosemia and Biochemistry of Galactose Metabolism

Classic galactosemia (CG) is a rare autosomal recessive disorder, affecting approximately 1 in 50,000 live births screened in the United States.[1] This metabolic disorder arises from a profound deficiency of galactose-1-phosphate uridylyltransferase (GALT), which is the second enzyme in the Leloir pathway of galactose metabolism (Figure 1).[2,3] In this pathway, GALT catalyzes the transfer of a uridine monophosphate group from UDP-glucose to galactose-1-phosphate via a double displacement reaction to form glucose-1-phosphate and UDP-galactose.[3] Profound deficits in or complete absence of GALT activity result in the accumulation of upstream metabolites in the pathway or products of alternate pathways, especially after dietary galactose exposure.[4] These include galactose, galactose-1-phosphate, galactitol, and in some tissues, galactonate.[4]

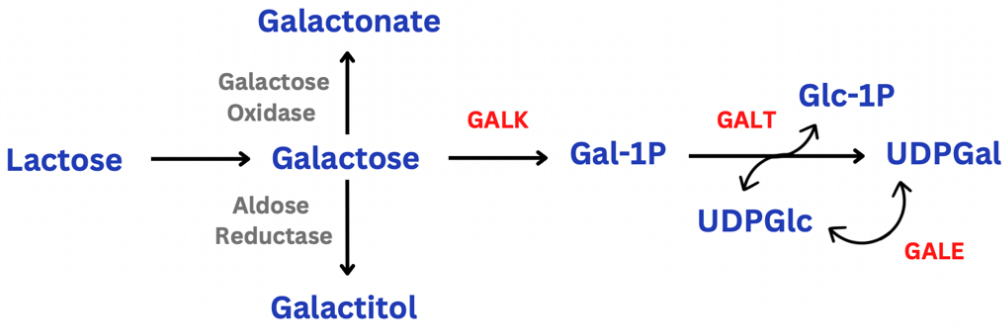


Figure 1. The Leloir Pathway of galactose metabolism in humans. The enzymes in the pathway are showcased in red and the metabolites are in blue. Enzymes that do not directly play a role in the pathway but catalyze reactions with metabolites in the pathway are presented in grey.

As of 2022, there have been 363 different variants of the *GALT* gene reported to the University of Utah's *GALT* database, a majority of which are missense mutations.[5] Frequently reported mutations of this gene include Q188R, K285N, and S135L; most other mutations observed are rare, at least in the genotyped population.[6] The severity of outcomes may be in part dependent on the severity of the mutations present in the *GALT* gene. The Q188R mutant allele, for instance, is particularly prevalent among patients of northern European descent, and in a study on ovarian function of females with CG conducted by Guerrero et al. (2000), Q188R contributed to greater risk of ovarian failure.[6,7] Similarly, Robertson et al. (2000) found homozygosity of Q188R was associated with increased likelihood of experiencing developmental verbal dyspraxia.[8] Yet, Kaufman et al. (1994) concluded that the presence of the Q188R allele was not solely responsible for variations observed in these outcomes.[9]

Method of Detection and Adverse Outcomes in Galactosemia

CG emerges clinically in newborns after exposure to dietary galactose from consumption of breastmilk or formula containing lactose.[2] The acute symptoms experienced by infants with CG during the neonatal period can include vomiting, diarrhea, poor weight gain, cataracts, jaundice, *E. coli* sepsis, lethargy, or even neonatal death.[2,10] These adverse outcomes can be minimized or even prevented by immediate removal of sources of high galactose from the infant's diet, which is frequently accomplished by replacing milk with lactose-free or soy-based formulas.[2]

In the United States and many other countries, diagnosis of CG typically occurs via newborn screening.[1] The first instance of newborn screening for CG in the United States

occurred in 1963, but it wasn't until 2004 that galactosemia was present on all newborn screening panels within the United States.[1] To screen for CG in newborns, typically a coupled assay is conducted to measure GALT enzyme activity in a dried blood spot.[2] In infants with CG, GALT activity is often absent or less than 1% of GALT enzyme activity detected in controls.[1,4] The levels of galactose metabolites, like galactose, galactose-1-phosphate, and galactitol, can also be measured from blood samples to screen for CG.[2] Given that other metabolic conditions that alter galactose metabolism can also contribute to elevated levels of galactose metabolites in the blood, the most reliable method for diagnosing CG remains measuring GALT enzyme activity.[11]

Early detection of the disease in newborns and immediate restriction of dietary galactose following diagnosis can avert the acute and often-times lethal symptoms experienced by patients with CG during the neonatal period.[2] Even with dietary restriction of high galactose foods, however, individuals can still encounter CG-associated complications long-term.[12] This is perhaps due to endogenous galactose synthesis, which can contribute to continued elevated levels of galactose metabolites.[13] Endogenous galactose synthesis is highest in children, producing 0.53-1.05 mg/kg of galactose per hour, and results in concentrations of the sugar that exceed concentrations of galactose consumed by individuals with CG actively restricting high galactose foods from their diet.[14]

The adverse outcomes experienced by individuals with CG throughout their lifetime are extensive and include cognitive deficits (45%), impaired motor function (estimated percent affected ranges from 18% to 52%), speech difficulties (56%), cataracts (30%), and primary ovarian insufficiency in girls and women (at least 80%).[9,12,15] Moreover, the plethora of long-

term outcomes experienced by patients with CG reveal incomplete penetrance and variable expressivity.[9,12] This means that not all patients experience all of the adverse outcomes and among those who do experience a given outcome, severity can differ.

Existing Literature Concerning Whether Outcomes in Galactosemia Are Progressive

In the existing literature concerning CG, there remains no clear consensus as to whether the long-term outcomes associated with CG are progressive. According to Merriam-Webster Dictionary, progressive is an adjective that can be used to describe when a medical condition is increasing in extent or severity, with other definitions including that such changes develop gradually.[16] There are many genetic disorders that are known to be progressive, such as Duchenne muscular dystrophy, which is characterized by the degeneration of muscle function over time.[17] The progressive nature of this disorder can affect the function of many muscles at different time points.[17] For instance, a young child with Duchenne muscular dystrophy may be able to walk. Yet, once the disorder affects their leg muscles, they may lose that ability and require a wheelchair. Still, an abundance of genetic disorders are not progressive at all.

In 1995, Kaufman et al. reported the results of a cross-sectional study with 45 patients with CG and found that older participants had lower IQ scores than younger participants in the study.[18] This difference in IQ scores was also observed in other cross-sectional studies during this time period.[9] Kaufman et al. did acknowledge, however, that the accessibility of special educational services changed dramatically in the years leading up to the study, so it is likely that older and younger participants enrolled in Kaufman et al.'s (1995) and other studies during this time period had different educational opportunities as children.[18] Conversely, a more recent

cross-sectional study published in 2012 by Waisbren et al. found no significant difference between the cognitive function, specifically IQ scores, of younger and older participants with CG.[15] Longitudinal approaches to studying cognitive function of patients with CG in Europe by Schadewaldt et al. (2010) also found no differences in IQ scores by age.[19] However, these longitudinal studies were limited by small cohort sizes.[19]

With respect to speech outcomes in CG, most studies concluded that speech did not decline over time and in some studies, speech actually improved.[9] It is important to note that children with CG have an increased likelihood of having a speech disorder compared to other children.[9,12,15] Thus, instances where speech improved over time could be due to speech interventions. Similarly, a case study focusing on language outcomes in CG conducted by Lewis et al. (2013) utilized a longitudinal approach to track the language skills of a toddler patient as they developed over the course of 7 years and concluded that the language skills of the child did not get worse over time.[20]

With regard to motor function, MacWilliams et al. (2021) collected and analyzed digital spiral drawings from 57 participants with CG and 80 controls.[21] The cross-sectional results of this study, with participants' ages ranging from 6 to 65 years, suggested that hand fine motor control does not decline with age.[21] As for ovarian function, Spencer et al. (2013) studied 158 girls and women with CG and cross-sectionally determined that anti-Mullerian hormone (AMH) was present in lower concentrations than usual across all ages included in the study, indicating that ovarian function does not worsen over time, but rather is always impaired in most girls and women with CG, even before that impairment may be clinically visible.[22]

Thus, the existing data on long-term complications of CG from cross-sectional and longitudinal studies suggest that the disorder is not progressive. However, after interviewing 12 adult patients with CG participating in the Applied Therapeutics AT-007 clinical trial and 8 caregivers, a recent study by Randall et al. (2022) concluded that these outcomes do get progressively worse over time.[23] It is unclear why the results of the Randall et al. (2022) study do not align with those of prior larger studies examining whether symptoms of CG do or do not change over time. Perhaps the small sample size included in the Randall et al. (2022) study contributed to this result, as the cohort studied may not have been representative of the total population of adults with CG.

Assessments For Long-Term Outcomes in Galactosemia

One method for quickly gathering data regarding long-term health outcomes in galactosemia is through distributing surveys for enrolled study participants to complete. Surveys are advantageous because they are standardized in that every participant is asked the same set of questions, can result in a large sample for a given study, and are able to be accessed with ease. Yet, there are also many drawbacks to only utilizing surveys for research purposes, notably when tracking health outcomes over a long period of time. Given that data collected from surveys are exclusively based on the respondents' perspectives, survey results can also be subject to various biases that impact the accuracy and reliability of the results.[24] For instance, when inquiring about long-term outcomes in galactosemia, a CG case may believe their health is better than their medical history would suggest.

Formal assessments that measure development and behavior can be used to investigate long-term health outcomes in galactosemia as well. The Vineland Adaptive Behavior Scales, for instance, is an instrument available through Pearson Assessments that tests domains of adaptive behaviors identified by the American Association on Intellectual and Developmental Disabilities (AAIDD), which includes communication, daily living skills, socialization, and motor skills.[25] Each domain is comprised of 2 to 3 subdomains that individually assess numerous items, which are each assigned a ranking corresponding to how often the participant chooses to perform the behavior independently.[26] The most recent edition, Vineland-3, has many applications, often being utilized in the diagnosis of intellectual and developmental disabilities.[26] However, this assessment is equally valuable for research purposes.

The communication domain assesses how well an individual reads, writes, listens, understands, and speaks.[25] Within the communication domain, there are three subdomains, including receptive, expressive, and written.[25] The next domain assessed, daily living skills, is comprised of the personal, domestic, and community subdomains, which together determine how well the individual can complete everyday tasks that are appropriate for their age.[25] The socialization domain measures how the individual functions in social situations and includes the interpersonal relationships, play and leisure, as well as coping skills subdomains.[25] Lastly, the motor skills domain assesses both gross motor and fine motor subdomains.[25]

Such assessment instruments, including the Vineland-3, are generally completed by parents, caregivers, or teachers, so they are valuable in that the scores are not a reflection of the participant's own perspective, and can be adjusted as either raw scores or normed scores.[26] Normed scores are adjusted according to the corresponding scores of controls matched by age

and sex who are believed to be developing normally.[25] These tests are also beneficial because they often evaluate a plethora of skills within one assessment to provide an abundance of information about participants and can be completed repeatedly over months or years to track progress over a given period of time.[25]

When respondents are completing the Vineland-3 assessment, they are able to indicate if scores assigned to specific items within each subdomain were estimates.[27] The assessment quantifies this for each subdomain raw score by providing a value for percent of estimated items, which represents the number of estimated items within a subdomain divided by the total number of items answered within a subdomain.[27] Thus, the Vineland-3 does assess the extent to which scores were based on estimates to maintain accuracy of the results. Just as with survey responses, however, scores obtained from these developmental and behavioral assessments are still largely subjective measures of outcomes.

In light of this, another method for gathering data on long-term outcomes in galactosemia is by collecting the medical records of study participants. Medical records, such as clinical values, growth charts, as well as scholastic records, provide a direct glimpse into the medical history of study participants and are especially useful for longitudinal analyses of health outcomes over time. There are, however, some difficulties with obtaining medical records, as they require the study participant or healthcare providers to authorize the release of medical records regardless of whether protected health information is included and actually receiving these records could be subject to delays. Nevertheless, medical records represent a critical source of data for studying long-term health outcomes, including those experienced by patients with CG.

Longitudinal Study of Outcomes in Classic Galactosemia

The ongoing longitudinal study of outcomes in CG based in REDCap utilizes a combination of the methods described above to obtain information regarding the long-term health outcomes of study participants. Both patients with CG and controls, which are often unaffected siblings, have been enrolled in this observational study. All participants were invited in 2022 to complete a series of surveys that inquired about infant health, diet and dietary restrictions, speech and language outcomes, cognitive development and educational experiences, motor and neurological function, growth and bone health, as well as puberty and reproductive health. Survey updates have also been sent out to participants within 6 months to 1 year of completion of the initial surveys in order to collect longitudinal data.

Additionally, participants enrolled in the observational study have been invited to complete the third edition of the Vineland Adaptive Behavior Scales and any results from additional assessments are also being collected. Participants will be sent invitations to conduct Vineland-3 assessment updates as well to monitor outcomes longitudinally. Finally, study participants are able to provide medical records, including clinical values for galactose metabolites, hormone values, growth charts, DEXA scans for bone density, and *GALT* sequencing for the identification of genotypes, upon authorization of the release of medical records by participants.

To address the controversy in the existing literature regarding whether or not outcomes in CG are progressive, data from the completion of the initial round of surveys and the first completion of the Vineland-3 within the longitudinal study will be analyzed using both case-control and cross-sectional methods. In particular, study participants who completed every

initial survey they received and were eligible to complete were included in this investigation. It is important to note that the longitudinal study currently has over 500 participants enrolled, so this sample is only a subset of the larger sample that constitutes the ongoing observational study.

While medical records are a valuable source of longitudinal data regarding patient health outcomes, the study remains in the process of collecting medical records for participants who are willing to do so. When participants had medical records available, they were included in the scoring of health outcomes experienced. Patient medical records will continue to be collected, as well as survey response updates. Beyond the scope of this thesis, such data will be scrutinized using longitudinal analyses to further approach this research question and in time produce a manuscript for publication.

CHAPTER 2: CROSS-SECTIONAL ANALYSIS OF THE SEVERITY OF HEALTH OUTCOMES EXPERIENCED BY INDIVIDUALS WITH CLASSIC GALACTOSEMIA OVER TIME

Introduction

It is evident that the existing literature concerning outcomes in CG has not reached a consensus as to whether outcomes are progressive. In analyzing the survey responses of participants enrolled in the longitudinal study, this study aimed to determine how participants perceived their health outcomes changed over time. Another objective of this investigation was to generate summary scores for the severity of each health outcome experienced. These summary scores would not only allow for the determination of whether the scores assigned for a participant's given health outcome aligned with their own perceptions but also enable cross-sectional analyses for the severity of these outcomes to address the broader research question.

Methods

All methods involving human subjects were approved under Emory Institutional Review Board (IRB) protocol 00024933.

Recruitment of Study Participants

Participants in the ongoing longitudinal study of outcomes in CG within REDCap were recruited by targeting patients with CG and family members connected through the Galactosemia Foundation's social media platforms. In addition, patients with CG and family members who attended the Galactosemia Foundation 2022 Conference were recruited to enroll in the study. Finally, a portion of study participants were recruited by contacting metabolic specialists internationally.

Upon enrolling in the study, all volunteers had to complete an informed consent form and assent if volunteers were less than 18 years old. In addition, study participants signed addendum forms for the completion of any assessments within the study. Medical records for study participants were collected after receiving authorization from study participants for the disclosure of their health information or following the completion of a questionnaire by healthcare providers, which enabled them to provide non-protected health information for participants' medical records that pertained to the study.

Study Participants

All volunteer participants enrolled in the ongoing longitudinal study that had completed every initial survey by November 2022 were included in this investigation, which represents a subset of the total participants currently enrolled in the observational study. The remaining participants will be included in the coming months as initial surveys continue to be completed, and together with this investigation, the results will be used to produce a manuscript for publication. Table 1 below summarizes the demographic characteristics of the subset of study participants whose data are described in the chapter below.

