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.

Nichole M. Stettner March 17, 2023

Exploring Potential Correlations Among Developmental Deficits Experienced by Patients with Classic Galactosemia

by

Nichole M. Stettner

Dr. Judith Fridovich-Keil
Adviser

Center for the Study of Human Health

Dr. Judith Fridovich-Keil

Adviser

Dr. Jill Welkley

Committee Member

Dr. Myra Woodworth-Hobbs

Committee Member

2023

Exploring Potential Correlations Among Developmental Deficits Experienced by Patients with Classic Galactosemia

By

Nichole M. Stettner

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 Arts with Honors

Center for the Study of Human Health

2023

Abstract

Exploring Potential Correlations Among Developmental Deficits Experienced by Patients with Classic Galactosemia

By Nichole M. Stettner

Classic galactosemia (CG) is an inborn error in metabolism of galactose, which results from a deficiency of galactose-1-phosphate uridylyltransferase (GALT) activity. Without treatment, CG gives rise to acute symptoms such as poor feeding, frequent vomiting, frequent diarrhea, poor weight gain, jaundice, and other acute symptoms including *E. coli* sepsis, which can be fatal. Adverse long-term outcomes in the speech, cognitive, motor/neurological, female puberty/reproduction, and growth/ bone health domains are still probable despite acute symptom prevention and dietary treatment. These long-term complications demonstrate incomplete penetrance and variable expressivity among patients – even those with exactly the same *GALT* genotype. The primary goal of this study is to determine whether the long-term outcomes in CG cluster. Results indicated that among the cognitive, speech, and motor/neurological outcome domains for both cases and controls, a strong clustering effect was observed in both the summary score and Vineland Test data. Although similar clustering effects were present, the CG cases experienced more severe adverse outcomes overall. Thus, if a participant with a CG diagnosis had an adverse outcome in a cognitive, speech, or motor/neurological domain, it is more likely that the participant also experienced one or more adverse outcomes in the other domains. These results suggest that shared modifiers may underlie multiple outcome domains in CG.

Exploring Potential Correlations Among Developmental Deficits Experienced by Patients with Classic Galactosemia

By

Nichole M. Stettner

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 Arts with Honors

Center for the Study of Human Health

2023

Acknowledgements

I would like to express my sincere appreciation to all the members of the Fridovich-Keil lab, particularly my advisor, Dr. Judith Fridovich-Keil, for her invaluable guidance and unwavering support. I am also grateful for the work by Danielle Orloff, Nicole Smith, Nathan Paull, Emma Hendrickson, and the rest of the human subjects team for their help inputting and organizing participant records. I would also like to acknowledge the enthusiastic participation from my committee members, Dr. Jill Welkley and Dr. Myra Woodworth-Hobbs, who provided continued encouragement throughout this process as well. Additionally, I would like to thank Shauna Rasmussen PhD, for her long-standing support and mentorship. Finally, I am grateful to my family and friends for their tremendous encouragement and support.

Table of Contents

Tables

Figures

CHAPTER 1: Background and Introduction to Classic Galactosemia

What is Classic Galactosemia?

Classic galactosemia (CG) is an inborn error in metabolism of galactose, which results from a deficiency of galactose-1-phosphate uridylyltransferase (GALT) activity (Berry, 2021). Diagnosis is confirmed by checking for GALT activity and confirmation by genetic testing (Berry, 2021). CG is often detected through newborn screening, symptoms, and/or suspicion from family history (CDC, 2017; Gitzelmann, 1980; Welling et al. 2017). There are different approaches to classifying CG, and this study follows the convention from the United States in which GALT enzyme activity is undetectable or less than 1% normal level (Berry, 2021). GALT activity among patients diagnosed with CG is a continuum, and many patient cohorts include patients who have clinical variant galactosemia with cryptic residual activity instead of CG, which may explain the variability in the results (Ryan et al., 2013).

Genetics and Prevalence

The *GALT* gene encodes the galactose-1-phosphate uridylyltransferase (GALT) enzyme (Berry, 2021). Individuals diagnosed with CG are deficient in GALT enzyme activity, caused by mutations in both their inherited copies of the *GALT* gene (Los & Ford, 2022; Berry, 2021). In CG, allelic heterogeneity in *GALT* can contribute to a varied range of clinical presentations (Berry, 2021; McCorvie et al., 2016). A decrease in GALT enzyme activity leads to increased levels of erythrocyte (RBC) galactose-1-phosphate (Gal-1-P) (Abi-Hann & Saavedra, 1998; Berry, 2021). As

illustrated in the Leloir pathway (Figure 1), increased RBC Gal-1-P levels manifest – notably when milk is consumed – and can exceed 10 mg/dL (Timson, 2016; Berry, 2021).

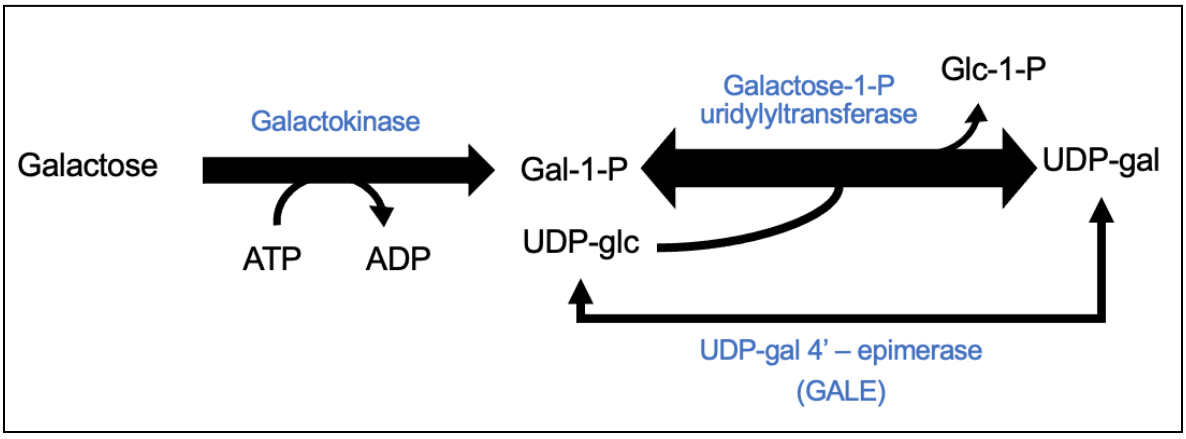

Figure 1. The Leloir pathway of galactose metabolism. Enzymes are presented in blue type.

Metabolic Findings

Individuals with CG are also likely to have increased Gal-1-P and galactitol levels when compared to controls (Berry, 2021; Hagen-Lillevik et al., 2021). As a standard of care, individuals diagnosed with CG are recommended to restrict galactose from their diets for life (Bosch, 2006; Welling et al. 2017).

Acute Symptoms

Without treatment, CG gives rise to acute symptoms (Coelho et al., 2017). When newborns diagnosed with CG are exposed to galactose after birth, they are more likely to develop symptoms such as poor feeding, frequent vomiting, frequent diarrhea, poor weight

gain, jaundice, and other acute symptoms including *E. coli* sepsis, which can be fatal (Bech et al., 2018; Rubio-Gozalbo et al., 2019).

Long-Term Complications Despite Standard of Care Treatment

Despite dietary treatment and prevention of acute symptoms with management of the diet, it is still likely for individuals diagnosed with CG to experience adverse long-term outcomes in the following domains, among others (Antshel et al., 2004; Hoffmann et al., 2011; Schweitzer et al., 1993; Waggoner et al., 1990; Waisbren et al., 2012).

- Speech
- Cognitive development
- Motor/neurological
- Female puberty/reproduction
- Growth & bone health

Current Understanding of the Onset, Prevalence, & Mechanisms of Long-Term Complications in Relation to Possible Modifiers

Cognitive Defects

According to Waggoner (1990), cognitive developmental delays had an overall prevalence of 45%, with the greatest drop in cross-sectional IQ scores occurring between the age ranges of 3-5 and 6-9. The prevalence of cognitive delays in the patient population of the Rubio-Gozalbo study was 39.5% (2019). Furthermore, participation in special education

programs was more prevalent in the patients with cognitive delays (Rubio-Gozalbo et al., 2019). Both the Rubio-Gozalbo and Waggoner studies had similar prevalence rates for cognitive developmental delays, with differing patient cohorts producing slightly varied results. This variability may arise since the Rubio-Gozalbo study has a predominantly European patient population and was published more recently; thus, special education service availability may have changed from the time the Waggoner study was published.

Speech Defects

The 1990 Waggoner study found that CG patients who were at least three years old had a higher prevalence of speech defects at 56%. Similarly, the Rubio-Gozalbo study found that speech defects were reported in 41% of the patient population, occurring more frequently in younger male patients (2019).

Motor/Neurological Defects

Waggoner's study observed a prevalence in motor defects in 18% of their population of patients with an age greater than 3.5 years (1990). According to the Rubio-Gozalbo study, the prevalence of motor/neurological problems was 52% of participants between the ages of 0 and 65 years. Neurological problems overall were more prevalent among those greater than 18 years; CG diagnosis via newborn screening may contribute to this trend, as it subsequently leads to a faster change in diet (Rubio-Gozalbo et al., 2019).

Compared to neurological problems that may present in any age group (such as ataxia and seizures), motor abnormalities were mainly observed in preschool-aged patients

(Rubio-Gozalbo et al., 2019). Contrastingly, tremors were first observed more frequently in patients following the second decade of life (41.3%).

Puberty/Reproduction

Primary ovarian insufficiency (elevated FSH levels) was observed in 80% of the female patient population from the Waggoner study; the onset of raised FSH levels was also observed in some cases following the onset of puberty (1990). Two cases in the Waggoner study changed from an elevated FSH level to a normal level; there was no explicit indication whether these participants' levels changed due to beginning hormone replacement therapy. Despite this, only about 30% of the women in Waggoner's patient population experienced normal menstruation. A majority of the women in the Waggoner study regulated FSH levels with hormone therapy.

The Rubio-Gozalbo study indicated that primary ovarian insufficiency was present in 79.7% of women in their patient population. In their study, a majority of women with primary ovarian insufficiency utilized hormone replacement therapy (83.5%); the median age at which hormone replacement therapy commenced was 16 years old (Rubio-Gozalbo et al., 2019).

Regarding hormone replacement therapy, Spencer et al. found a prevalence of 39.3% (n=28) of the studied population required it in order to reach menarche (Spencer et al., 2013). The Frederick study indicated that 14.2 years was the average age in which menarche was achieved following the initiation of hormone replacement therapy (Frederick et al, 2018). Frederick et al. also found that 30.8 was the average age at which hormone replacement therapy was stopped (2018). Moreover, they also demonstrated that there was an association

between detectable (greater than 0.04 ng/mL) plasma anti-Müllerian hormone and spontaneous menarche (Frederick et al, 2018).

