DYSMORPHOLOGY, ABNORMAL GROWTH AND COPY NUMBER VARIANTS IN AUTISM SPECTRUM DISORDER

by

Norazlin Kamal Nor

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

December 2017

Abstract

Background and Purpose: Autism Spectrum Disorders (ASD) is a complex persistent neurodevelopmental disorder characterized by impairment in social interaction, communication difficulties and repetitive or stereotypic behaviors. In the past four decades, the prevalence of ASD has increased dramatically. The risk factors associated with ASD include genetic, environmental and possibly gene-environment interactions. Although the core features of ASD are well characterized, ASD presents heterogeneously with a wide spectrum of manifestations. These overlapping features or phenotypes are co-morbidities that occur in ASD and span developmental, medical, behavioral and psychiatric conditions. These co-morbid features can include dysmorphology and growth abnormalities. It is postulated that children with significant dysmorphology are more likely to have an underlying genetic etiology and may have a higher load of genes controlling risk to ASD, including a higher copy number variant (CNV) burden and more single location variants. CNVs are alterations of the DNA resulting in structural variants, including deletion and duplication of genome sequence. ASD sub-phenotypes (such as dysmorphology) offer the potential of determining distinct genetic etiologies and enhancing genotype-phenotype correlations in ASD. In this study, we first characterized growth abnormalities in children with ASD in the Study to Explore Early Development (SEED) study. To investigate genotype-phenotype associations in ASD, we determined the association between genome-wide CNV burden with dysmorphology and abnormal growth in SEED children and tested for association between ASD-associated CNVs with dysmorphology and growth abnormalities in children in the SEED study.

Methods: The study population was drawn from the SEED Study, which was developed to identify risk factors for ASD in the prenatal and early post-natal period. To characterize abnormal growth patterns associated with ASD, we assessed growth abnormalities for all available anthropometric measures of growth (height, weight and head circumference), the bi-dimensional measure of body mass index (BMI), and a tri-dimensional growth measure of growth phenotype assessing the symmetry of growth involving all three modalities in a single individual. We examined genotype-phenotype associations between genome-wide estimated CNV burden with dysmorphology and abnormal growth. Finally, we investigated the association of specific CNVs reportedly associated with ASD for possible association with dysmorphology and abnormal growth in children in the SEED study.

Results: Assessment of growth abnormalities in SEED 1 study showed females with ASD had short stature and a combination of short stature, microcephaly and normal weight compared to typically developing or control females. We found genome-wide CNV burden was negatively associated with dysmorphology, and CNV burden in recognized ASD genes was negatively associated with tall stature and macrocephaly; these associations varied by sex. Investigation of association between CNVs associated with ASD and abnormal growth revealed a potential shared genetic risk for ASD and short stature at CNV region 15p11.2, which was significantly associated only in females.

Conclusions: These findings demonstrate the importance of taking into consideration the influence of sex in characterizing any association between growth abnormalities and/or dysmorphology and risk of ASD, as well as variability of reported genetic risk factors in ASD. There is potential shared genetic risk for ASD and growth abnormality that differs by sex, and this may lead to potential future clinical application in diagnosing of ASD that could be tailored to the child's needs.

Future directions: These results should be replicated in a different population, while expanding measurments of growth assessment to incorporate longitudinal change to better characterize growth abnormalities in ASD. Using a larger sample size and with parental genotyping information would enable CNV burden for *de novo* and rare CNVs to be considered. Whole exome sequencing would be a useful in excluding chromosomal abnormalities and non-chromosomal genetic syndromes when considering CNV burden associations.

Thesis Readers:

Margaret Daniele Fallin, Ph.D. (Co-Advisor)

Terri H. Beaty, Ph.D. (Co-Advisor)

Christine Ladd-Acosta, Ph.D.

Julie Hoover-Fong, M.D., Ph.D.

Brion Maher, Ph.D.

Abbreviations

ADOS – Autism Diagnostic Observation Schedule
ADI-R – Autism Diagnostic Interview- Revised
ASD – Autism Spectrum Disorder
BMI – Body Mass Index
CAGS – Chromosomal Abnormalities and Non-chromosomal Genetic Syndromes
CGI – Caregiver's Interview
CNV – Copy Number Variants
CNV – Copy Number Variant
DD – Children with Developmental Delay
DRF – Dysmorphology Review Form
DSM – Diagnostic and Statistical Manual of Mental Disorders
HC – Head Circumference
ID – Intellectual disability
MCA – Multiple Congenital Anomalies
POP – Typically developing children in the SEED Study
SCQ – Social Communication Questionnaire
SEED – Study to Explore Early Development
SFARI – Simons Foundation Autism Research Initiative
TGP – Trivariate Growth Phenotype

ACKNOWLEDGEMENTS

In the name of God, Most Gracious and Most Merciful, firstly I give my deepest gratitude to the Almighty, without whom none of this would be possible. I also give my greatest thanks to my wonderful parents Kamal Nor Mohd Yusof and Latifah Norma Mohamed, whose love and unwavering support carries me through the darkest hours. The one person who has helped me the most in giving encouragement, support and unwavering belief is my beloved son, Majid Abouzid, who has cajoled, rallied, bullied and cheered me in equal measure. I also give thanks to my husband, whose kindness and help has made the task of trying to juggle academic pursuit, work and family life lighter. I would like to thank my lovely siblings, Farah and Norazlan, and their families, for their love and for always being there. I would also like to thank my sponsors, the Malaysian Government Higher Education Ministry, and my university in Malaysia, the National University of Malaysia, for their financial support throughout these doctoral years. I need to say a special thanks to my department, the Paediatrics Department at the Faculty of Medicine at the National University, for supporting my academic endeavours.

I am grateful for friends and extended family across continents from Malaysia to the United States. In this doctoral journey, a great many people have made an impression on me, and have been integral to any success I hope to achieve. Many of these people I met throughout the doctoral program at the Johns Hopkins Bloomberg School of Public Health. While it may not be possible to adequately thank and even name everyone who has provided support and encouragement throughout this program, I would like to highlight certain individuals to whom I owe immeasurable debt. Firstly, I would like to

give my sincerest thanks to Dr. Dani Fallin, my advisor and mentor, who has, from the beginning of my journey at Hopkins, been a source of inspiration and admiration. She has taught with wisdom and kindness, with fairness and integrity, and if there is one thing I hope to take away from this experience, it is her lesson of reaching for the highest of scientific standards while developing and maintaining a supportive and collaborative research network. In addition, Dr. Terri Beaty has been an invaluable source of support, pragmatic wisdom and endless patience, and has made me feel able to straighten my spine and get through another day even when the challenges were daunting. In addition, I have learnt a lot of valuable lessons and tools from Dr. Julie Hoover-Fong, who is an inspiring advisor always ready to give her encouragement and support. I am also very grateful to have Dr. Chris Ladd-Acosta as a committee member and frequent advisor, and she has been unwavering in her help, kindness and willingness to guide. I also would like to thank the staff at both the Epidemiology and Mental Health departments at the Johns Hopkins School of Public Health, as well as the Institute of Genetic Medicine at the School of Medicine. Fran Burman and Matt Miller from the Epidemiology department are always ever helpful and supportive, and Michelle Maffett from the Mental Health department has been extremely kind and efficient in helping organize scheduling. I am deeply grateful to the SEED Study for the privilege of working on SEED data, and to participants and their families for their contribution to the advancement of research in the field of autism. I would like to extend my thanks also to the SEED Dysmorphology group and Dr. Stuart Shapira, for helping me better understand the dysmorphology study within SEED and for the use of dysmorphology data generated by the group. I would also like to

thank Brooke Sheppard, who has since graduated from the school, for the painstaking groundwork she undertook that later became part of the foundation for my own work.

I have met an amazing group of people at Hopkins, who have helped me immeasurably not only to survive the doctoral program, by giving invaluable advise and support, but also in making the day-to-day worthwhile, with their friendship, encouragement and kindness. In the Genetic Epidemiology track, I would like to thank Samantha Bomotti, Weiyan Li, Fei Chen, and Yanzi Xiao. I am very grateful for how incredibly helpful and supportive they have been throughout this process. Sam and Weiyan especially, have spent countless hours helping and commiserating with me, and having them in my life has made a tremendous difference. I also want to thank the rest of the Genetic Epidemiology cohort, students and staff alike. In addition to the Genetic Epidemiology track, I was lucky enough to have developed close friendships with my other fellow doctoral students. I have to give special thanks to two people especially, Nicole Armstrong and Sheriza Baksh, who have been my comrade-in-arms from the beginning, and without whom I doubt I could have made it through. Both have been invaluable friends and allies, necessary critiques and an immense source of emotional support. Their friendship and support has made incredibly challenging situations bearable.

Lastly, I want to give tribute to my youngest sister Fazrin, who is my inspiration and who I deeply love. I dedicate this Ph.D. to her, and to my son Majid. They both inspire me to be the best that I can be, and to dedicate my life to making things better.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The purpose of this thesis is to explore certain aspects related to the biological underpinnings of autism spectrum disorder (ASD) by examining alternative phenotypes related to ASD based on dysmorphological features, which may better inform etiologic discovery, prognosis, or early identification of this complex and heterogeneous neurodevelopmental disorder. In this work, we investigated overall dysmorphology and abnormal physical growth specifically as an ASD sub-phenotypes or subgroup. We examined the association between these phenotypes and ASD, and also examined the genetic risk for these phenotypes using estimated copy number variant (CNV), and their tested for their potential overlap with other genetic findings. Below, we describe what is currently known about ASD prevalence, potential causes, and outcomes, and how this work adds to knowledge and opportunities for the field.

BACKGROUND

Prevalence of Autism Spectrum Disorder

ASD is a complex neurodevelopmental disorder characterized by impairment in social interaction, communication difficulties and delays and repetitive or stereotypic behaviors. Autism is diagnosed clinically, with the American Psychiatric Association's Diagnostic and Statistical Manual (DSM). In 2013, the field moved from using DSM-IV-R, which considered autism spectrum disorders as an umbrella grouping of several specific disorders including Asperger's syndrome, Autistic disorder, and Pervasive Development Disorder – Not otherwise specified, to the DSM-5 criteria, which designated a single

diagnostic group, “Autism Spectrum Disorder”. The DSM-5 diagnostic criteria for ASD include abnormalities in two main categories: firstly, persistent deficits in social communication and social interaction across multiple contexts, and secondly, restricted, repetitive patterns of behaviors, activities or interests. The deficits in the category of communication and social interactions include poor non-verbal communicative behaviors, difficulty in developing and maintaining relationships, as well as a lack of reciprocity of socio-emotional interactions. Restricted and repetitive behaviors include stereotyped and repetitive motor movements, highly restricted, fixated interests with an abnormally high degree of intensity or focus, and unusual responses to sensory input, including both hypo- and hyper-reactivity. The degree of severity in ASD is taken into account for both these main areas of neurodevelopmental impairment, as well as the determination that the difficulties observed cannot be better explained by intellectual disability alone. In addition, these difficulties must be present early in development and impact current functioning in important areas, including socialization. In clinical practice, the deficits as laid out in DSM-5 manifest differently in different individuals, with their varying degrees of severity (American Psychiatric Association, 2013).

Since the earliest epidemiologic studies in the late 1960s and early 1970s, the global prevalence of ASD has been increasing by over thirty-fold in the past five decades. The prevalence estimates from European studies were then one in 2,500 children, or approximately 0.04% (Gillberg & Wing, 1999). The prevalence increased to approximately 1-2% of all children by the turn of the 21st century (Schieve et al., 2012; Blumberg et al., 2013; Wingate et al., 2014; Christensen et al., 2016). The most recent

report from the Centers for Disease Control and Prevention (CDC) surveillance program estimated in 2012, 1 in 68 (1.5%) 8-year-old children in the US had some form of ASD (CDC, 2016). In comparison, the childhood prevalence of intellectual disability is approximately 1-2% and has remained stable over time (Leonard & Wen, 2002). The increasing prevalence of ASD makes understanding, managing and preventing this condition a major public health issue. ASD leads to an enormous emotional burden to the family and a substantial economic burden to the community. In 2011, the total societal cost for autism in the U.S. was $11.5 billion, and it is projected to increase to $300 billion by 2030 (Lavelle et al., 2014).

Etiology of Autism Spectrum Disorder

The exact etiologies and pathways leading to ASD are not well established. Research in autism has revealed associations with various risk factors (both environmental and genetic in origin), and raise the possibility of gene-environment interactions. These risk factors include increasing parental age, male gender, and medical conditions such as neurofibromatosis and tuberous sclerosis (Geschwind, 2011; Hallmayer et al., 2014; Harris, 2012; Risch et al., 2014; Sandin et al., 2014). Various non-inherited risk factors have been implicated with ASD, such as maternal peri-conceptual folic acid intake and SSRI (selective serotonin reuptake inhibitor) use in pregnancy (Schmidt et al., 2012; Suren et al., 2013; Kaplan et al., 2017; Mezzacappa et al., 2017). In addition, Caesarean section delivery, short inter-pregnancy interval, very low birth weight and maternal fever during pregnancy have all been associated with an increase risk of ASD (Atladottir et al.,

2012; Yip et al., 2016; Durkin, DuBois & Maenner, 2015; Joseph et al., 2017). ASD has also been linked to environmental risk factors such as traffic-related air pollution and gestational pesticide exposure to organochlorine compounds (Volk et al., 2013; Shelton et al., 2014). The precise contribution of genes and environment is still debated. Early reports showed a high degree of heritability for ASD, with twin studies showing concordance as high as 0.77 in monozygotic compared to 0.31 in dizygotic male twins (Folstein & Rutter, 1977; Bailey et al, 1995; Rosenberg et al, 2009), and the sibling recurrence risk has been estimated to be between 2-19% (Chakrabarti & Fombonne, 2001; Lauritsen, Pederson & Mortensen, 2005; Ozonoff et al., 2011). In addition, first-degree relatives of individuals with ASD are more likely to display sub-threshold traits of autistic phenotypes, suggesting autistic traits occur within a spectrum and may be strongly correlated with genetic factors (Constantino, 2014; Robinson et al., 2014). However, more recent heritability studies have estimated only 50% heritability (Hallmayer et. al. 2011), although other studies continue to find higher estimates as well (Sandin et al., 2014; Colvert et al., 2015; Lyall et al., 2017).

Autism Spectrum Disorder Phenotypes

ASD is one of the most heterogeneous neurodevelopmental disorders, with a greater diagnostic complexity beyond the impairments described in DSM-5. While the diagnosis of ASD is based on the observation of atypical behaviors, the specific constructs in ASD remain unclear. The two behavioral dimensions specified in DSM-5 represent the core defining features of ASD, with intellectual and language ability adding another

dimension to ASD heterogeneity. In addition, the characterization of ASD sub-groups, defined by the presence of known medical, developmental, neurologic, genetic, or other psychiatric disorders, furthers our understanding of the heterogeneity in ASD (Levy et al., 2010). Some examples of conditions that co-occur with ASD are epilepsy, dysmorphology, gastrointestinal abnormalities and sleep disturbance. The disorder is now widely accepted as a complex, pervasive, heterogeneous condition with multiple etiologies, sub-types, and developmental trajectories. Characterizing these co-occurring issues as ASD sub-groups or sub-phenotypes could improve our understanding of the nuances of autism biology and potentially unveil distinct associations with risk factors. The heterogeneity within ASD is postulated to be partly attributable to genetic risk factors (Miles, 2011). When ASD is classified as one homogenous category, the determination of true causal factors would be diluted, as different sub-groups of ASD may well have differing etiologies. Using ASD phenotypes as outcomes or sub-groups may enable delineation of possible distinct etiologies and therefore enhance power to detect associations with possible etiologic factors, including genes. The over-arching objective of this study is thus to investigate whether consideration of dysmorphology and growth, as ASD sub-phenotypes, can reduce heterogeneity and increase elucidation of ASD risk factors, particularly genetic risk factors.

Genetics of Autism Spectrum Disorder

Given the generally high estimated heritability for ASD, genetic studies in ASD have been a major focus of the field. There are notable autistic features among individuals

with known chromosomal abnormalities and syndromes such as Fragile X, Rett syndrome and Down syndrome (Hall, Lightbody & Reiss, 2008; DiGiuseppi et al., 2010). In about 10% of ASD cases, genetic, neurologic and metabolic conditions have been identified as either leading to, or associated with, ASD-like characteristics (Kielinen et al., 2004; Cohen et al., 2005; Bolton, 2009).

More recently, studies investigating genetic risk factors in ASD have found associations with both inherited and *de novo* mutations (Geschwind, 2011; Robinson et al., 2014). Inherited genetic risk factors include both common and rare variants, and may involve single base pair changes, small insertions or deletions, or copy number variants (CNVs). Common single base pair variants are called single nucleotide polymorphisms (SNPs), while rare variants are simply termed "single nucleotide variants" (SNVs). CNVs are "a segment of DNA $\geq$ 1 kilobase in size that differs in copy number compared with a representative reference genome". Alterations in CNVs resulting in structural variants, including deletion and duplication of genome sequence (DiGiuseppi et al., 2016). Some studies have attributed at least 20% of ASD liability to common SNP variants (Robinson et al., 2014; Gaugler et al., 2014), although few particular SNPs have been convincingly implicated as controlling risk, likely due to underpowered genome wide association studies in ASD (Pinto et al., 2010; Gaugler et al., 2014; Iossifov et al. 2014; Bralten et al., 2017). The majority of findings in ASD genetics have been for rare variants, including CNVs (Prasad et al., 2012; Griswold et al., 2012; Malhotra & Sebat, 2012; Leppa et al., 2016). Many of these CNVs are *de novo,* meaning they occurred spontaneously in an individual, rather than being inherited from his/her parents. These are by definition rare

individually, but several studies have found numerous *de novo* rare variants across at least 400 different genes that may represent approximately 30% of genetic liability to ASD (Iossifov et al., 2014). The majority of specific genetic findings now thought to contribute significantly to ASD are variants that are rare, *de novo* and likely to disrupt normal gene function (Pinto et al., 2010; Grayton et al., 2012; Shishido, Aleksic & Ozaki, 2014; Iossifov et al. 2014). Many common and rare, inherited and *de novo*, variant associations in ASD have involved CNVs, as discussed further below.

Copy Number Variants in Autism Spectrum Disorder

Recent studies have found CNVs are a genetic risk factor of considerable importance in ASD. Studies of CNVs in ASD have shown both single location associations between specific CNVs and ASD, as well as an increased CNV burden by size or count across the entire genome. Pinto et al. in 2010 reported a higher global burden of rare CNVs in individuals with ASD, as well as a higher burden of ASD and intellectual disability (ID) genes and enrichment of CNV deletions in ASD compared to controls (Pinto et al., 2010). Others also see higher CNV burden in terms of CNV size (in kilobase) as well as counts of several CNVs, particularly for deletion CNVs and for rare CNVs (Vulto-van Silfhout et al., 2013). ASD cases have also been reported to have an increased CNV burden over the whole genome for *de novo* CNV deletions and duplications (Luo et al., 2012; Ericksson et al, 2015). Leppa et al. in 2016 found multiplex families (i.e. those with multiple individuals affected with ASD) there is a higher burden of large, rare, inherited mutations, while in simplex families (i.e. those with only one affected child), there is a

higher burden of large, rare, *de novo* mutations (Leppa et al., 2016). Unpublished data from the SEED study – the sample upon which this dissertation is based, shows greater CNV burden by size in children with ASD compared to typically developing controls, with odds of ASD being 37% larger than controls in children whose cumulative CNV length was more than 1 standard deviation above the mean. This effect size was stronger for large CNVs (>400kb), particularly when considering CBVs overlapping with previously implicated as influencing risk to ASD.