Table 1. Demographic Characteristics of Study Participants

		CG Cases [*n*(%)]	**Controls [*n*(%)]**	**Total [*N*(%)]**
Age (years)	0 to 9	23 (28.05)	18 (36.00)	41 (31.06)
	10 to 19	28 (34.14)	22 (44.00)	50 (37.88)
	20 to 29	11 (13.41)	3 (6.00)	14 (10.61)
	30 to 39	9 (10.98)	3 (6.00)	12 (9.09)
	40 to 79	11 (13.41)	4 (8.00)	15 (11.36)
Sex	Male	41 (50.00)	20 (40.00)	61 (46.21)
	Female	40 (48.78)	30 (60.00)	70 (53.03)
	Other	1 (1.22)	0 (0.00)	1 (0.76)
Total		82	50	132

Assigning Summary Scores to Health Outcomes

A participant summary data instrument was created using REDCap software to summarize responses from each survey for individual participants and assign summary scores to the different health outcomes the surveys focused on. Separate surveys within the longitudinal study were devoted to questions asking about infant health, diet and dietary restrictions, speech and language outcomes, cognitive development and educational experiences, motor and neurological function, growth and bone health, as well as puberty and reproductive health of study participants.

Summary scores were created to rank the duration the health outcome was experienced or to rank the severity of the health outcome experienced, which was the case for most of the summary scores. Scores with multiple levels were established to summarize the severity of neonatal outcomes, the duration of adverse speech outcomes, the severity of adverse cognitive outcomes and motor/neurological outcomes, as well as the level of ovarian function within female participants. These summary scores were assigned primarily based off of survey responses, however if medical records were available for that participant, such records were also considered when scoring the health outcomes.

As mentioned above, only a subset of participants in the longitudinal study were included in this investigation, so study participants that completed every initial survey by November 2022 were prioritized when completing summary scores to ensure their scores could be included in this investigation. As more participants complete surveys, more participant summary data instruments with summary scores will be completed and they will be compiled in

future publications of the longitudinal study. In addition, as more medical records become available, summary scores will be updated to best reflect the medical history of participants.

Statistical Analyses

All statistical analyses were performed in R version 4.2.2 via RStudio. Figures were created to visualize how participants perceived their outcomes had changed over time. This data was available from surveys inquiring about adverse speech outcomes, cognitive difficulties, and motor/neurological outcomes. Moreover, figures were created to showcase the proportion of child participants that had a certain health outcome and the proportion of adult participants that had also reported experiencing that health outcome by CG status. Child participants are any participants less than 18 years of age and adult participants are any participants between 18 and 79 years of age. This comparison was done for adverse speech outcomes, cognitive difficulties, and motor/neurological outcomes.

Pearson's chi-squared (X^2) tests of independence with a significance level of $\alpha = 0.05$ were conducted to compare the proportions by assessing the cell count frequencies in a contingency table. All expected cell counts were greater than 5, so a Fisher's exact test of independence was not used. The chi-squared test of independence between two categorical variables compares the observed frequencies with the expected frequencies based on the total values in a 2x2 contingency table. If two variables are considered independent, or not associated, the observed and expected frequencies will be similar. Significant p-values from all comparisons of proportions are summarized in Table 4.

To further assess how these health outcomes are changing over time among CG cases, for all health outcomes that had summary scores created (neonatal outcomes, dietary

restrictions, adverse speech outcomes, cognitive difficulties, motor/neurological outcomes, and ovarian function), a boxplot was created for each outcome to visualize the distribution of the ordinal summary scores among CG cases and controls by age. Ages were grouped by decade and the final age group encompassed participants from 40 to 79 years of age to account for the small sample size of older adults enrolled in the Fridovich-Keil Lab's longitudinal study that had completed all surveys at the time data was collected for this investigation.

When assigning summary scores within the participant summary data instrument, the summary scores were ordinal categories with descriptive labels, such as moderate or severe. For the purposes of statistically analyzing the summary scores, these descriptive labels were transformed into categories with numeric labels in R. The higher the number, the longer the health outcome was experienced or the more severe the health outcome was. Table 2 summarizes the original summary score labels for each health outcome being assessed and the corresponding numeric label that was created when the data was transformed.

Table 2. Descriptive and Numeric Labels for Summary Scores by Health Outcome

Health Outcome	**Descriptive Labels**	**Numeric Labels**
Neonatal Outcomes	No symptoms	0
	Mild symptoms	1
	Moderate symptoms	2
	Severe symptoms	3
Dietary Restrictions	No restrictions	0
	Restricted only dairy milk	1
	Restricted dairy milk and legumes	2
	Restricted dairy milk, legumes, and more	3
Adverse Speech Outcomes	No problems	0
	Problems in childhood	1
	Problems persisted into adulthood	2
Adverse Cognitive Outcomes	No problems	0
	Isolated or mild problems	1
	Moderate to severe problems	2
Adverse Motor/Neurological Outcomes	No problems	0
	Isolated or mild problems	1
	Moderate to severe problems	2
Ovarian Function	Apparently normal	0
	At least some ovarian function	1
	Clear primary ovarian insufficiency	2

A Shapiro-Wilk normality test was conducted for each outcome with a significance level of $\alpha = 0.05$ to determine if the distributions were normal (Table 3). All p-values were less than $\alpha = 0.05$, indicating that the distributions were non-normal. Given this, a non-parametric Wilcoxon rank sum test was utilized to compare scores for CG cases and controls within each age group. Any significant p-values obtained from these Wilcoxon rank sum tests comparing summary scores between CG cases and controls by age group are summarized in Table 5.

Table 3. Shapiro-Wilk Normality Test P-values for Outcome Summary Scores

Health Outcome	**P-value**
Neonatal Outcomes	1.122e-10
Dietary Restrictions	5.683e-12
Adverse Speech Outcomes	8.784e-15
Adverse Cognitive Outcomes	7.898e-15
Adverse Motor/Neurological Outcomes	<2.2e-16
Ovarian Function	4.271e-07

*P-values are considered significant if p<0.05, which provides strong evidence the data are not normally distributed.

Results

Neonatal Outcomes

A comparison between scores representing the severity of neonatal outcomes experienced by CG cases and controls found that there was a statistically significant difference between the ages of 0 to 9, 10 to 19, and 40 to 79 (Figure 2).

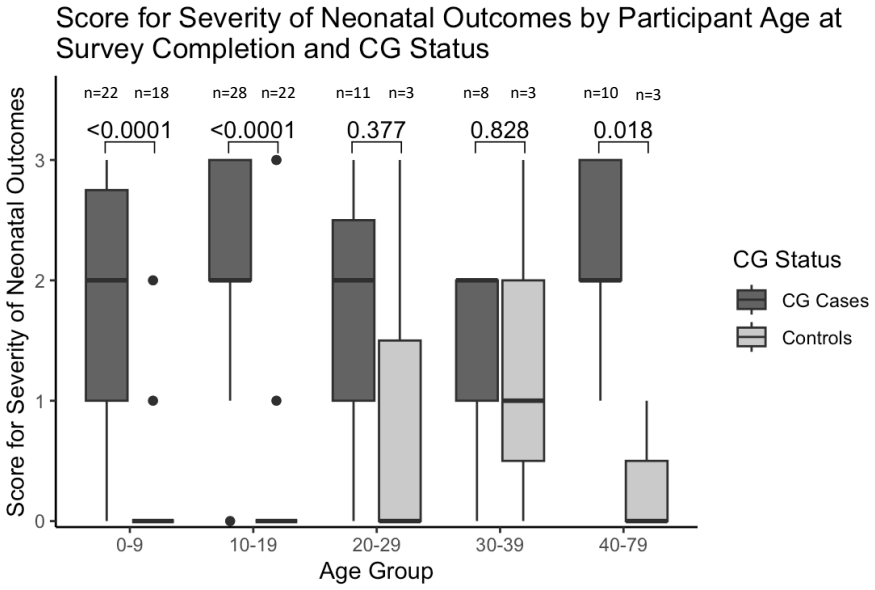


Figure 2. Score for Severity of Neonatal Outcomes by Participant Age at Survey Completion and CG Status. A score of 0 indicates no adverse symptoms during the neonatal period were reported, a score of 1 represents those who experienced mild symptoms, a score of 2 corresponds to participants whose neonatal symptoms were moderate, and a score of 3 was assigned to those who experienced severe symptoms during the neonatal period. P-values are significant if p<0.05.

Dietary Restrictions

When comparing the extent to which participants restricted their diets until at least age 5, there was a statistically significant difference between CG cases and controls in all age groups tested (Figure 3). When CG cases were asked about how their diet changed over time, 46% reported that their diet had liberalized over time (Figure 4). 40% of those reported that they had liberalized their diet over time but continued to restrict high galactose dairy products, while the remaining 6% liberalized their diet over time to include some high galactose dairy products.

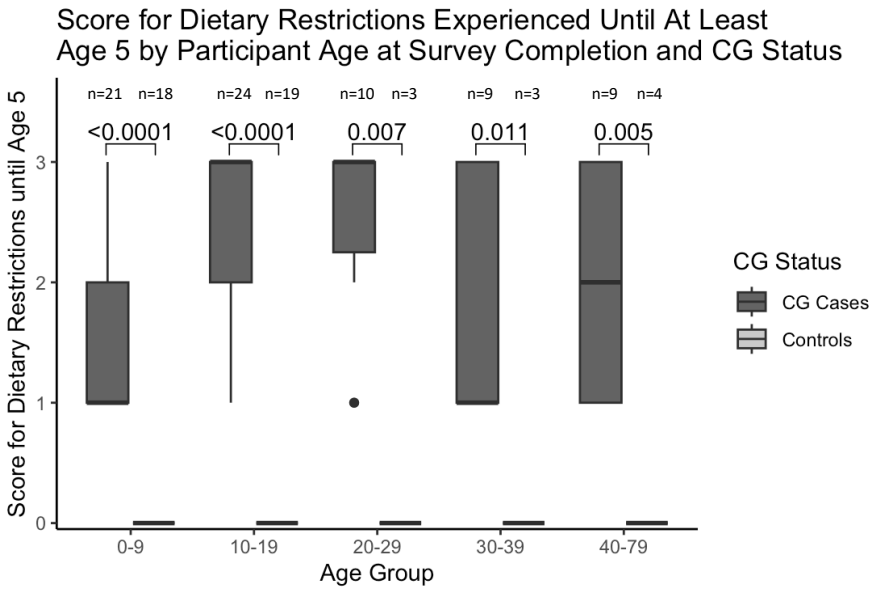


Figure 3. Score for Dietary Restrictions Experienced Until At Least Age 5 by Participant Age at Survey Completion and CG Status. A score of 0 indicates no galactose restriction to the present age, and a score of 1 represents those who restricted only dairy milk/high galactose dairy products until at least age 5. A score of 2 corresponds to participants who restricted dairy milk/high galactose dairy products and legumes until at least age 5. Finally, a score of 3 was assigned to those who restricted dairy milk/high galactose dairy products, legumes and some other foods until at least age 5. P-values are significant if $p<0.05$.

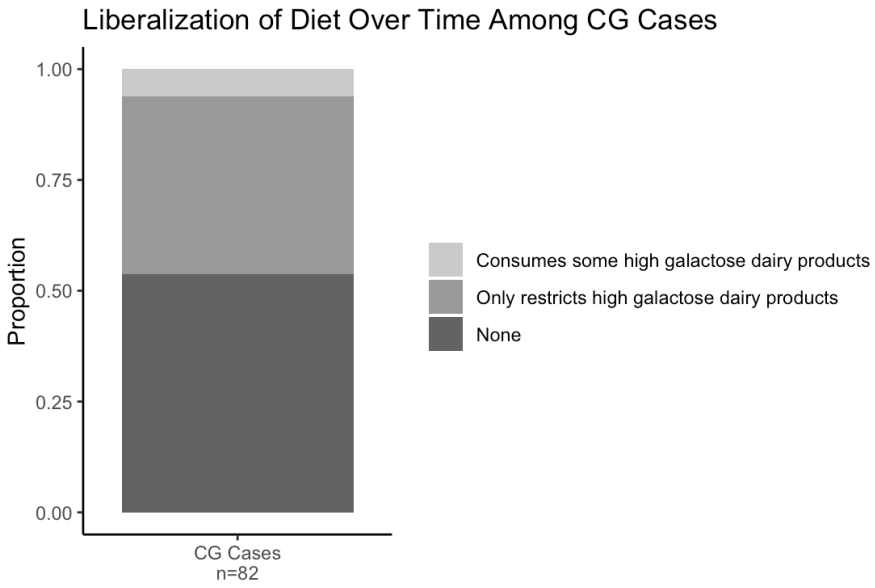


Figure 4. Liberalization of Diet Over Time Among CG Cases. 46% of CG Cases reported that their diet had liberalized over time. No controls indicated dietary restrictions to limit galactose exposure, so only CG cases were included in this figure.

Adverse Speech Outcomes

Out of the survey respondents that had reported experiencing adverse speech outcomes at some point in their life, 72% of CG cases and 80% of controls reported that these problems had improved over time (Figure 5). There was a significant difference between the proportion of child participants who had adverse speech outcomes when comparing CG cases and controls (Figure 6). However, the difference in the proportion of adult CG cases and controls with adverse speech outcomes was not significant (Figure 7). A comparison between scores for the severity of adverse speech outcomes experienced by CG cases and controls found a significant difference in scores by CG status for participants 0 to 9 years of age (Figure 8). A visual analysis of the summary scores by age group in Figure 8 highlights that the scores for CG cases are decreasing as age increases, suggesting the severity of outcomes decreases with time.

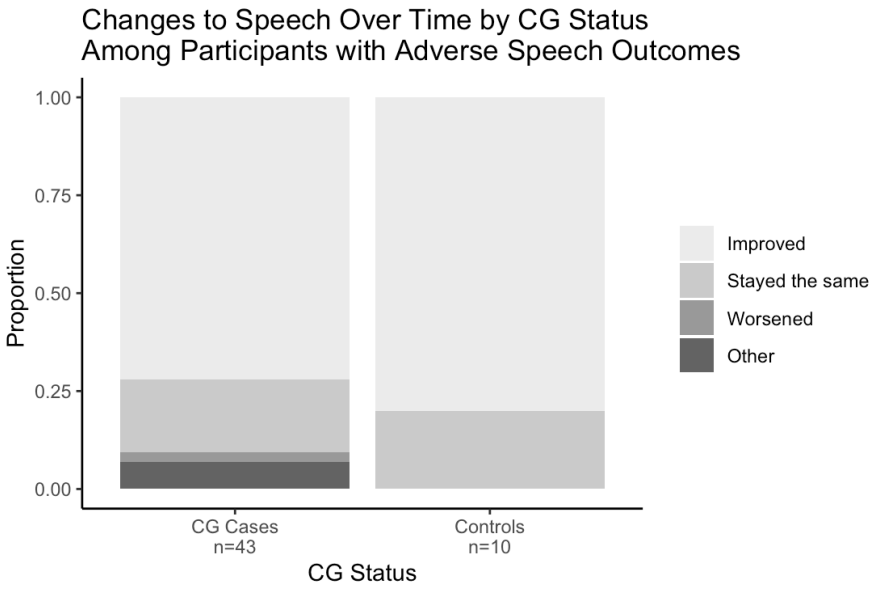


Figure 5. Changes to Speech Over Time by CG Status Among Participants Reporting Adverse Speech Outcomes. This characterization is solely from perceptions of study participants in survey responses. 72% of CG cases and 80% of controls with adverse speech outcomes reported such problems had improved over time.

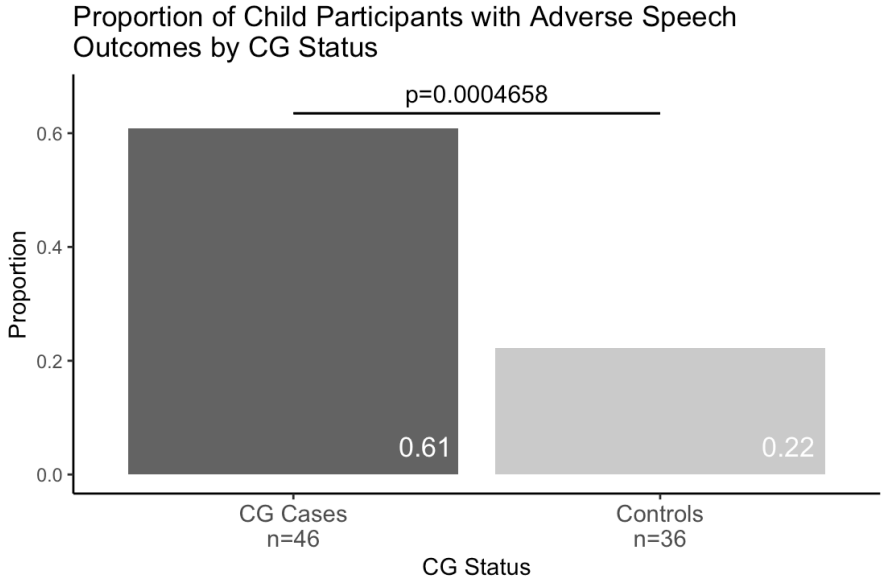


Figure 6. Proportion of Child Participants with Adverse Speech Outcomes by CG Status. According to a report from the Centers for Disease Control and Prevention (CDC) in 2012, the national percentage of children with any communication disorder was 7.7%.[28] P-values are significant if $p<0.05$.