Growth/Bone Health

Waggoner reported the prevalence of growth delay in females as 33% and males as 12% of a patient population that had an age range of 5 to 16 years (1990). A patient's growth was reported as delayed if their height-for-age was less than the 3rd percentile (Waggoner et al. 1990). The Rubio-Gozalbo study reported that the median age of patients that experienced bone fractures was 24, which is younger than the general population's (Rubio-Gozalbo et al., 2019).

Socioemotional/Behavioral Problems

Socioemotional and behavioral problems have been demonstrated to be more likely to occur at an older age (Rubio-Gozalbo et al., 2019); depression was more present after the second decade of life and had a prevalence of 12.5% in all patients. Contrastingly, anxiety had a prevalence of 36.8% in patients among those at the preschool age and those in the second decade of life (Rubio-Gozalbo 2019). The age distribution of the participants in the Rubio-Gozalbo study may have led to differing results in this area. The onset of psychiatric problems differs from that of behavioral problems in that the former is less likely to occur prior to the second decade of life (Rubio-Gozalbo et al., 2019).

Table 1. Prevalence of long-term outcomes examined in previous studies

Outcome Domain	Waggoner (1990)	Rubio-Gozalbo (2019)	Frederick (2018)	Spencer (2013)	Schweitzer (1993)
Speech problems	56%	41%	--	--	65%
Cognitive deficits	45%	39.5%	--	--	45%
Motor skill deficits	18%	52%	--	--	--
Growth delays	33.3% *(females)* & 12% *(males)*	--	--	--	--
Primary ovarian insufficiency / delayed puberty	80%	79.7%	≥90%	>73%	55%
Socioemotional / behavioral	--	44.4%	--	--	--

Possible Modifiers

Waggoner et al. looked at correlations between the age when diet switch began and the prevalence of cognitive delays, speech defects, abnormal ovarian function, and growth delays in their patient population (1990). They found a significant correlation between patients who had a cognitive delay and altered their diet after 2 months (Waggoner et al. 1990). In terms of the relationship between diet and cognition, no significant correlations were found between adherence to a strict or a non-strict (fair) diet and growth or the use of speech therapy or

special educational interventions in the studied patient population (Frederick et al., 2017). The Rubio-Gozalbo study supports the findings of the Frederick study that a diet with moderate liberalization should be recommended (Rubio-Gozalbo et al., 2019). Furthermore, when examining the relationship of neurological problems with regards to diet, it was demonstrated that there was a higher prevalence of symptoms associated with a strict diet than a non-strict diet (Rubio-Gozalbo et al., 2019).

Research cites Waggoner et. al (1990), N.L. Potter et al. (2008), and N.L. Potter et al. (2013) to indicate that speech problems are more likely to be diagnosed initially in childhood (Fridovich-Keil & Berry, 2022). Individuals who are diagnosed with a speech problem have been shown to improve this outcome via speech therapy (Peter et al., 2022).

Relevance of Investigating Possible Associations and Independence of Different Outcomes

Fridovich-Keil and Berry again cite Waggoner et. al (1990), N.L. Potter et al. (2008), and N.L. Potter et al. (2013) to demonstrate that their research has given some indication that there may be some correlations among the long-term outcomes of CG (2022). To validate their findings, more tests and studies are needed to be completed, which is part of what this thesis intends to examine.

Waggoner's paper also examined associations between speech and intelligence quotient (IQ) (Waggoner et. al, 1990). They demonstrated the trend that CG patients in the age range of 6-9 years with higher IQs were more likely to have normal speech relative to their counterparts with lower IQs (n=88). Furthermore, the study also established that patients with cognitive developmental delays and lower IQ scores had some correlation with speech defects, motor

deficits, and growth delays (Waggoner et. al, 1990). Participants who tested in the 3-5 and 6-9 age range, demonstrated a mean 6.2 point decrease in IQ from the former to the latter (Waggoner et. al, 1990). Similarly, from the 6-9 age group to the 10-16 age group, there was a mean 4.4 point decrease in IQ (Waggoner et. al, 1990). Moreover, Waggoner reported that for participants in the 10-16 age group, males had a significantly higher mean IQ score when compared with females (Waggoner et. al, 1990). Waggoner's research was later re-considered by Francine Kaufman, as Kaufman noted that older CG patients might show lower IQ scores because they did not have access to the early and preventative intervention that was available to the younger participants (Kaufman et al., 1995).

Thus, this study seeks to determine whether long-term outcomes in CG cluster. This study's findings may differ as the population is not confounded, since a majority of the study's population was born after early interventions were available. Thus, this study's cohort was predominantly identified by newborn screening and had the benefit of early intervention.

N.L. Potter (2013) found that individuals between the ages of 4-16 years diagnosed with CG are likely to have problems in speech in tandem with coordination and strength disorders. Specifically, they investigated potential correlations between motor skills (tongue and hand strength) and speech disorders; their results demonstrated that in their population of CG patients, there may be a common factor that leads to these outcomes to occur simultaneously (N.L. Potter et al., 2013).

The literature therefore has questioned whether clustering in the long-term outcomes of classic galactosemia is present. Patients diagnosed with classic galactosemia have incomplete penetrance, meaning that not every patient expresses every long-term adverse outcome. They

also demonstrate variable expressivity, in which patients will experience the same outcome but with varying severity.

Thus, since not every patient is affected the same, this study also questions whether the long-term outcomes cluster or if the outcomes are randomly distributed within the CG patient population.

CHAPTER 2: Study Methods

Study Participants

Data has been accumulated from family-response surveys as well as formal data (i.e., medical and school records, such as speech assessments and standardized test results) that has been entered into the Emory IRB-approved protocol 00024933 (PI: JL Fridovich-Keil) lab REDCap records of classic galactosemia patients (Table 2). Participants with a CG diagnosis served as the cases.

Unaffected siblings' data represented children without a CG diagnosis and served as a control for comparisons of prevalence of adverse outcomes. The unaffected siblings served as the controls for this project as they are matched from the same ancestry as the cases; they most often have similar –if not the same– lifestyle factors as their CG-diagnosed siblings such as attending the same school districts, which, in effect, controls for the long-term outcomes with respect to preventive treatment options. Pairwise comparisons between cases and controls were not conducted.

TABLE 2. Age and female/male distribution of the custom survey summary score data and Vineland-3 Test data

Dataset	**Participant Status**	**Sample Size**	**Age Range**	**Percent Female & Male**	
				Female	*Male*
Summary Score Data	Case	74	1-77 years	47%	51%
	Control	50	3-71 years	54%	46%
Vineland-3 Test Data	Case	120	0-77 years	56%	41%
	Control	78	0-71 years	55%	45%

Statistical Methods

A cross-sectional study was conducted to analyze and draw comparisons from the patient data. Outcome domains were assigned ordinal summary scores. The following includes the outcome domain assignments:

- Cognitive
    0. No cognitive problems (may have received preventative intervention)
    1. Mild or isolated problems with cognitive development (formal assessment result and/or received some responsive intervention for cognitive challenge, e.g., IEP in school)
    2. Moderate to severe problems with cognitive development (formal assessment result and/or received substantial prolonged responsive intervention for cognitive challenge in school)
- Speech
    0. No problems with speech (may have received preventative speech therapy)
    1. Documented problems with speech at least in childhood
    2. For adult participant, documented problems with speech that persisted to adulthood
- Motor/Neurological
    0. No motor/neurological problems documented
    1. Mild or isolated motor/neurological problems
    2. Moderate to severe motor/neurological problems

Vineland-3 Test Data

The Vineland-3 Test allows for behaviors related to the cognitive, speech, and motor/neurological outcome domains of childhood development to be evaluated (US, 2018). Scores are based on interview responses from caregivers, teachers, or interviewees (US, 2018). For this study, all Vineland-3 surveys were completed by parents or caregivers. Vineland-3 interview questions vary by the participant's age and developmental stage.

The standardized Vineland-3 Test Data was analyzed as a separate data source to compare against the results of the summary score data. The Vineland-3 Test domains were selected based on their similarities to the cognitive, speech, and motor/neurological outcome domains of the custom survey summary score data (Table 3). The following domains were tested for associations using raw scores (not normed to a reference population):

- Adaptive Behavior Composite Score
- Communication Score
- Motor Skills Score

TABLE 3. Vineland-3 Test sample scoring criteria and interview questions for the selected domains

	Participant Scoring Criteria Example (Sparrow et al., n.d.)	**Sample Question**
Adaptive Behavior Composite Score	Abilities to respond to unexpected changes	How does the participant respond to unanticipated changes in tasks?
Communication Score	Ability to utilize verbs in the past tense	How does the participant discuss events that took place in the past?
Motor Skills Score	Ability to draw with pen	Is the participant able to utilize a pen or similar writing utensil without difficulty?

Clustering effects were also tested for among the following subdomains:

- Receptive Communication
- Expressive Communication
- Written Communication
- Gross Motor Skills
- Fine Motor Skills

Statistical Tests Completed

Statistical tests were completed utilizing *RStudio* statistical software. Participants without sufficient data or who were too young to assess for a given domain were excluded from analysis.

The Shapiro-Wilk test was utilized to test the normality of the summary score outcome domains. After rejecting the null hypothesis of the Shapiro-Wilk test, the data was presumed as nonparametric. Following, the Wilcoxon Rank Sum Test with Continuity Correction was then utilized to determine whether there was a significant difference between the cognitive, speech, and motor/neurological domains relative to participant status (case/control) as well as reported sex.

Linear regression analyses were used to determine whether correlations of outcome domains were significant among the individuals diagnosed with CG who experienced adverse long-term outcomes. Thus, to determine any clustering effects relative to participant status, both the summary score and Vineland-3 Test data were subjected to linear regression analyses. Possible association of the cognitive, speech, and motor/neurological domain ordinal score

were separately determined among cases and controls. Linear regression analyses were also performed on the selected Vineland Test domain and subdomain raw scores to test for significant linear associations between the selected domains.

CHAPTER 3: Results

Summary Score Analyses – Cases & Controls

Shapiro-Wilk Test for Normality – Outcome Domains

The Shapiro-Wilk test was conducted to test the normality of the cognitive, speech, and motor/neurological domains summary score data (Table 4). The cognitive, speech, and motor/neurological domains were each found to be significantly different from a normal distribution; the results indicated that the cognitive (p=6.815e-14), speech (p=6.061e-14), and motor/neurological (p=3.787e-16) domains are nonparametric datasets (p-values < α=0.05).

Table 4. Shapiro-Wilk Test p-values for cognitive, speech, and motor domains

Summary Score Outcome Domain	Shapiro-Wilk Test p-value
Cognitive	6.815e-14
Speech	6.061e-14
Motor	3.787e-16

Wilcoxon Rank Sum Test with Continuity Correction – Cases & Controls

The Wilcoxon Rank Sum Test with Continuity Correction was used to test for a significant difference in the medians of the individual outcome domains relative to case/control status (Table 5; Table 6). The results indicated that a participant's score significantly differed on the basis of case versus control status (p-values < α=0.05). Thus, the null hypothesis that there is no true location shift of each outcome domain relative to the diagnostic status of the participant was rejected for the cognitive (Figure 2; p=7.975e-07), speech (Figure 3; p=1.337e-05), and

motor/neurological (Figure 4; p=1.923e-05) outcome domains. Although the medians for the motor/neurological outcome scores were the same for the case and control group, there continues to be a significant difference because the Wilcoxon Rank Sum Test is a non-parametric test, meaning that a normal distribution is not assumed. The case diagnostic group, therefore, has statistically significantly more severe results overall than the control group.