In addition to increased genome-wide burden, location-specific CNVs have also been reported to be associated with ASD. The most consistent of these include duplications at 7q11.23, and duplication or deletion at 16p11.2 (Weiss et al., 2008; Merla et al., 2010; Sanders et al., 2011; Green Snyder et al., 2016). Weiss et al. identified 16p11.2 as a novel, recurrent microdeletion and a reciprocal microduplication CNV region that accounted for approximately 1% of all ASD cases (Weiss et al., 2008). ASD affects 15-25% of carriers of such 16p11.2 deletions (Moreno-De-Luca et al., 2015). Complementary sub-phenotypes have been observed in individuals with copy number changes at 16p11.2, which has been associated with ASD, abnormal head size and weight abnormalities. Individuals with 16p11.2 deletions are more likely to have ASD, overweight and macrocephaly, and those with 16p11.2 duplications have increased risk of ASD and schizophrenia, underweight and microcephaly (Shinawi et al., 2010; Jacquemont et al., 2011; Qureshi et al., 2014; Maillard et al., 2014; Stein, 2015). While deletions in16p11.2 has been associated with a shift of IQ and social responsiveness, 16p11.2 duplications have a wider variability in presentation (Moreno-De-Luca et al.,

2015; Green Snyder et al., 2016). Other regions such as chr. 7q11.23 and 1q21.1 have also been implicated with ASD and abnormal head size (Merla et al., 2010; Sanders et al., 2011).

The Simons Foundation Autism Research Initiative (SFARI) maintains the searchable SFARI Gene database, from which all curated genetic associations in ASD can be pulled. SFARI gene designation gives confidence in identifying CNVs associated with ASD at locations including 1q21.1, 3q29, 15q11.2-13 and 22q11.2 regions (SFARI Gene https://gene.sfari.org/autdb/CNVHome.do).

Studies showing association between CNVs and the development and function of cells and neurons suggest potential biological pathways for ASD. Enrichment analyses of CNV findings in ASD have implicated neuronal signaling, cell projection and motility, microtubule cytoskeleton, chromatin remodeling and kinase activity, and neuronal degeneration and regulation (Pinto et al., 2010; 2014). These links give insight into how genetic risk factors may affect and disrupt biological pathways fundamental for normal development. Another important finding from recent studies is the integrated networks found for autism and related disorders (such as intellectual impairment and neuropsychiatric conditions like schizophrenia) (Torres, Barbosa & Maciel, 2015). Recurrent CNVs have been associated with ASD, however they are often not specific to ASD, and also occur in other neurodevelopmental conditions such as intellectual impairment, suggesting a pleiotropic effect for certain causal genes. CNVs may also play some role in gene-environment interaction (Freitag et al., 2010; Mazina et al., 2015).

To further understand ASD genetics, in this study we assess genotype-phenotype associations between overall CNV burden and the ASD sub-phenotypes of dysmorphology and abnormal growth, and consider the association between dysmorphology (and growth abnormalities) with ASD-associated CNVs previously reported in literature.

Dysmorphology and Autism Spectrum Disorder

David W. Smith first proposed the term dysmorphology in 1966, referring to "the study of abnormalities of structural development regardless of severity, timing, or etiology" (Smith, 1966). Today, dysmorphology is understood to be the study of structural defects either genetic or idiopathic in origin that result in the development of physical abnormalities during the fetal or embryogenic stages of development. These include congenital malformations such as dysplastic ears, as well as abnormal anthropometric findings such as macrocephaly (enlarged head circumference) or short stature. Dysmorphic features can be categorized into measurement abnormalities or descriptive traits. Both are features at the extremes of expectation, with observations that are markedly higher or lower compared to age-specific population means for measurement abnormalities (for example macrocephaly), and physical features at the extreme range of variability for descriptive traits (for example clinodactyly) (Zahnleiter et al., 2013). Individually, dysmorphic features occur in approximately ≤4% of the general population (Aase 1990; Merks et al. 2003). Multiple dysmorphic features may indicate abnormal

development. In general, the presence of multiple dysmorphic features rarely occurs without co-existing genetic conditions or teratogenic exposure. Dysmorphology involves recognition and identification of patterns of structural malformations to elucidate etiologies and potential developmental trajectories plus outcomes of specific conditions.

Various studies indicate the proportion of children with autism who have physical signs of some alteration in early development range between 5-30% (Ozgen et al., 2010; Angkutsiri et al., 2011). Children with ASD are more likely to have major congenital anomalies compared to the general population (Wier et al., 2016). Children with ASD and dysmorphic features also have a greater probability of a structural cranial abnormality or a known genetic syndrome (Ozgen et al., 2011). In 2011, Angkutsiri et al. described clinical heterogeneity of physical features in ASD, and reported significantly more children with ASD were classified as dysmorphic compared to typically developing children (Angkutsiri et al., 2011). Various studies have described correlations between morphological abnormalities and ASD (Schendel et al., 2009; Miles et al. 2000; 2005; Ozgen et al., 2011; 2013; Wier et al., 2016;). A child with dysmorphology has a higher likelihood of carrying detectable genetic aberrations. By focusing on this ASD sub-phenotype, it is possible there may be a greater likelihood of finding an association between specific ASD sub-phenotypes and identify underlying genetic mutations.

The clinical diagnosis of dysmorphology is based on qualitative observations, for which there is wide variability amongst clinicians and no specified gold standard. There is no general consensus of a recognized rigorous method to classify dysmorphology for

research purposes. There have been previous studies that developed scales to assess dysmorphism among children with ASD. One of the most prominent recent scales was one pioneered by Miles et al., who developed a dysmorphology scale for use by clinicians without extensive dysmorphology training (Miles et al., 2008). However, it is likely this scale would have a high degree inconsistency due to the varying skill levels of the user. The SEED Dysmorphology Group also found that this algorithm had shortcomings in consistently identifying dysmorphism in children with ASD. To develop a better scale for quantifying these essentially qualitative observations, the SEED Dysmorphology Group led by Dr. Stuart Shapira developed a custom dysmorphology measure to identify and summarize dysmorphic features and classify dysmorphism among SEED participants. Shapira's study found children with ASD have significantly higher prevalence of dysmorphology compared to controls (Shapira et al., 2014). Children with multiple dysmorphic features were also more likely to have an underlying genetic condition or exposures to teratogens affecting normal developmental processes (Christensen et al., 2013). In this dissertation, we extend this work on dysmorphology to examine whether the specific subtype of dysmorphism (or abnormal growth) might be associated with ASD itself, and whether genetic findings for either dysmorphism specifically can further inform ASD genetic investigations.

Copy Number Variants in Dysmorphology

There is a dearth of studies testing for associations between CNVs and dysmorphology. This could be due to the absence of a recognized gold standard for dysmorphology

classification, as well as the perceived lack of utility of dysmorphism as a clinical entity itself. Instead, dysmorphism, and various permutations (indicating dysmorphic features of some type, including congenital malformations) are often used in tandem with more clinically useful or relevant phenotypes such as intellectual disability, autistic features and neurological conditions such as epilepsy. A recent study on CNV-phenotype association using Winter-Baraitser Dysmorphology Database/London Dysmorphology Database reported rare, *de novo* and familial CNVs associated with cranial and forehead abnormalities (Qiao et al., 2014). Most studies in this area have examined associations between CNVs and intellectual disability (ID) with/without structural congenital malformations. A study in 2013 using the De Vries score, a composite measure of ID, growth retardation, ≥2 dysmorphic features and congenital anomalies, found an association between *de novo* and familial CNVs with De Vries score >3 (Vulto-van Silfhout et al., 2013). Cooper et al. in 2011 reported large CNVs (>400kb) were more prevalent in children with severe developmental phenotypes associated with multiple congenital anomalies (Cooper et al., 2011). **Table 1** is a summary of recently published literature on CNV burden and dysmorphology.

Copy Number Variants, Dysmorphology and Autism Spectrum Disorder

Previous research assessing the intersection between CNVs for dysmorphology and ASD have focused on CNV associations among ASD individuals with comorbidities such as dysmorphic features, congenital anomalies and intellectual impairment. Eriksson et al. analyzed a population-based cohort of 162 children with ASD, and reported rare CNVs

were detected in 8.6%, with a higher likelihood found in children who also had dysmorphic features or congenital malformations (Ericksson et al., 2015). Dysmorphic features have also been identified in patients with ASD in targeted interrogation of CNVs of specific candidate genes (Nava et al., 2014). Al-Mamari found clinically significant CNVs were detected in 27% of individuals with ASD using chromosomal microarray analysis of 100 ASD patients from a highly consanguineous population, and that patients with dysmorphic features and congenital anomalies were statistically more likely to carry CNVs (Al-Mamari et al., 2015).

The association between dysmorphism and ASD could potentially be developed into a diagnostic clinical tool for early intervention through laboratory-based and clinical methods (utilizing a reliable dysmorphism algorithm or measure). Understanding dysmorphic phenotypes may lead to more focused and potentially earlier provision of intervention in sub-groups of ASD children with abnormal growth, which could lead to better outcomes (Dawson, 2008; Boyd et al., 2010).

Growth Abnormalities in Children

Growth in children is influenced by both genetic and environmental factors. The anthropometric growth measurements used in assessing growth in children include height, weight and head circumference. Multiple factors influence these different aspects of growth, with different effects at different stages of growth. As an example, height is a complex phenotype with multiple genetic factors influencing it. In addition, there are bi-

dimensional aspects of growth based on two growth measures, the most frequently used is the Body Mass Index (BMI). Finally, the three types of growth (height, weight, head circumference) may be combined to check for growth symmetry, whereby any deviation from symmetry would indicate disproportionate growth.

For growth measurements to be interpretable across children, specific growth measures are typically plotted on age and sex standardized growth percentile curves based on large, population based samples. The CDC growth chart, which has separate charts for boys and girls, and accommodations for prematurity, is a commonly used reference (Kuczmarski, Ogden & Guo, 2002).

Population-based values for each growth modality typically exhibit a Gaussian distribution. Extremes of these distributions are declared "growth abnormalities", although different cut-offs are used for different measures of growth. Thresholds for extremes of growth are often set at growth percentiles or based on extremes of standard deviations from the mean. For example, abnormally large head circumference, termed macrocephaly, is often recognized clinically as head circumference greater or equal to the 97^{th} percentile for sex and age, and abnormally small head circumference (microcephaly) is less than the 3^{rd} percentile for sex and age. For height, abnormally tall stature is defined by height greater or equal to 3 standard deviations above the mean based on percentiles for sex and age, and abnormally short stature as less than the 10^{th} percentile for sex and age. For weight, overweight is defined by weight greater or equal to the 97^{th} percentile for sex and age, and underweight as weight less than the 10^{th} percentile for sex and age.

The somewhat arbitrary nature of these cut-offs makes it challenging to compare growth measurements to each other, and to use the measures for association analyses in genetic studies. One way to standardize assessment of these extremes of growth measures is to set a uniform threshold across growth modalities, for example by using the top 10^{th} and bottom 10^{th} percentile (decile) of the growth modality. This may be more useful for comparing different aspects of growth, as well as assessing the composite of all growth modalities. In this dissertation, we explored results based on definitions of growth abnormalities using both a clinically-derived threshold of extremes, as well as a simple decile definition.

Growth Abnormalities and Autism Spectrum Disorder

Children with ASD have been recognized to have abnormalities in several parameters of growth. These abnormalities include accelerated overgrowth of the head in the first year of life, leading to macrocephaly in some younger children with ASD (Courchesne, Campbell & Solso, 2011). Leo Kanner was the first to make observe some autistic children had macrocephaly (Kanner, 1943). The distribution of head circumference in ASD was quite wide, and ASD children have larger head circumference relative to height by the age of 9-10 years (Lainhart et al., 2016). Post-mortem studies of autistic brains have shown increased gray and white matter volume, in addition to enlarged head circumference.

In addition to abnormal head growth, children with ASD have also been reported to have both abnormal weight and height. Curtin et al. (2005) reported adolescents with ASD have an increased prevalence of being overweight (Curtin et al., 2005).

In 2007, van Daalen et al. postulated ASD is associated with a general growth dysregulation, with increased rate of macrocephaly (11.3%) in the first year of life and accelerated growth of body length, or height (van Daalen et al., 2007). A study of physical growth in 429 children with autism in China (Xiong et. al., 2009) reported ASD children had above-average height, weight and BMI, with 17% of those aged between 2 to 5 years being overweight, rising to 21.8% among 6-11 year-olds (Xiong et al., 2009).

A measure of growth in ASD where there is a paucity of data is the simultaneous consideration of the all three growth measures: head circumference, height, and weight, which we term 'trivariate growth phenotype'. Consideration of trivariate growth allows assessment of symmetry, whether children are small on all three measures, normal on all three, or large on all three. Deviations from symmetry, where one or more modalities of growth are not in proportion to the other, suggest a potentially pathological process, or an intrinsic cause of growth abnormality, for example endocrinopathies or specific syndromes.

Copy Number Variants in Growth Abnormalities

Macrocephaly and Microcephaly: Several studies have assessed ASD, growth abnormalities and potential genetic factors. One example is the association between autism, macrocephaly and the *PTEN* gene, which has a number of genetic mutations found to be associated with ASD with macrocephaly (Conti et al., 2012; Klein et al., 2013). *PTEN* (phosphatase and tensin homolog on chromosome 10q23.31) is a tumor suppressor and has been implicated in tumor syndromes including Cowden syndrome, and is often mutated in neoplasms affecting the central nervous system, the intestines, as well as specific conditions such as small-cell lung cancer and endometrial cancer (Conti et al., 2012). Klein et al. in 2013 confirmed the association of *PTEN* mutations and extreme macrocephaly (>3 s.d.), identified mutations in 22% of patients with ASD, and suggested different phenotypic groups based on patterns of growth, including general overgrowth and disproportionate or relative macrocephaly (Klein et al., 2013). Children with ASD who have concurrent macrocephaly and detected to be carriers of *PTEN* may be screened for malignancies later in life, as this association has been linked to tumor syndromes.

Weight and BMI: The phenomenon of "mirror phenotypes", macrocephaly/microcephaly and high BMI/low BMI, have been associated with differential expression of the 16p11.2 CNV region, which is associated with ASD. As noted in the CNV section above, this region is one of the most commonly reported ASD-associated CNVs, and is estimated to contain CNVs in up to 1% of all ASD cases. Deletions in this chromosomal region have been associated with risk to ASD, as well as obesity and macrocephaly, while reciprocal duplications have been associated with

underweight and microcephaly, as well as schizophrenia (Qureshi et al., 2014; Maillard et al., 2015; Kummer et al., 2015). It has been suggested chromatin modification in the differentially expressed 16p11.2 region (for example in lymphoblastoid cell lines) may be one possible mechanism for this "mirror phenotype" observation; this is an interesting finding as chromatin remodeling is one of the network of clusters of CNVs implicated in ASD (Loviglio et al., 2017).

Tall and Short Stature: Height is a polygenic trait known to be highly heritable, with over 700 common variants identified thus far. A study by Dauber et al. (2011) found short stature is associated with an increased burden of CNVs over the genome for both combined counts of CNVs and their length ($p<0.002$) in low-frequency (<5%) and rare (<1%) CNV deletions. No significant association for tall stature was found (Dauber et al., 2011). In 2013, Zahnleiter et al. showed that rare CNVs are a common cause of short stature, with patients with short stature having significantly larger CNVs statistically and 55% of these CNVs enriched with known syndromes associated with short stature. Somewhere you must provide a review of CNVs. How they are *estimated* from chip data. Why it is necessary to sum counts or size of estimated CNVs. Define your use of 'genome-wide' only considering the autosomes. Define deletions and duplications clearly (n.b. duplications may/may not be perfectly contiguous). Give some sense of prevalence of CNVS, overall and common vs. rare variants. Clearly state you cannot identify de novo CNVs because you don't have genotypes on parents. The stuff on page 8 is not adequate.

OVERVIEW OF DISSERTATION

Study Population: SEED 1

This study will use data from the Study to Explore Early Development (SEED) Phase I, a multisite research collaboration under the auspices of the centers for Autism and Developmental Disabilities Research and Epidemiology (CADDRE) Network and funded by the CDC. The SEED 1 network consisted of study sites in six states: California, Colorado, Georgia, Maryland, North Carolina and Pennsylvania, a data coordinating center (DCC) in Michigan, and a central laboratory and bio-sample repository (CLBR) in Maryland. The SEED Study is an ASD case-control study with population-based ascertainment of cases and controls, with the objective of characterizing ASD cases between 2-5 years of age and identifying risk factors. Children aged 2-5 years were recruited into one of three groups: i) children with ASD (ASD group); ii) children with other developmental disabilities (DD group); and iii) children born in the same birth years and same zip codes from the general population (POP group). There were three criteria for eligibility of children into the study: i) born in the study catchment area during the period between September 1st, 2003 to August 31st 2006, ii) reside in the area at the time of first contact, and iii) live with a knowledgeable caregiver who was able to communicate orally in English or Spanish competently and provide informed consent. The enrolled children also had to be between the ages of 30 and 68 months of age at the completion of the in-person clinical developmental assessment. ASD and DD subjects were recruited from clinical and educational service providers from the study areas, and population-based children were recruited through state vital statistics. Upon screening

and clinical evaluation, children were given a final classification of ASD, non-ASD DD, typically developing population control (POP), or ambiguous phenotype.

Data collection included biosampling (blood, saliva, hair), phenotypic data from caregiver interviews and questionnaires, and in-person developmental assessments and physical examination including dysmorphology measures. For all eligible children, a brief screening interview, the Social Communication Questionnaire (SCQ; Rutter et al. (2003)), was administered to the primary caregiver to identify children who required clinical diagnostic assessment to determine final ASD status. For SEED, a positive screen was defined as an SCQ score ≥11. Final classification was assigned using a SEED-specific research algorithm based on Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview-Revised (ADI-R) and clinical judgment (Wiggins et al., 2015). Regardless of ascertainment source, any eligible children with a previous ASD diagnosis, who were receiving special education services, and who had a positive screen, were assigned to the ASD workflow. This determined which instruments were administered and the type of diagnostic evaluation the child received during the data collection phase. Tools for ASD assessment included the ADOS and ADI-R (Falkmer et al., 2013; Lord et al., 2000). Based on previous diagnosis and SCQ screening, DD and POP children with negative SCQ screens were assigned to the DD or POP workflow, respectively. If a clinician suspects ASD during the clinical evaluation of a child in the DD or POP workflow, the child would be moved into the ASD workflow (Schendel et al., 2012; DiGiuseppe et al., 2016). Blood and saliva biosamples shipped to the SEED biosample repository were used to isolate DNA for genetic and epigenetic studies.

Study Design

Each paper of this dissertation is a cross-sectional analysis embedded within the case-control study design of SEED 1. Although ascertainment was based on ASD status, Aim 1 treats ASD as the independent variable, and three measures of growth abnormalities as dependent outcomes. Aims 2 and 3 focus on dysmorphology and abnormal growth specifically as outcomes, and CNV burden and candidate regions as independent variables. In all three aims, the dependent and independent variables were assessed at one point in time.