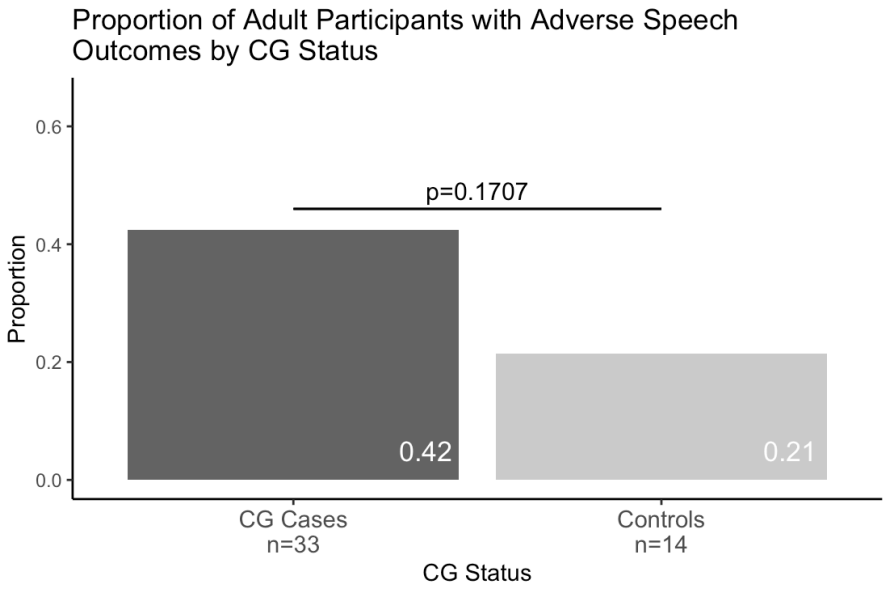


Figure 7. Proportion of Adult Participants with Adverse Speech Outcomes by CG Status. This proportion includes both those who reported speech problems in childhood and those whose problems persisted into adulthood. This p-value is not significant because p>0.05.

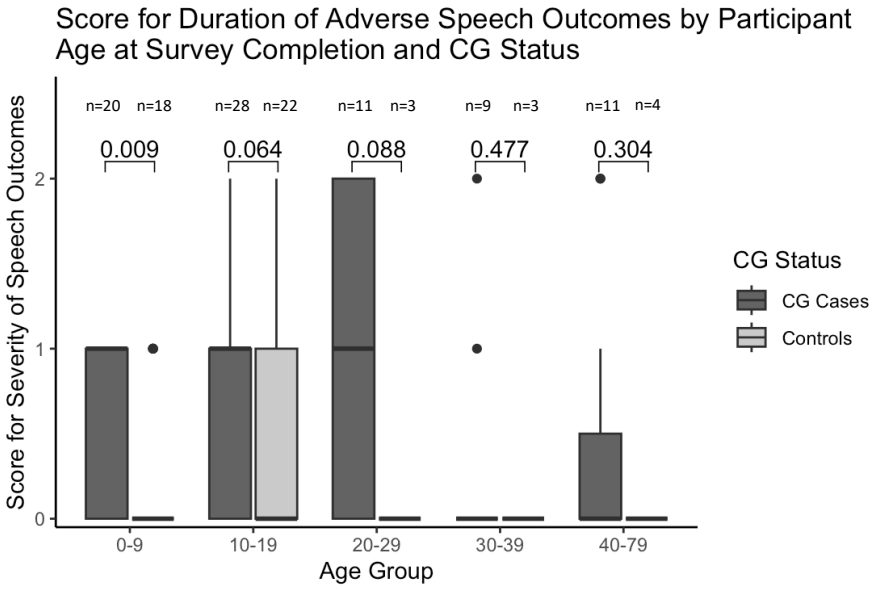


Figure 8. Score for Duration of Adverse Speech Outcomes by Participant Age at Survey Completion and CG Status. A score of 0 indicates no speech problems were reported, a score of 1 represents those who experienced speech problems in childhood, and a score of 2 corresponds to participants whose speech problems that persisted into adulthood. P-values are significant if p<0.05.

Adverse Cognitive Outcomes

All participants whose cognitive outcomes were attributed solely to social, emotional, or behavioral issues, such as anxiety, depression, or ADHD diagnoses, were considered to have no adverse cognitive outcomes. Among participants experiencing adverse cognitive outcomes, 84% of CG cases and 80% of controls reported that their cognitive challenges had stayed the same or improved over time (Figure 9). Among pediatric CG cases and controls, there was a significant difference in the proportion of participants experiencing adverse cognitive outcomes (Figure 10). A significant difference was not obtained for a comparison between adult CG cases and controls, though the p-value was close to 0.05. (Figure 11). There was also a significant difference between the severity of adverse cognitive outcomes experienced by CG cases and controls in the 0 to 9 and 10 to 19 age groups (Figure 12).

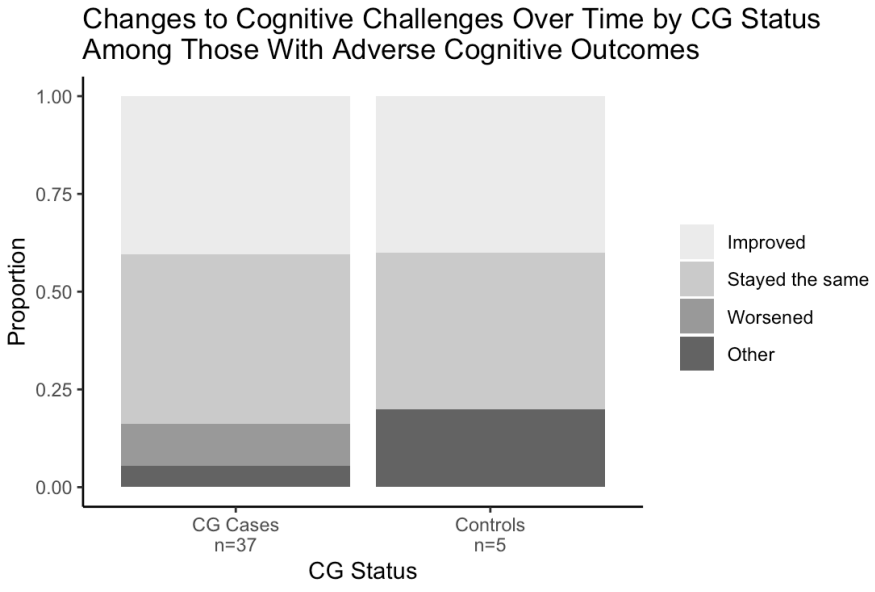


Figure 9. Changes to Cognitive Challenges Over Time by CG Status Among Participants Reporting Adverse Cognitive Outcomes. This characterization is solely from perceptions of study participants in survey responses. 84% of CG cases and 80% of controls with adverse cognitive outcomes reported such problems had either stayed the same or improved over time.

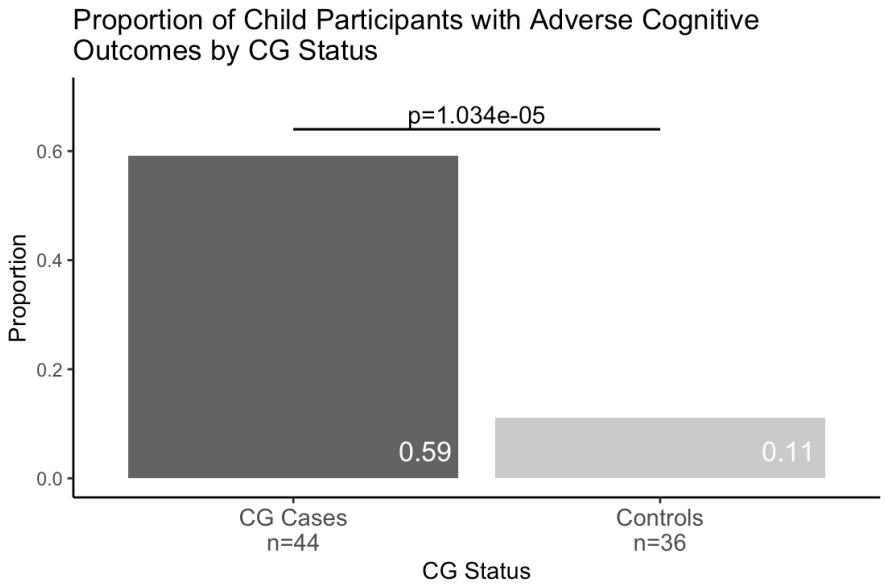


Figure 10. Proportion of Child Participants with Adverse Cognitive Outcomes by CG Status. This proportion includes all child participants that were scored with isolated, mild, moderate, or severe cognitive difficulties and responded to this question within the Cognitive and Educational Experiences Survey. P-values are significant if $p<0.05$.

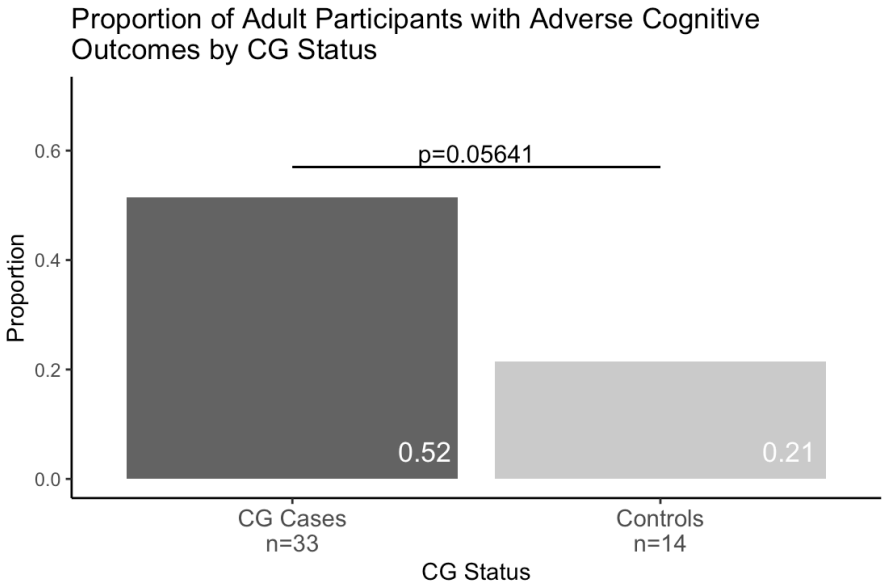


Figure 11. Proportion of Adult Participants with Adverse Cognitive Outcomes by CG Status. This proportion includes all adult participants that were scored with isolated, mild, moderate, or severe cognitive difficulties and responded to this question within the Cognitive and Educational Experiences Survey. This p-value is not significant because $p>0.05$.

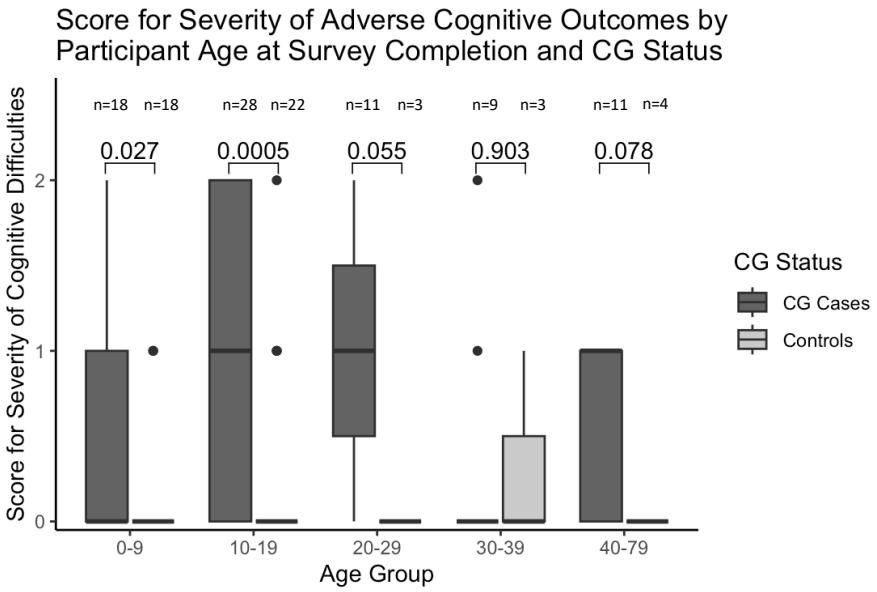


Figure 12. Score for Severity of Adverse Cognitive Outcomes by Participant Age at Survey Completion and CG Status. A score of 0 indicates no cognitive problems were reported, a score of 1 represents those who experienced isolated or mild cognitive problems, and a score of 2 corresponds to participants whose cognitive problems were moderate to severe. P-values are significant if $p<0.05$.

Adverse Motor and Neurological Outcomes

50% of CG cases and 50% of controls who had reported experiencing adverse motor or neurological outcomes indicated that no issues had changed over time or at least some had changed over time (Figure 13). When comparing the proportion of child participants with CG who experienced adverse motor or neurological outcomes with controls in the same age range, CG cases had a significantly higher proportion (Figure 14). On the other hand, the difference in the proportion of adult participants with adverse motor or neurological outcomes by CG status was not significant (Figure 15). An analysis of the differences between CG cases and controls in the severity of the motor or neurological outcomes experienced found significant differences only among participants within 0 to 9 and 10 to 19 years of age (Figure 16).

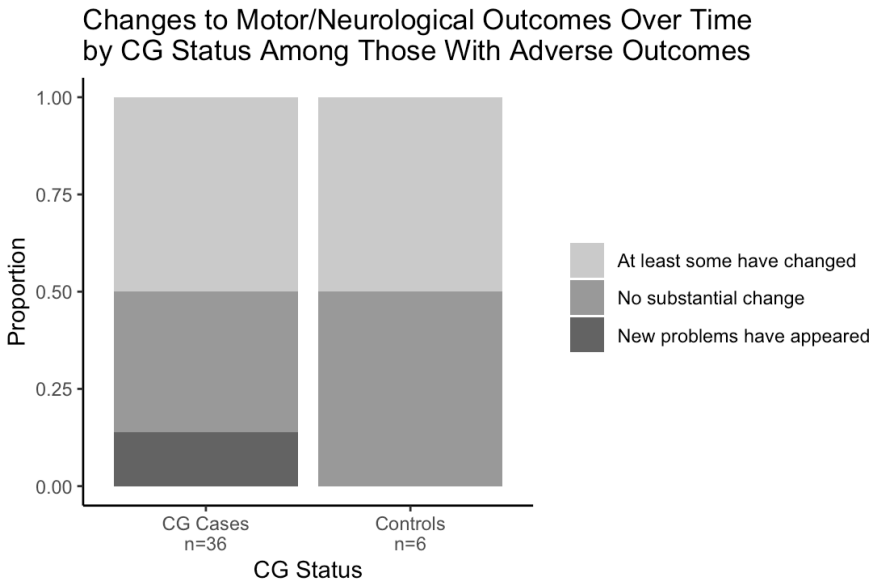


Figure 13. Changes to Motor/Neurological Outcomes Over Time by CG Status Among Participants Reporting Adverse Outcomes. This characterization is solely from perceptions of study participants in survey responses. 50% of CG cases and 50% of controls with adverse motor and neurological outcomes reported such problems had not substantially changed or at least some had changed over time.