Table 5. Medians for the cognitive, speech, and motor/neurological outcome domain summary scores (cases & controls)

Outcome Domain	Cases	Controls
Cognitive	1	0
Speech	1	0
Motor/Neurological	0	0

Table 6. Wilcoxon Rank Sum Test with Continuity Correction p-values for cognitive, speech, and motor/neurological domains by status

Summary Score Outcome Domain	Wilcoxon p-value
Cognitive vs Diagnostic Status	7.975e-07
Speech vs Diagnostic Status	1.337e-05
Motor/Neurological vs Diagnostic Status	1.923e-05

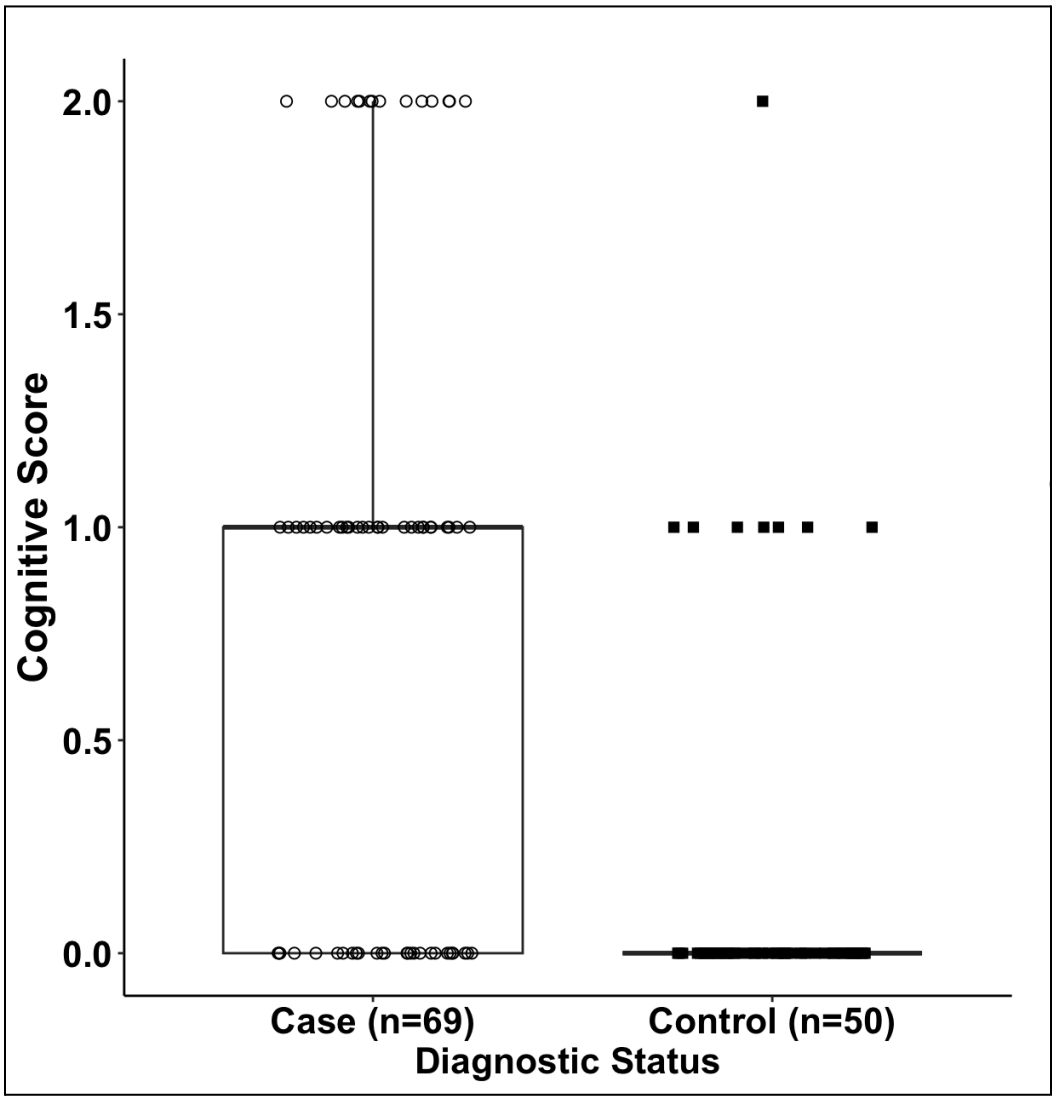

Figure 2. Distribution of cognitive outcome summary score by diagnostic status

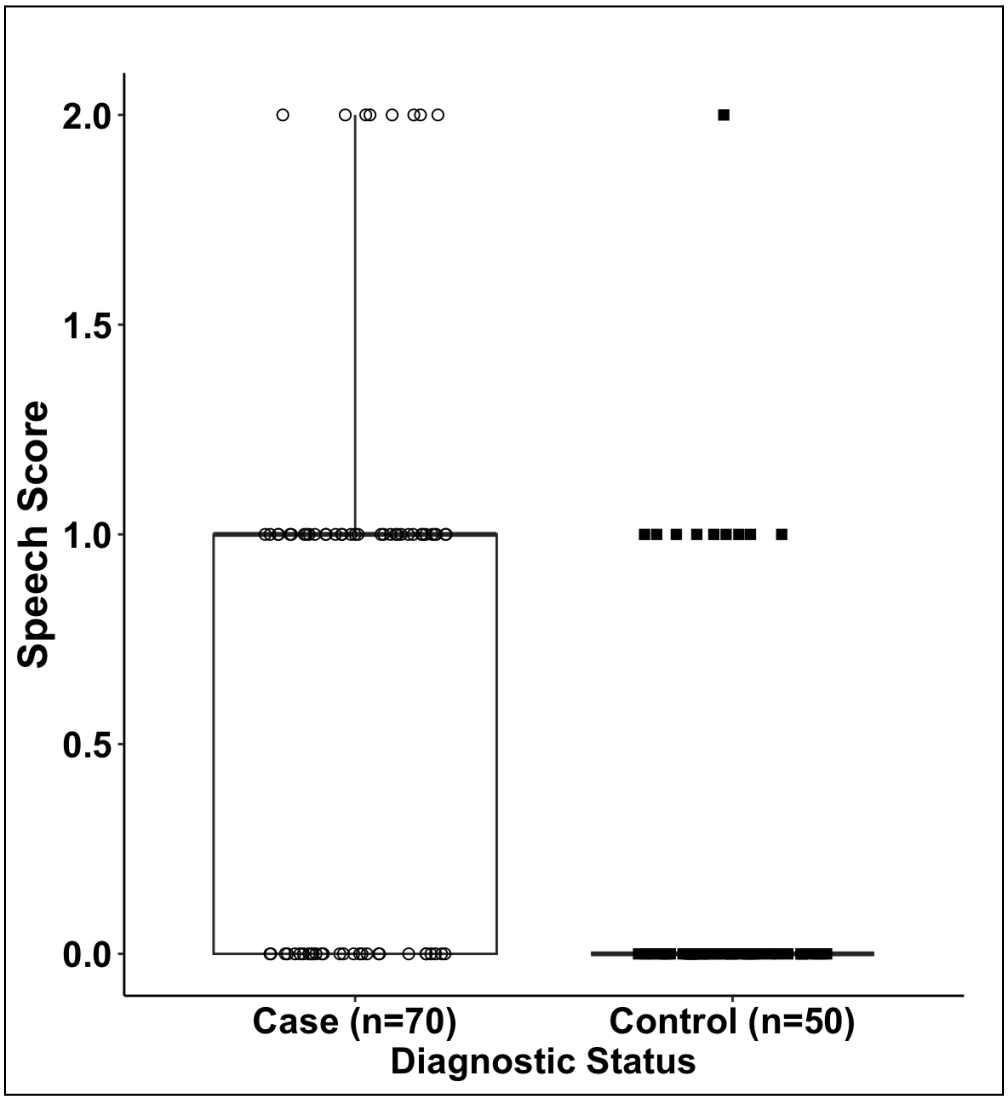

Figure 3. Distribution speech outcome summary score by diagnostic status

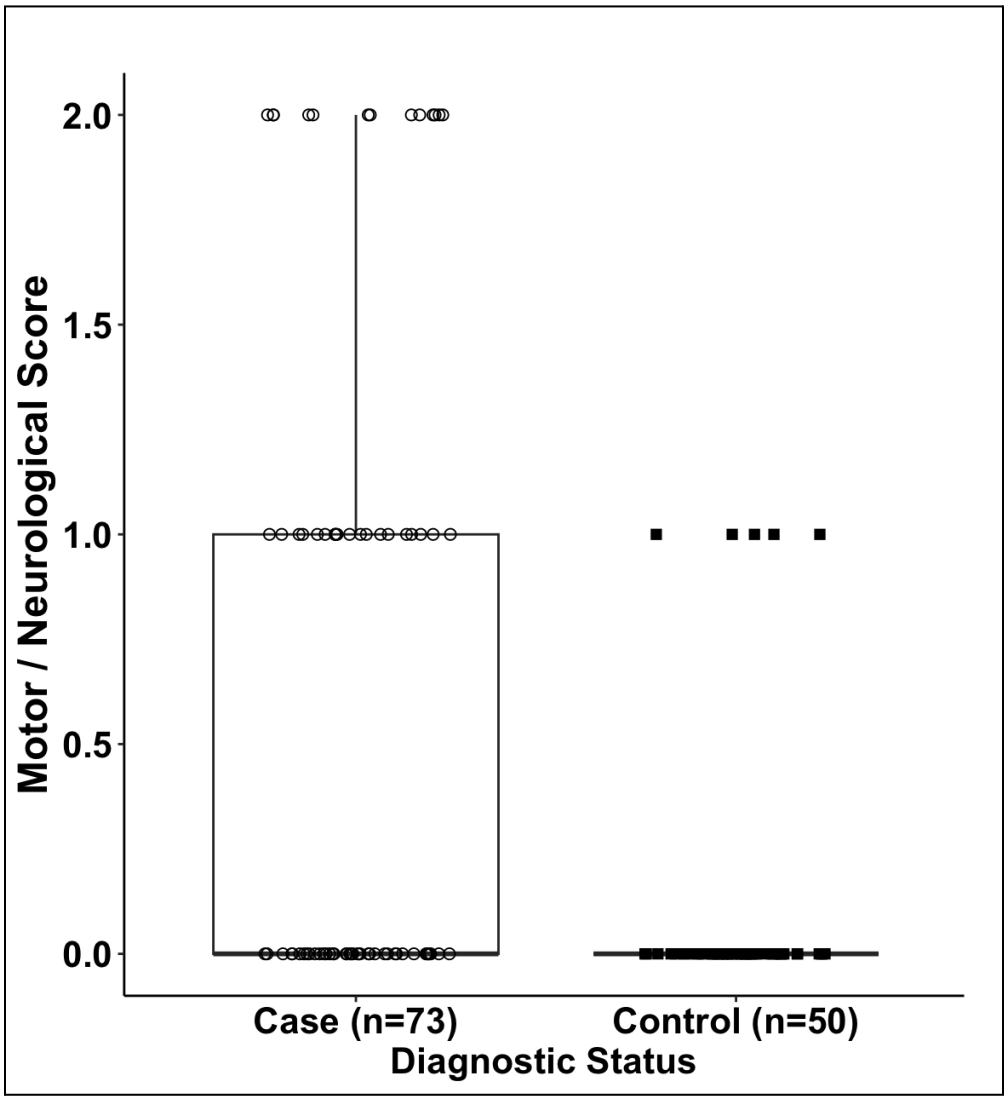

Figure 4. Distribution of motor/neurological outcome summary score by diagnostic status

The Wilcoxon Rank Sum Test with Continuity Correction failed to reject the null hypothesis that there is no true location shift of the medians in each outcome domain relative to the reported sex of the participant for both case and control groups (Tables 7A, 7B, and 7C). Thus, there is insufficient evidence to demonstrate a significant difference in the outcome domain scores relative to the reported sex of the participant (p-values > α=0.05).