Specific Aims

Specific Aim 1:

To estimate the association between ASD and abnormal growth measures in children of preschool age.

H1: Children with ASD have abnormal growth and abnormal measures of growth compared to typically developing children.

Specific Aim 2:

a. To estimate the association between genome-wide (specifically, autosome-wide) CNV burden and dysmorphology among preschool children in the SEED Study.

b. To assess the association between CNVs recognized to be associated with ASD and dysmorphology.

H2.a: Children with dysmorphic features will have greater burden of CNVs across the genome than children without such features.

H2.b: ASD-associated CNVs will show an association with dysmorphology.

Specific Aim 3:

a. To estimate the association between genome-wide (specifically, autosome-wide) CNV burden and growth abnormalities among preschool children in the SEED Study.

b. To assess association between CNVs recognized to be associated with ASD and growth abnormalities.

H3.a: Children with growth abnormalities will have greater burden of CNVs across the genome than children without such features.

H3.b: CNVs recognized to be associated with ASD will show association with growth abnormalities.

The specific aims correlate with the labeled aims on the conceptual framework in Figure 1.

Impact Statement

ASD is a condition of considerable public health significance, with increasing prevalence reported over the past few decades. Comprehensive characterization of ASD phenotypes (such as abnormal growth) is a valuable contribution to improving in our understanding

of the heterogeneity and complexity of ASD. Exploration of genotype-phenotype associations may allow clear elucidation of the link between genetic risk factors and specific ASD phenotypes, and this could improve discovery of potential risk factors for ASD. These findings may open avenues leading towards increasing awareness of ASD sub-phenotypes, improving early detection of these sub-phenotypes and potentially allowing for earlier and timelier intervention.

References

Aase, J. M. (1990). The physical examination in dysmorphology. In *Diagnostic dysmorphology* (pp. 33-42). New York: Plenum Medical Book Co.

Abrahams, B. S., & Geschwind, D. H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nature Reviews. Genetics*, *9*(May), 341–355.

Abrahams, B. S., Arking, D. E., Campbell, D. B., Mefford, H. C., Morrow, E. M., Weiss, L. a, … Packer, A. (2013). SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Molecular Autism*, 4, 36.

Al-Mamari, W., Al-Saegh, A., Al-Kindy, A., Bruwer, Z., Al-Murshedi, F., Al-thihli, K. (2015). Diagnostic yield of chromosomal microarray analysis in a cohort of patients with autism spectrum disorders from a highly consanguineous population. *J. Autism Dev Disord,* 45:2323-2328.

Atladottir, H. O., Henriksen, T. B., Schendel, D. E., & Parner, E. T. (2012). Autism After Infection, Febrile Episodes, and Antibiotic Use During Pregnancy: An Exploratory Study. *Pediatrics*, 130(6), e1447–e1454.

American Psychiatric Association's Diagnostic and Statistical Manual, Fifth Edition (DSM-5) http://www.dsm5.org

Angkutsiri, K., Krakowiak, P., Moghaddam, B., Wardinsky, T., Gardner, J., Kalamkarian, N., Hertz-Picciotto, I., Hansen, R.L. Minor physical anomalies in children with autism spectrum disorders. *Autism*. 15(6):746-60 (2011).

Ashwood, P., Wills, S., & Van de Water, J. (2006). The immune response in autism: a new frontier for autism research. *Journal of Leukocyte Biology*, *80*(July), 1–15.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological Medicine*, *25*, 63–77.

Blumberg, S.J., Bramlett, M.D., Kogan, M.D., et al. (2013). Changes in prevalence of parent-reported autism spectrum disorder in school-aged U.S. children: 2007 to 2011–2012. *Natl Health Stat Rep*. 65:1–11.

Bolton, P. F. (2009). Medical conditions in autism spectrum disorders. *Journal of Neurodevelopmental Disorders*, *1*(2), 102–13.

Boyd, B. A., Odom, S. L., Humphreys, B. P., & Sam, A. M. (2010). Infants and Toddlers With Autism Spectrum Disorder: Early Identification and Early Intervention. *Journal of Early Intervention*. 32:75-98.

Bralten, J., van Hulzen, K. J., Martens, M. B., Galesloot, T. E., Arias Vasquez, A., Kiemeney, L. A., … Poelmans, G. (2017). Autism spectrum disorders and autistic traits share genetics and biology. *Mol Psych*, (March), 1–8.

Centers for Disease Control and Prevention (CDC). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *Surveillance Summaries*. April 1, 2016 / 65(3); 1–23 http://www.cdc.gov/mmwr/volumes/65/ss/ss6503a1.htm

Centers for Disease Control and Prevention (CDC) Growth Percentile Charts. http://www.cdc.gov/growthcharts/charts.htm

Chakrabarti, S., Fombonne , E. (2001). Pervasive Developmental Disorder in Preschool Children, *Journal of American Medical Association, 285*(24), 3093–3099.

Christensen, J., Grønborg, T. K., Sørensen, M. J., Schendel, D., Parner, E. T., Pedersen, L. H., & Vestergaard, M. (2013). Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA : The Journal of the American Medical Association*, *309*, 1696–703.

Christensen, D.L., Baio, J., Braun, K.V., et al. (2016) Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. MMWR Surveill Summ. 65(No. SS-3)(No. SS-3):1–23.

Cohen, D., Pichard, N., Tordjman, S., Baumann, C., Burglen, L., Excoffier, E., … Héron, D. (2005). Specific genetic disorders and autism: Clinical contribution towards their identification. *Journal of Autism and Developmental Disorders*, *35*(1), 103–116.

Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S. R., Woodhouse, E., … Bolton, P. (2015). Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. *JAMA Psychiatry*, 1–9.

Constantino JN. Recurrence Rates in Autism Spectrum Disorders. *JAMA*. 2014;312(11):1154-1155.

Conti, S., Condò, M., Posar, A., Mari, F., Resta, N., Renieri, A., Neri, I., Patrizi, A., Parmeggiani, A. (2012). Phosphatase and tensin homolog (PTEN) gene mutations and autism: literature review and a case report of a patient with Cowden syndrome, autistic disorder, and epilepsy. *J Child Neurol*. 27(3):392-7.

Cooper, G. M., Coe, B. P., Girirajan, S., Rosenfeld, J. a, Vu, T. H., Baker, C., … Eichler, E. E. (2011). A copy number variation morbidity map of developmental delay. *Nature Genetics*, 43(9), 838–846.

Courchesne, E., Campbell, K., & Solso, S. (2011). Brain growth across the life span in autism: Age-specific changes in anatomical pathology. *Brain Research*, *1380*, 138–145.

Curtin, C., Bandini, L.G., Perrin, E.C., Tybor, D.J., Must, A. (2005). Prevalence of overweight in children and adolescents with attention deficit hyperactivity disorder and autism spectrum disorders: a chart review. *BMC Pediatr*. (21) 5:48.

D'Angelo, D., Lebon, S., Chen, Q., Martin-Brevet, S., Snyder, L. G., Hippolyte, L., … Cardiff University Experiences of Children With Copy Number Variants (ECHO) Study, the 16p11.2 European Consortium, and the S. V. in I. P. (Vip) C. (2015). Defining the effect of the 16p11.2 duplication on cognition, behavior, and medical comorbidities. *JAMA Psychiatry*, 10032, 1–11.

Dauber, A., Yu, Y., Turchin, M.C., Chiang, C.W., Meng, Y.A., Demerath, E.W., Patel, S.R., Rich, S.S., Rotter, J.I., Schreiner, P.J., Wilson, J.G., Shen, Y., Wu, B.L., Hirschhorn, J.N. (2011). Genome-wide association of copy-number variation reveals an association between short stature and the presence of low-frequency genomic deletions. *Am J Hum Genet.* 9;89(6):751-9.

Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Development and Psychopathology*, *20*, 775–803.

DiGuiseppi C, Hepburn S, Davis JM, Fidler DJ, Hartway S, Lee NR, Miller L, Ruttenber M, Robinson C. (2010). Screening for autism spectrum disorders in children with Down syndrome: population prevalence and screening test characteristics. *J Dev Behav Pediatr.* 31(3):181-91.

DiGuiseppi, C. G., Daniels, J. L., Fallin, D. M., Rosenberg, S. a., Schieve, L. a., Thomas, K. C., … Schendel, D. E. (2016). Demographic profile of families and children in the Study to Explore Early Development (SEED): Case-control study of autism spectrum disorder. *Disability and Health Journal*. 9(3):544-51.

Durkin, M. S., DuBois, L. A., & Maenner, M. J. (2015). Inter-Pregnancy Intervals and the Risk of Autism Spectrum Disorder: Results of a Population-Based Study. *Journal of Autism and Developmental Disorders*, 45(7), 2056–2066.

Eriksson, M.A., Lieden, A., Westerlund, J., Bremer, A., Wincent, J. Sahlin, E., Gillberg, C., Fernell, E., Anderlid, B-M. (2015). Rare Copy Number Variants are common in young children with autism spectrum disorder. *Acta Paediatrica*. 104: 610-618.

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: a systematic literature review. *European Child & Adolescent Psychiatry*, 22, 329–340.

Freitag, C. M., Staal, W., Klauck, S. M., Duketis, E., & Waltes, R. (2010). Genetics of autistic disorders: review and clinical implications. *European Child & Adolescent Psychiatry*, *19*, 169–178.

Folstein, S., Rutter, M. (1977). Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry,*18(4):297-321.

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., … Buxbaum, J. D. (2014). Most genetic risk for autism resides with common variation. *Nat Genet*, 46(8), 881–885.

Geschwind, D. H. (2011). Genetics of autism spectrum disorders. *Trends in Cognitive Sciences*. doi:10.1016/j.tics.2011.07.003

Gillberg,C., Wing L. (1999). Autism: not an extremely rare disorder. *Acta Psychiatr* Scand 1999;99:399–406.

Girirajan, S., Brkanac, Z., Coe, B. P., Baker, C., Vives, L., Vu, T. H., … Eichler, E. E. (2011). Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genetics*, *7*(11).

Girirajan, S., Rosenfeld, J. A., Coe, B. P., Parikh, S., Friedman, N., Goldstein, A., … Eichler, E. E. (2012). Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *The New England Journal of Medicine*, *367*(14), 1321–31.

Grayton, H.M., Fernandes, C., Rujescu, D., Collier, D.A. (2012). Copy number variations in neurodevelopmental disorders. *Prog Neurobiol*. 99(1):81-91.

Green Snyder, L., D'Angelo, D., Chen, Q., Bernier, R., Goin-Kochel, R. P., Wallace, A. S., … Hanson, E. (2016). Autism Spectrum Disorder, Developmental and Psychiatric Features in 16p11.2 Duplication. *Journal of Autism and Developmental Disorders*, 46(8), 2734–2748.

Griswold, A. J., Ma, D., Cukier, H. N., Nations, L. D., Schmidt, M. a., Chung, R. H., … Pericak-Vance, M. a. (2012). Evaluation of copy number variations reveals novel candidate genes in autism spectrum disorder-associated pathways. *Human Molecular Genetics*, 21(15), 3513–3523.

Hall, S. S., Lightbody, A. a, & Reiss, A. L. (2008). Compulsive, self-injurious, and autistic behavior in children and adolescents with fragile X syndrome. *American Journal of Mental Retardation : AJMR*, *113*(1), 44–53.

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., Miller, J., Fedele, A., Collins, J., Smith, K., Lotspeich, L., Croen, L.A., Ozonoff, S., Lajonchere, C., Grether, J.K., Risch, N. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry* 68(11):1095-102.

Harris, J. C. (2012). Autism risk factors: Moving from epidemiology to translational epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(5), 461–463.

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., … Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, *515*(7526), 216–221.

Jacquemont, S., Reymond, A., Zufferey, F., Harewood, L., Walters, R. G., Kutalik, Z., … Froguel, P. (2011). Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature*, *478*(7367), 97–102.

Joseph, R. M., Korzeniewski, S. J., Allred, E. N., O'Shea, T. M., Heeren, T., Frazier, J. A., … Vogt, K. (2017). Extremely low gestational age and very low birthweight for gestational age are risk factors for autism spectrum disorder in a large cohort study of 10-year-old children born at 23-27 weeks' gestation. *American Journal of Obstetrics and Gynecology*, 216(3), 304.e1-304.e16.

Kanner, L. (1943). Autistic disturbance of affective contact. *Nervous Child*, 2, 217–250.

Kaplan, Y. C., Keskin-Arslan, E., Acar, S., & Sozmen, K. (2017). Maternal SSRI discontinuation, use, psychiatric disorder and the risk of autism in children: A meta-analysis of cohort studies. *British Journal of Clinical Pharmacology*, 1–9.

Kearney, H.M., Thorland, E.C., Brown K.K. (2011). Working Group of the American College of Medical Genetics Laboratory Quality Assistance Committee. American College of Medical genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med*, 13(7): 680-5

Kielinen, M., Rantala, H., Timonen, E., Linna, S.-L., & Moilanen, I. (2004). Associated medical disorders and disabilities in children with autistic disorder: a population-based study. *Autism : The International Journal of Research and Practice*, *8*, 49–60.

Klein, S., Sharifi-Hannauer, P., & Martinez-Agosto, J. a. (2013). Macrocephaly as a clinical indicator of genetic subtypes in autism. *Autism Research*, *6*(January), 51–56.

Kummer, A., Barbosa, I. G., Rodrigues, D. H., Rocha, N. P., Rafael, M. D. S., Pfeilsticker, L., … Teixeira, A. L. (2015). Frequency of overweight and obesity in children and adolescents with autism and attention deficit/hyperactivity disorder. *Revista Paulista de Pediatria : Orgao Oficial Da Sociedade de Pediatria de Sao Paulo*, 34(1), 71–77.

Kuczmarski, R.J., Ogden, C.L., Guo, S.S., et al. (2002). 2000 CDC growth charts for the United States: methods and development. *Vital Health Stat,*11. 246:1

Lainhart, J.E., Bigler, E.D., Bocian, M., Coon, H., Dinh, E., Dawson, G., Deutsch, C.K., Dunn, M., Estes, A., Tager-Flusberg, H., Folstein, S., Hepburn, S., Hyman, S., McMahon, W., Minshew, N., Munson, J., Osann, K., Ozonoff, S., Rodier, P., Rogers, S., Sigman, M., Spence, M.A., Stodgell, C.J., Volkmar, F. (2016). Head circumference and height in autism: a study by the Collaborative Program of Excellence in Autism. *Am J Med Genet A,* 2140(21):2257-74.

Lauritsen, M. B., Pedersen, C. B., & Mortensen, P. B. (2005). Effects of familial risk factors and place of birth on the risk of autism: A nationwide register-based study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*, 963–971.

Lavelle, T.A., Weinstein, M.C., Newhouse, J.P., Munir, K., Kuhlthau, K.A., Prosser, L.A. (2014). Economic Burden of Childhood Autism Spectrum Disorders. *Pediatrics* 133(3):e520-9

Leonard, H., & Wen, X. (2002). The epidemiology of mental retardation: Challenges and opportunities in the new millennium. *Mental Retardation and Developmental Disabilities Research Reviews*, *8*(May), 117–134.

Leppa, V.M., Kravitz, S.N., Martin, C.L., Andrieux, J., Le Caignec, C., Martin-Coignard, D., DyBuncio, C. Sanders, S., Lowe, J., Cantor, R.M., Geschwind, D.H. (2016). Rare Inherited De Novo CNVs Reveal Complex Contributions to ASD Risk in Multiplex Families. *Am. J. Hum. Gen*. 99:540-554.

Levy, S.E., Giarelli, E., Lee, L.C., Schieve, L.A., Kirby, R.S., Cunniff, C., Nicholas, J., Reaven, J., Rice, C.E. (2010). Autism spectrum disorder and co-occurring developmental, psychiatric, and medical conditions among children in multiple populations of the United States. *J Dev Behav Pediatr*. 31(4):267-75.

Lord, C., Risi, S., Lambrecht, L., et al. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord*. 30(3):205–223

Loviglio, M.N., Leleu, M., Männik, K., Passeggeri, M., Giannuzzi, G., van der Werf, I., Waszak, S.M., Zazhytska, M., Roberts-Caldeira, I., Gheldof, N., Migliavacca, E., Alfaiz, A.A., Hippolyte, L., Maillard, A.M.; 2p15 Consortium; 16p11.2 Consortium, Van Dijck, A., Kooy, R.F., Sanlaville, D., Rosenfeld, J.A., Shaffer, L.G., Andrieux, J., Marshall, C., Scherer, S.W., Shen, Y., Gusella, J.F., Thorsteinsdottir, U., Thorleifsson, G., Dermitzakis, E.T., Deplancke, B., Beckmann, J.S., Rougemont, J., Jacquemont, S., Reymond, A. (2017). Chromosomal contacts connect loci associated with autism, BMI and head circumference phenotypes. *Mol Psychiatry*. 22(6):836-849.

Luo, R., Sanders, S. J., Tian, Y., Voineagu, I., Huang, N., Chu, S. H., … Geschwind, D. H. (2012). Genome-wide transcriptome profiling reveals the functional impact of rare

de novo and recurrent CNVs in autism spectrum disorders. *American Journal of Human Genetics*, *91*, 38–55.

Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., … Newschaffer, C. (2017). The Changing Epidemiology of Autism Spectrum Disorders. *Annual Review of Public Health*, 38(1), 81–102.

Maillard, A.M., Ruef, A., Pizzagalli, F., Migliavacca, E., Hippolyte, L., Adaszewski, S., Dukart, J., Ferrari, C., Conus, P., Männik, K., Zazhytska, M., Siffredi, V., Maeder, P., Kutalik, Z., Kherif, F., Hadjikhani, N., Beckmann, J.S., Reymond, A., Draganski, B., Jacquemont, S.; 16p11.2 European Consortium. (2015). The 16p11.2 locus modulates brain structures common to autism, schizophrenia and obesity. *Mol Psychiatry*. 20(1):140-7.

Malhotra, D., & Sebat, J. (2012). CNVs: Harbingers of a rare variant revolution in psychiatric genetics. *Cell*, 148(6), 1223–1241.

Mazina, V., Gerdts, J., Trinh, S., Ankenman, K., Ward, T., Dennis, M. Y., … Bernier, R. (2015). Epigenetics of Autism-related Impairment : Copy Number Variation and Maternal Infection, *36*(2), 61–67.

Merks, J.H., van Karnebeek, C.D., Caron, H.N., Hennekam, R.C. (2003). Phenotypic abnormalities: terminology and classification. *Am J Med Genet A.,* 123A(3):211-30.

Merla, G., Brunetti-Pierri, N., Micale, L., & Fusco, C. (2010). Copy number variants at Williams-Beuren syndrome 7q11.23 region. *Human Genetics*, 128(1), 3–26.

Mezzacappa, A., Lasica, P.-A., Gianfagna, F., Cazas, O., Hardy, P., Falissard, B., … Gressier, F. (2017). Risk for Autism Spectrum Disorders According to Period of Prenatal Antidepressant Exposure. *JAMA Pediatrics*, 171(6), 555.