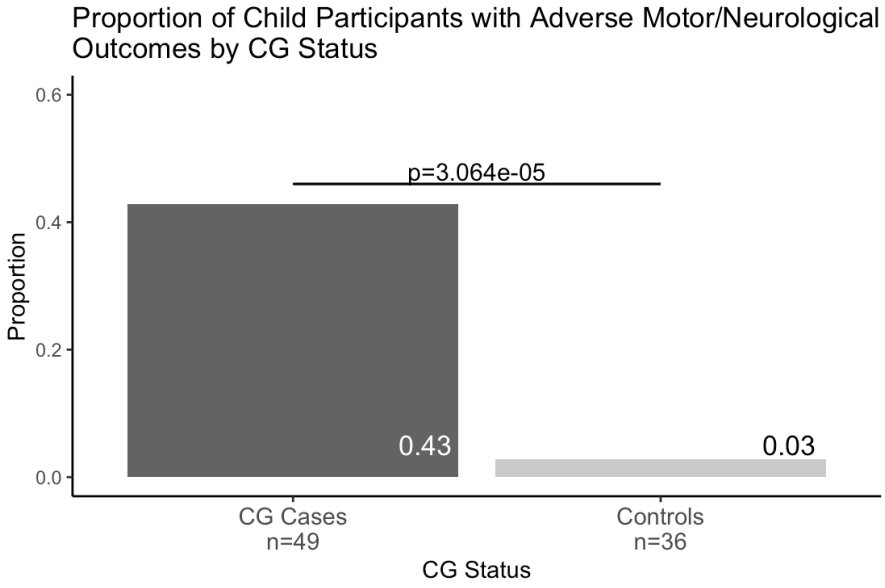


Figure 14. Proportion of Child Participants with Adverse Motor/Neurological Outcomes by CG Status. This proportion includes all child participants that were scored with isolated, mild, moderate, or severe motor/neurological outcomes. P-values are significant if $p<0.05$.

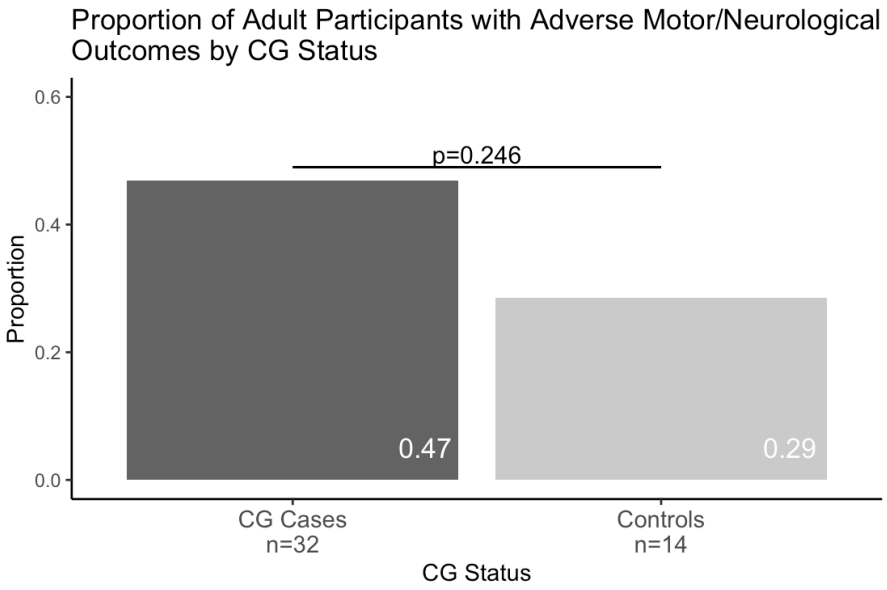


Figure 15. Proportion of Adult Participants with Adverse Motor/Neurological Outcomes by CG Status. This proportion includes all adult participants that were scored with isolated, mild, moderate, or severe motor/neurological outcomes. This p-value is not significant because p>0.05.

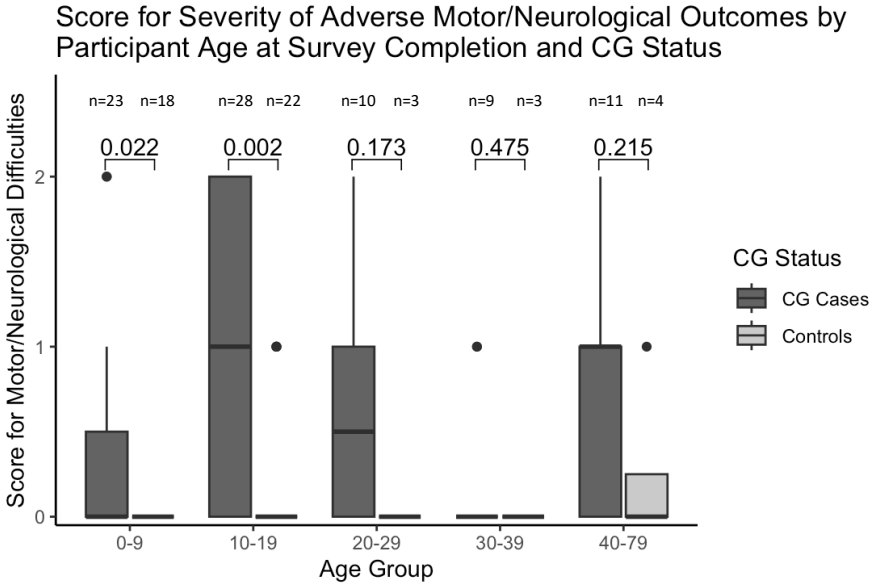


Figure 16. Score for Severity of Adverse Motor/Neurological Outcomes by Participant Age at Survey Completion and CG Status. A score of 0 indicates no motor/neurological problems were reported, a score of 1 represents those who experienced isolated or mild motor/neurological problems, and a score of 2 corresponds to participants whose motor/neurological problems were moderate to severe. P-values are significant if p<0.05.

Ovarian Function

A comparison of ovarian function scores of females with CG and controls found a significant difference in ovarian function for those between 10 to 19 years of age (Figure 17).

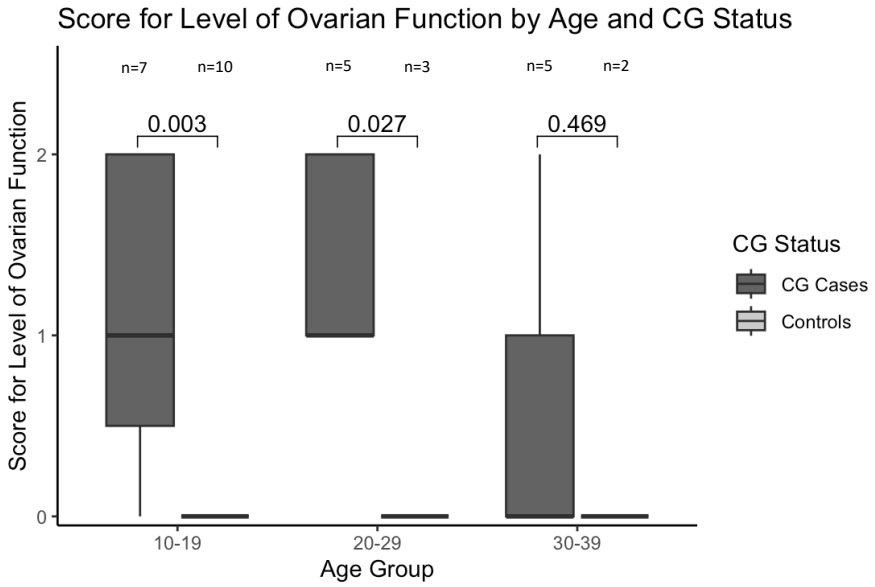


Figure 17. Score for Level of Ovarian Function by Female Participant Age at Survey Completion and CG Status. A score of 0 indicates the participant had apparently normal ovarian function and a score of 1 represents those with at least some ovarian function. A score of 2 corresponds to participants who had clear primary ovarian insufficiency. Participants whose ovarian function could not be assessed, such as participants who had not yet begun puberty, were excluded from this figure. In addition, participants older than 40 were excluded, given that impaired ovarian function any time after this age is considered a normal component of menopause.[29] P-values are significant if $p<0.05$.

Table 4 below summarizes the results of chi-squared (X^2) tests of independence from comparing cell count frequencies to determine any significant differences in prevalence of health outcomes among CG cases and controls. In addition, Table 5 summarizes significant p-values comparing the distribution of scores for health outcomes by age among CG cases and controls.

Table 4. Pearson's Chi-Squared Test P-values for Outcome Summary Scores by Age Group

Health Outcome	Age at Time of Survey	P-value
Adverse Speech Outcomes	Child participants	0.0004658
	Adult participants	0.1707
Adverse Cognitive Outcomes	Child participants	1.034e-05
	Adult participants	0.05641
Adverse Motor/Neurological Outcomes	Child participants	3.064e-05
	Adult participants	0.246

*P-values are considered significant if p<0.05.

Table 5. Significant P-values and Age Group for Summary Scores Using Wilcoxon Test

Health Outcome	Age Group	P-value
Neonatal Outcomes	0 to 9	<0.0001
	10 to 19	<0.0001
	40 to 79	0.018
Dietary Restrictions	0 to 9	<0.0001
	10 to 19	<0.0001
	20 to 29	0.007
	30 to 39	0.011
	40 to 79	0.005
Adverse Speech Outcomes	0 to 9	0.009
Adverse Cognitive Outcomes	0 to 9	0.027
	10 to 19	0.0055
Adverse Motor/Neurological Outcomes	0 to 9	0.022
	10 to 19	0.002
Ovarian Function	10 to 19	0.003
	20 to 29	0.027

*P-values are considered significant if p<0.05.

Discussion

The survey responses from participants enrolled in the observational study provided insight into how adverse health outcomes experienced by individuals with CG are changing over time. According to the participants' own characterizations of their adverse speech, cognitive, and motor/neurological outcomes, most CG case and controls indicated that they were experiencing no changes or some improvements over time. Although these are only from a

participants' perspective, such responses do provide a glimpse at how these health outcomes are experienced over time.

The prevalence of adverse outcomes was also analyzed cross-sectionally by CG status for child and adult participants. Child participants with CG had a significantly higher prevalence of adverse speech, cognitive, and motor/neurological outcomes compared to controls. It is important to note, however, that the proportion of pediatric controls reporting adverse speech outcomes was also above the national percentage of children with communication disorders according to the CDC.[28] Among adult participants, the differences in prevalence of any of these adverse health outcomes by CG status was not significant. When comparing the proportion of adult participants that experienced adverse cognitive outcomes, the p-value was 0.05641, which is close to the significance level of $\alpha = 0.05$, so perhaps the difference would be significant with a larger sample size.

In addition, when comparing the differences in prevalence of these adverse health outcomes between child and adult participants with CG, the prevalence either decreases or remains relatively stagnant between these two groups. This suggests that CG cases as children experience more challenges than their peers, yet these health outcomes are not perceived to worsen over time, as the difference was no longer significant among adult participants. For both adverse cognitive and motor/neurological outcomes, it appears that the prevalence of controls experiencing these outcomes increases between child and adult participants. Given that declining cognitive or motor/neurological function is also experienced as individuals age, this increase among adult controls could perhaps be due to the wide range of ages included in the adult participant group.

For the comparison of the severity of summary scores by CG status and age for these three adverse health outcomes, there were statistically significant differences within age group comparisons. Such significant differences were mostly found when comparing CG cases and controls of younger age groups, such as 0 to 9 and 10 to 19 years old. As age increased, the general trend was that the severity of outcomes was decreasing as well. These findings indicate that for adverse speech, cognitive, and motor/neurological outcomes, children with CG do experience adverse outcomes. Yet, the trends over time suggest that their adverse outcomes either remain stagnant or improve with time, perhaps due to responsive therapy.

Despite these general trends, specifically when scoring the severity of adverse cognitive and motor/neurological outcomes, it appears that median score actually increased among the final age group, 40 to 79 years. This increase in the severity of adverse outcomes among the oldest age group studied could be attributed to the small sample size of adult participants included in the study, given that the adult participants included in this investigation represent only a subset of the total adult participants enrolled in the ongoing observational study.

Moreover, this apparent increase in the severity of could be due to differences in the level of special educational services available to CG cases during their childhood. As mentioned with regard to Kaufman et al.'s study published in 1995, worse outcomes may be observed among adult participants simply due to the fact that special educational services were not widely accessible to them during their educational years.[18] It is also important to recognize that even with more severe scores among the oldest age group, the median score was never more extreme than any median scores obtained for the previous age groups.

The survey responses also provided insight into additional experiences of CG cases, including the severity of neonatal outcomes, dietary restrictions experienced, as well as ovarian function in females. As for neonatal outcomes, there were statistically significant differences in the mean severity of neonatal outcomes by CG status for all age groups except for 20 to 29 and 30 to 39 years of age. Significant differences by CG status are not unusual, as CG cases prior to diagnosis can experience a range of acute symptoms.[2,9] But, it is also not surprising that some age groups had no significant differences, as many symptoms experienced by CG cases during the neonatal period, such as jaundice, are quite common among newborns.[30]

Given that no controls indicated any restrictions of their diet and all CG cases had dietary restrictions until at least 5 years of age, all age groups compared had significant differences. Yet, 46% of CG cases indicated that they were able to liberalize their diet over time, indicating that for many their health outcomes were favorable enough to allow for liberalization of diet over time. Furthermore, for all age groups except for one, ovarian function was significantly impaired compared to controls. Although there was no significant difference for the 30 to 39 age group, the consistently high scores for impaired ovarian function across all age groups indicates that ovarian function is not slowly deteriorating with age. Rather, once there is diminished ovarian function it remains impaired.

CHAPTER 3: CROSS-SECTIONAL ANALYSIS OF ADAPTIVE BEHAVIORS OF INDIVIDUALS WITH CLASSIC GALACTOSEMIA MEASURED USING THE VINELAND-3

Introduction

While our custom survey responses do provide valuable information addressing the experiences of participants with various health outcomes, the scoring of these health outcomes was primarily based off of the subjective survey responses. Given the abundance of participants within the study who had also been evaluated by the Vinland-3, the domain normed scores and subdomain raw scores were also analyzed to address the research question at hand. Although the Vineland-3 can also be characterized as a subjective measure that is based off of the respondents perceptions of the participant's behaviors, the assessment does quantify the percentage of scores that the respondent estimated. Incorporating the Vineland-3 data into this investigation will provide more insight into how the behaviors of CG cases are scored by different age groups.

Methods

All methods involving human subjects were approved under Emory IRB protocol 00024933. The study participants who were scored using the Vineland-3 assessment were enrolled in the observational study as described in Chapter 2.

Administration of Vineland-3 and Standardization of Scores

Every volunteer participant in the observational study was invited to designate someone to complete the third edition of the Vineland Adaptive Behavior Scales. The Vineland-3

assessment can be administered using an interview form, parent/caregiver form, or teacher form depending on who is to complete the survey, with comprehensive or domain-level forms available for each method.[25] All Vineland-3 assessments completed for the purpose of this investigation utilized the comprehensive parent/caregiver form, meaning the surveys were completed by parents, caregivers, or other close associates of study participants.

As described earlier, the Vineland-3 assesses four domains (communication, daily living skills, socialization, and motor skills), each consisting of 2 to 3 subdomains.[24] The subdomain scores are raw scores while the domain scores are standardized by comparing to a normed sample which closely reflects the U.S. population, resulting in domain distributions with a mean = 100 and standard deviation = 15.[25] Table 6 below summarizes the demographic information of the normed sample that is used for the parent/caregiver form of Vineland-3 to create domain standard scores.[25] The Adaptive Behavior Composite (ABC) score is determined based on the individual domain standard scores for communication, daily living skills, and socialization.[25]

Table 6. Demographic Characteristics of Total Norm Sample for Vineland-3 Parent/Caregiver Form Compared to U.S. Population

		Vineland-3 Control Group (U.S. Population)
Race/Ethnicity	White	53.8 (52.9)
	African American	13.6 (13.2)
	Asian	4.0 (4.7)
	Hispanic	23.2 (23.7)
Maternal Education Level	Less than High School	13.3 (14.0)
	High School Diploma	23.9 (23.3)
	Some College/Tech/Assoc	32.3 (31.7)
	Bachelor's Degree/more	30.5 (31.0)

Note. Adapted from *Vineland Adaptive Behavior Scales, Third Edition* by C. A. Saulnier, 2017.[26]

Study Participants

All volunteer participants enrolled in the observational study for whom a Vineland-3 assessment was completed were included in this investigation, with the exception of participants older than 45 years of age. These older study participants were excluded from this investigation because the sample size was very small (CG cases n = 5; controls n = 1). Thus, any conclusions regarding whether or not CG is progressive from cross-sectional analyses with a minute sample of older adults could be anti-factual. Table 7 summarizes the demographic characteristics of the study participants who were included in this study.