Table 7A. Wilcoxon Rank Sum Test with Continuity Correction p-values for cognitive, speech, and motor/ neurological domains by reported sex (cases & controls); n=123 *(one participant excluded)*

Summary Score Outcome Domain	p-value
Cognitive by Sex	0.6065
Speech by Sex	0.2034
Motor/Neurological by Sex	0.3216

Table 7B. Wilcoxon Rank Sum Test with Continuity Correction p-values for cognitive, speech, and motor/neurological domains by reported sex (cases); n=73 *(one participant excluded)*

Summary Score Outcome Domain	p-value
Cognitive by Sex	0.3254
Speech by Sex	0.05072
Motor/Neurological by Sex	0.1261

Table 7C. Wilcoxon Rank Sum Test with Continuity Correction p-values for cognitive, speech, and motor/ neurological domains by reported sex (controls); n=50

Summary Score Outcome Domain	p-value
Cognitive by Sex	0.5818
Speech by Sex	0.6438
Motor/Neurological by Sex	0.2307

Summary Score Linear Regression Results – Cases & Controls

To lower the probability of committing a type I error, the Bonferroni correction was applied to ensure the appropriate interpretation of any p-values. The Bonferroni correction for an α=0.05 significance level given 3 tests conducted equated to an α=0.01667 (=0.05/3).

A statistically significant linear relationship (Figure 5) was found among the cases for the cognitive and speech summary score domains (Table 8). Although the R^2 value for relationship between the cognitive and speech summary score domains was closer to 0 than to 1, the sample size (n=69) was large enough so that the p-value holds significance (p-value=0.0003636).

The cognitive and motor/neurological summary score domains demonstrated a moderately strong and positive relationship among the cases (Table 8). A statistically significant linear relationship (Figure 6) was found between the cognitive and motor/neurological summary score domains (p-value=5.446e-08).

Among the cases, the speech and motor/neurological summary score domains also demonstrated a statistically significant linear relationship (Figure 7; Table 8). The R^2 value for relationship between the speech and motor/neurological summary score domains demonstrated a moderately strong and positive relationship. The p-value (p-value=1.579e-05) is significant due to the large sample size (n=69).

TABLE 8. Adjusted R-squared values, p-values, and residual standard errors (RSE) for the case comparisons of the selected summary score outcome domains. **5 observations deleted due to missingness; **6 observations deleted due to missingness; the degrees of freedom were therefore adjusted.*

Summary Score Outcome Domain Comparisons	Sample Size (n)	Adjusted R-Squared	p-value	RSE (degrees of freedom)
Cognitive vs Speech	69	0.1616	0.0003636	0.6102 (67)*
Cognitive vs Motor	68	0.3534	5.446e-08	0.6281 (66)**
Speech vs Motor	69	0.2331	1.579e-05	9.428 (67)*

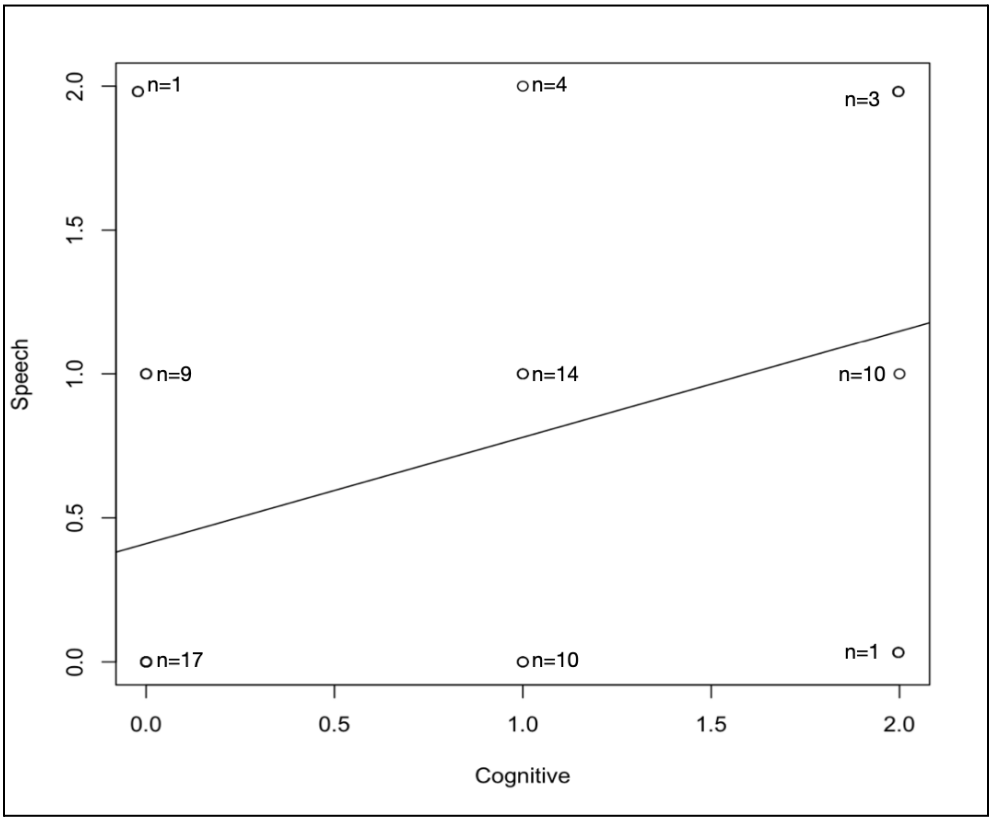

Figure 5. Linear regression of the cognitive and speech summary score for cases (n=69; age range: 1-77 years (median age: 14 years); 32 male, 36 female, 1 other/unknown); p-value=0.0003636

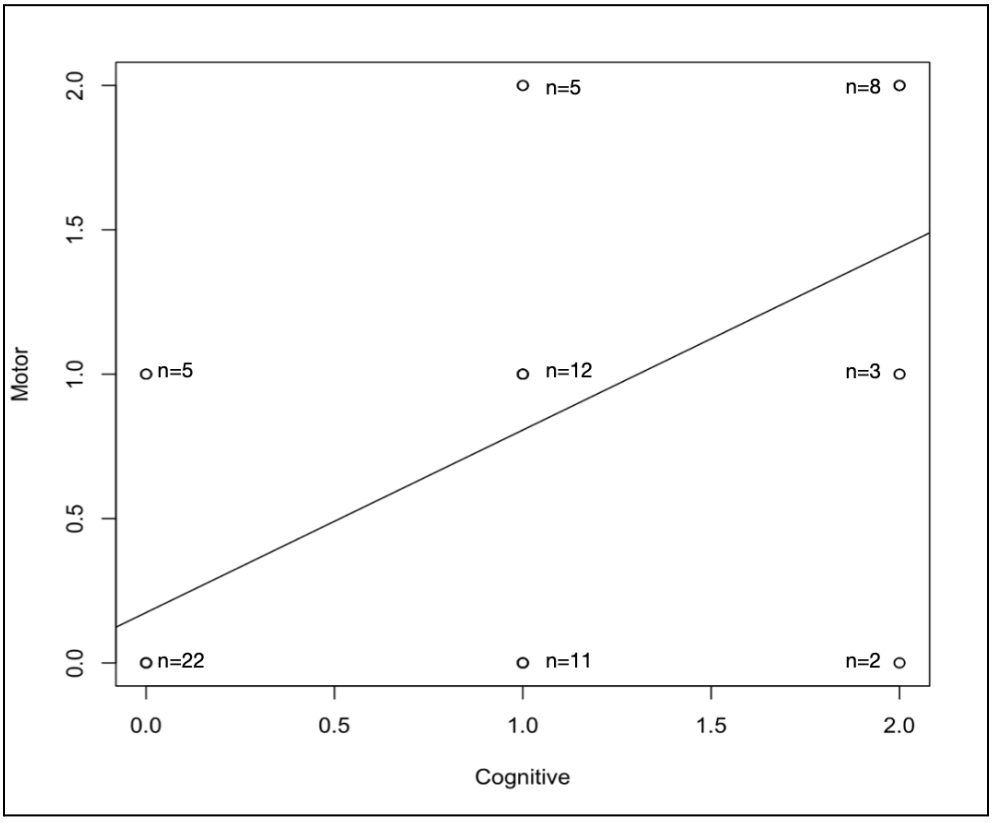

Figure 6. Linear regression of the cognitive and motor/neurological summary score for cases (n=68; age range: 1-77 years (median age: 14 years); 31 male, 36 female, 1 other/unknown); p-value=5.446e-08