Miles, J. H., & Hillman, R. E. (2000). Value of a clinical morphology examination in autism. *American Journal of Medical Genetics*, *91*(January 1999), 245–253.

Miles, J. H., Takahashi, T. N., Bagby, S., Sahota, P. K., Vaslow, D. F., Wang, C. H., … Farmer, J. E. (2005). Essential versus complex autism: definition of fundamental prognostic subtypes. *American Journal of Medical Genetics. Part A*, *135*(2), 171–80.

Miles, J.H., Takahasi, T.N., Hong, J., Munden, N., Flournoy, N., Braddock, S.R., et al. Development and Validation of a Measure of Dysmorphology: Useful for Autism Subgroup Classification. *American Journal of Medical Genetics* Part A, 146:1101-16 (2008)

Miles, J. (2011). Autism spectrum disorders- A genetics review *Genetics In Medicine,* 13:4.

Moreno-De-Luca, A., Evans, D. W., Boomer, K. B., Hanson, E., Bernier, R., Goin-Kochel, R. P., … Ledbetter, D. H. (2015). The Role of Parental Cognitive, Behavioral, and Motor Profiles in Clinical Variability in Individuals With Chromosome 16p11.2 Deletions. *JAMA Psychiatry*, 72(2), 119.

Nava, C., Keren, B., Mignot, C., Rastetter, A., Chantot-Bastaraud, S., Faudet, A..et. al. Depiene, C. (2014). Prospective diagnostic analysis of copy number variants using SNP microarrays in individuals with autism spectrum disorders. *Eur. J. Hum. Gen*., 22: 71-78.

Ozgen, H. M., Hop, J. W., Hox, J. J., Beemer, F. a, & van Engeland, H. (2010). Minor physical anomalies in autism: a meta-analysis. *Molecular Psychiatry*, *15*(3), 300–7.

Ozgen, H., Hellemann, G. S., Stellato, R. K., Lahuis, B., van Daalen, E., Staal, W. G., … van Engeland, H. (2011). Morphological features in children with autism spectrum disorders: a matched case-control study. *Journal of Autism and Developmental Disorders*, *41*(1), 23–31.

Ozgen, H., Hellemann, G. S., de Jonge, M. V, Beemer, F. a, & van Engeland, H. (2013). Predictive value of morphological features in patients with autism versus normal controls. *Journal of Autism and Developmental Disorders*, *43*(1), 147–55.

Ozonoff, S., Young, G. S., Carter, a., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Stone, W. L. (2011). Recurrence Risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium Study. *Pediatrics*.

Prasad, A., Merico, D., Thiruvahindrapuram, B., Wei, J., Lionel, A. C., Sato, D., … Scherer, S. W. (2012). A Discovery Resource of Rare Copy Number Variations in Individuals with Autism Spectrum Disorder, Genes|Genomes|Genetics, 2(12), 1665–1685.

Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., … Betancur, C. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, *466*(7304), 368–372.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., … Scherer, S. W. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *American Journal of Human Genetics*, 94, 677–694.

Qiao, Y., Mercier, E., Dastan, J., Hurlburt, J., McGillivray, B., Chudley, A. E., … Rajcan-Separovic, E. (2014). Copy number variants (CNVs) analysis in a deeply phenotyped cohort of individuals with intellectual disability (ID). *BMC Medical Genetics*, *15*(1), 82.

Qureshi, A.Y., Mueller, S., Snyder, A.Z., Mukherjee, P., Berman, J.I., Roberts, T.P., Nagarajan, S.S., Spiro, J.E., Chung, W.K., Sherr, E.H., Buckner, R.L.; Simons VIP

Consortium. (2014). Opposing brain differences in 16p11.2 deletion and duplication carriers. *J Neurosci*. 34(34):11199-211.

Risch, N., Hoffmann, T. J., Anderson, M., Croen, L. A., Grether, J. K., & Windham, G. C. (2014). Familial recurrence of autism spectrum disorder: Evaluating genetic and environmental contributions. *American Journal of Psychiatry*, 171(11), 1206–1213.

Robinson, E. B., Samocha, K. E., Kosmicki, J. a., McGrath, L., Neale, B. M., Perlis, R. H., & Daly, M. J. (2014). Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. *Proceedings of the National Academy of Sciences*, 111(42), 15161–15165.

Rosenberg, R. E., Law, J. K., Yenokyan, G., McGready, J., Kaufmann, W. E., & Law, P. a. (2009). Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Archives of Pediatrics & Adolescent Medicine*, *163*(10), 907–914.

Rosenbloom, K. R., Armstrong, J., Barber, G. P., Casper, J., Clawson, H., Diekhans, M., … Kent, W. J. (2015). The UCSC Genome Browser database: 2015 update. *Nucleic Acids Research*, 43(November 2015), D670–81. https://genome.ucsc.edu/cgi-bin/hgGateway

Rutter, M., Caspi, A., Moffitt, T.E. (2003). Using sex differences in psychopathology to study causal mechanisms: unifying issues and research strategies. *J Child Psychol Psychiatry*. 44(8):1092–115.

Sanders, S.J., Ercan-Sencicek, A.G., Hus, V., Luo, R., Murtha, M.T., Moreno-De-Luca, D., Chu, S.H., Moreau, M.P., Gupta, A. R., Thomson, S.A., et. al. (2011). Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron,* 70: 863-885.

Sanders, S. J., He, X., Willsey, a. J., Ercan-Sencicek, a. G., Samocha, K. E., Cicek, a. E., … State, M. W. (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron*, 87(6), 1215–1233.

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Larsson, H., Hultman, C. M., & Reichenberg, A. (2014). The familial risk of autism. *JAMA : The Journal of the American Medical Association*, *311*(17), 1770–7.

Schendel, D. E., Autry, A., Wines, R., & Moore, C. (2009). The co-occurrence of autism and birth defects: Prevalence and risk in a population-based cohort. *Developmental Medicine and Child Neurology*, *51*(1), 779–786.

Schendel, D. E., Diguiseppi, C., Croen, L. a, Fallin, M. D., Reed, P. L., Schieve, L. a, … Yeargin-Allsopp, M. (2012). The Study to Explore Early Development (SEED): a multisite epidemiologic study of autism by the Centers for Autism and

Developmental Disabilities Research and Epidemiology (CADDRE) network. *Journal of Autism and Developmental Disorders*, *42*(10), 2121–40.

Schieve, L.A., Rice, C., Yeargin-Allsopp, M., et al. (2012). Parent-reported prevalence of autism spectrum disorders in U.S.-born children: an assessment of changes within birth cohorts from the 2003 to the 2007 National Survey of Children's Health. *J. Matern Child Health,* 16:S151–7.

Schmidt, R. J., Tancredi, D. J., Ozonoff, S., Hansen, R. L., Hartiala, J., Allayee, H., … Hertz-Picciotto, I. (2012). Maternal periconceptional folic acid intake and risk for developmental delay and autism spectrum disorder: A case-control study. *American Journal of Clinical Nutrition*, 96(3), 80–9.

Serra-Juhé, C., Rodríguez-Santiago, B., Cuscó, I., Vendrell, T., Camats, N., Torán, N., & Pérez-Jurado, L. A. (2012). Contribution of Rare Copy Number Variants to Isolated Human Malformations. *PLoS ONE*, *7*(10).

Shinawi, M., Liu, P., Kang, S.-H. L., Shen, J., Belmont, J. W., Scott, D. A., Lupski, J. R. (2010). Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *Journal of Medical Genetics*, *47*(5), 332–341.

Shelton, J. F., Geraghty, E. M., Tancredi, D. J., Delwiche, L. D., & Schmidt, R. J. (2014). Research | Children ' s Health Neurodevelopmental Disorders and Prenatal Residential Proximity to Agricultural Pesticides: The CHARGE Study, 122(10), 1103–1110.

Shoukier, M., Klein, N., Auber, B., Wickert, J., Schr??der, J., Zoll, B., … Zirn, B. (2013). Array CGH in patients with developmental delay or intellectual disability: Are there phenotypic clues to pathogenic copy number variants? *Clinical Genetics*, *83*(1), 53–65.

Surén, P., Roth, C., Bresnahan, M., & et al. (2013). Association between maternal use of folic acid supplements and risk of autism spectrum disorders in children. *JAMA*, 309(6), 570–577.

SFARI Gene https://gene.sfari.org/autdb/CNVHome.do

Shapira, K., Tian, L., Aylsworth, A., Elias, E., Hoover-Fong, J., Meeks, N., Souders, M., Tsai, C., Alexander, A., Schieve, L. (2014). Development of a Novel Protocol for Characterizing Dysmorphology to Enhance the Phenotypic Classification of Autism Spectrum Disorders, presented at International Meeting for Autism Research, Atlanta, 2014. Georgia: INSAR.

Sheppard, B., Benke, K.S., Croen, L., Daniels, J., Newschaffer, C., Reynolds, A., Schendel, D., Schieve, L., Ladd-Acosta, C., Fallin, D. (2015). Polygene –by–Prenatal

Environment Interaction in Autism Spectrum Disorder using Copy Number Variant Burden, presented at IGES, Baltimore, 2015. Maryland: IGES.

Smith, D. W. (1966). Dysmorphology (teratology). *J. Pediatr* 69:1150-1169.

Shishido, E., Aleksic, B., & Ozaki, N. (2014). Copy-number variation in the pathogenesis of autism spectrum disorder. *Psychiatry and Clinical Neurosciences*. 68(2):85-95.

Stein, J. L. (2015). Copy number variation and brain structure: lessons learned from chromosome 16p11.2. *Genome Medicine*, *7*(1), 13.

Torres, F., Barbosa, M., & Maciel, P. (2015). Recurrent copy number variations as risk factors for neurodevelopmental disorders: critical overview and analysis of clinical implications. *Journal of Medical Genetics*, 53(2):73-90.

US News Ranking of Best Graduate Schools. http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-health-schools/johns-hopkins-university-162928

van Daalen, E., Swinkels, S.H., Dietz. C., van Engeland, H., Buitelaar, J.K. (2007). Body length and head growth in the first year of life in autism. *Pediatr Neurol*. 37(5):324-30.

Volk, H. E., Lurmann, F., Penfold, B., Hertz-Picciotto, I., & McConnell, R. (2013). Traffic-Related Air Pollution, Particulate Matter, and Autism. *JAMA Psychiatry*, 70(1), 71.

Vulto-van Silfhout, A. T., Hehir-Kwa, J. Y., van Bon, B. W. M., Schuurs-Hoeijmakers, J. H. M., Meader, S., Hellebrekers, C. J. M., … de Vries, B. B. a. (2013). Clinical Significance of De Novo and Inherited Copy-Number Variation. *Human Mutation*, 34, 1679–1687.

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F. a, … Bucan, M. (2007). PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research*, 17, 1665–1674.

Weiss, L. A., Shen, Y., Korn, J. M., Arking, D. E., Miller, D. T., Fossdal, R., … Daly, M. J. (2008). Association between Microdeletion and Microduplication at 16p11.2 and Autism. *New England Journal of Medicine*, 358(7), 667–675.

Wier, M. L., Yoshida, C. K., Odouli, R., Grether, J. K., & Croen, L. a. (2006). Congenital anomalies associated with autism spectrum disorders. *Developmental Medicine and Child Neurology*, *48*, 500–507.

Wiggins, L. D., Reynolds, A., Rice, C. E., Moody, E. J., Bernal, P., Blaskey, L., … Levy, S. E. (2015). Using Standardized Diagnostic Instruments to Classify Children with Autism in the Study to Explore Early Development. Journal of Autism and Developmental Disorders, 45(5), 1271–1280.

Wingate, M., Kirby, R. S., Pettygrove, S., Cunniff, C., Schulz, E., Ghosh, T., … Autism Dev Disabilities, M. (2014). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. MMWR Surveillance Summaries, 63(2), 22.

Wong, V. C. N., Fung, C. K. Y., & Wong, P. T. Y. (2014). Use of dysmorphology for subgroup classification on autism spectrum disorder in Chinese children. *Journal of Autism and Developmental Disorders*, *44*(1), 9–18.

Xiong, N., Ji, C., Li, Y., He, Z., Bo, H., Zhao, Y. (2009). The physical status of children with autism in China. *Res Dev Disabil*. 30(1):70-6.

Yip, B. H. K., Leonard, H., Stock, S., Stoltenberg, C., Francis, R. W., Gissler, M., … Sandin, S. (2016). Caesarean section and risk of autism across gestational age: a multi-national cohort study of 5 million births. *International Journal of Epidemiology*, 46(2):429-439.

Zahnleiter, D., Uebe, S., Ekici, A.B., Hoyer, J., Wiesener, A., Wieczorek, D., Kunstmann, E., Reis, A., Doerr, H.G., Rauch, A., Thiel, C.T. (2013). Rare copy number variants are a common cause of short stature. *PLoS Genet*. 9(3):e1003365.

Tables

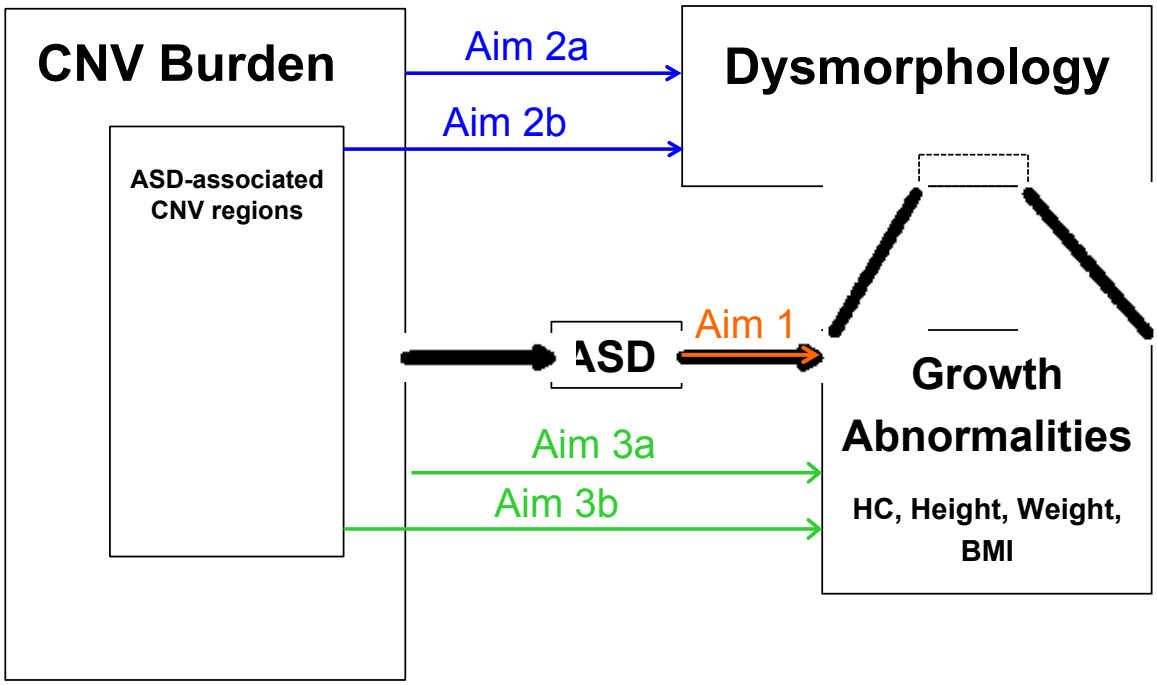
Table 1: Recent published literature on CNV burden and dysmorphology

Author, Year	Outcome assessed, sample size	Findings
Girirajan et al., 2011	Intellectual disability (ID), ASD, Dyslexia, Multiple Congenital Anomalies (MCA). 1,227 individuals with neurological deficits and 337 controls.	• Large CNV burden correlated positively with severity of childhood disability: ID and MCA most severely affected. • Greater burden of rare, *de novo* CNVs in ASD and ID. • Increased frequency of large CNVs (>1Mb) in ID compared to ASD. • Increased burden of large CNVs in ID with MCA compared to ID without MCA.
Cooper et al., 2011	ID with or without MCA. 15,767 children with ID & MCA and 8,329 unaffected adult controls.	• Large CNVs (>400kb) more prevalent in more severe developmental phenotypes associated with multiple congenital anomalies. • Greater enrichment of CNVs in individuals with craniofacial anomalies and cardiovascular defects compared to those with epilepsy or autism.
Girirajan et al., 2012	ID and MCA. – 2312 children with ID and MCA.	• Multiple, large CNVs associated with increasing clinical severity. • Total number of CNVs distinguishes those with syndromic disorders from those without.
Serra-Juhé et al., 2012	Congenital malformations not ascribed to a specific syndrome 95 fetuses – 68: isolated malformations – 27: multiple malformations	• Rare, deletion CNVs (>100kb), mostly inherited but also *de novo* was associated with congenital malformations, especially heart hypoplasia and brain malformations.

Author, Year	Outcome assessed, sample size	Findings
Vulto-van Silfhout et al., 2013	ID/ MCA 5,531 well-phenotyped patients 5,531 with ID/MCA. *De Vries score (to assess phenotype severity): Intellectual disability (ID), Growth retardation, ≥2 dysmorphic features and congenital anomalies	• Increased frequency of *de novo* CNVs in those with MCA and dysmorphism. • Patients with severe phenotypes, including organ malformations and abnormal head circumference, had more *de novo* CNVs, whereas patient groups with milder phenotypes, such as facial dysmorphisms, were enriched for both *de novo* and inherited CNVs. • Multiple CNVs were associated with a more severe phenotype than single CNV. • *De novo* and familial CNVs associated with greater severity score. • CNV deletions were more likely to result in severe phenotypes than CNV duplications. • CNV size was correlated with the phenotype severity.
Shoukier et al., 2013	DD/ID with congenital anomalies 342 DD/ID cases	• Congenital anomalies, especially heart defects, as well as primary microcephaly, short stature and failure to thrive were more frequent in children with pathogenic CNVs compared with children with normal array CGH results.
Qiao et al., 2014	ID with phenotypic abnormalities 78 ID subjects with phenotypic abnormalities classified using the Winter-Baraitser Dysmorphology Database (WBDD)	• CNV/phenotype correlation analysis showed rare, *de novo* and familial CNVs were associated with cranial and forehead abnormalities.

Figures

Figure 1. Conceptual Framework



Chapter 2: Abnormal Growth in Autism Spectrum Disorder in the SEED Study

Norazlin Kamal Nor, MBBS[1], Julie Elizabeth Hoover-Fong, MD, PhD[3],

Christine Ladd-Acosta, PhD[1], Terri H. Beaty, PhD[1],

Margaret Daniele Fallin, PhD[2]

Department of Epidemiology[1] and Mental Health[2],

Johns Hopkins Bloomberg School of Public Health and Department

of Medicine[3], Johns Hopkins University School of Medicine, Baltimore, MD, USA

Abstract

Introduction:

Autism spectrum disorder (ASD) has a heterogeneous presentation with multiple subtypes, and is frequently found with co-morbid conditions ranging from neurodevelopmental to physical abnormalities. Greater understanding of autism and its sub-phenotypes may improve identification of risk factors, allow better prognostication, and potentially allow earlier, targeted, interventions. Here we investigate phenotypes of somatic growth, specifically for the anthropometric measures of head circumference, stature, weight, and body mass index (BMI) in a sample of ASD children compared to typically developing children from Phase 1 of the SEED Study. In addition, we assessed the combination of height, weight and head circumference simultaneously, which we term trivariate growth phenotype, to examine differences in growth symmetry in children with ASD compared to typically developing children.