Table 7. Demographic Characteristics of Vineland-3 Study Participants

		CG Cases [*n*(%)]	**Controls [*n*(%)]**	**Total [*N*(%)]**
Age (years)	0 to 9	44 (43.14)	27 (40.30)	71 (42.01)
	10 to 19	36 (35.29)	33 (49.25)	69 (40.83)
	20 to 29	12 (11.76)	4 (5.97)	16 (9.47)
	30 to 45	10 (9.80)	3 (4.48)	13 (7.69)
Sex	Male	41 (40.20)	28 (41.79)	69 (40.83)
	Female	60 (58.82)	39 (58.21)	99 (58.58)
	Other	1 (0.98)	0 (0.00)	1 (0.59)
Total		102	67	169

Statistical Analyses

All statistical analyses were performed in R version 4.2.2 via RStudio. Simple linear regression was used to assess the relationship between domain/composite standard scores and age by CG status (Figures 18-25). Specifically, the r-squared and slope values were applied to investigate the association between these two variables, which are summarized in Table 10. Separate figures were created for CG cases and controls to highlight their respective domain/composite score distributions and the adaptive levels outlined in the Vineland-3 assessment for ranges of domain/composite scores were included as shaded regions.[31]

Figures were also created to summarize raw subdomain scores across age groups (Figures 29-39). A Shapiro-Wilk normality test was conducted with a significance level of α = 0.05 to assess the distribution of scores for each subdomain and all p-values were significant, indicating that the distribution of scores for each subdomain is non-normal (Table 8). As such, a non-parametric Wilcoxon rank sum test was utilized to compare subdomain scores of CG cases and controls within each age group using a significance level of α = 0.05. Significant p-values from these comparisons are summarized in Table 11. Ages at time of completion of the Vineland-3 were grouped by decade and given that the maximum age cut off was 45 years of age, the last age group was 30 to 45 years.

Table 8. Shapiro-Wilk Normality Test P-values for Subdomain Score Distributions

Subdomain	**P-value**
Receptive	<2.2e-16
Expressive	<2.2e-16
Written	2.247e-11
Personal	<2.2e-16
Domestic	2.069e-09
Community	2.043e-05
Interpersonal Relationships	2.74e-14
Play and Leisure	1.659e-14
Coping Skills	2.865e-10
Gross Motor	<2.2e-16
Fine Motor	<2.2e-16

*P-values are considered significant if p<0.05, which provides strong evidence the data are not normally distributed.

As discussed above when outlining the administration of Vineland-3, the assessment quantifies a percent estimating value for each subdomain raw score based on how many item scores the respondents characterized as estimates.[27] According to Pearson, which provides the

instrument, if a given subdomain has a percent estimating value of at least 25%, indicating the respondent estimated for at least 25% of item scores, the results should not be interpreted due to the large proportion of estimated scores.[27] Given this suggestion, subdomain raw scores for participants were excluded if the percent estimating value was at least 25%. The count of raw scores excluded for each subdomain is summarized in Table 9.

Table 9. Count of Raw Scores Excluded for Each Subdomain with Percent Estimating Values At Least 25 Percent

Subdomain	**Count of Raw Scores Excluded**
Receptive	2
Expressive	4
Written	8
Personal	1
Domestic	1
Community	6
Interpersonal Relationships	6
Play and Leisure	4
Coping Skills	10
Gross Motor	7
Fine Motor	4

Results

Domain and Composite Standard Scores

Simple linear regression comparing the domain/composite normed scores for CG cases and controls by age indicated that as participants' ages increased, the domain/composite scores decreased marginally among CG cases (Figures 18, 20, 22, 24) and increased marginally among controls (Figures 19, 21, 23, 25), evidenced by negative and positive slopes, respectively. This suggests that with age, the gap between CG cases and their peers in both the large

comparison sample and unaffected siblings control sample widens with regard to the adaptive behaviors assessed, with CG cases to some extent lagging behind their peers.

However, all of the slopes were close to 0, ranging from -0.37 to 0.46, so any negative or positive linear associations were weak. All R^2 values ranged from 0.27% to 9.6%, suggesting that very little variation in the domain/composite scores can be explained by age. In addition, a majority of CG cases and controls were within the adequate or moderately high adaptive levels, represented by the yellow and light green shaded regions on Figures 18-25.

Table 10. Relationship Between Standard Scores and Age for Cases and Controls

Domain/Composite		CG Cases	Controls
Communication			
	R^2	0.0027	0.096
	Slope	-0.072	0.35
Daily Living Skills			
	R^2	0.025	0.095
	Slope	-0.25	0.46
Socialization			
	R^2	0.069	0.0029
	Slope	-0.37	0.065
Adaptive Behavior Composite	R^2	0.022	0.082
	Slope	-0.2	0.36

*R^2 describes the amount of variation in domain/composite standard scores that is accounted for by age. The slope is a result of calculating the change in domain/composite standard scores divided by the change in age.

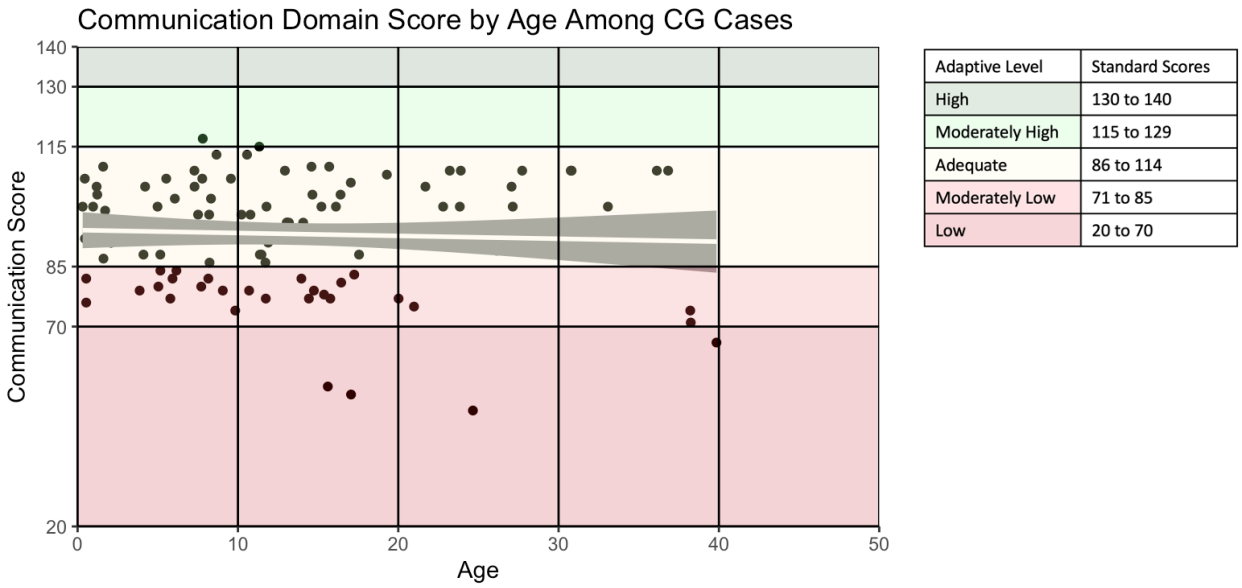


Figure 18. Communication Domain Score by Age Among CG Cases. The shaded regions correspond to the qualitative adaptive level for a range of standard scores.[31] R^2 = 0.0027 and slope = -0.072.

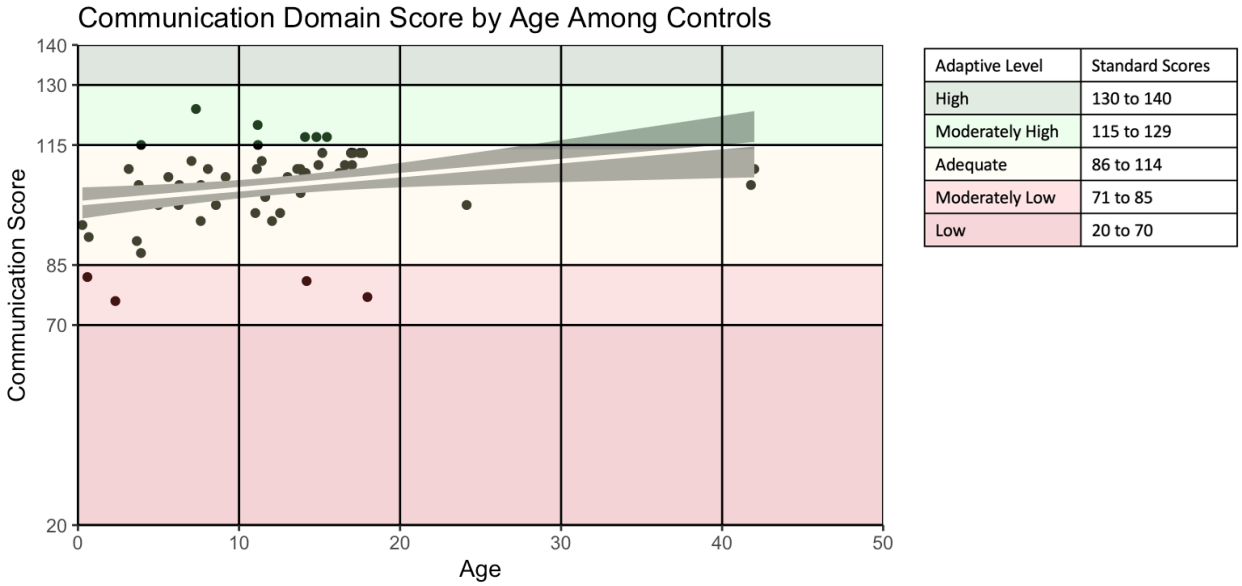
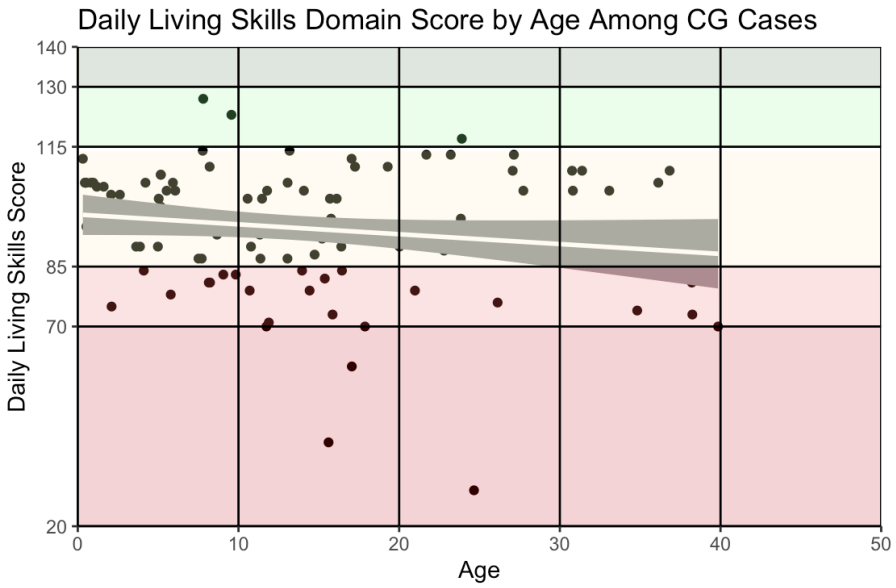
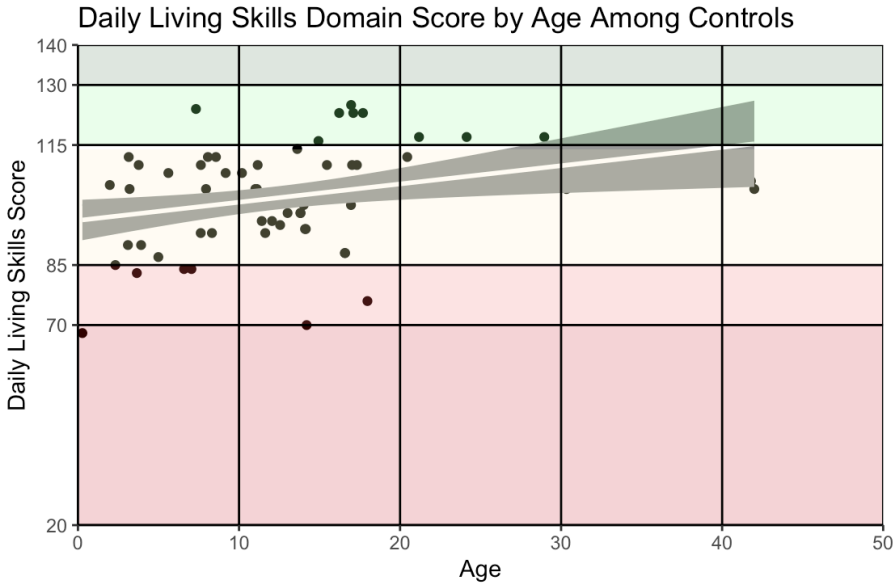


Figure 19. Communication Domain Score by Age Among Controls. The shaded regions correspond to the qualitative adaptive level for a range of standard scores.[31] R^2 = 0.096 and slope = 0.350.



Adaptive Level	Standard Scores
High	130 to 140
Moderately High	115 to 129
Adequate	86 to 114
Moderately Low	71 to 85
Low	20 to 70

Figure 20. Daily Living Skills Domain Score by Age Among CG Cases. The shaded regions correspond to the qualitative adaptive level for a range of standard scores.[31] R^2 = 0.025 and slope = -0.25.



Adaptive Level	Standard Scores
High	130 to 140
Moderately High	115 to 129
Adequate	86 to 114
Moderately Low	71 to 85
Low	20 to 70

Figure 21. Daily Living Skills Domain Score by Age Among Controls. The shaded regions correspond to the qualitative adaptive level for a range of standard scores.[31] R^2 = 0.095 and slope = 0.46.

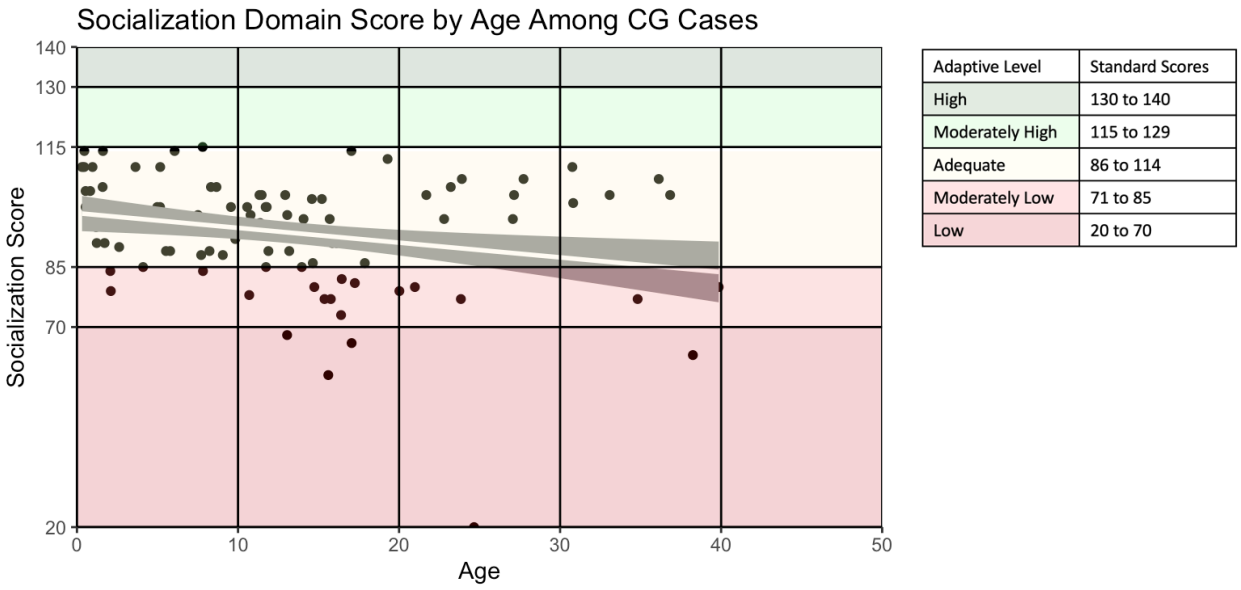


Figure 22. Socialization Domain Score by Age Among CG Cases. The shaded regions correspond to the qualitative adaptive level for a range of standard scores.[31] R^2 = 0.069 and slope = -0.37.