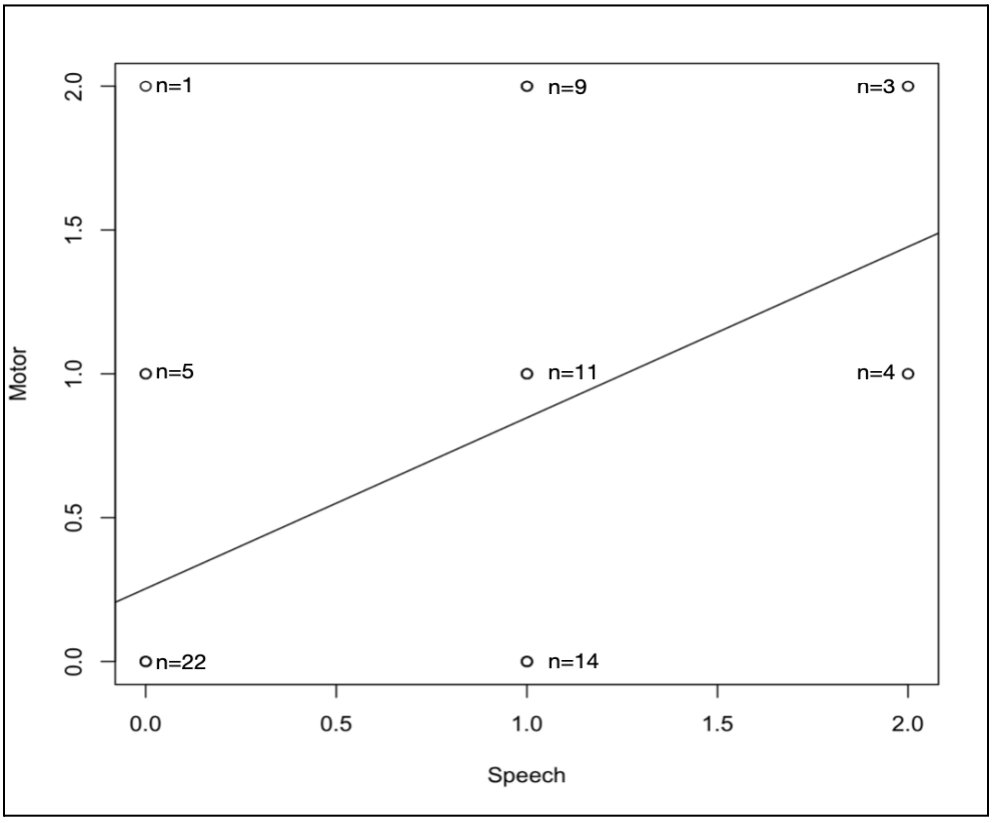

Figure 7. Linear regression of the speech and motor summary score for cases (n=69; age range: 1-77 years (median age: 14 years); 32 male, 36 female, 1 other/unknown); p-value=1.579e-05

The Bonferroni correction for an α=0.05 significance level was once again conducted to lower the chance of a type I error occurring in the statistical tests.

A statistically significant linear relationship (Table 9) was found between the cognitive and speech (Figure 8; p=0.0001909), cognitive and motor/neurological (Figure 9; p=0.0004591), and speech and motor/neurological (Figure 10; p=1.898e-05) summary score domains (p-values < α=0.01667). The large sample size of the controls (n=50) overcame the slightly weak R^2 values found between the outcome domains.

Table 9. Adjusted R-squared values, p-values, and residual standard errors for the control comparisons of the selected summary score outcome domains.

Summary Score Outcome Domain Comparisons	Sample Size (n)	Adjusted R-Squared	p-value	Residual Standard Error (degrees of freedom)
Cognitive vs Speech	50	0.2384	0.0001909	0.4055 (48)
Cognitive vs Motor	50	0.2116	0.0004591	0.2691 (48)
Speech vs Motor	50	0.3053	1.898e-05	0.2526 (48)

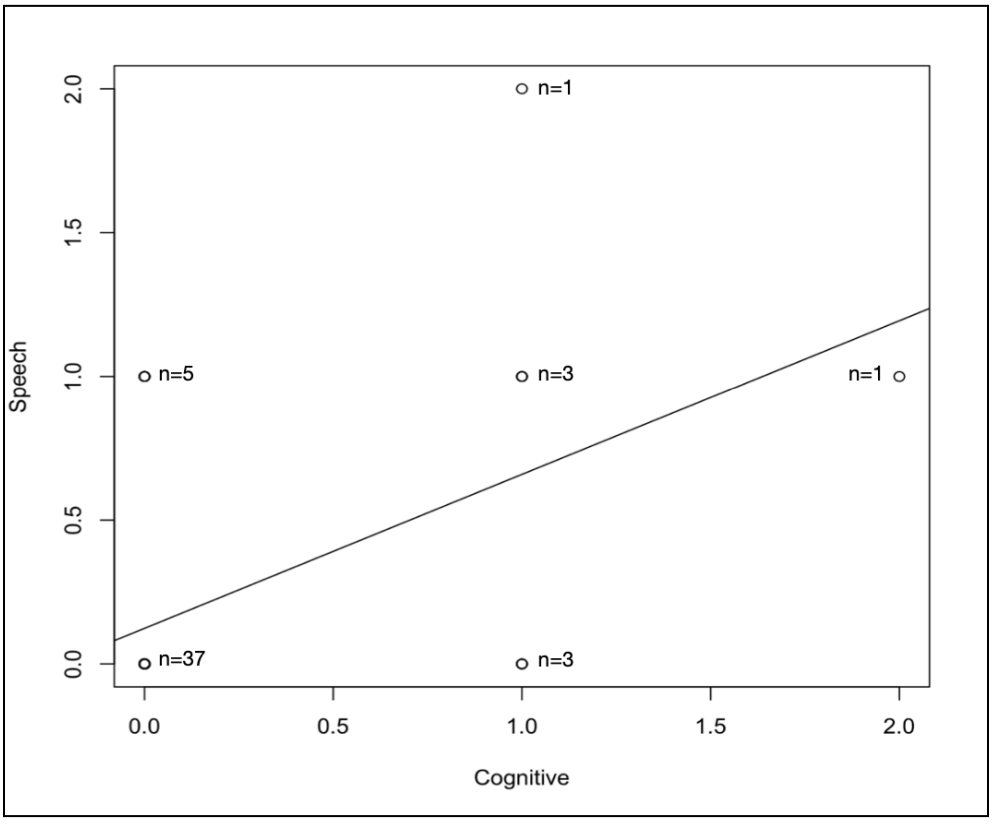

Figure 8. Linear regression of the cognitive and speech summary scores for controls (n=50; age range: 3-77 years (median age: 14 years); 23 male, 27 female); p-value=0.0001909

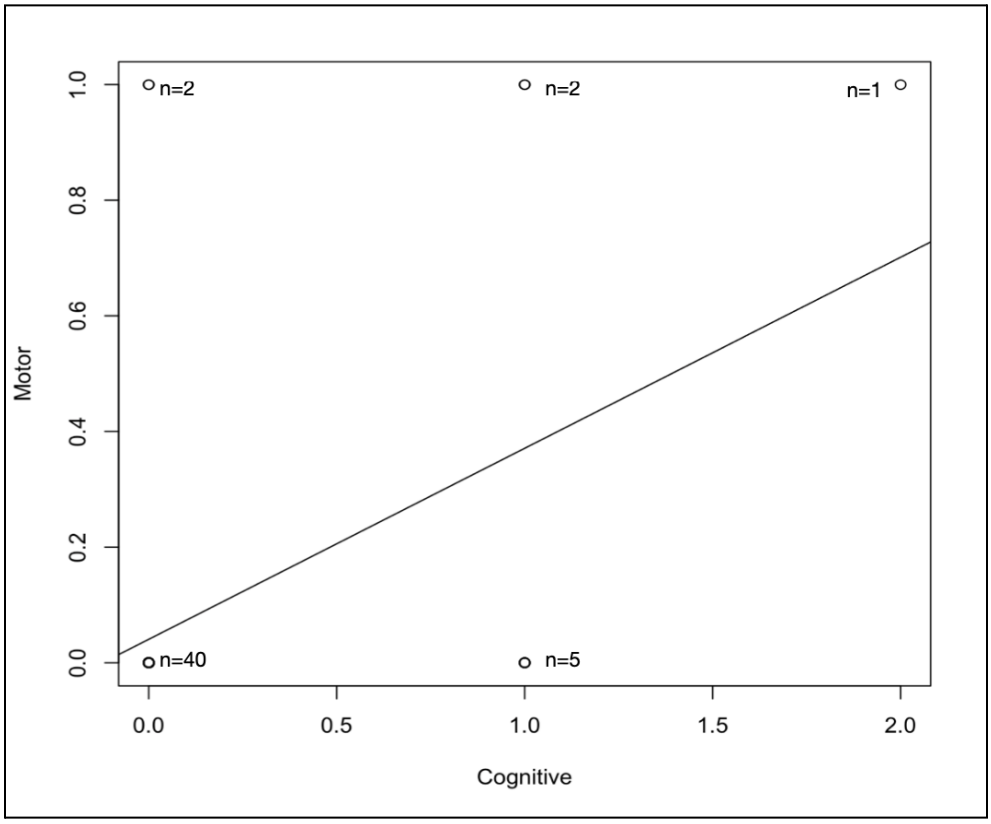

Figure 9. Linear regression of the cognitive and motor/neurological summary scores for controls (n=50; age range: 3-77 years (median age: 14 years); 23 male, 27 female); p-value=0.0004591

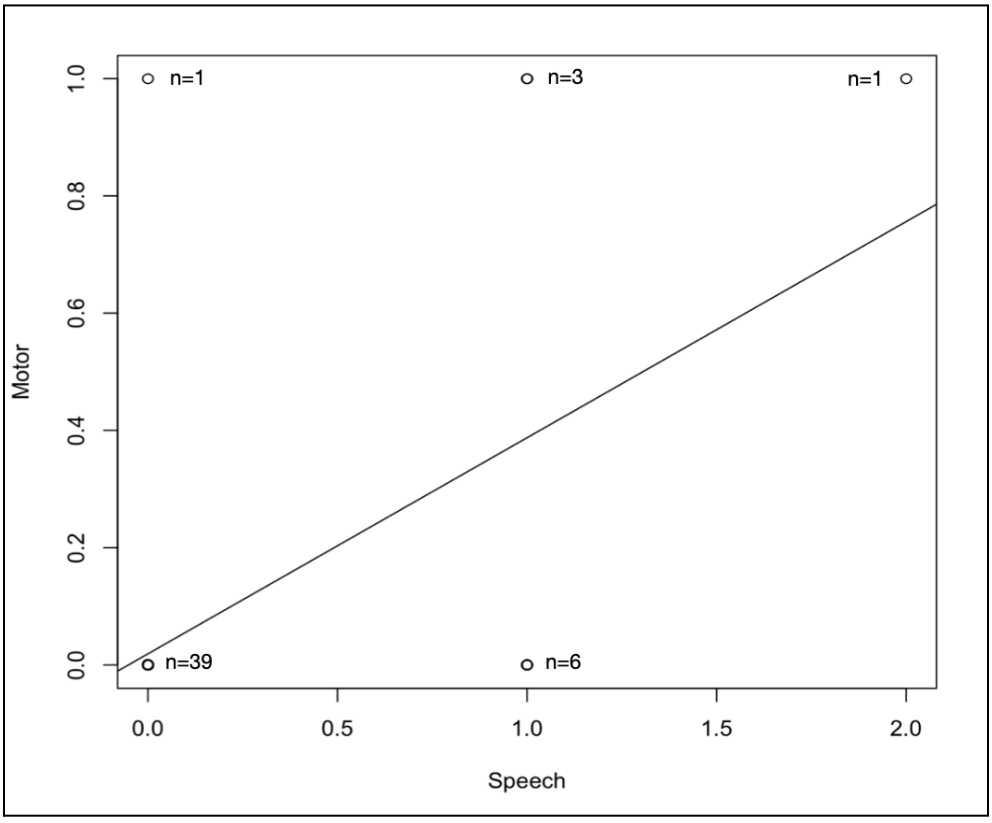

Figure 10. Linear regression of the speech and motor/neurological summary scores for controls (n=50; age range: 3-77 years (median age: 14 years); 23 male, 27 female); p-value=1.898e-05

Vineland-3 Data Analyses

Table 10. Vineland-3 Data parameters table for raw scores

Vineland-3 Domain	Description	Age Range Criteria
Adaptive Behavior Composite	A combination of the communication, daily living skills, and socialization domains	Birth to 90+ years
Communication Score	Indicates the participant's ability to listen and comprehend as well as to articulate notions via speech, reading, and writing	
Motor Skills	Ability of the participant to utilize gross and fine motor skills	

The Bonferroni correction value (α=0.01667) was also used when comparing the results of the Vineland-3 Test data (Table 10).