Methods:

Our study sample is comprised of 913 children aged 2 to 5 years from the Study to Explore Early Development (SEED) Phase 1, with 532 ASD and 381 typically developing children (POP). We excluded children with known chromosomal abnormalities and genetic syndromes. We selected two different thresholds to dichotomously classify growth abnormality, one using clinician derived thresholds, and another using top and bottom deciles. Dichotomous growth phenotypes were derived for macrocephaly, microcephaly, tall and short stature, overweight and underweight plus

high and low BMI. These were compared between ASD cases and controls using logistic regression of growth abnormality on ASD adjusting for self-reported race and stratifying by sex. We also compared the frequency of ASD and POP children with each possible combination of trivariate growth phenotype using Fisher's exact test.

Results:

Using clinical definitions for growth abnormalities, children with ASD had higher odds of short stature compared to typically developing POP children, adjusted for race (aOR=1.92, 95% CI: 1.06, 3.49; p=0.03). This odds ratio was attenuated using decile-defined short stature (aOR=1.47, C.I. 0.91, 2.37; p=0.10). Furthermore, this association was higher among girls (aOR = 5.56, C.I. 1.94, 15.97; p=0.001, using clinical definitions; aOR=3.52, C.I. 1.6, 7.76; p=0.002, using top deciles). Decile-defined tall stature was also decreased in ASD girls (aOR = 0.47, C.I. 0.22, 0.97; p=0.04). High BMI, based on the top decile from reference data, was again associated with ASD status (aOR=1.44, CI: 1.04, 2.00; p=0.03).

ASD children also had a higher frequency of the trivariate growth combination of microcephaly, short stature and normal weight (clinical definitions: Fisher's p=0.01; decile definitions: Fisher's p=<0.01). This association was strongest and statistically significant among girls, regardless of the definition used, but was not statistically significant among boys.

Conclusions:

Our study showed sex-specific differences in growth abnormalities among young children with ASD. ASD girls were found to have greater odds of short stature compared to control girls, and to have a higher prevalence of the combined growth phenotype involving microcephaly and short stature, but normal weight. This study illustrates the importance of considering how sex may influence the presentation of growth abnormalities among ASD childrene and how crucial it is to include girls in all future ASD research.

Keywords: autism spectrum disorder, abnormal growth, macrocephaly, microcephaly, tall stature, short stature, overweight, underweight, high BMI, low BMI

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition with impairments in socialization, communication and restricted, repetitive behaviors and stereotypies (American Psychiatric Association, 2013). Although the specific causes of ASD are not yet established, ASD has been associated with multiple risk factors encompassing both genetic and environmental factors. The prevalence of autism is currently approximately 1 in 68 children in the United States (CDC, 2016). It has a greater prevalence in boys compared to girls, with 3-5 times more boys affected (Werling & Geschwind, 2013; Loomes et al., 2017). Although there are common core features of ASD, considerable heterogeneity has been observed. Autistic individuals often have a wide variability of clinical features as well as differences in the degree of severity and developmental trajectories. Girls with ASD have been reported to have more severe presentations with co-existing conditions such as intellectual disabilities (Werling & Geschwind, 2013). Studying these autism phenotypes can improve our biological understanding of ASD, particularly if different risk factor constellations are reflected in different phenotype presentations. This may be particularly helpful in parsing out different genetic risk factors. In addition, characterization of specific autistic phenotypes may aid in early diagnosis and prediction of outcomes (Walsh et al., 2011). For example, epilepsy, a co-morbidity observed in 8-20% of all children with ASD (Berg, Plioplys & Tuchman, 2011; Christensen et al., 2016), can inform treatment and prognosis for those with both conditions. Growth abnormalities, including macrocephaly, (i.e. the presence of a large head) have also been associated with ASD, and identification of children with

overlapping growth abnormalities and ASD may help guide identification and early intervention.

Growth in early life is affected by both genetic and environmental factors, with different factors having varying influences at specific stages of development. Normal growth in children consists of 3 phases: prenatal, childhood and pubertal and relies on three uni-dimensional growth parameters (height, weight and head circumference). These measures are considered individually, and also as bi-dimensional traits such as Body Mass Index (BMI), which incorporates both weight and height. Finally, the three modalities of growth may be combined to symmetry of overall growth in an individual. In this paper, we term a three-dimensional measure of 'Trivariate Growth Phenotype' (TGP) based on symmetry or asymmetry of head circumference, height, and weight. To allow comparability of anthropometric measurements, specific growth measures are typically standardized to age and sex specific population-based growth percentile curves (Kuczmarski et al., 2000; Wells, 2007; WHO, 2006).

Growth "abnormalities" consider the extremes of growth, with different cut-offs used for different measures of growth. For example, abnormally large head circumference (macrocephaly) is often recognized clinically as head circumference >= the 97th sex and age-specific percentile, while microcephaly defined as head circumference <= the 3rd percentile. For height, tall stature is defined by height >= 3 standard deviations above the mean for sex and age, but short stature defined as height less < the 10th percentile for sex and age. For weight, overweight is defined by weight >= the 97th percentile for sex and

age, and underweight as weight < the 10th percentile for sex and age. The arbitrary nature of these cut-offs, in terms of comparability to each other, makes it challenging to directly compare all three growth measurements. In this work, in addition to these clinical thresholds, we have also used the top and bottom 10% of the population to define abnormality for comparison of results across all growth parameters.

Growth Abnormalities and Autism Spectrum Disorder

Children with ASD have been recognized to have abnormalities in several parameters of physical growth. In fact, Kanner first made the observation that 5 of the 11 autistic children in his original clinical sample had abnormally large heads (i.e. macrocephaly) (Kanner, 1943). Accelerated overgrowth of the head early in development, leading to macrocephaly, has been frequently observed in more recent literature (Lainhart et al., 1997; Fombonne et al., 1999; Miles et al., 2000; Courchesne et al., 2001; Courchesne et al., 2003; Redcay & Courchesne, 2005; Lainhart et al. 2006). Approximately 15-20% of children with autism have been reported to have macrocephaly, with some variability in these estimates (Fombonne et al., 1999; Lainhart et al., 2006; Dementieva et al., 2005;). Enlarged head circumference has been found to be associated with increased total brain volumes in some neuroimaging studies (Hazlett et al., 2005, Tate et al., 2007), and post-mortem studies of autistic brains have shown increased gray and white matter volume in addition to enlarged head circumference (Redcay & Courchesne, 2005). A recent systematic review and meta-analysis lead by Sacco in 2015 based on 27 studies of brain size in autism, concluded 15.7% of autistic individuals have macrocephaly (Sacco,

Gabriele & Persico, 2015). Publications assessing macrocephaly stratified by sex have reported a higher prevalence of macrocephaly in ASD boys compared to ASD girls (Amaral et al., 2017). In the majority of studies on abnormal growth in ASD, however, girls are poorly represented due to the lower prevalence of ASD in girls compared to boys.

In addition to abnormal head growth, children with ASD have also been reported to have abnormal weight and height. Autism has been associated with obesity in numerous studies (Ahearn et al., 2001; Evans et al., 2012; Curtin et al. 2005; Rimmer et al. 2010; Xiong et al., 2009; Broder-Fingert et al., 2014; Zuckerman & Fombonne, 2014; Must et al., 2017), including research in large databases such as the National Survey of Children's Health (Chen et al., 2009; Curtin et al., 2010). Excessive weight among people with ASD trends with increasing age, especially among adolescents with ASD (Hill, Zuckerman & Fombonne, 2015, Bicer & Alsaffar, 2013). Higher rates of obesity have also been observed among ASD boys (Must et al., 2017).

Some studies have described what is termed abnormal 'early generalized overgrowth' in children with ASD. Growth trajectories among children with ASD have revealed an enlarged head circumference, taller stature and higher weight among autistic boys compared to typically developing controls in the first year of life (Chawarska et al., 2011). It was postulated children with autism show a general growth dysregulation instead of simply overgrowth, as certain growth modalities were found to be asymmetric (Curtin, 2005 et al.; van Daalen et al., 2007). Rarely are combinations across all three

measures of physical growth considered in studies of growth and ASD. Here, we use the term the 'trivariate growth phenotype' to capture the simultaneous combination of head circumference, height, and weight per child, allowing consideration of growth symmetry and estimation of whether some particular asymmetric combinations are more prevalent among ASD cases. Asymmetric growth, where one or more modalities of growth are not in proportion to the other, may suggest a more specific pathological process, as seen in for example endocrinopathies or other recognized syndromes.

Here, we estimate the association between abnormal growth and ASD among 2-5 year old children who participated in the SEED 1 study, a national ASD case-control study supported by the CDC. We considered head circumference, height, weight, BMI, and the trivariate combination of the first three. Characterization of specific growth differences in cases and controls, can inform ASD phenotyping that may enable improved risk factor discovery, as well as better prediction of outcomes and perhaps targeted early interventions for children with ASD.

Materials and Methods

Study Population

The Study to Explore Early Development (SEED), phase I, is a national multi-site case control study funded by the CDC. Children between ages 2-5 years old were recruited and evaluated at one of six sites: California, Colorado, Georgia, Maryland, North Carolina and Pennsylvania. Potential cases, born between 2003 and 2006 were recruited through partnerships with developmental disability service providers, including healthcare and educational systems. Population-based control children born in the same years from the same catchment areas were recruited through vital statistics. After phone-based screening and in-person evaluations, children were classified as ASD, other non-ASD developmental disabilities, and children without developmental disabilities from the general population (POP) (Schendel et al., 2012). Eligibility included birth in the study catchment area during the period between September 1st, 2003 to August 31st 2006, current residence in the area at the time of first contact, and child living with a knowledgeable caregiver who was able to communicate orally in English or Spanish competently and gave informed consent for participation. The enrolled children also had to be between the ages of 30 and 68 months of age at the completion of the in-person clinical developmental assessment.

There were 3,899 children recruited into the SEED 1 study. Only ASD and POP children were considered in these analyses. There were 1,130 children who underwent

dysmorphology assessment. Children with complete dysmorphology assessment information were obtained after filtering out those with a high proportion of dysmorphic features categorized as 'not scored' (80 or more, out of a possible 397), either because the assessment was not completed by assessor or due to poor photograph quality. We restricted the sample to complete-case analyses, where data were available for growth abnormalities measures, sex, maternal education and race (N=958). We then excluded all children with chromosomal abnormalities and recognized genetic syndromes (N=45), where data were derived from parent report during the original clinic visit, recorded on the Dysmorphology Exam Form, and verified by the Clinical Geneticist in the SEED Dysmorphology Working Group whenever pediatric medical records were available.

ASD Assessment

For all eligible children, a brief screening interview, the Social Communication Questionnaire (SCQ; Rutter et al. (2003)), was administered to the primary caregiver to identify children who required clinical diagnostic assessment to determine final ASD status. For SEED, a positive screen was defined as an SCQ score ≥11. Tools for ASD assessment included the Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview-Revised (ADI-R) (Falkmer et al., 2013; Lord et al., 2000). Final classification was assigned using a SEED-specific research algorithm based on ADOS, ADI-R and clinical judgment (Wiggins et al., 2015). Regardless of ascertainment source, any eligible children with a previous ASD diagnosis, who were receiving special education services, and who had a positive screen, were assigned to the ASD workflow.

This determined which instruments were administered and the type of diagnostic evaluation the child received during data collection. Based on previous diagnosis and SCQ screening, POP children with negative SCQ screens were assigned to the POP workflow. If a clinician suspects ASD during the clinical evaluation of a child in the POP workflow, the child would be moved into the ASD workflow (Schendel et al., 2012; DiGiuseppe et al., 2016).

Growth Measures

SEED participants underwent anthropometric measurements and had standardized photographs taken of specific regions of the body. The specific growth measures of interest here were collected using a standardized procedure by trained clinic staff using standardized supplies including the tape measure for head circumference, the stadiometer and the weighing scale. For head circumference, a non-stretchable, plasticized measuring tape was used to measure head circumference for maximum circumference of the head. The tape was placed just above the eyebrows, above the ears and around the most protuberant part of the back of the head (occiput), pulled snugly to compress hair and read to the nearest 0.1cm, after which the measurement was recorded on the Dysmorphology Review Form (DRF) and repeated, so repeated measurements were within 0.2cm from each other. For height measurement, an accurate and appropriate stadiometer was used: a vertical board with an attached metric rule and a horizontal headpiece was brought into direct contact with the most superior (top) part of the head, and read to the closest 0.1cm. Height measurement was performed for all children with

hair accessories removed and without shoes. The child was measured standing with heels, buttocks, shoulders and head touching a flat upright surface. The arms were held on the side, with shoulders relaxed and legs straight, and heels close together. The child was asked to look straight ahead and the perpendicular headpiece lowered to the crown of the head snugly to compress the hair. The measurer's eyes were parallel with the headpiece. The measurement was repeated, with agreement to within 1cm, and recorded on a growth chart appropriate for the child's age and sex. The raw measurement and percentile growth from the percentile chart was then transferred to the DRF form. The stadiometer position was standardized, so there were no attachments to the wall and no underlying carpet. The stadiometer was also calibrated monthly using the SEED Equipment Checklist and Equipment Calibration Forms. For the measurement of weight, a safe and accurate scale with a wide enough platform to support the child being weighed was used. The scale was required to be calibrated with standard weights, able to be zeroed, and was not on a carpeted surface. The child stood on the weighing platform without assistance and wearing only light undergarments or gown. The reading was recorded, and repeated until agreement within 0.1kg.

Clinical geneticists in the SEED Dysmorphology Group assessed these measures and photographs using the customized DRF tool. The SEED Dysmorphology Group developed the DRF to quantify the intrinsically qualitative observations used to classify dysmorphology. This tool consisted of clinical observations of 397 specific physical features in seven body regions divided into: Head/Hair/Face/Neck, Ears, Eyes/Eyebrows, Nose/Philtrum, Mouth/Lips/Teeth, Hands/Feet, and Growth/Skin. The clinical geneticists

were blinded to the child's case status (ASD vs POP) and were each assigned a particular body region for consistency in rating individual regions. The DRF had specified cut-offs for defining abnormal growth, generally aligned with CDC guidelines.

For these analyses, raw values of head circumference, height, and weight were converted to z-scores and percentiles using the 2000 CDC Head Circumference-for-Age Growth Charts for girls and boys ages 2–20, the Stature for-Age Growth Charts for girls and boys ages 2–20, the Weight-for-Age Growth Charts for girls and boys ages 2–20 and the BMI-for-Age Growth Charts for girls and boys ages 2–20.

Dichotomous abnormal growth outcomes were defined for each child for: large and small head circumference (i.e. macrocephaly and microcephaly), tall stature, short stature, overweight, and underweight. Two alternative threshold strategies were used to define growth abnormalities. First, abnormalities were defined based on the SEED Dysmorphology Group thresholds, empirically derived through examination of SEED data, literature, and existing clinical thresholds. These cut-offs are implicit in the definitions in the DRF, and capture extreme dysmorphology. As a second strategy, children were also classified simply by being in the top or bottom decline for head circumference, height, and weight. This includes less extreme features, but allows comparability across all three growth modalities. The bi-dimensional measure for weight and height: body mass index (BMI), was measured using the formula (BMI= Weight (kilograms)/ Height^2 (meters)). A trivariate growth phenotype was derived as the simultaneous combination of dichotomous growth classification across the three

measures of head circumference, stature and weight. This resulted in 27 trivariate groups. Children with the same category of growth across all three (e.g., macrocephaly, tall stature, overweight) were considered as having symmetric growth. All others were considered asymmetric (Shapira, personal communication, SEED Dysmorphology Group unpublished manuscript).

Other Covariates

Participant self-identified race and sex, as well as parental race/ethnicity and maternal education were obtained from the Caregiver Interview (CGI), a computer-assisted telephone interview with mothers or other knowledgeable caregivers. This tool was used to acquire information about the child, family and caregiver, and quality assurance measures were used to ascertain their reliability (Schendel et al., 2012). Maternal and paternal race were used in an algorithm to determine child race/ethnicity as coded in the variable 'DR_DRF_DysmRace' in the dysmorphology analysis. Maternal education was also obtained from the CGI to ascertain the mother's highest attained educational level.

Statistical Analysis

We characterized the analytic sample using counts and percentages for ASD and typically developing (POP) children. We compared percentile means in growth modalities for both groups and repeated the analysis after stratifying by sex.

Odds ratios comparing the odds of each growth abnormality between children with ASD and POP children were estimated using the *logistic* command in STATA (MP12.1), modeling the log probability of each specific growth abnormality as a function of ASD status, adjusted for race. Overall, and sex-stratified models were run; models based on abnormalities defined by the DRF as well as by decile were also run separately. Frequencies of trivariate growth patterns were estimated in STATA, with Fisher's exact test used to examine differences between ASD and POP children. For all abnormality analyses – univariate, divariate, and trivariate – both nominal and multiple test-corrected (Bonferroni) p values were calculated.

Results

Sample Characteristics

Of the 2595 SEED 1 case (n=1,306) and POP (n=1,289) children, 1,130 had complete dysmorphology information, and of these, 913 were available for this analysis (see **Figure 1**). **Table 1** shows the sample characteristics of 532 children classified with ASD and 381 POP children. The study population included the following demographic breakdown: by ancestry 55% Non-Hispanic White, 22% African-American, 19% Hispanic and 4% other; by sex 70% boys and 30% girls; and by ASD status 44% ASD and 56% non-ASD, including 36% POP and 20% DD subjects. The higher percentage in boys overall is due to greater numbers of boys compared to girls in the ASD group, reflecting the higher preponderance of ASD in males. There were more males in the ASD group than POP group, as expected given the established sex bias in ASD. Maternal education and maternal self-reported race were not statistically significantly different between groups.

Mean Differences in Growth Percentiles

Distributions of reference-standardized percentiles among SEED children for head circumference, stature, weight and BMI for ASD and POP children are illustrated in **Figures 2** (overall) and **3** (sex-stratified). There is generally an inflation of higher percentiles among the SEED sample. Comparison of mean percentile values between

groups for each growth modality are shown in **Table 2,** and stratified by sex in **Table 3**. Mean estimates in ASD children were larger for head circumference and BMI, and lower for weight and height, than in POP children, although neither of these differences reached statistical significance. ASD versus POP means were not statistically significantly different for boys or girls for head circumference, BMI or weight. ASD girls did show statistically significantly shorter mean height than seen in control (POP) girls (p=0.04).

Differences in Growth Abnormalities

The counts and proportions of ASD and POP children with each of the 8 growth abnormalities considered are shown in **Tables 4 and 5.** Results are shown using the clinically-informed SEED DRF thresholds, as well as using top and bottom deciles. In general, DRF-defined macrocephaly and tall stature were rarely observed (5.1%, 1.6%, respectively). Using the reference-based decile cut-offs, more than the expected 10% were observed in the top reference decile for microcephaly, tall stature, overweight, and high BMI. Sex-stratified results were only possible for decile-defined abnormalities, because DRF-defined frequencies were too low. There were 640 males (67.6% with ASD) and 273 females (36.2% with ASD). Using decile cut-offs, the numbers of females with growth abnormalities were quite low, for example only 10 females were identified with low BMI (n=6 and n=4 for POP and ASD, respectively). Overall, there were a higher percentage of ASD males with high BMI compared to POP males (24.5% vs. 18.4%). For height in females, a higher percentage of ASD females were found to have short stature compared to POP females (19.2% vs. 6.3%). A smaller proportion of ASD females were tall compared to POP females (11% vs. 20.7%). Within the male ASD

subgroup, the growth abnormality with the highest number of males identified was high BMI, and the growth abnormality with the lowest number of males identified was low BMI. For the female ASD subgroup, the growth abnormality with the highest number of females identified was microcephaly, and the growth abnormality with the lowest number of females identified was low BMI.