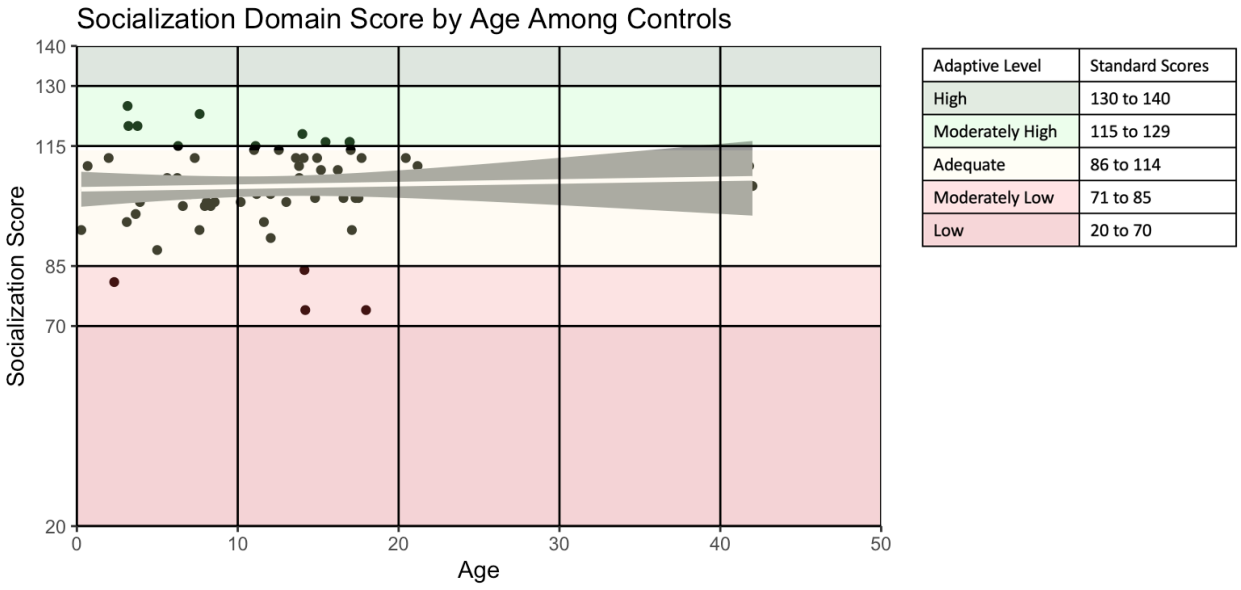
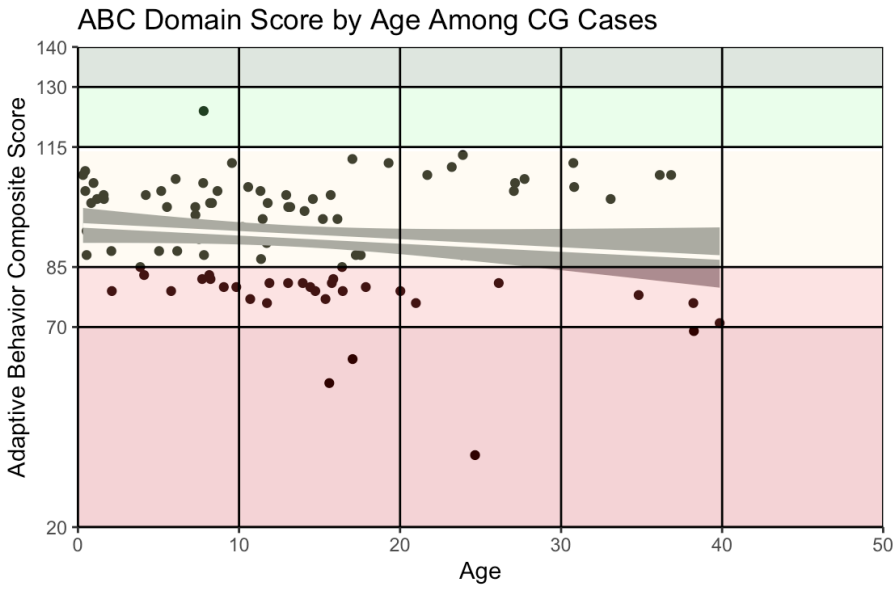
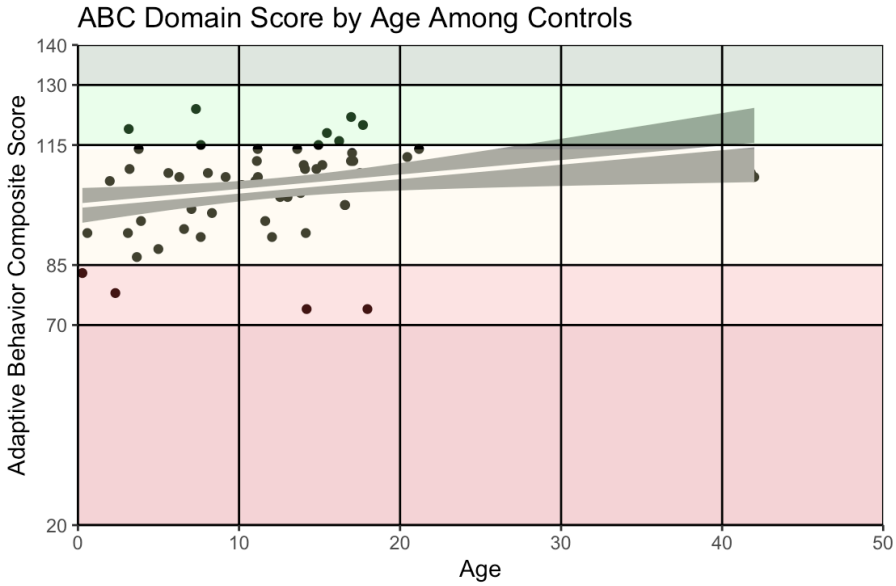


Figure 23. Socialization Domain Score by Age Among Controls. The shaded regions correspond to the qualitative adaptive level for a range of standard scores.[31] R^2 = 0.0029 and slope = 0.065.



Adaptive Level	Standard Scores
High	130 to 140
Moderately High	115 to 129
Adequate	86 to 114
Moderately Low	71 to 85
Low	20 to 70

Figure 24. Adaptive Behavior Composite (ABC) Score by Age Among CG Cases. The shaded regions correspond to the qualitative adaptive level for a range of standard scores.[31] R^2 = 0.022 and slope = -0.2.



Adaptive Level	Standard Scores
High	130 to 140
Moderately High	115 to 129
Adequate	86 to 114
Moderately Low	71 to 85
Low	20 to 70

Figure 25. Adaptive Behavior Composite (ABC) Score by Age Among Controls. The shaded regions correspond to the qualitative adaptive level for a range of standard scores.[30] R^2 = 0.082 and slope = 0.36.

These data suggest that relative to a large “normal” population, CG cases gain milestones more slowly, while their unaffected siblings gain milestones faster than the “normal” population.

Subdomain Raw Scores

There was a statistically significant difference between raw (not normed) scores of CG cases and controls for all subdomains when participants were between 10 and 19 years of age when the Vineland-3 was administered (Figures 26-36), with the exception of the coping skills subdomain (Figure 34). There was also a statistically significant difference between scores of CG cases and controls for the domestic as well as play and leisure subdomains when participants were between 20 and 29 years of age at the time the Vineland-3 assessment was administered (Figures 30, 33). Yet, beyond 29 years of age, there were no statistically significant differences between scores of CG cases and controls for any of the subdomains assessed (Figures 26-36).

Table 11. Significant P-values and Age Group for Raw Scores Using Wilcoxon Test

Subdomain	Age Group	P-value
Receptive	10 to 19	0.003
Expressive	10 to 19	<0.0001
Written	10 to 19	0.002
Personal	10 to 19	0.006
Domestic	10 to 19	0.029
	20 to 29	0.03
Community	10 to 19	0.003
Interpersonal relationships	10 to 19	<0.0001
Play and Leisure	10 to 19	<0.0001
	20 to 29	0.015
Gross Motor	10 to 19	0.001
Fine Motor	10 to 19	<0.0001

*P-values are considered significant if $p<0.05$.

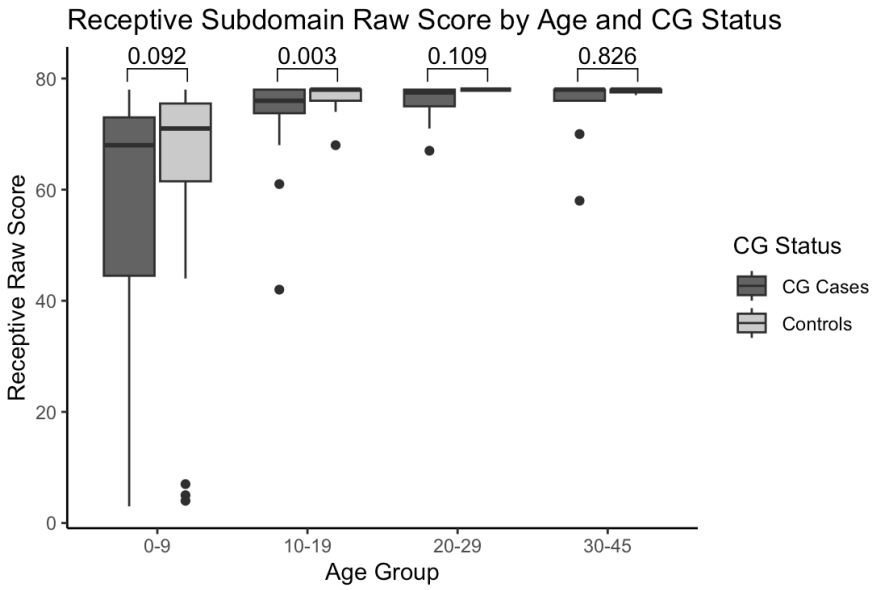


Figure 26. Receptive Subdomain Raw Score by Age and CG Status. The receptive subdomain is a component of the communication domain in the Vineland-3 assessment. There were 2 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if p<0.05.

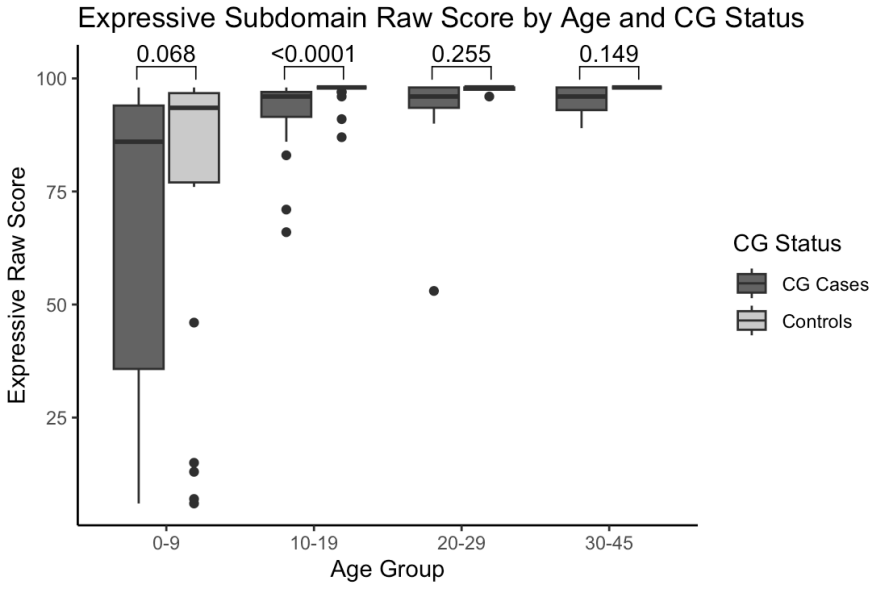


Figure 27. Expressive Subdomain Raw Score by Age and CG Status. The expressive subdomain is a component of the communication domain in the Vineland-3 assessment. There were 4 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if p<0.05.

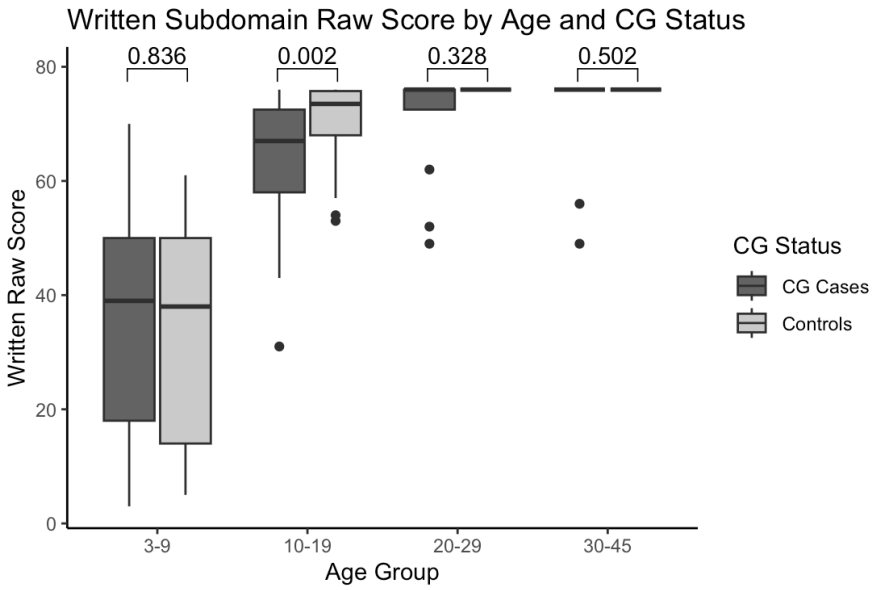


Figure 28. Written Subdomain Raw Score by Age and CG Status. The written subdomain is a component of the communication domain in the Vineland-3 assessment. Written raw scores are only collected for participants at least 3 years of age, so the first age group is 3 to 9 years of age. There were 8 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if $p<0.05$.

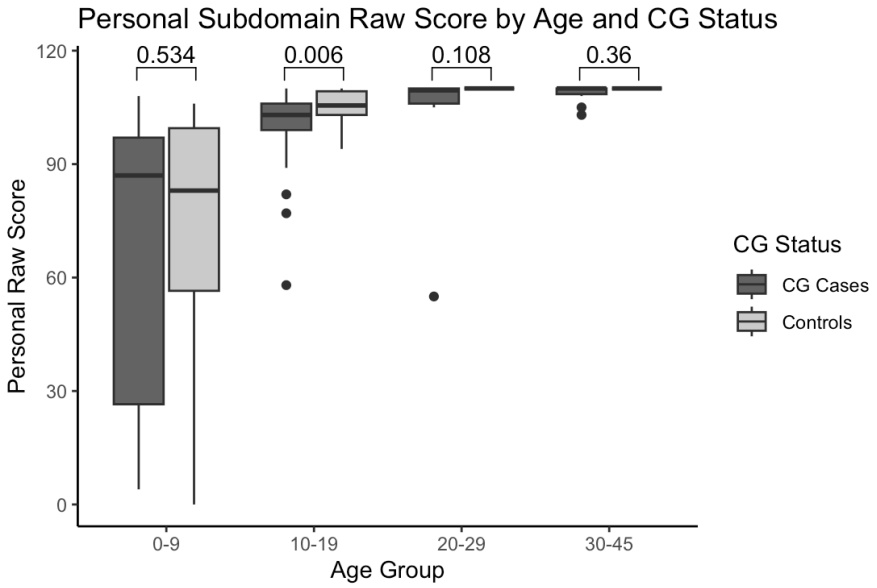


Figure 29. Personal Subdomain Raw Score by Age and CG Status. The personal subdomain is a component of the daily living skills domain in the Vineland-3 assessment. There was 1 participant whose raw score was excluded due to a percent estimating value of at least 25%. P-values are significant if $p<0.05$.