The Adaptive Behavior Composite Score and Communication Score domains (n=120) demonstrated a fairly strong and positive relationship among the cases (Figure 11). A statistically significant linear relationship (Table 11A) was found between the Adaptive Behavior Composite Score and Communication Score domains (p-value < α=0.01667).

Table 11A. Adjusted R-squared values, p-values, and residual standard errors (RSE) for the case comparisons of the selected domains of the Vineland-3 Test **63 observations deleted due to missingness; the degrees of freedom were therefore adjusted.*

Domain Comparisons	Sample Size (n)	Adjusted R-Squared	p-value	RSE (degrees of freedom)
Adaptive Behavior Composite Score vs Communication Score	120	0.7637	< 2.2e-16	6.573 (118)
Adaptive Behavior Composite Score vs Motor Skills Score	57	0.1531	0.00153	9.16 (55)*
Communication Score vs Motor Skills Score	57	0.1424	0.002221	9.217 (55)*

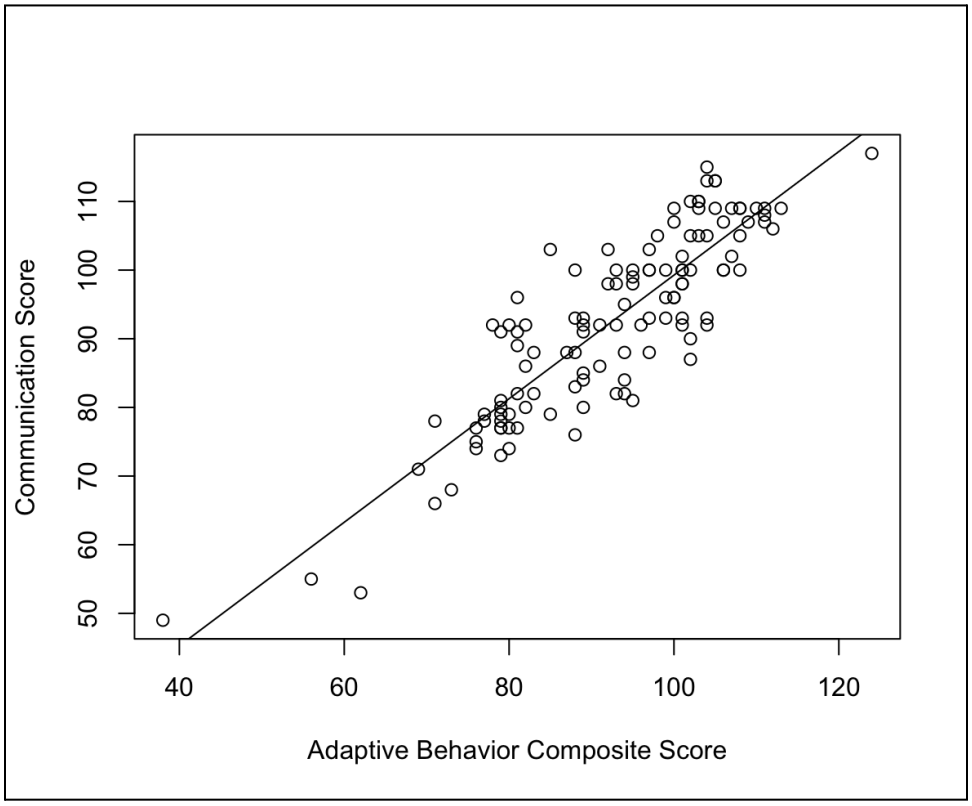

Figure 11. Linear regression of the Adaptive Behavior Composite Score and the Communication Score for cases (n=120; age range: 0-77 years (median age: 10.7 years); 49 male, 67 female, 4 other/unknown); p-value= < 2.2e-16

A statistically significant linear relationship was found among the cases for the Adaptive Behavior Composite Score and Motor Skills domains (Table 11A). Although the R^2 value for relationship between the Adaptive Behavior Composite Score and Motor Skills (Figure 12) domains was closer to 0 than to 1, the sample size (n=57) was large enough so that the p-value holds significance (p-value < α=0.01667).

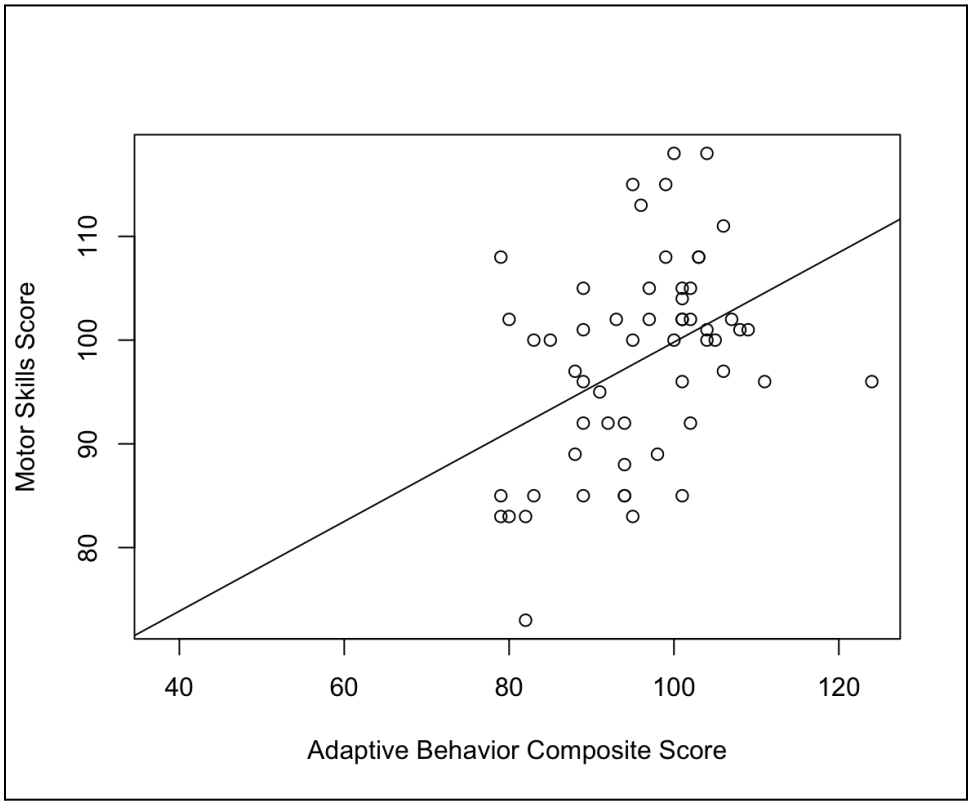

Figure 12. Linear regression of the Adaptive Behavior Composite Score and the Motor Skills Score for cases (n=57; age range: 0-9 years (median age: 10.7 years); 22 male, 32 female, 3 other/unknown); p-value=0.00153

The Communication Score and Motor Skills case domains also demonstrated a slightly strong and positive linear relationship (Figure 13) with a statistically significant linear relationship (Table 11A). The R^2 value for relationship between the Communication Score and Motor Skills was once again closer to 0 than to 1; however, the p-value (p-value < α=0.01667) is significant due to the large sample size (n=57). (For full analysis of Vineland subdomain scores of the cases, see Supplemental Table 1).

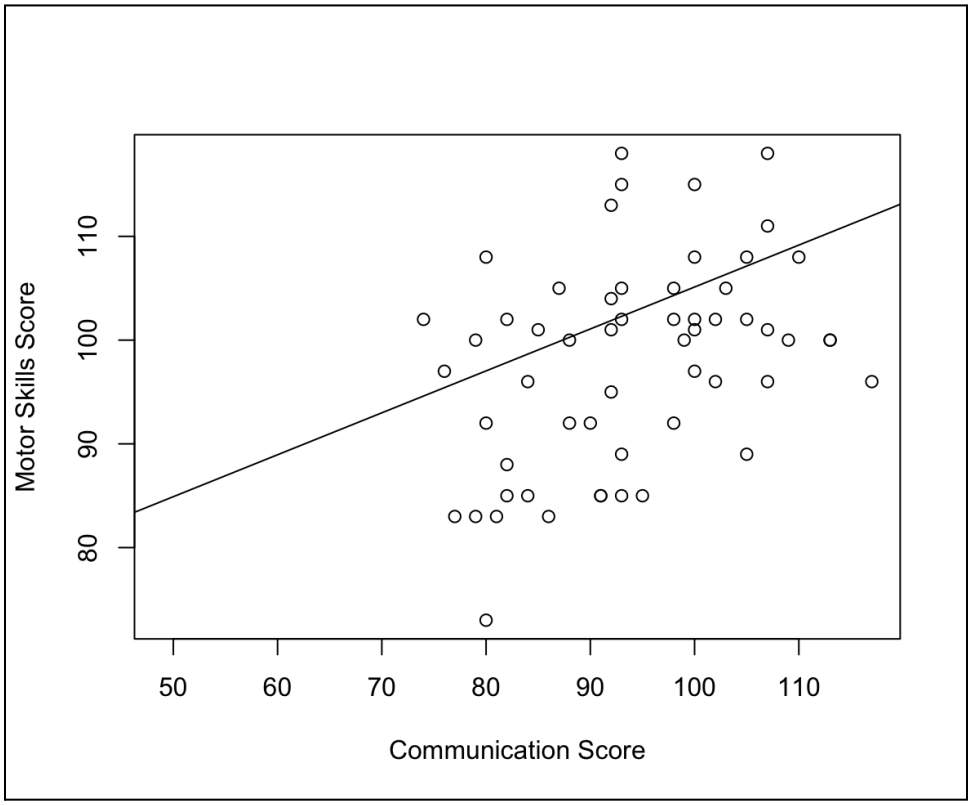

Figure 13. Linear regression of the Communication Score and the Motor Skills Score for cases (n=57; age range: 0-9 years; 22 male, 32 female, 3 other/unknown); p-value=0.002221

A statistically significant linear relationship (Table 11B) was found between the Adaptive Behavior Composite Score and Communication Score (p-value < 2.2e-16; Figure 14), Adaptive Behavior Composite Score and Motor Skills (p-value=1.479e-05; Figure 15), and Communication Score and Motor Skills (p-value=0.006427; Figure 16) domains for the controls (p-values < α=0.01667). (For full analysis of Vineland-3 subdomain scores of the controls, see Supplemental Table 2).