Results from logistic regression analyses estimating the association between each growth abnormality (defined by DRF) with ASD status, adjusted for race, are shown in **Table 6**. The unadjusted odds ratio for short stature among children with ASD was 1.90 (C.I.: 1.05, 3.45; p=0.03) compared to POP children. Adjusting for race, the estimated odds ratio remained consistent at 1.92 (CI: 1.06, 3.49; p=0.03). However, this association did not survive correction for multiple testing. When using decile-defined abnormalities, the association with short stature was attenuated and not statistically significant (aOR=1.47, C.I. 0.91, 2.37; p=0.10; **Table 8**). However, high BMI showed increased odds among ASD cases (aOR=1.44, C.I. 1.04-2.0, p=0.03), although this did not survive multiple testing correction.

Similar analyses stratified by sex showed higher odds of DRF-defined short stature (**Table 7)** among girls with ASD compared to POP girls (aOR = 5.56, C.I. 1.94, 15.97; p=0.001). This was also observed using the decile definition of short stature (aOR=3.52, C.I. 1.6, 7.76; p=0.002, **Table 9**). Decile-defined tall stature was also observed to be decreased in ASD girls (aOR = 0.47, C.I. 0.22, 0.97; p=0.04).

Trivariate Growth Phenotypes

Frequencies of ASD and POP children with each of the 27 possible combinations of 3 growth modalities are show in **Tables 10 and 11**, reflecting both DRF-defined and decile-defined abnormalities, respectively. As expected, most children for both groups had normal head circumference, normal height and normal stature. Tests of overall differences in symmetry between ASD and POP children were not statistically significant using either abnormality definition.

Using the DRF cut-offs for growth abnormalities, there were 12 cells in which no children were observed. Only one pattern of DRF-defined trivariate growth showed a statistically significant difference between ASD and POP: the combination of microcephaly, short stature and normal weight, although this association did not survive correction for multiple testing. This may be driven by girls, where 7% of 99 ASD girls (compared to 1.1% of 174 POP girls) had this combination ($p=0.01$), yet only 2.7% of 433 ASD boys and 0.9% of 207 POP boys had this combination ($p=0.24$). Using decile-defined thresholds, two patterns of trivariate growth showed statistically significant differences between ASD and POP: the combination of microcephaly, short stature and normal weight, as seen with DRF definitions, and also the combination of microcephaly, normal stature and overweight, although neither remained significant after correction for multiple testing. The association with microcephaly, short stature, normal weight phenotype and ASD was again statistically significant in girls (10.1% ASD vs 3.4% POP in girls, $p=0.03$), but not in boys (3.4% ASD vs 1.9% POP in boys, $p=0.33$).

Discussion

Abnormal growth in various forms, i.e. macrocephaly, has been recognized as a sub-phenotype in autism spectrum disorder (ASD), a complex and heterogeneous group of neurological developmental disabilities where both genetic and environmental risk factors contribute to the etiology. Considering sub-groups of ASD such as growth abnormalities holds some potential for improving not only our understanding of autism spectrum disorder by providing a clearer picture of the complexities of clinical presentations of autism, but may also point towards important etiological risk factors, specifically genetic risk factors, as well as possibly being useful in screening for ASD and its phenotypes, and managing therapeutic strategies.

In this work, we found increased prevalence of short stature among ASD children ages 2-5, compared to population-based controls, and this association was stronger among girls. ASD children also had a higher frequency of the trivariate growth combination of microcephaly, short stature and normal weight (based on clinical definitions: Fisher's $p=0.01$, Decile definitions: Fisher's $p=<0.01$). This association was strongest and statistically significant among girls, regardless of the abnormality definition used, but was not statistically significant among boys. We also observed a greater proportion of ASD cases above the 90^{th} percentile for BMI. Notably, we did not observe statistically significant associations between head circumference and ASD, although ASD children had a slightly higher mean head circumference than POP children, and higher frequencies of both macrocephaly and microcephaly.

Previous studies of height and stature abnormalities in ASD cases have shown inconsistent results. Some studies found short stature and microcephaly among ASD children, but included some children with syndromic diagnosis, for example Smith-Lemli-Opitz, where short stature is a recognized feature (Goldenberg et al., 2003). Most studies have found increased, rather than decreased, height among ASD children, particularly in boys, and often corresponding to increased head circumference (Davidovitch et al., 1996; Dissanayake et al., 2006). The observation of an isolated association between short stature and ASD in girls has not been as widely reported. A study by Lainhart et al. (2006) reported a higher percentage of short stature in females compared to males with ASD (26.7% vs. 7%), although the sub-sample of subjects with short stature (n=13) was modest (Lainhart et al., 2006). The paucity of such reports may be due to the greater preponderance of ASD among boys, making it much more difficult to obtain cohorts of autistic girls with growth abnormalities. In fact, a number of studies concerning growth in ASD children, excluded girls due to inadequate sample sizes (Suren et al., 2013).

Our findings also showed the combination of microcephaly, short stature and normal weight was associated with ASD. Interestingly, no POP children were found to have this particular combination of growth abnormalities, and this association appears to be driven by growth abnormalities in girls only. This particular combination of growth pattern is unusual. Clinically, children are frequently observed to have short stature and microcephaly following a prolonged period of failure to thrive, which is typically accompanied with these children being underweight. Microcephaly with short stature and

the absence of decreased body weight represents a deviation in expected patterns of growth, even in children with failure to thrive. Previous studies have observed associations between ASD and combinations of head circumference, weight, and height, particularly for symmetric overgrowth (Torrey et al., 2004; Dissanayake et al., 2006; Mraz et al., 2007; Fukumoto et al., 2008, Zwaigenbaum et al., 2014). A recent population-based longitudinal study of 376 children culled from the Norwegian Mother and Child Cohort assessed these three modalities of growth in autistic children (Suren et al., 2013) and reported symmetric overgrowth in ASD boys, but a trend towards undergrowth in ASD girls compared to control children. However, these investigators remarked there was an inadequate number of ASD girls in their study to support meaningful inferences. These authors found the growth acceleration among boys became apparent from 6 months of age onwards, and ASD girls had decreased head circumference and weight compared to control girls. They postulated their findings among girls with autism were driven by concurrent epilepsy, other forms of intellectual disability or genetic disorders, but that these comorbidities did not impact findings on head growth trajectory among boys. In our study, we excluded all children with chromosomal abnormalities and recognized genetic syndromes, but not those with reported epilepsy, and we had a reasonable, but modest, number of girls with ASD. Another longitudinal study in 347 ASD children found a child's age and sex influenced growth abnormalities (Campbell et al., 2014). ASD boys were symmetrically larger across the three modalities. These changes were attributed to early, generalized overgrowth starting around 6 months of age. Girls with ASD did not show abnormal patterns of growth, but had a general trend of a smaller head circumference.

Previous studies of weight abnormalities in autism have also shown variable results. The prevalence of overweight and obesity in children with autism have been reported to be higher than in control children (de Vinck-Baroody et al., 2013; Egan et al., 2013; Evans et al., 2012; Hyman et al., 2012; Rimmer et al., 2010; Cheung et al., 2016; Phillips et al., 2014). Two studies comparing ASD children directly to control children found higher rates of obesity, approaching statistical significance (Curtin et al., 2010; Evans et al., 2012). However, these were relatively small studies (Curtin et al., 2005; Egan et al., 2013; Evans et al., 2012; Ho et al., 1997), relied on parent-reported weight instead of direct measurements (Chen et al., 2010; Curtin et al., 2005), and used unconventional definitions of obesity (Ho et al., 1997). Our finding of increased high BMI among children with ASD compared to control children may also reflect other co-morbidities such as sleep disorders (Broder-Fingert et al., 2014; Hill et al., 2015) and poorer psychosocial functioning (Hill et al., 2015) in the ASD group.

We did not see statistically significant differences in head circumference between ASD and POP children, although ASD children had slightly higher means and slightly greater prevalence of both macrocephaly and microcephaly. This is different from many previous studies suggesting macrocephaly is associated with ASD (Courchesne et al., 2003; Fombonne et al., 1999; Lainhart et al., 2003; Lainhart et al., 1997; Barnard-Brak et al., 2011, Chaste et al., 2013); although more recent research has also shown inconsistencies in this association (Suren et al., 2013; Dinstein et al., 2017). This may reflect study selection and design, such as cross-sectional versus longitudinal data, or inclusion of

children with co-morbidities (including chromosomal abnormalities or genetic syndromes). Our finding of no statistically significant differences in head circumference and ASD may be due to our exclusion of children with chromosomal abnormalities and recognized genetic syndromes from our analysis, or our use of DRF and population-reference definitions for abnormal growth. Recent cross-sectional studies comparing head circumference in ASD children and controls often have found no significant difference (Raznahan et al., 2013; Dinstein et al., 2017).

When considering growth abnormalities, we used two alternative definitions, one based on a clinically-informed algorithm implemented in the SEED study (DRF), and another based simply on population reference deciles. The DRF definitions captured children at the extreme ends of specific growth modality distributions, while decile definitions were more inclusive. Further, the DRF thresholds were not consistent across the three modalities, while the decile definitions allowed standardized comparison across modalities. For example, the threshold for tall stature under the DRF is height three standard deviations above the mean (≥99.9th percentile), while that for short stature is height 1.25 standard deviations below the mean (<10th percentile). Comparing these stringent and more liberal definitions allowed consistent patterns to emerge. For example, results for short stature were consistent among analyses using both definitions, but stronger when based on the more stringent DRF definitions.

In our study, we excluded children with chromosomal abnormalities and recognized genetic syndromes. ASD is a complex condition for which no specific etiology has been

determined, although genetic, environmental and gene-by-environment risk factors have been suggested (Iossifov et al., 2014; Krum et al., 2015; Karimi et al., 2017). We therefore wanted to exclude growth abnormalities likely caused by single chromosomal disorders or recognized genetic syndromes with Mendelian inheritance. In addition, increased prevalence of chromosomal abnormalities and recognized genetic syndromes in ASD children compared to control children may affect the distribution of growth abnormalities between ASD and POP children because these chromosomal defects and genetic syndromes increase likelihood of growth abnormalities and thus may confound any association between ASD and abnormal growth.

Previous studies reporting rapid acceleration of growth early in life, particularly in autistic boys, suggested this may be an indicator of abnormal neural development and somatic growth dysregulation (Hazlett et al. 2005), potentially mediated via neurotrophins (a group of proteins involved in neurodevelopment, including neurotrophic factor and insulin-like growth factor, and neuropeptides such as vasoactive intestinal polypeptide [VIP] and calcitonin gene-related protein [CGRP] (Akshoomoff et al., 2002; Courchesne et al., 2001)). Although previous hypotheses have considered effects of neurotrophins on somatic overgrowth, and here we observed undergrowth among ASD girls, it is still possible that neurotrophin-mediated mechanisms are acting in opposing directions. It is also possible co-occurring plasma growth hormone dysregulation and pituitary-hypothalamic dysfunction may play a role. This has been previously observed in children with autism (Deutsch et al., 1985; 1986).

Whether growth abnormalities represent a true neurobiological sub-type of autism is not yet established. While some have postulated growth abnormalities (e.g. macrocephaly) may be endophenotype-specific etiopathogenic factors (Sacco et al., 2007), others suggest a more general tendency towards growth dysregulation (Fombonne et al., 1999; Lainhart et al., 2006; Van Daalen et al., 2007). Here, the preponderance of growth abnormalities in girls suggests a potential association with sex-specific risk for autism. Sex-specific findings in previous growth studies of ASD have been observed (Campbell et al., 2014; Suren et al., 2013; Campbell et al., 2014; Must et al., 2017). Campbell et al. (2014) suggested generalized somatic overgrowth, seen only in autistic boys, is part of the sexual dimorphism for autism. The observation in a longitudinal study of overgrowth in autistic boys, and microcephaly and underweight among autistic girls, gave rise to the hypothesis that growth trajectories in autism are sex-specific (Suren et al., 2013). There may be differences in growth regulation and hence growth abnormalities based on sex and the child's age (Campbell et al., 2014; Suren et al., 2013). Our study further contributes to the hypothesis of growth abnormalities are sexual dimorphisms in autism. We describe a distinctive pattern of abnormal growth in autistic girls, specifically short stature in ASD girls, and the combination of short stature, microcephaly and normal weight. Compared to the overgrowth seen in austic boys, girls with autism appear to present with the opposite growth abnormality phenotype of undergrowth, at least for short stature with microcephaly. Other studies have reported sex-specific differences in physical brain changes in ASD. A recent study showed autistic boys have small callosal regions projecting to the orbitofrontal cortex, and autistic girls have smaller callosal region projecting to the anterior frontal cortex (Nordahl et al., 2015). In addition,

comparing autistic males and females, there are substantial differences in the overall pattern of changes in gray and white matter volume across the brain (Lai et al., 2013). These differences in brain regions may be part of the neurodevelopmental basis for the differences in ASD phenotypes between the sexes.

One of the complexities of studying sex-specific effects in autism is that sex has sometimes been treated as a covariate rather than simply doing a stratified analysis. In addition, to reduce variability and to better characterize autistic core symptoms or phenotypes, many studies have restricted all analyses to boys. For example, there is an 8:1 male bias in brain volumetric studies and a 15:1 bias in functional neuroimaging studies (Lai et al., 2015). Studies of autistic females, even if it may be limited by smaller samples, are therefore important to identify possible sex-specific risk factors. Sex gives a unique perspective to understand the underlying etiologies in autism, and should become as a core principle in autism research to further explore the heterogeneity of this neurodevelopmental condition (Rutter et al., 2003; Lai et al., 2015; Loomes et al., 2017; Ecker et al., 2017). Understanding sexual dimorphism in ASD could also potentially lead to new and targeted treatment strategies.

In Chapter 4, we explore the association between CNV burden with growth abnormalities in the SEED Study, and examine the potential association of growth abnormalities with autism-associated CNVs. Lai et al. (2015) suggested early growth trajectories in autism may show differential trajectories, where ASD females should be considered different, instead of simply more severe, than ASD males.

Study Limitations and Strengths

This is a cross-sectional study, and children were only assessed once between the ages of 2 and 5 years. Unfortunately, we were unable to draw a more comprehensive characterization of developmental changes pertaining to physical growth as in a longitudinal study design. Growth in children is a developmental process with sensitive periods and trajectories of change over age. As described earlier, growth trajectories (and their deviations) are time-sensitive, with abnormal growth changes occurring often from 6 months and starting to decline in severity after 2 years of age. Thus, whenever a growth measure is performed, it is integral to not just the results of any one analysis, but also the interpretation of these results. Absence of any evidence of overgrowth in autistic boys may reflect this limitation in our study design.

Also, despite this being a reasonably-sized study of children with ASD, for certain growth abnormalities examined here, e.g. tall stature based on the DRF cut-offs and low BMI using the decile cut-offs (stratifying by sex), we had limited statistical power due to the small numbers of children with abnormalities. This was also an issue with the assessment of certain patterns of trivariate growth phenotype, where the stringent DRF cut-off, resulted in a number of empty cells for both ASD and control children, (e.g. the combination of macrocephaly, overweight and short stature). Using the less stringent decile thresholds, each cell was populated by at least one individual, although there were some cells without representation for one or another group, (e.g. the combination of

macrocephaly, underweight and normal height in ASD children). Thus, this study may be underpowered to test certain patterns of trivariate growth phenotype associated with ASD.

An important limitation is the possibility there may be undiagnosed chromosomal abnormalities and recognized genetic syndromes left among our study sample. The exclusion of these children was based only on parental report; children did not undergo genetic evaluation. These children with undetected genetic syndromes could influence the effect estimate assessed as these conditions are also associated with ASD and may be the cause of growth abnormalities. This important issue could be addressed in future work.

Despite the limitations described above, this study also has many strengths. The study ascertainment was population-based, and involved individuals from six sites across the United States, with various racial groups reflecting the racial composition of the source populations (Schendel et al., 2012). Large numbers of individuals with ASD were assessed, including 99 ASD girls. Including girls in ASD studies can be challenging due to lower prevalence in females, and stratified analysis by sex can become more difficult. Stratification was not attempted in several studies of autism and growth due to sample size issues. Some studies do not even include autistic girls in their final analysis. Inclusion of autistic girls in ASD studies is increasingly recognized as quite valuable, and may give rare insights into potential risk factors. Growth abnormality in autistic phenotypes is also an important area to further explore, especially in terms of the variable

sex-specific presentations, and whether this may potentially be associated with in terms of etiology, management and prognosis.

SEED also provided rigorously obtained exposure and outcome data. All SEED sites conducted a uniform assessment of ASD based on a standard protocol across sites delivered and adjudicated by research-reliable professionals, within ongoing quality control (QC) throughout the data collection phase. Physical measurements were taken using standardized growth measurements across all sites, with frequent QC. These provided data on multiple growth modalities. Previous studies often concentrated on just one aspect of growth abnormality, for example, macrocephaly, in ASD children.

This study adds to the literature in ASD research, increasing our understanding of the autism phenotypes related to somatic growth, and highlighting the importance of studying growth in ASD girls. These findings will hopefully contribute to our understanding of ASD, and with future studies, offer hope for early detection, intervention and potentially prevention. These autistic sub-phenotypes may help parse out risk factors, particularly genetic risk factors leading to autism and particular autistic sub-phenotypes. Identifying these individuals may improve detection of these genetic risk factors, by examining the confluence of genetic abnormalities leading to both autism and specific sub-phenotypes.

For future work, it would be useful to expand this study to a longitudinal analysis, to further understand developmental trajectories for somatic growth. Another avenue to be explored is considering assessing other ASD co-morbidities (e.g. intellectual impairment)

as a covariate for certain growth abnormalities, and potentially other covariates. Finally, for future projects in this area, the design of the study could be improved by obtaining parental growth measurements, and genetic analyses incorporating whole exome sequencing and whole genome sequencing on these subjects. Sequencing information would have helped to eliminate undiagnosed chromosomal abnormalities from our study population, to better delineate growth phenotypes associated with ASD as opposed to single chromosomal abnormalities.

References

Akshoomoff, N., Pierce, N.K., Courchesne, E. (2002). The neurobiological basis of autism from a developmental perspective. *Dev Psychopathol*, 14:613-634

Ahearn, W.H., Castine, T., Nault, K., Green, G. (2001) An assessment of food acceptance in children with autism or pervasive developmental disorder-not otherwise specified. *J Autism Dev Disord*., 31(5):505-11

Amaral, D. G., Li, D., Libero, L., Solomon, M., Van de Water, J., Mastergeorge, A., … Wu Nordahl, C. (2017). In pursuit of neurophenotypes: The consequences of having autism and a big brain. *Autism Research*, 10(5), 711–722.