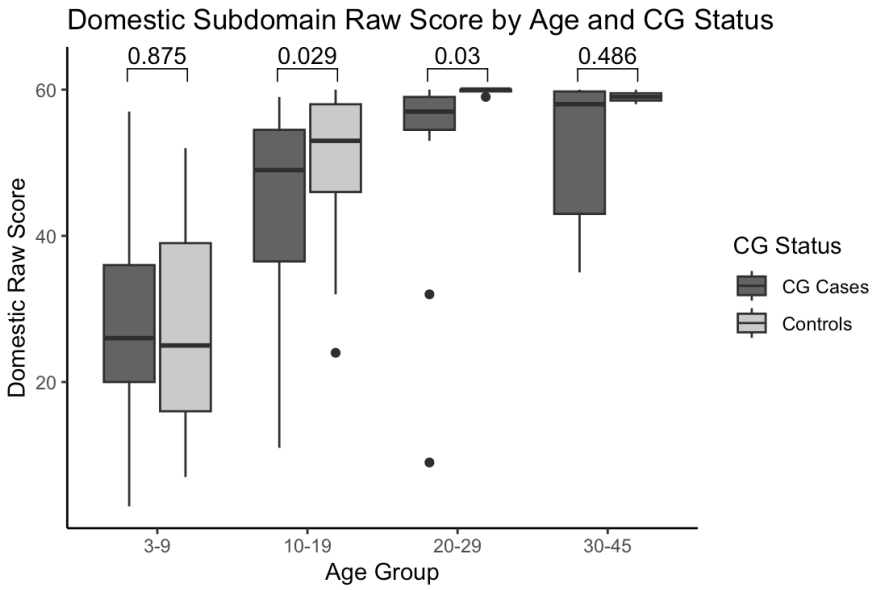


Figure 30. Domestic Subdomain Raw Score by Age and CG Status. The domestic subdomain is a component of the daily living skills domain in the Vineland-3 assessment. Domestic raw scores are only collected for participants at least 3 years of age, so the first age group is 3 to 9 years of age. There was 1 participant whose raw score was excluded due to a percent estimating value of at least 25%. P-values are significant if $p<0.05$

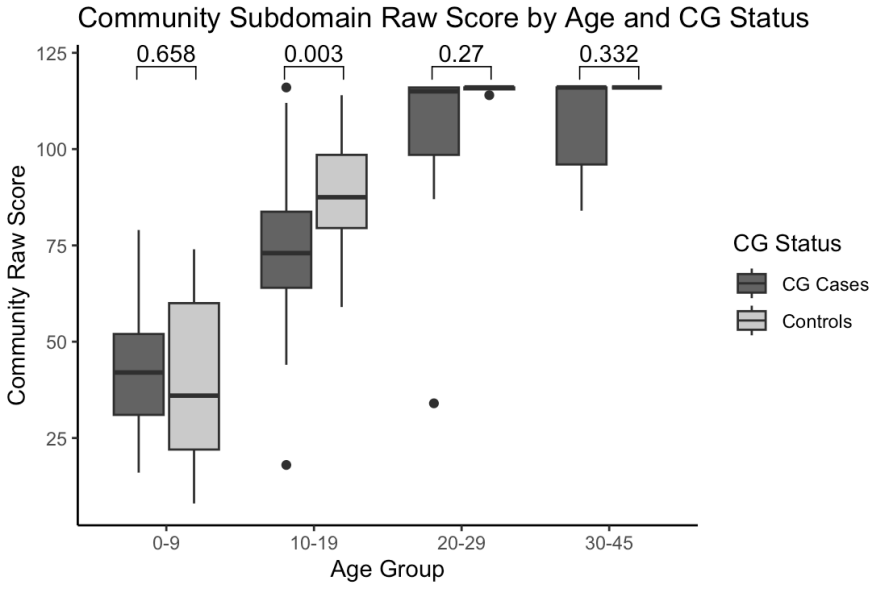


Figure 31. Community Subdomain Raw Score by Age and CG Status. The community subdomain is a component of the daily living skills domain in the Vineland-3 assessment. Community raw scores are only collected for participants at least 3 years of age, so the first age group is 3 to 9 years of age. There were 6 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if $p<0.05$.

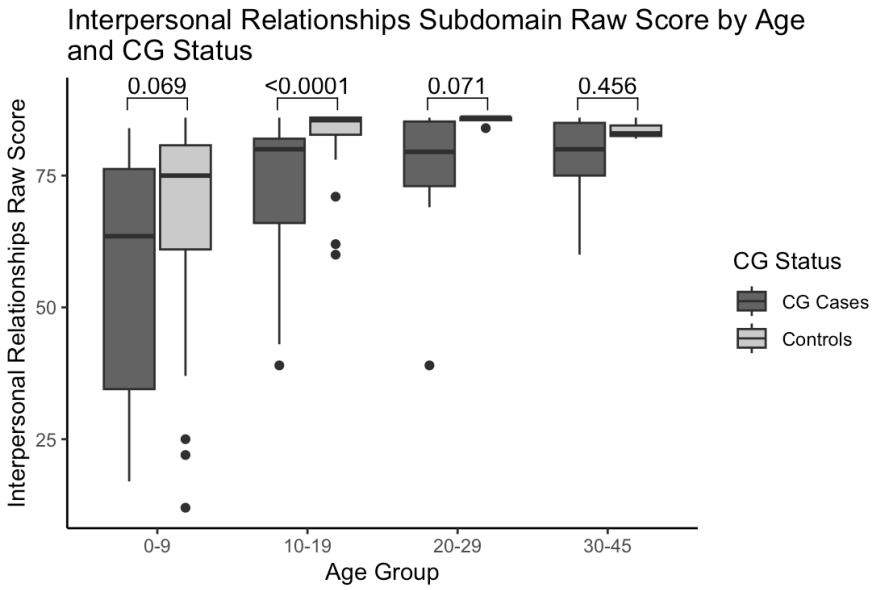


Figure 32. Interpersonal Relationship Subdomain Raw Score by Age and CG Status. The interpersonal relationship subdomain is a component of the socialization domain in the Vineland-3 assessment. There were 6 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if $p<0.05$.

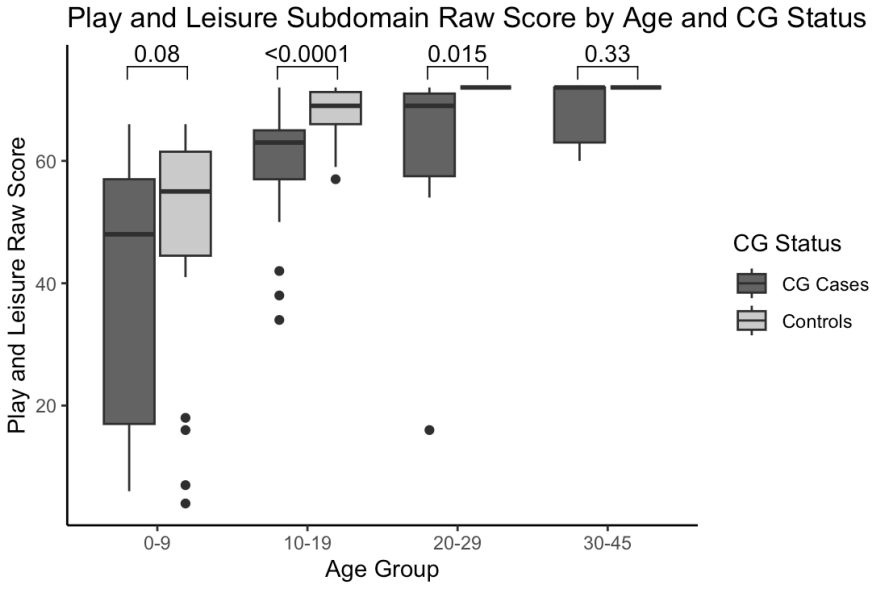


Figure 33. Play and Leisure Subdomain Raw Score by Age and CG Status. The play and leisure subdomain is a component of the socialization domain in the Vineland-3 assessment. There were 4 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if $p<0.05$.

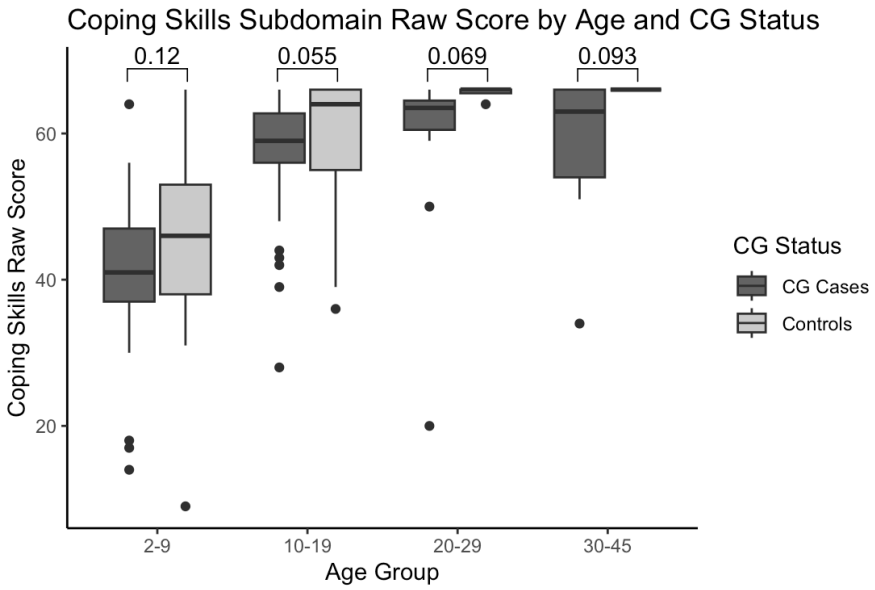


Figure 34. Coping Skills Subdomain Raw Score by Age and CG Status. The coping skills subdomain is a component of the socialization domain in the Vineland-3 assessment. Coping skills raw scores are only collected for participants at least 2 years of age, so the first age group is 2 to 9 years of age. There were 10 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if $p<0.05$.

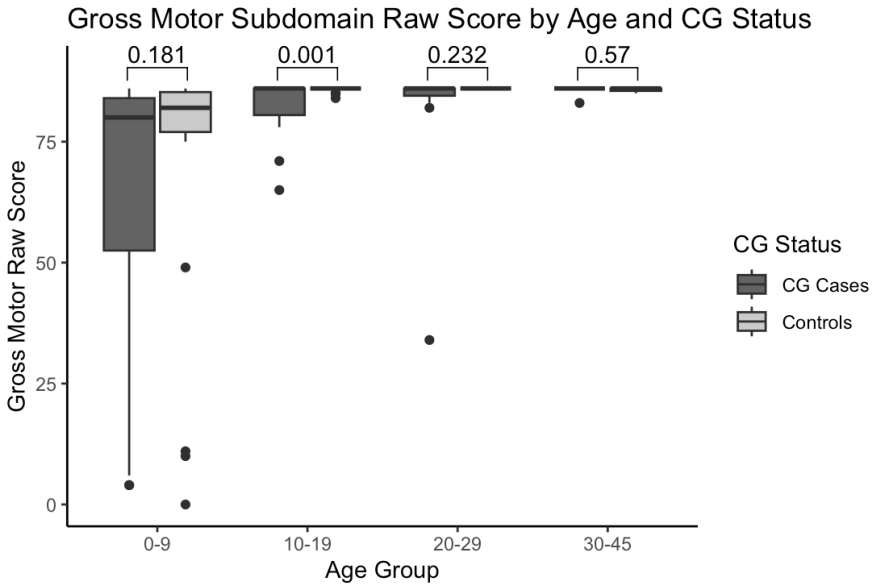


Figure 35. Gross Motor Subdomain Raw Score by Age and CG Status. The gross motor subdomain is a component of the motor skills domain in the Vineland-3 assessment. There were 7 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if $p<0.05$.

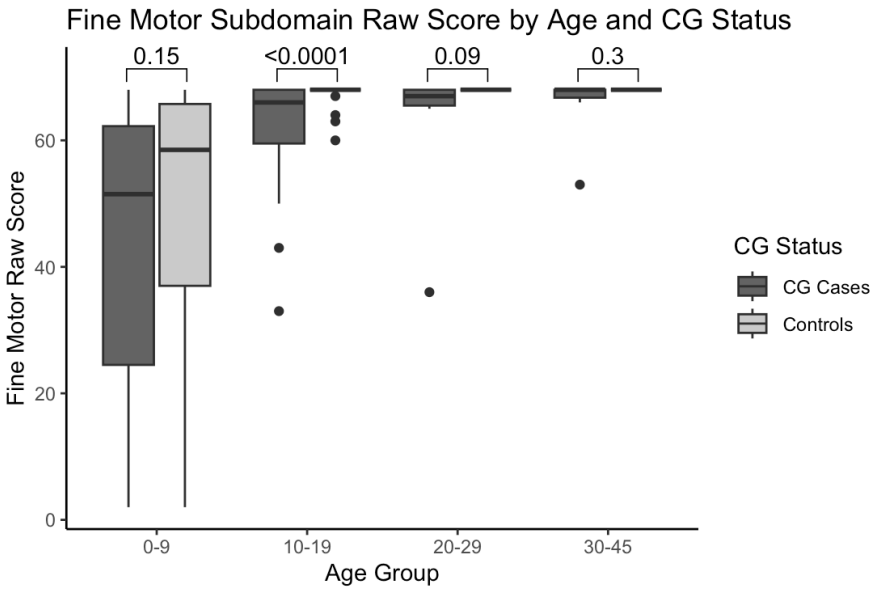


Figure 36. Fine Motor Subdomain Raw Score by Age and CG Status. The fine motor subdomain is a component of the motor skills domain in the Vineland-3 assessment. There were 4 participants whose raw scores were excluded due to percent estimating values of at least 25%. P-values are significant if $p<0.05$.

Discussion

An analysis of simple linear regression cross-sectionally comparing the domain and composite normed scores for CG cases by age highlighted that there was a marginal decrease in domain and composite normed scores as age increased. This contrasted the slight increase in domain and composite normed scores for sibling controls as age increased, evidenced by the positive slope of the linear trendline. This positive difference for controls, specifically, does not align with the expected horizontal linear trendline, given that these scores were normed to those of peers of the same age using a sample representative of the U.S. population.

However, it is not surprising. Recent research has suggested that families who choose to participate in medical research have a higher level of education, which has also been associated with having children that are high achievers.[32,33] Still, the slight decrease in domain and

composite scores of CG cases emphasizes that individuals with CG are facing challenges and are gaining milestones more slowly from these same families than both the general population and their unaffected siblings. The fact that a majority of CG cases were still within the adaptive level ranges of adequate or moderately high suggests that CG cases are reaching milestones, just less quickly than their unaffected siblings and unrelated healthy peers.

With regard to subdomain raw scores, CG cases had significantly lower scores than controls for a majority of subdomains between the ages of 10 to 19. Significantly lower scores were also observed for CG cases compared to controls in the domestic as well as play and leisure subdomains between the ages of 20 to 29. Significantly lower adaptive functioning among CG cases is not out of the ordinary, given that the existing literature has emphasized CG cases can experience more difficulties than their peers.[12] Yet, after 29 years old, there were no significant differences in subdomain raw scores by CG status.

The prevalence of significant differences only among CG cases and controls in younger age groups suggests that CG case may be reaching milestones in adaptive behavior but are struggling to meet these milestones at the same pace as their peers. Once CG cases reach adulthood, however, many disparities are no longer significant, suggesting CG cases are able to overcome the challenges they experienced and reach an adaptive functioning level that parallels that of their peers. A visual analysis of the data alone also highlights that there is a general increase in raw scores for both CG cases and controls across all subdomains as age increased, further emphasizing that the behaviors of CG cases are not worsening with age.

CHAPTER 4: DISCUSSION

CG is a rare autosomal recessive disorder that arises from a profound deficiency of GALT, an enzyme necessary for galactose metabolism.[2,3] Despite early detection of the disease and immediate restriction of dietary galactose, individuals with CG can still encounter long-term complications.[11] A majority of the existing cross-sectional and longitudinal studies of speech, cognitive, motor, and reproductive outcomes in CG suggest that the disorder is not progressive. However, this point remains controversial.