Table 11B. Adjusted R-squared values, p-values, and residual standard errors (RSE) for the control comparisons of the selected domains of the Vineland-3 Test. **41 observations deleted due to missingness; the degrees of freedom were therefore adjusted.*

Domain Comparisons	Sample Size (n)	Adjusted R-Squared	p-value	RSE (degrees of freedom)
Adaptive Behavior Composite Score vs Communication Score	76	0.6644	< 2.2e-16	6.201 (74)
Adaptive Behavior Composite Score vs Motor Skills Score	35	0.4213	1.479e-05	7.92 (33)*
Communication Score vs Motor Skills Score	35	0.1801	0.006427	9.428 (33)*

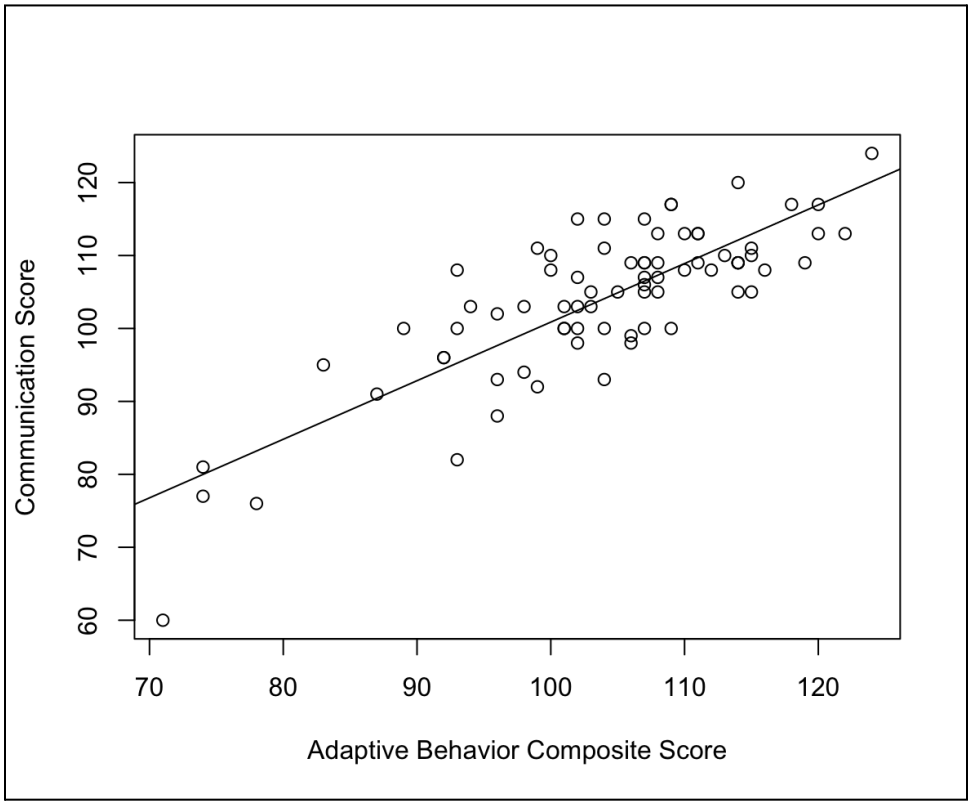

Figure 14. Linear regression of the Adaptive Behavior Composite Score and the Communication Score for controls (n=76; age range: 0-71 years (median age: 11.1 years); 34 male, 42 female); p-value = < 2.2e-16

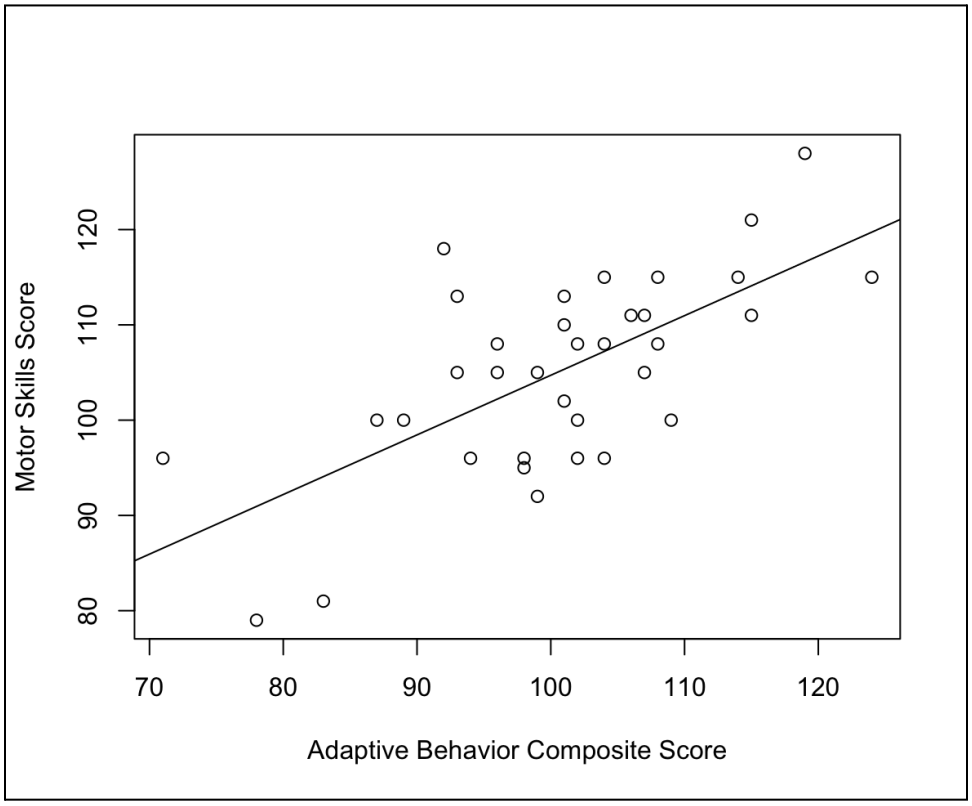

Figure 15. Linear regression of the Adaptive Behavior Composite Score and the Motor Skills Score for controls (n=35; age range: 0-71 years (median age: 13.6 years); 17 male, 18 female); p=1.479e-05

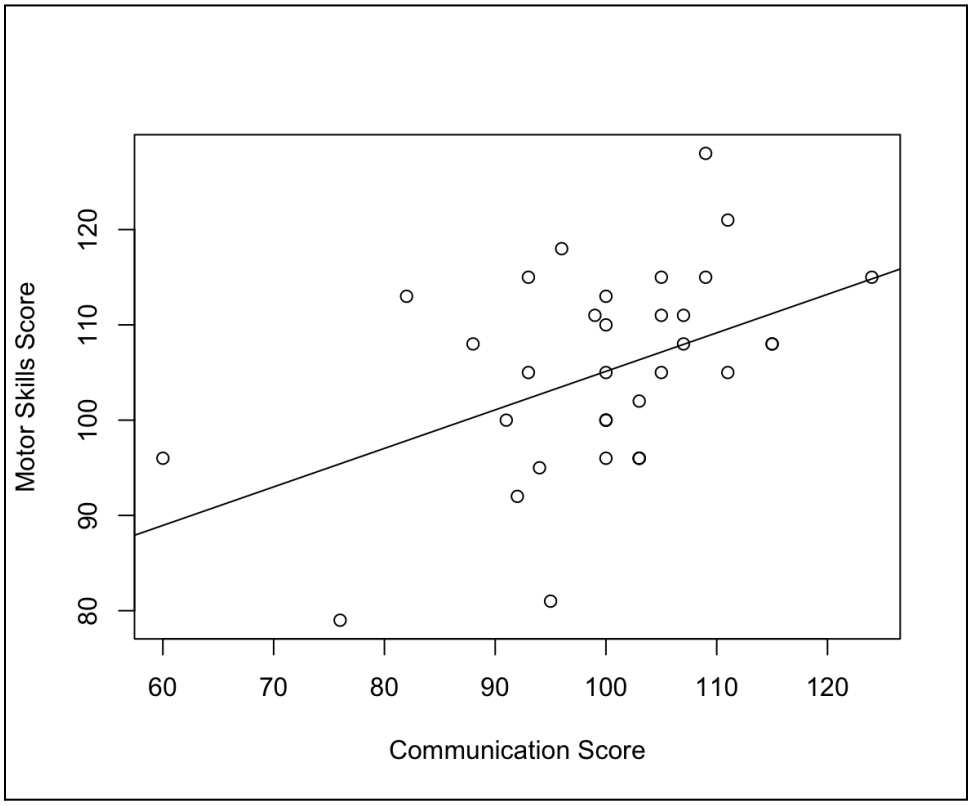

Figure 16. Linear regression of the Communication Score and the Motor Skills Score for controls (n=35; age range: 0-71 years (median age: 13.6 years); 17 male, 18 female); p-value=0.006427

CHAPTER 4: Discussion and Conclusion

The summary score data analyzed in this study demonstrate a strong clustering effect among the cognitive, speech, and motor/neurological outcome domains for both CG cases and controls. Vineland Test data similarly demonstrate a clustering effect among the outcome domains. Thus, it is more likely that if a participant diagnosed with CG has an adverse outcome in either the cognitive, speech, or motor/neurological domains, then the participant is also more likely to experience an adverse outcome in one or more of the other domains.

Given that the outcome scores have similar associations for both the case and control groups, the clustering we observed may not be specific to classic galactosemia. Rather, the similar clustering effects present in the case and control groups indicate that the factors that lead to the different outcomes may be shared within both groups. Of course, despite the similar clustering effects, the cases experienced much more prevalent and severe outcomes overall. Thus, there seems to be an additional factor that impacts all the domains and dictates the adverse outcomes for CG cases.

Although effective dietary treatment can prevent acute symptoms of disease, there are currently no interventions known that can improve the adverse long-term outcomes in CG. Further, there are few meaningful prognostic markers in CG. Thus, it is essential to study the interrelationships between the long-term outcomes in the cognitive, speech, and motor/neurological domains, to identify whether shared or independent risk factors mediate the various adverse long-term outcomes and to help families know that if their child with CG experiences an adverse outcome in on domain, they are also at increased risk to experience

problems in other domains. Our data presented here are consistent with the conclusion that the risk factors are shared.

Why is this important?

The data presented here support the conclusion that adverse long-term outcomes in CG cluster as a result of one or more shared modifiers. Determining clustering among CG adverse outcomes is also important as it helps families know whether their child with CG is at an increased risk of experiencing more than one adverse long-term outcome. Further, in planning for eventual clinical trials of novel interventions for CG, knowing that outcomes tend to cluster also suggests that study participants who may only demonstrate one negative outcome at the time of enrollment are also at increased risk for experiencing more adverse outcomes over time. In short, knowledge of clustering will help families and health care providers make more informed risk-benefit decisions.