Barnard-Brak, L., Sulak, T., & Ivey Hatz, J. K. (2011). Macrocephaly in children with autism spectrum disorders. *Pediatric Neurology*, 44(2), 97–100.

Berg, A. T., Plioplys, S., & Tuchman, R. (2011). Risk and correlates of autism spectrum disorder in children with epilepsy: a community-based study. *Journal of Child Neurology*, 26, 540–547.

Bicer, A. H., & Alsaffar, A. A. (2013). Body mass index, dietary intake and feeding problems of Turkish children with autism spectrum disorder (ASD). *Research in Developmental Disabilities*, 34(11), 3978–3987.

Broder-Fingert, S., Brazauskas, K., Lindgren, K., Iannuzzi, D., & Van Cleave, J. (2014). Prevalence of overweight and obesity in a large clinical sample of children with autism. *Academic Pediatrics*, 14(4), 408–414.

Buescher, A. V. S., Cidav, Z., Knapp, M., & Mandell, D. S. (2014). Costs of Autism Spectrum Disorders in the United Kingdom and the United States. *JAMA Pediatrics*, 168(8), 721.

Campbell, D. J., Chang, J., & Chawarska, K. (2014). Early generalized overgrowth in autism spectrum disorder: Prevalence rates, gender effects, and clinical outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, 53(10), 1063–1073.

CDC clinical growth charts. https://www.cdc.gov/growthcharts/clinical_charts.htm#Set2

CDC 2000 Growth Charts for the United States: Methods and development. (2000). *Vital and Health Statistics*, 11:246. https://www.cdc.gov/growthcharts/2000growthchart-us.pdf

The Centers for Disease Control and Prevention (CDC). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States.

Surveillance Summaries. (2012). 65(3); 1–23
http://www.cdc.gov/mmwr/volumes/65/ss/ss6503a1.htm

Carayol, J., Schellenberg, G.D., Dombroski, B., Genin, E., Rousseau, F., Dawson, G. (2011). Autism risk assessment in siblings of affected children using sex-specific genetic scores. *Mol Autism*. 2(1):17.

Chaste, P., Klei, L., Sanders, S. J., Murtha, M. T., Hus, V., Lowe, J. K., … Kim, S. J. (2013). Adjusting head circumference for covariates in autism: Clinical correlates of a highly heritable continuous trait. *Biological Psychiatry*, 74(8), 576–584.

Chawarska, K., Campbell, D., Chen, L., Shic, F., Klin, a., & Chang, J. (2011). Early Generalized Overgrowth in Boys With Autism. *Archives of General Psychiatry*, 68(10), 1021–1031.

Chen, A.Y., Kim, S.E., Houtrow, A.J., Newacheck, P.W. (2010) Prevalence of obesity among children with chronic conditions. Obesity (Silver Spring).18(1):210-3.

Cheung, P. C., Cunningham, S. A., Naryan, K. M., & Kramer, M. R. (2016). Childhood obesity incidence in the United States: a systematic review. *Childhood Obesity*, 12(1), 1–11.

Christensen, J., Overgaard, M., Parner, E. T., Vestergaard, M. and Schendel, D. (2016), Risk of epilepsy and autism in full and half siblings—A population-based cohort study. *Epilepsia*, 57: 2011–2018.

Courchesne, E., Karns, C. M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D., … Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study.*Neurology*, 57(2), 245–254.

Courchesne, E., Carper, R., & Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. *JAMA*, 290(3), 337–44.

Curtin, C., Bandini, L.G., Perrin, E.C., Tybor, D.J., Must, A. (2005). Prevalence of overweight in children and adolescents with attention deficit hyperactivity disorder and autism spectrum disorders: a chart review. *BMC Pediatr.*, 5:48.

Curtin, C., Anderson, S. E., Must, A., & Bandini, L. G. (2010). The prevalence of obesity in children with autism: a secondary data analysis using nationally representative data from the National Survey of Children's Health. *BMC Pediatrics*, 10, 11.

Davidovitch, M., Golan, D., Vardi, O., Lev, D., & Lerman-Sagie, T. (2011). Israeli Children With Autism Spectrum Disorder Are not Macrocephalic. Journal of *Child Neurology*, 26(5), 580–585.

Dawson, G., Munson, J., Webb, S.J., Nalty, T., Abbott, R., Toth, K.. (2007) Rate of head

growth decelerates and symptoms worsen in the second year of life in autism. *Biol. Psychiatry*. 15;61(4):458-64

Dementieva, Y.A., Vance, D.D., Donnelly, S.L., Elston, L.A., Wolpert, C.M., Ravan, S.A., DeLong, G.R., Abramson, R.K., Wright, H.H., Cuccaro, M.L. (2005). Accelerated head growth in early development of individuals with autism. *Pediatr Neurol*.,32(2):102-8

De Onis, M., Garza, C., Victora, C.G., Bhan, M.K., Norum, K.R., eds (2004a). WHO Multicentre Growth Reference Study (MGRS): Rationale, planning and implementation. *Food and Nutrition Bulletin*, 25(Suppl 1):S1–S89.

De Onis, M., Garza, C., Onyango, A. W., & Borghi, E. (2007). Comparison of the WHO Child Growth Standards and the CDC 2000 Growth Charts 1. *J. Nutr*, 137, 144–148.

Deutsch, S.I., Campbell, M., Sachar, E.J., Green, W.H., David, R. (1985). Plasma growth hormone response to oral l-dopa in infantile autism. *J Autism Dev Disord*.15(2):205-12.

de Vinck-Baroody, O., Shui, A., Macklin, E.A., Hyman, S.L., Leventhal, J.M., Weitzman, C. (2015). Overweight and Obesity in a Sample of Children With Autism Spectrum Disorder. *Acad Pediatr*. 15(4):396-404.

Dibley, M.J., Goldsby, J.B., Staehling, N.W., Trowbridge, F.L. (1987). Development of normalized curves for the international growth reference: Historical and technical considerations. *Am J Clin Nutr* 46:736–48.

Diguiseppi, C. G., Daniels, J. L., Fallin, D. M., Rosenberg, S. A., Schieve, L. A., Thomas, K. C., … Schendel, D. E. (2016). Demographic profile of families and children in the Study to Explore Early Development (SEED): Case-control study of autism spectrum disorder. *Disability and Health Journal*, *9*(3), 544–551.

Dinstein, I., Haar, S., Atsmon, S., & Schtaerman, H. (2017). No evidence of early head circumference enlargements in children later diagnosed with autism in Israel. *Molecular Autism*, 8(1), 15.

Dissanayake, C., Bui, Q. M., Huggins, R., & Loesch, D. Z. (2006). Growth in stature and head circumference in high-functioning autism and Asperger disorder during the first 3 years of life. Development and Psychopathology, 18(2), 381–93.

Ecker, C., Andrews, D. S., Gudbrandsen, C. M., Marquand, A. F., Ginestet, C. E., Daly, E. M., … Murphy, D. G. M. (2017). Association Between the Probability of Autism Spectrum Disorder and Normative Sex-Related Phenotypic Diversity in Brain Structure. *JAMA Psychiatry*, 74(4), 329.

Egan, A. M., Dreyer, M. L., Odar, C. C., Beckwith, M., & Garrison, C. B. (2013). Obesity in Young Children with Autism Spectrum Disorders: Prevalence and Associated Factors. *Childhood Obesity*, 9(2), 125–131.

Eisenmajer, R., Prior, M., Leekam, S., Wing, L., Gould, J., Welham, M., Ong, B. (1996). Comparison of clinical symptoms in autism and Asperger's disorder. *J Am Acad Child Adolesc Psychiatry*. 35(11):1523-31.

Evans, E. W., Must, A., Anderson, S. E., Curtin, C., Scampini, R., Maslin, M., & Bandini, L. (2012). Dietary patterns and body mass index in children with autism and typically developing children. *Research in Autism Spectrum Disorders*, 6(1), 399–405.

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: a systematic literature review. *European Child & Adolescent Psychiatry*, 22, 329–340.

Fidler, D.J., Bailey, J.N., Smalley, S.L. (2000). Macrocephaly in autism and other pervasive developmental disorders. *Dev Med Child Neurol*. 42(11):737-40.

Fombonne, E., Rogé, B., Claverie, J., Courty, S., Frémolle, J. (1999). Microcephaly and macrocephaly in autism. *J Autism Dev Disord*. 29(2):113-9.

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: an update. *J Autism Dev Disord*. 33(4):365-82.

Fombonne, E. Epidemiology of pervasive developmental disorders. *Pediatr Res*. 2009 Jun; 65(6):591-8

Fukumoto, A., Hashimoto, T., Ito, H., Nishimura, M., Tsuda, Y., Miyazaki, M., … Kagami, S. (2008). Growth of head circumference in autistic infants during the first year of life. *Journal of Autism and Developmental Disorders*, 38(3), 411–418.

Geraedts, E. J., van Dommelen, P., Caliebe, J., Visser, R., Ranke, M. B., van Buuren, S., Wit, J. M., Oostdijk, W. (2011). Association between Head Circumference and Body Size. *Horm Res Paediatr*. 75:213-219

Ghaziuddin, M., Zaccagnini, J., Tsai, L., Elardo, S. (1999). Is megalencephaly specific to autism? *J Intellect Disabil Res*.43 (Pt 4):279-82.

Gillberg, I. C., Gillberg, C., & Kopp, S. (1992). Hypothyroidism and autism spectrum disorders. *J Child Psychol Psychiatry*, 33(3), 531–542.

Goldenberg, A., Chevy, F., Bernard, C., Wolf, C., & Cormier-Daire, V. (2003). Circonstances cliniques du diagnostic du syndrome de Smith-Lemli-Opitz et

tentatives de corrélation phénotype-génotype: À propos de 45 cas. Archives de Pediatrie, 10(1), 4–10.

Gillberg, I. C., Gillberg, C., & Kopp, S. (1992). Hypothyroidism and autism spectrum disorders. *J Child Psychol Psychiatry*, 33(3), 531–542.

Growth Parameters Measurements in SEED Q:\DDB BANCH|GA CADDRE\dysmorphology\April 07 – dys training and manual\Manual\04b-Growth Parameters APR 07.doc

Hazlett, H., Poe, M., Gerig, G., & al, et. (2005). Magnetic resonance imaging and head circumference study of brain size in autism: Birth through age 2 years. *Archives of General Psychiatry*, 62(12), 1366–1376.

Head, A.M., McGillivray, J.A., Stokes, M.A. (2014). Gender differences in emotionality and sociability in children with autism spectrum disorders. *Mol Autism*. 28;5(1):19.

Hill, A. P., Zuckerman, K. E., & Fombonne, E. (2015). Obesity and Autism. *Pediatrics*, 136(6), 1051–1061.

Ho, H.H., Eaves, L.C. (1997). Kabuki make-up (Niikawa-Kuroki) syndrome: cognitive abilities and autistic features. *Dev Med Child Neurol*. 39(7):487-90.

Hyman, S.L., Stewart, P.A., Schmidt, B., Cain, U., Lemcke, N., Foley, J.T., Peck, R., Clemons, T., Reynolds, A., Johnson, C., Handen, B., James, S.J., Courtney, P.M., Molloy, C., Ng, P.K. (2012). Nutrient intake from food in children with autism. *Pediatrics*.130 Suppl 2:S145-53.

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, 515(7526), 216-221.

Jacquemont, S., Coe, B. P., Hersch, M., Duyzend, M. H., Krumm, N., Bergmann, S., et al. (2014). A higher mutational burden in females supports a "female protective model" in neurodevelopmental disorders. *American Journal of Human Genetics*, 94(3), 415-425.

Kanner, L. (1943). Autistic disturbance of affective contact. *Nervous Child*, 2, 217–250.

Karimi, P., Kamali, E., Mousavi, S. M.,& Karahmadi, M. (2017). Environmental factors influencing the risk of autism. *Journal of Research in Medical Sciences*, 22, 27.

Krumm, N., Turner, T. N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., et al. (2015). Excess of rare, inherited truncating mutations in autism. *Nature Genetics*, 47(6), 582-588.

Kuczmarski, R.J., Ogden, C.L., Grummer-Strawn, L.M., Flegal, K.M., Guo, S.S., Wei, R., Mei, Z., Curtin, L.R., Roche, A.F., Johnson, C.L. (2000). CDC growth charts: United States. Adv Data. 314:1-27.

Kuczmarski, R.J., Ogden, C.L., Guo, S.S., et al. (2002). 2000 CDC growth charts for the United States: methods and development. *Vital Health Stat* 11. 246:1

Lai, M.-C., Lombardo, M. V., Suckling, J., Ruigrok, A. N. V, Chakrabarti, B., Ecker, C., … Baron-Cohen, S. (2013). Biological sex affects the neurobiology of autism. *Brain*, 136(9), 2799–2815.

Lai, M.-C., Baron-Cohen, S., & Buxbaum, J. D. (2015). Understanding autism in the light of sex/gender. *Molecular Autism*, 6(1), 24.

Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *Journal of the American Academy of Child and Adolescent Psychiatry*, 54(1), 11–24.

Lai, M.-C., Lerch, J. P., Floris, D. L., Ruigrok, A. N.V., Pohl, A., Lombardo, M. V. and Baron-Cohen, S. (2017), Imaging sex/gender and autism in the brain: Etiological implications. *Journal of Neuroscience Research*, 95: 380–397.

Lainhart, J.E., Piven, J., Wzorek, M., Landa, R., Santangelo, S.L., Coon, H., Folstein, S.E.(1997). Macrocephaly in children and adults with autism. *J Am Acad Child Adolesc. Psychiatry.* 36(2):282-90.

Lainhart, J. E., Bigler, E. D., Bocian, M., Coon, H., Dinh, E., Dawson, G., Deutsch, C. K., Dunn, M., Estes, A., Tager-Flusberg, H., Folstein, S., Hepburn, S., Hyman, S., McMahon, W., Minshew, N., Munson, J., Osann, K., Ozonoff, S., Rodier, P., Rogers, S., Sigman, M., Spence, M. A., Stodgell, C. J. and Volkmar, F. (2006), Head circumference and height in autism: A study by the collaborative program of excellence in autism. *Am. J. Med. Genet*., 140A: 2257–2274.

Lavelle, T.A., Weinstein, M.C., Newhouse, J.P., Munir, K., Kuhlthau, K.A., Prosser, L.A. (2014). Economic Burden of Childhood Autism Spectrum Disorders. *Pediatrics,* 133(3):e520-9

Lettre, G. (2009). Genetic regulation of adult stature. *Curr Opin Pediatr*. 21(4):515-22.

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(6), 466–474.

Lord, C., Risi, S., Lambrecht, L., et al. (2000). The autism diagnostic observation

schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord*, 30(3):205–223

Lord, C., Rutter, M., Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*, 24: 659-685

Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., … Newschaffer, C. (2017). The Changing Epidemiology of Autism Spectrum Disorders. *Annual Review of Public Health*, 38(1), 81–102.

Marouli, E., Graff, M., Medina-Gomez, C., Lo, K. S., Wood, A. R., Kjaer, T. R., … Lettre, G. (2017). Rare and low-frequency coding variants alter human adult height. *Nature*, 542(7640), 186–190.

Memari, A. H., Kordi, R., Ziaee, V., Mirfazeli, F. S., & Setoodeh, M. S. (2012). Weight status in Iranian children with autism spectrum disorders: Investigation of underweight, overweight and obesity. *Research in Autism Spectrum Disorders*, 6(1), 234–239.

Miles, J.H., Hadden, L.L., Takahashi, T.N., Hillman, R.E. (2000). Head circumference is an independent clinical finding associated with autism. *Am J Med Genet*.11;95(4):339-50.

Mraz, K. D., Green, J., Dumont-mathieu, T., Makin, S., & Fein, D. (2007). Correlates of Head Circumference Growth, 7–9.

Must, A., Eliasziw, M., Phillips, S. M., Curtin, C., Kral, T. V. E., Segal, M., … Bandini, L. G. (2016). The Effect of Age on the Prevalence of Obesity among US Youth with Autism Spectrum Disorder. *Childhood Obesity* (Print), X(X), 1–11.

Nordahl, C. W., Iosif, A.-M., Young, G. S., Perry, L. M., Dougherty, R., Lee, A., … Amaral, D. G. (2015). Sex differences in the corpus callosum in preschool-aged children with autism spectrum disorder. *Molecular Autism*, 6(1), 26.

Phillips, K. L., Schieve, L. A., Visser, S., Boulet, S., Sharma, A. J., Kogan, M. D., … Yeargin-Allsopp, M. (2014). Prevalence and Impact of Unhealthy Weight in a National Sample of US Adolescents with Autism and Other Learning and Behavioral Disabilities. *Maternal and Child Health Journal*, 18(8), 1964–1975.

Raznahan, A., Wallace, G. L., Antezana, L., Greenstein, D., Lenroot, R., Thurm, A., … Giedd, J. N. (2013). Compared to what? Early brain overgrowth in autism and the perils of population norms. *Biological Psychiatry*, 74(8), 563–575.

Redcay, E., Courchesne, E. (2005). When Is the Brain Enlarged in Autism? A Meta-Analysis of All Brain Size Reports. *Biological Psychiatry*, Volume 58 , Issue 1 ,

1 - 9

Retico, A., Giuliano, A., Tancredi, R., Cosenza, A., Apicella, F., Narzisi, A., … Calderoni, S. (2016). The effect of gender on the neuroanatomy of children with autism spectrum disorders: a support vector machine case-control study. *Molecular Autism*, 7, 5.

Rimmer, J.H., Yamaki, K., Lowry, B.M., Wang, E., Vogel, L.C. (2010). Obesity and obesity-related secondary conditions in adolescents with intellectual/developmental disabilities. *J Intellect Disabil Res*. 54(9):787-94.

Robinson, E. B., Lichtenstein, P., Anckarsäter, H., Happé, F., & Ronald, A. (2013). Examining and interpreting the female protective effect against autistic behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 110(13), 5258-5262.

Rodier, P.M., Bryson, S.E., Welch, J.P. (1997). Minor malformations and physical measurements in autism: data from Nova Scotia. *Teratology*. 55(5):319-25.

Rutter, M., Caspi, A., Moffitt, T.E. (2003). Using sex differences in psychopathology to study causal mechanisms: unifying issues and research strategies. *J Child Psychol Psychiatry*. 44(8):1092–115.

Sacco, R., Gabriele, S., & Persico, A. M. (2015). Head circumference and brain size in autism spectrum disorder: A systematic review and meta-analysis. *Psychiatry Research - Neuroimaging*, 234(2), 239–251.

Sanders, S., Ercan-sencicek, A. G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., & State, M. (2011). Multiple recurrent de novo copy number variations (CNVs), including duplications of the 7q11.23 Williams-Beuren syndrome region, are strongly associated with autism. *Neuron*, *70*(5), 863–885.

Scheepens, A., Moderscheim, T.A., Gluckman, P.D. (2005). The role of growth hormone in neural development. *Horm Res*, 64:66-72

Schendel, D. E., Diguiseppi, C., Croen, L. a, Fallin, M. D., Reed, P. L., Schieve, L. a, … Yeargin-Allsopp, M. (2012). The Study to Explore Early Development (SEED): a multisite epidemiologic study of autism by the Centers for Autism and Developmental Disabilities Research and Epidemiology (CADDRE) network. *Journal of Autism and Developmental Disorders*, 42(10), 2121–40.