One objective of this thesis was to identify how study participants enrolled in the ongoing longitudinal study of outcomes in CG characterized changes to their health outcomes over time. When participants experiencing adverse speech, cognitive, and motor/neurological outcomes were asked how their symptoms have changed over time, a majority indicated that their adverse outcomes had stayed the same or improved with time. Although this characterization is solely based on the perceptions of study participants, it provides a glimpse of participants' experiences. These characterizations alone suggest that for most CG cases, outcomes are not worsening over time.

The prevalence of adverse speech, cognitive, and motor/neurological outcomes was significantly higher in child participants with CG compared to controls. Among adult participants, however, the prevalence of these same adverse outcomes by CG status was not

statistically different. Thus, children with CG are experiencing more adverse outcomes than their peers, yet when they reach adulthood, the prevalence of adverse outcomes experienced among CG cases is no different than that of their peers. Additionally, when comparing the prevalence of these adverse outcomes between child and adult participants with CG, the prevalence either decreased or remained relatively stagnant as age group increased.

Another aim of this study was to score the severity of each outcome experienced by participants in the observational study and cross-sectionally analyze the severity of health outcomes among CG cases over time. A majority of the significant differences were found when comparing the severity of health outcomes by CG status between younger age groups. Moreover, the severity of adverse speech, cognitive, and motor/neurological outcomes generally decreased as the age of the cohorts analyzed increased.

These findings do seem to align with the trend of participants' perceptions of how their adverse health outcomes have evolved over time, as the comparison of the severity of scores suggest that adverse outcomes either do not change or improve as time progresses. The high prevalence of liberalization of CG case diets over time as well as consistent scores for severe impairment of ovarian function in females with CG across age groups further highlight this conclusion that such health outcomes associated with CG are not worsening with age.

Lastly, this thesis intended to analyze cross-sectionally how the adaptive behaviors of CG cases, measured via the Vineland-3, change over time. The figures plotting domain normed

scores, as well as composite scores, by age found a slight decrease among CG cases and slight increase among controls as age increased. This was true for all domains evaluated. Such trends emphasize that individuals with CG are not reaching milestones as quickly as others. Yet, a majority were still within adequate/moderately high adaptive levels, indicating CG cases are still achieving milestones.

The subdomain scores found significantly lower raw scores for CG cases in all subdomains between 10 to 19 years of age and no significant differences in raw scores by CG status beyond 29 years of age. As such, although younger cases with CG may struggle to reach milestones at the same pace as their peers, in adulthood, CG cases are considered to have developed adaptive behaviors that are no different than those of their peers. The general increase in scores for CG cases among the subdomains with increasing age further emphasizes that CG cases are not losing adaptive functions they had previously achieved as they age.

Limitations

One limitation to this study was the small sample sizes for both CG cases and controls with regard to survey responses and Vineland-3 assessments, particularly among adult participants. When analyzing the Vineland-3 data, the maximum age of participants included in the investigation had to be set at 45 years of age, despite having had participants older than 45

years of age complete the assessment. This was due to the fact that the sample size of adult participants greater than 45 years of age was quite small (CG cases n = 5; controls n = 1), and any cross-sectional comparisons by CG status would not be adequate with only one control beyond 45 years of age. Excluding participants due to percent estimating values for Vineland-3 respondents also decreased the sample size.

Similarly, the sample size for summary scores was only a subset of the total sample of participants enrolled in the observational study, as participants who had completed every initial survey were prioritized. The sample size for adult participants was especially small, with some age groups only having 3 controls, so any statistical analyses comparing adult participants by CG status and age group may not be representative of the population. Larger sample sizes for comparing both summary scores of health outcomes and Vineland-3 subdomain raw scores would increase the validity of the results and reduce the impact of outliers on the conclusions.

Another considerable limitation to this study was the absence of a longitudinal component. The investigation focused on cross-sectional analyses due to the abundance of available responses from the initial surveys and first round of the Vineland-3 assessment within the ongoing study. Yet, as a result of solely applying cross-sectional analyses of survey responses and behavioral assessment results to address this research question, all sources of

data were based off of the perceptions of participants and respondents. Likewise, the summary scores were created based off of interpretations of survey responses.

Furthermore, despite attempts to create clear guidelines for assigning summary scores, the scores were assigned by two researchers, myself and another honors thesis candidate, so there could potentially be differences based on who scored the outcomes. Nevertheless, a study like this one seeking to determine if outcomes are progressive would greatly benefit from the use of contemporaneous data, such as medical or scholastic records, to provide a validated longitudinal aspect to the analysis.

Future Directions

The survey responses and Vineland-3 scores of participants included in this thesis represent only a fraction of the total participants enrolled in our observational study of outcomes in CG, so the ongoing investigation to determine whether outcomes in CG are progressive will include a larger sample. The aforementioned limitations are actively being addressed, as more study participants are completing initial surveys and Vineland-3 assessments, to ensure comparisons reflect the experiences of individuals with CG. In addition, medical and scholastic records are currently being compiled for study participants to obtain a

more accurate measure of how the long-term outcomes in CG are or are not changing over time. These next steps will provide a necessary longitudinal component to the study.

The results from this thesis and future investigations regarding whether outcomes in CG are progressive will have several implications. First, the conclusions will allow for more accurate prognostic information that will help families with CG in planning for the future. Moreover, the study will contribute to furthering the understanding of the natural history of CG as well as inform how future interventions would best be used to treat patients with the disorder.

REFERENCES

1. Pyhtila, B. M., Shaw, K. A., Neumann, S. E., & Fridovich-Keil, J. L. (2015). Newborn screening for galactosemia in the United States: looking back, looking around, and looking ahead. *JIMD Rep*, *15*, 79-93. https://doi.org/10.1007/8904_2014_302

2. Berry, G. (2020). Classic Galactosemia and Clinical Variant Galactosemia. In E. D. Adam MP, Mirzaa GM, et al. (Ed.), *GeneReviews® [Internet]*. University of Washington. https://www.ncbi.nlm.nih.gov/books/NBK1518/

3. Holden, H. M., I. Rayment and J. B. Thoden, 2003 Structure and function of enzymes of the Leloir pathway for galactose metabolism. J Biol Chem 278**:** 43885-43888.

4. Coelho, A. I., Rubio-Gozalbo, M. E., Vicente, J. B., & Rivera, I. (2017). Sweet and sour: an update on classic galactosemia. *J Inherit Metab Dis*, *40*(3), 325-342. https://doi.org/10.1007/s10545-017-0029-3

5. Calderon, F. R., Nelson, L., Dobrowolski, P., Sinitsyna, I., Phansalkar, A., Longo, N., Pasquali, M., & Mao, R. (2007). Combination of enzyme analysis, allele-specific PCR and sequencing to detect mutations in the GALT gene. *J Inherit Metab Dis*, *30*(5), 818. https://doi.org/10.1007/s10545-007-0461-x

6. Tyfield, L., Reichardt, J., Fridovich-Keil, J., Croke, D. T., Elsas, L. J., Strobl, W., Kozak, L., Coskun, T., Novelli, G., Okano, Y., Zekanowski, C., Shin, Y., & Boleda, M. D. (1999). Classical galactosemia and mutations at the galactose-1-phosphate uridyl transferase (GALT) gene. *Human Mutation*, *13*(6), 417-430. https://doi.org/Doi 10.1002/(Sici)1098-1004(1999)13:6<417::Aid-Humu1>3.0.Co;2-0

7. Guerrero, N. V., Singh, R. H., Manatunga, A., Berry, G. T., Steiner, R. D., & Elsas, L. J., 2nd. (2000). Risk factors for premature ovarian failure in females with galactosemia. *J Pediatr*, *137*(6), 833-841. https://doi.org/10.1067/mpd.2000.109148

8. Robertson, A., Singh, R. H., Guerrero, N. V., Hundley, M., & Elsas, L. J. (2000). Outcomes analysis of verbal dyspraxia in classic galactosemia. *Genet Med*, *2*(2), 142-148. https://doi.org/10.1097/00125817-200003000-00005

9. Kaufman, F. R., Reichardt, J. K., Ng, W. G., Xu, Y. K., Manis, F. R., McBride-Chang, C., & Wolff, J. A. (1994). Correlation of cognitive, neurologic, and ovarian outcome with the Q188R mutation of the galactose-1-phosphate uridyltransferase gene. *J Pediatr*, *125*(2), 225-227. https://doi.org/10.1016/s0022-3476(94)70197-0

10. Waggoner, D. D., Buist, N. R., & Donnell, G. N. (1990). Long-term prognosis in galactosaemia: results of a survey of 350 cases [Research Support, U.S. Gov't, P.H.S.]. *J Inherit Metab Dis*, *13*(6), 802-818. https://doi.org/10.1007/BF01800204

11. Fridovich-Keil, J., Bean, L., He, M., & Schroer, R. (2011). Epimerase Deficiency Galactosemia. In R. Pagon, T. Bird, C. Dolan, & K. Stephens (Eds.), *GeneReviews*. Seattle (WA): University of Washington, Seattle.

12. Rubio-Gozalbo, M. E., Haskovic, M., Bosch, A. M., Burnyte, B., Coelho, A. I., Cassiman, D., Couce, M. L., Dawson, C., Demirbas, D., Derks, T., Eyskens, F., Forga, M. T., Grunewald, S., Haberle, J., Hochuli, M., Hubert, A., Huidekoper, H. H., Janeiro, P., Kotzka, J., . . . Berry, G. T. (2019). The natural history of classic galactosemia: lessons from the GalNet registry. *Orphanet J Rare Dis*, *14*(1), 86. https://doi.org/10.1186/s13023-019-1047-z

13. Berry, G. T., Nissim, I., Lin, Z. P., Mazur, A. T., Gibson, J. B., & Segal, S. (1995). Endogenous Synthesis of Galactose in Normal Men and Patients with Hereditary Galactosemia. *Lancet*, *346*(8982), 1073-1074. https://doi.org/Doi 10.1016/S0140-6736(95)91745-4

14. Berry, G. T., Nissim, I., Gibson, J. B., Mazur, A. T., Lin, Z., Elsas, L. J., Singh, R. H., Klein, P. D., & Segal, S. (1997). Quantitative assessment of whole body galactose metabolism in galactosemic patients. *Eur J Pediatr*, *156 Suppl 1*, S43-49. https://doi.org/10.1007/pl00014271

15. Waisbren, S. E., Potter, N. L., Gordon, C. M., Green, R. C., Greenstein, P., Gubbels, C. S., Rubio-Gozalbo, E., Schomer, D., Welt, C., Anastasoaie, V., D'Anna, K., Gentile, J., Guo, C. Y., Hecht, L., Jackson, R., Jansma, B. M., Li, Y., Lip, V., Miller, D. T., . . . Berry, G. T. (2012). The adult galactosemic phenotype [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. *J Inherit Metab Dis*, *35*(2), 279-286. https://doi.org/10.1007/s10545-011-9372-y

16. Definition of progressive. (2023). In *Merriam-Webster Dictionary*. https://www.merriam-webster.com/dictionary/progressive

17. Duan, D., Goemans, N., Takeda, S., Mercuri, E., & Aartsma-Rus, A. (2021). Duchenne muscular dystrophy. *Nat Rev Dis Primers*, *7*(1), 13. https://doi.org/10.1038/s41572-021-00248-3

18. Kaufman, F. R., McBride-Chang, C., Manis, F. R., Wolff, J. A., & Nelson, M. D. (1995). Cognitive functioning, neurologic status and brain imaging in classical galactosemia. *Eur J Pediatr*, *154*(7 Suppl 2), S2-5. https://doi.org/10.1007/BF02143794

19. Schadewaldt, P., Hoffmann, B., Hammen, H. W., Kamp, G., Schweitzer-Krantz, S., & Wendel, U. (2010). Longitudinal assessment of intellectual achievement in patients with classical galactosemia [Research Support, Non-U.S. Gov't]. *Pediatrics*, *125*(2), e374-381. https://doi.org/10.1542/peds.2008-3325

20. Lewis, F. M., Coman, D. J., Syrmis, M., Kilcoyne, S., & Murdoch, B. E. (2013). Charting a seven-year trajectory of language outcomes for a child with galactosemia. *J Dev Behav Pediatr*, *34*(6), 414-418. https://doi.org/10.1097/DBP.0b013e31829a7be1

21. MacWilliams, J., Patel, S., Carlock, G., Vest, S., Potter, N. L., & Fridovich-Keil, J. L. (2021). Hand fine motor control in classic galactosemia. *J Inherit Metab Dis*, *44*(4), 871-878. https://doi.org/10.1002/jimd.12376

22. Spencer, J. B., Badik, J. R., Ryan, E. L., Gleason, T. J., Broadaway, K. A., Epstein, M. P., & Fridovich-Keil, J. L. (2013). Modifiers of ovarian function in girls and women with classic galactosemia [Research Support, N.I.H., Extramural]. *J Clin Endocrinol Metab*, *98*(7), E1257-1265. https://doi.org/10.1210/jc.2013-1374

23. Randall, J. A., Sutter, C., Wang, S., Bailey, E., Raither, L., Perfetti, R., Shendelman, S., & Burbridge, C. (2022). Qualitative interviews with adults with Classic Galactosemia and their caregivers: disease burden and challenges with daily living. *Orphanet J Rare Dis*, *17*(1), 138. https://doi.org/10.1186/s13023-022-02287-9

24. Jones, T. L., Baxter, M. A., & Khanduja, V. (2013). A quick guide to survey research. *Ann R Coll Surg Engl*, *95*(1), 5-7. https://doi.org/10.1308/003588413X13511609956372

25. Black, D., & Flowers, A. (2019). *The Newly Revised Vineland-3: An Overview of Essential Changes and Case Study Applications* [The Texas Association of School Psychologists 24th Annual Convention]. Pearson. https://tasp.memberclicks.net/assets/conference-materials/2016_Annual_Convention/Handouts/vineland-3_handout_tasp2016.pdf

26. *Vineland Adaptive Behavior Scales.* (2019). Shirley Ryan AbilityLab. https://www.sralab.org/rehabilitation-measures/vineland-adaptive-behavior-scales

27. Saulnier, C. A. (2017). *Vineland Adaptive Behavior Scales, Third Edition.* Pearson. http://downloads.pearsonclinical.com/videos/101817-vineland3/Vineland-3-Overview-Webinar-Handout-2017.pdf

28. Black, L., Vahratian, A., & Hoffman, H. (2015). Communication disorders and use of intervention services among children age 3-17 years: United States, 2012. *NCHS Data Brief*, *205*.

29. Gold E. B. (2011). The timing of the age at which natural menopause occurs. *Obstetrics and gynecology clinics of North America*, *38*(3), 425–440. https://doi.org/10.1016/j.ogc.2011.05.002

30. van der Geest, B. A. M., de Mol, M. J. S., Barendse, I. S. A., de Graaf, J. P., Bertens, L. C. M., Poley, M. J., Ista, E., Kornelisse, R. F., Reiss, I. K. M., Steegers, E. A. P., Been, J. V., & Group, S. S. (2022). Assessment, management, and incidence of neonatal jaundice in

healthy neonates cared for in primary care: a prospective cohort study. *Sci Rep*, *12*(1), 14385. https://doi.org/10.1038/s41598-022-17933-2

31. Farmer, C., Adedipe, D., Bal, V. H., Chlebowski, C., & Thurm, A. (2020). Concordance of the Vineland Adaptive Behavior Scales, second and third editions. *J Intellect Disabil Res*, *64*(1), 18-26. https://doi.org/10.1111/jir.12691

32. Kripalani, S., Goggins, K., Couey, C., Yeh, V. M., Donato, K. M., Schnelle, J. F., Wallston, K. A., Bell, S. P., Harrell, F. E., Mixon, A. S., Rothman, R. L., & Schildcrout, J. S. (2021). Disparities in Research Participation by Level of Health Literacy. *Mayo Clinic Proceedings*, *96*(2), 314-321. https://doi.org/10.1016/j.mayocp.2020.06.058

33. Tamayo Martinez, N., Xerxa, Y., Law, J., Serdarevic, F., Jansen, P. W., & Tiemeier, H. (2022). Double advantage of parental education for child educational achievement: the role of parenting and child intelligence. *European Journal of Public Health*, *32*(5), 690-695. https://doi.org/10.1093/eurpub/ckac044