Limitations & Future Directions

Although we made every effort to remain unbiased when scoring the data, individual bias is always a risk when scoring each participant's qualitative survey and formal data. Furthermore, the scores in some cases may reflect incomplete or potentially inaccurate information shared by a parent or caregiver in surveys on behalf of the participant. Thus, using similar methods in different study cohorts while simultaneously demonstrating similar outcomes is essential to confirming the external validity of this study.

Limitations and future directions: One of the most significant limitations to this study is that the summary score data are largely dependent on survey data and are not yet completely substantiated by medical records. However, to address this issue in the future, we plan to increase the internal validity of the study by determining the fraction of each outcome domain scores upheld by formal data such as medical records, scholastic assessments, and speech pathology data. We then plan to run the same statistical tests that were conducted in this study on the participant subset that has medical record support. We would thus be able to determine whether there are significant differences in results that are either dependent on the survey or record support.

In addition to the cognitive, speech, and motor/neurological outcome domains, additional long-term outcomes can be studied and tested for significant clustering effects. Finally, a continuation of this project could also include separating statistical tests for clustering by age group.

APPENDIX

Supplemental Table 1. Adjusted R-squared values, p-values, and residual standard errors for the case comparisons of the selected subdomains of the Vineland-3 Test **22 observations deleted due to missingness; the degrees of freedom were therefore adjusted.*

Subdomain Comparisons	Adjusted R-Squared	p-value	Residual Standard Error (degrees of freedom)
Receptive vs Expressive	0.9253	< 2.2e-16	7.779 (118)
Receptive vs Written	0.2896	6.616 e-09	18.02 (96)*
Expressive vs Written	0.2948	4.601e-09	17.96 (96)*
Receptive vs Gross Motor	0.8284	< 2.2e-16	9.398 (118)
Receptive vs Fine Motor	0.8168	< 2.2e-16	8.134 (118)
Expressive vs Gross Motor	0.8607	< 2.2e-16	8.469 (118)
Expressive vs Fine Motor	0.8821	< 2.2e-16	6.527 (118)
Gross Motor vs Fine Motor	0.8634	< 2.2e-16	7.024 (118)
Written vs Fine Motor	0.3714	1.664e-11	6.882 (96)*
Written vs Gross Motor	0.0564	0.01058	8.226 (96)*

Supplemental Table 2. Adjusted R-squared values, p-values, and residual standard errors for the control comparisons of the selected subdomains of the Vineland Test. **9 observations deleted due to missingness; the degrees of freedom were therefore adjusted.*

Subdomain Comparisons	Adjusted R-Squared	p-value	Residual Standard Error (degrees of freedom)
Receptive vs Expressive	0.9351	< 2.2e-16	5.904 (76)
Receptive vs Written	0.6565	< 2.2e-16	13.9 (67)*
Expressive vs Written	0.5849	1.229e-14	15.28 (67)*
Receptive vs Gross Motor	0.8799	< 2.2e-16	5.823 (76)
Receptive vs Fine Motor	0.8586	< 2.2e-16	6.159 (76)
Expressive vs Gross Motor	0.7991	< 2.2e-16	7.531 (76)
Expressive vs Fine Motor	0.8669	< 2.2e-16	5.975 (76)
Gross Motor vs Fine Motor	0.713	< 2.2e-16	8.775 (76)
Written vs Fine Motor	0.5781	2.135e-14	5.168 (67)*
Written vs Gross Motor	0.7057	< 2.2e-16	1.589 (67)*

References

Abi-Hanna, A., & Saavedra, J. M. (1998). GALACTOSE. In *Encyclopedia of Human Nutrition* (pp. 377–383). Elsevier.

Antshel, K. M., Epstein, I. O., & Waisbren, S. E. (2004). Cognitive strengths and weaknesses in children and adolescents homozygous for the galactosemia Q188R mutation: a descriptive study. *Neuropsychology*, *18*(4), 658–664. https://doi.org/10.1037/0894-4105.18.4.658

Bech, L. F., Donneborg, M. L., Lund, A. M., & Ebbesen, F. (2018). Extreme neonatal hyperbilirubinemia, acute bilirubin encephalopathy, and kernicterus spectrum disorder in children with galactosemia. *Pediatric Research*, *84*(2), 228–232. https://doi.org/10.1038/s41390-018-0066-0

Berry, G. T. (2021). Classic Galactosemia and Clinical Variant Galactosemia. In *GeneReviews® [Internet]*. University of Washington, Seattle.

Bosch, A. M. (2006). Classical galactosaemia revisited. *Journal of Inherited Metabolic Disease*, *29*(4), 516–525. https://doi.org/10.1007/s10545-006-0382-0

Centers for Disease Control and Prevention (CDC). (2012). CDC Grand Rounds: Newborn screening and improved outcomes. *MMWR. Morbidity and Mortality Weekly Report*, *61*(21), 390–393. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6121a2.htm

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

Frederick, A. B., Cutler, D. J., & Fridovich-Keil, J. L. (2017). Rigor of non-dairy galactose restriction in early childhood, measured by retrospective survey, does not associate with severity of five long-term outcomes quantified in 231 children and adults with classic galactosemia. *Journal of Inherited Metabolic Disease*, *40*(6), 813–821. https://doi.org/10.1007/s10545-017-0067-x

Frederick, A. B., Zinsli, A. M., Carlock, G., Conneely, K., & Fridovich-Keil, J. L. (2018). Presentation, progression, and predictors of ovarian insufficiency in classic galactosemia. *Journal of Inherited Metabolic Disease*, *41*(5), 785–790. https://doi.org/10.1007/s10545-018-0177-0

Fridovich-Keil, J. L., & Berry, G. T. (2022). Pathophysiology of long-term complications in classic galactosemia: What we do and do not know. *Molecular Genetics and Metabolism*, *137*(1–2), 33–39. https://doi.org/10.1016/j.ymgme.2022.07.005

Gitzelmann, R. (1980). Newborn screening for inherited disorders of galactose metabolism. In *Neonatal Screening for Inborn Errors of Metabolism* (pp. 67–79). Springer Berlin Heidelberg.

Hagen-Lillevik, S., Rushing, J. S., Appiah, L., Longo, N., Andrews, A., Lai, K., & Johnson, J. (2021). Pathophysiology and management of classic galactosemic primary ovarian insufficiency. *Reproduction & Fertility*, *2*(3), R67–R84. https://doi.org/10.1530/RAF-21-0014

Hoffmann, B., Wendel, U., & Schweitzer-Krantz, S. (2011). Cross-sectional analysis of speech and cognitive performance in 32 patients with classic galactosemia. *Journal of Inherited Metabolic Disease*, *34*(2), 421–427. https://doi.org/10.1007/s10545-011-9297-5

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. *European Journal of Pediatrics*, *154*(7 Suppl 2), S2-5. https://doi.org/10.1007/bf02143794

Los, E., & Ford, G. A. (2022). *Galactose 1 Phosphate Uridyltransferase Deficiency*. StatPearls Publishing.

McCorvie, T. J., Kopec, J., Pey, A. L., Fitzpatrick, F., Patel, D., Chalk, R., Shrestha, L., & Yue, W. W. (2016). Molecular basis of classic galactosemia from the structure of human galactose 1-phosphate uridylyltransferase. *Human Molecular Genetics*, *25*(11), 2234–2244. https://doi.org/10.1093/hmg/ddw091

Peter, B., Davis, J., Finestack, L., Stoel-Gammon, C., VanDam, M., Bruce, L., Kim, Y., Eng, L., Cotter, S., Landis, E., Beames, S., Scherer, N., Knerr, I., Williams, D., Schrock, C., & Potter, N. (2022). Translating principles of precision medicine into speech-language pathology: Clinical trial of a proactive speech and language intervention for infants with classic galactosemia. *HGG Advances*, *3*(3), 100119. https://doi.org/10.1016/j.xhgg.2022.100119

Potter, N. L., Lazarus, J.-A. C., Johnson, J. M., Steiner, R. D., & Shriberg, L. D. (2008). Correlates of language impairment in children with galactosaemia. *Journal of Inherited Metabolic Disease*, *31*(4), 524–532. https://doi.org/10.1007/s10545-008-0877-y

Potter, Nancy L., Nievergelt, Y., & Shriberg, L. D. (2013). Motor and speech disorders in classic galactosemia. *JIMD Reports*, *11*, 31–41. https://doi.org/10.1007/8904_2013_219

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., Häberle, 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 Journal of Rare Diseases*, *14*(1), 86. https://doi.org/10.1186/s13023-019-1047-z

Ryan, E. L., Lynch, M. E., Taddeo, E., Gleason, T. J., Epstein, M. P., & Fridovich-Keil, J. L. (2013). Cryptic residual GALT activity is a potential modifier of scholastic outcome in school age children with classic galactosemia. *Journal of Inherited Metabolic Disease*, *36*(6), 1049–1061. https://doi.org/10.1007/s10545-012-9575-x

Schweitzer, S., Shin, Y., Jakobs, C., & Brodehl, J. (1993). Long-term outcome in 134 patients with galactosaemia. *European Journal of Pediatrics*, *152*(1), 36–43. https://doi.org/10.1007/bf02072514

Sparrow, S. S., Cicchetti, D. V., Saulnier, C. A., & Smith, K. C. (n.d.). *Vineland adaptive behavior scales, third edition (vinelandTM-3) domain-level teacher form report*. Pearsonclinical.com. Retrieved March 22, 2023, from http://images.pearsonclinical.com/images/Assets/vineland-3/Vineland-3Domain-LevelTeacherFormSampleReport.pdf

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. *The Journal of Clinical Endocrinology and Metabolism*, *98*(7), E1257-65. https://doi.org/10.1210/jc.2013-1374

Timson, D. J. (2016). The molecular basis of galactosemia - Past, present and future. *Gene*, *589*(2), 133–141. https://doi.org/10.1016/j.gene.2015.06.077

US, P. A. [@PearsonAssessmentsUS]. (2018, August 8). *Vineland-3 Overview*. Youtube. https://www.youtube.com/watch?v=mvoT_D4A9nc

Waggoner, D. D., Buist, N. R., & Donnell, G. N. (1990). Long-term prognosis in galactosaemia: results of a survey of 350 cases. *Journal of Inherited Metabolic Disease*, *13*(6), 802–818. https://doi.org/10.1007/bf01800204

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. *Journal of Inherited Metabolic Disease*, *35*(2), 279–286. https://doi.org/10.1007/s10545-011-9372-y

Welling, L., Bernstein, L. E., Berry, G. T., Burlina, A. B., Eyskens, F., Gautschi, M., Grünewald, S., Gubbels, C. S., Knerr, I., Labrune, P., van der Lee, J. H., MacDonald, A., Murphy, E., Portnoi, P. A., Õunap, K., Potter, N. L., Rubio-Gozalbo, M. E., Spencer, J. B., Timmers, I., … Galactosemia Network (GalNet). (2017). International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. *Journal of Inherited Metabolic Disease*, *40*(2), 171–176. https://doi.org/10.1007/s10545-016-9990-5