Shapira, K., Tian, L., Aylsworth, A., Elias, E., Hoover-Fong, J., Meeks, N., Souders, M., Tsai, C., Alexander, A., Schieve, L. (2014). Development of a Novel Protocol for Characterizing Dysmorphology to Enhance the Phenotypic Classification of Autism Spectrum Disorders, presented at International Meeting for Autism Research, Atlanta, 2014. Georgia: INSAR.

Siervogel, R.M., Roche, A.F., Guo, S.M., et al. (1991). Patterns of change in weight/stature2 from 2 to 18 years: Findings from long-term serial data for children in the Fels Longitudinal Growth Study. Int J Obes 15:479–85.

Suren, P., Stoltenberg, C., Bresnahan, M., Hirtz, D., Lie, K. K., Lipkin, W. I., … Hornig, M. (2013). Early growth patterns in children with autism. *Epidemiology* (Cambridge, Mass.), 24(5), 660–670.

Szatmari, P., Jones, M.B. (1991). IQ and the genetics of autism. *J Child Psychol Psychiatry*. 32(6):897-908.

Tate, D.F., Bigler, E.D., McMahon, W., Lainhart, J. (2007). The relative contributions of brain, cerebrospinal fluid-filled structures and non-neural tissue volumes to occipital-frontal head circumference in subjects with autism. *Neuropediatrics,* 38(1):18-24.

Torrey, E. F., Dhavale, D., Lawlor, J. P., & Yolken, R. H. (2004). Autism and head circumference in the first year of life. *Biological Psychiatry*, 56(11), 892–894.

Tripi, G., Roux, S., Canziani, T., Brilhault, F. B., Barthélémy, C., & Canziani, F. (2008). Minor physical anomalies in children with autism spectrum disorder. Early Human Development, 84(4), 217–223.

van Daalen, E., Swinkels, S. H. N., Dietz, C., van Engeland, H., & Buitelaar, J. K. (2007). Body Length and Head Growth in the First Year of Life in Autism. *Pediatric Neurology*, 37(5), 324–330.

Walsh, P., Elsabbagh, M., Bolton, P., & Singh, I. (2011). In search of biomarkers for autism: scientific, social and ethical challenges. *Nat Rev Neurosci*, 12(10), 603–612.

Wells, J. (2007). Sexual dimorphism of body composition. *Best Practice & Research Clinical Endocrinology & Metabolism*. 21(3), 415-430.

Werling, D. M., & Geschwind, D. H. (2013). Sex differences in autism spectrum disorders. *Current Opinion in Neurology*, 26(2), 146–153.

Werling, D.M., Lowe, J.K., Luo, R,. Cantor, R.M., Geschwind, D.H. (2014). Replication of linkage at chromosome 20p13 and identification of suggestive sex-differential risk loci for autism spectrum disorder. *Mol Autism*. 5(1):13.

Werling, D. M. (2016). The role of sex-differential biology in risk for autism spectrum disorder. *Biology of Sex Differences*, 7, 58.

WHO (1995). Physical status: the use and interpretation of anthropometry. Report of a

WHO Expert Committee. Technical Report Series No. 854. Geneva: World Health Organization.

WHO Multicentre Growth Reference Study Group (2006a). Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. *Acta Paediatrica*, Suppl 450:57–66.

WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age (2006) www.who.int/childgrowth/standards/Technical_report.pdf

Wiggins, L. D., Reynolds, A., Rice, C. E., Moody, E. J., Bernal, P., Blaskey, L., … Levy, S. E. (2015). Using Standardized Diagnostic Instruments to Classify Children with Autism in the Study to Explore Early Development. *Journal of Autism and Developmental Disorders*, 45(5), 1271–1280

Xiong, N., Ji, C., Li, Y., He, Z., Bo, H., & Zhao, Y. (2009). The physical status of children with autism in China. *Research in Developmental Disabilities*, 30(1), 70–76.

van Daalen, E., Swinkels, S. H. N., Dietz, C., van Engeland, H., & Buitelaar, J. K. (2007). Body Length and Head Growth in the First Year of Life in Autism. *Pediatric Neurology*, 37(5), 324–330.

Vanhala, R., Turpeinen, U., Riikonen, R. (2001). Low levels of insulin-like growth factor-I in cerebrospinal fluid in children with autism. *Dev Med Child Neurol*, 43: 614-616.

Vargas, D.L., Nascimbene, C., Krishnan, C., Zimmerman, A.W., Pardo, P. A. (2005). Neuroglial activation and neuroinflammation in the brain of patients with autism *Ann Neurol*, 57:67-81

Zuckerman, K. E., & Fombonne, E. (2015). Obesity and Autism. *Pediatrics*, 136(6), 1051–1061.

Zwaigenbaum, L., Young, G. S., Stone, W. L., Dobkins, K., Ozonoff, S., Brian, J., … Messinger, D. (2014). Early head growth in infants at risk of autism: A baby siblings research consortium study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 53(10), 1053–1062.

Tables

Table 1: Baseline characteristics of study sample

Baseline Characteristics	**POP**	**ASD**	**P-value**
N (%)	381 (41.8)	532 (58.2)	
Sex			
Male	207 (54.3)	433 (81.4)	**<0.01**
Female	174 (45.7)	99 (18.6)	
Mother's Highest Education			
Less Than High School	19 (5.0)	36 (6.7)	0.53
High School	51 (13.5)	80 (15.0)	
Some College/ Trade	108 (28.3)	156 (29.4)	
Bachelor's Degree	121 (31.7)	164 (30.9)	
Advanced Degree	82 (21.5)	96 (18)	
Race/ Ethnicity			
Non-Hispanic White	216 (56.7)	336 (63.1)	0.23
Non-Hispanic Black	96 (25.2)	121 (22.7)	
Hispanic	66 (17.3)	72 (13.5)	
Others	3 (0.8)	3 (0.6)	

Bolded: $p<0.05$

Table 2: Percentile means overall (ASD and POP) for head circumference, weight, stature, and BMI

Overall N=913	ASD (N=532)			POP (N=381)		
	Mean %	**95% CI**	**SE**	**Mean %**	**95% CI**	**SE**
HC	44.81	42.06, 47.55	1.39	42.00	38.94, 45.06	1.55
Weight	55.91	53.33, 59.27	1.51	56.30	53.35, 58.46	1.30
Height	56.08	53.48, 58.69	1.32	59.34	56.43, 62.26	1.48
BMI	59.84	57.30, 62.37	1.29	58.04	55.1, 60.91	1.46

HC: Head Circumference

Table 3: Percentile means for head circumference, weight, stature and BMI stratified by sex

Males	**ASD (N=433)**			**POP (N=207)**			**P-value**
	Mean %	**95% CI**	**SE**	**Mean %**	**95% CI**	**SE**	
HC	45.30	42.31,48.28	1.52	43.29	39.20,47.37	2.08	0.44
Weight	55.81	52.97, 58.65	1.44	56.73	52.72,60.74	2.04	0.71
Height	57.39	54.57,60.21	1.43	57.66	53.70,61.61	2.01	0.91
BMI	58.92	56.06,61.78	1.45	57.12	53.10,61.15	2.05	0.47

Females	**ASD (N=99)**			**POP (N=174)**			**P-value**
	Mean %	**95% CI**	**SE**	**Mean %**	**95% CI**	**SE**	
HC	42.66	35.81,49.51	3.49	40.46	35.86,45.06	2.34	0.59
Weight	56.33	50.51,62.15	2.96	55.78	51.38,60.19	2.24	0.88
Height	50.35	43.79,56.91	3.34	61.37	57.08,65.66	2.18	**0.04**
BMI	63.76	58.40,69.13	2.73	59.15	55.09,63.20	2.06	0.18

Bolded: $p<0.05$

Table 4: Growth abnormalities in ASD and POP using DRF and Decile thresholds

Growth Abnormality	DRF			Decile		
	POP, N (%)	ASD, N (%)	Total, N (%)	POP, N (%)	ASD, N (%)	Total, N (%)
Total	381 (100)	532 (100)	913 (100)	381 (100)	532 (100)	913 (100)
Macrocephaly	19 (5.0)	30 (5.6)	54 (5.1)	35 (9.2)	61 (11.5)	96 (10.5)
Microcephaly	30 (7.9)	44 (8.3)	95 (8.9)	72 (18.9)	102 (19.2)	174 (19.1)
Tall Stature	5 (1.3)	8 (1.5)	17 (1.6)	69 (18.1)	84 (15.8)	153 (16.7)
Short Stature	16 (4.2)	41 (7.7)	80 (7.5)	28 (7.4)	56 (10.5)	84 (9.2)
Overweight	29 (7.6)	50 (9.4)	94 (8.8)	60 (15.7)	82 (15.4)	142 (15.5)
Underweight	19 (5.0)	31 (5.8)	65 (6.1)	37 (9.7)	57 (10.7)	94 (10.3)
High BMI	41 (10.8)	76 (14.3)	143 (13.4)	70 (18.4)	128 (24.1)	198 (21.7)
Low BMI	25 (6.6)	32 (6.0)	70 (6.6)	21 (5.5)	27 (5.1)	48 (5.2)

DRF: Dysmorphology Review Form. Thresholds based on these vary (refer to Figure 2)
Decile: Decile Percentile Curve thresholds at $\geq 90^{th}$ percentile for upper limit of growth (overgrowth) and $<10^{th}$ percentile for lower limit of growth (undergrowth)

Table 5: Growth abnormalities (count, (percent)) in ASD and POP using Decile thresholds stratified by sex

Growth Abnormality	Male			Female		
	POP, N= 207 (32.3%)	**ASD, N= 433 (67.6%)**	**Total, N= 640**	**POP, N= 174 (63.7%)**	**ASD, N= 99 (36.2%)**	**Total, N= 273**
Macrocephaly (≥90th)	19 (9.2)	46 (10.6)	65 (10.1)	16 (9.2)	15 (15.1)	31 (11.3)
Microcephaly (<10th)	37 (17.8)	79 (18.2)	116 (18.1)	35 (20.1)	23 (23.2)	58 (21.2)
Tall Stature (≥90th)	33 (15.9)	73 (16.8)	106 (16.5)	36 (20.7)	11 (11)	47 (17.2)
Short Stature (<10th)	17 (8.2)	37 (8.5)	54 (8.4)	11 (6.3)	19 (19.2)	30 (10.9)
Overweight (≥90th)	36 (17.4)	66 (15.2)	102 (15.9)	24 (13.8)	16 (16.1)	40 (14.6)
Underweight (<10th)	17 (8.2)	48 (11.1)	65 (10.1)	20 (11.5)	9 (9.1)	29 (10.6)
High BMI (≥90th)	38 (18.4)	106 (24.5)	144 (22.5)	32 (18.4)	22 (22.2)	54 (19.7)
Low BMI (<10th)	15 (7.2)	23 (5.3)	38 (5.9)	6 (3.4)	4 (4.0)	10 (3.6)

Table 6: Odds Ratios of growth abnormalities in ASD children compared to POP children using the DRF thresholds (unadjusted and adjusted for race)

ASD, N=532 POP, N=381	**Unadjusted Odds Ratio**	**95% Confidence Interval**	**p-value**	**Odds Ratio Adjusted for Race**	**95% Confidence Interval**	**p-value**
Macrocephaly	1.14	0.63,2.05	0.66	1.10	0.61,1.99	0.74
Microcephaly	1.05	0.65,1.71	0.83	1.13	0.69,1.84	0.62
Tall Stature	1.15	0.37,3.53	0.81	1.12	0.36,3.48	0.83
Short Stature	**1.90**	1.05, 3.45	0.03	**1.92**	1.06,3.49	0.03
Overweight	1.26	0.78,2.03	0.34	1.26	0.78,2.04	0.33
Underweight	1.18	0.65,2.12	0.58	1.19	0.66,2.15	0.55
High BMI	1.38	0.92,2.07	0.11	1.41	0.94,2.12	0.09
Low BMI	0.91	0.53,1.56	0.73	0.91	0.53,1.56	0.73

ORs are compared to POP

Bolded: p-value <0.05 before correcting for multiple testing

* p-value <0.003 after correcting for multiple testing

Table 7: Odds Ratios of growth abnormalities in ASD children compared to POP children using the DRF thresholds adjusted for race and stratified by sex

Using DRF	**Odds Ratio In Males**	**95% Confidence Interval**	**p-value**	**Odds Ratio In Females**	**95% Confidence Interval**	**p-value**
Macrocephaly	1.17	0.53,2.60	0.68	1.28	0.45,3.36	0.67
Microcephaly	0.92	0.49,1.72	0.81	1.98	0.86,4.54	0.10
Tall Stature	0.88	0.21,3.59	0.86	1.83	0.25,13.34	0.54
Short Stature	1.20	0.58,2.48	0.61	**5.56***	1.94,15.97	0.001
Overweight	1.59	0.85,2.98	0.14	0.56	0.20,1.62	0.29
Underweight	1.20	0.56,2.56	0.62	1.20	0.41,3.49	0.73
High BMI	1.68	0.97,2.54	2.90	1.15	0.56,2.37	0.69
Low BMI	0.68	0.36,1.29	0.24	1.58	0.55,4.49	0.39
	ASD, N= 433 POP, N= 207			ASD, N= 99 POP, N= 174		

ORs are compared to POP

Bolded: p-value <0.05 before correcting for multiple testing

* p-value <0.003 after correcting for multiple testing

Table 8: Odds Ratios of growth abnormalities in ASD children compared to POP children using the Decile thresholds (unadjusted and adjusted for race)

ASD, N= 532 POP, N= 381	Unadjusted Odds Ratio	95% Confidence Interval	p-value	Odds Ratio Adjusted for Race	95% Confidence Interval	p-value
Macrocephaly	1.28	0.82,1.98	0.27	1.25	0.80,1.94	0.32
Microcephaly	1.01	0.72,1.42	0.91	1.04	0.74,1.46	0.80
Tall Stature	0.84	0.59,1.20	0.35	0.84	0.59,1.20	0.35
Short Stature	1.48	0.92,2.38	0.10	1.47	0.91,2.37	0.10
Overweight	0.97	0.67,1.40	0.89	0.98	0.68,1.41	0.93
Underweight	1.11	0.72,1.72	0.62	1.12	0.72,1.74	0.59
High BMI	**1.40**	1.01,1.95	0.04	**1.44**	1.04,2.00	0.03
Low BMI	0.91	0.51,1.64	0.77	0.88	0.49,1.59	0.69

ORs are compared to POP
Bolded: p-value <0.05 before correcting for multiple testing
* p-value <0.003 after correcting for multiple testing

Table 9: Odds Ratios of growth abnormalities in ASD children compared to POP children using the Decile thresholds adjusted for race and stratified by sex

Using DRF	Odds Ratio In Males	95% Confidence Interval	p-value	Odds Ratio In Females	95% Confidence Interval	p-value
Macrocephaly	1.12	0.64,1.98	0.68	1.76	0.83,3.74	0.14
Microcephaly	1.05	0.68,1.62	0.82	1.22	0.67,2.22	0.51
Tall Stature	1.09	0.69,1.71	0.71	**0.47**	0.22,0.97	0.04
Short Stature	1.02	0.56,1.87	0.93	**3.52***	1.60,7.76	0.002
Overweight	0.86	0.55,1.33	0.51	1.21	0.61,2.41	0.58
Underweight	1.38	0.77,2.48	0.27	0.77	0.34,1.78	0.55
High BMI	1.48	0.97,2.24	0.06	1.28	0.69,2.37	0.42
Low BMI	0.69	0.35,1.35	0.28	1.16	0.32,4.23	0.81
	ASD, N= 433 POP, N= 207			ASD, N= 99 POP, N= 174		

ORs are compared to POP
Bolded: p-value <0.05 before correcting for multiple testing
* p-value <0.003 after correcting for multiple testing

Table 10: Trivariate Growth Phenotype Frequencies in ASD and POP groups using the DRF thresholds for growth abnormalities

ASD, N= 532 POP, N= 381	Overweight			Normal Weight			Underweight		
	Tall	Normal Height	Short	Tall	Normal Height	Short	Tall	Normal Height	Short
Macro-cephaly	3(0.5%) 2(0.5%) *1.00*	12(2.2%) 7(1.8%) *0.81*	0 0 --	0 0 --	15(2.8%) 10(2.6%) *1.00*	0 0 --	0 0 --	0 0 --	0 0 --
Normal Head Circum-ference	5(0.9%) 0 *0.08*	29(5.4%) 19(4.9%) *0.88*	0 0 --	0 3(0.8%) *0.07*	395(74.2%) 290(76.1%) *0.53*	12(2.2%) 8(2%) *1.00*	0 0 --	7(1.3%) 8(2.1%) *0.43*	10(1.8%) 4(1%) *0.41*
Micro-cephaly	0 0 --	1(0.2%) 1(0.2%) *1.00*	0 0 --	0 0 --	20(3.7%) 22(5.7%) *0.19*	9(1.7%) 0 ***0.01***	0 0 --	4(0.7%) 3(0.8%) *1.00*	10(1.8%) 4(1%) *0.41*

Table 11: Trivariate Growth Phenotype Frequencies in ASD and POP groups using the Decile thresholds for growth abnormalities

ASD, N= 532 POP, N= 381	Overweight			Normal Weight			Underweight		
	Tall	Normal Height	Short	Tall	Normal Height	Short	Tall	Normal Height	Short
Macro-cephaly	14(2.6%) 6(1.5%) *0.36*	4(0.7%) 1(0.2%) *0.40*	1(0.2%) 1(0.2%) *1.00*	5(0.9%) 3(0.8%) *1.00*	25(4.7%) 14(3.6%) *0.51*	0 2(0.5%) *0.17*	6(1.1%) 5(1.3%) *1.00*	0 1(0.2%) *0.41*	6(1.1%) 2(0.5%) *0.48*
Normal Head Circum-ference	18(3.3%) 22(5.7%) *0.10*	33(6.2%) 17(4.4%) *0.30*	5(0.9%) 3(0.8%) *1.00*	25(4.7%) 25(6.5%) *0.24*	257(48.3%) 186(48.8%) *0.89*	9(1.7%) 8(2.1%) *0.80*	5(0.9%) 4(1.0%) *1.00*	7(1.3%) 7(1.8%) *0.59*	10(1.8%) 2(0.5%) *0.08*
Micro-cephaly	5(0.9%) 0 *0.70*	1(0.2%) 6(1.5%) ***0.023***	1(0.2%) 2(0.5%) *0.57*	3(0.5%) 1(0.2%) *0.64*	58(10.9%) 45(11.8%) *0.67*	11(2%) 0 ***0.003***	3(0.5%) 1(0.2%) *0.64*	7(1.3%) 7(1.8%) *0.59*	13(2.4%) 8(2.1%) *0.82*

Key for Tables 10 and 11

Blue: counts and percentages of ASD children with the TGP combination described
Red: counts and percentages of POP children with the TGP combination described
Black Italic: Fisher's p-value
Bolded: significant for Fisher's $p<0.05$
Decile: Decile Percentile Curve cut-offs at $\geq 90^{th}$ percentile for upper limit of growth (overgrowth) and $\leq 10^{th}$ percentile for lower limit of growth (undergrowth)
TGP: Trivariate growth phenotype

Figures

Figure 1: Study Flow Cart

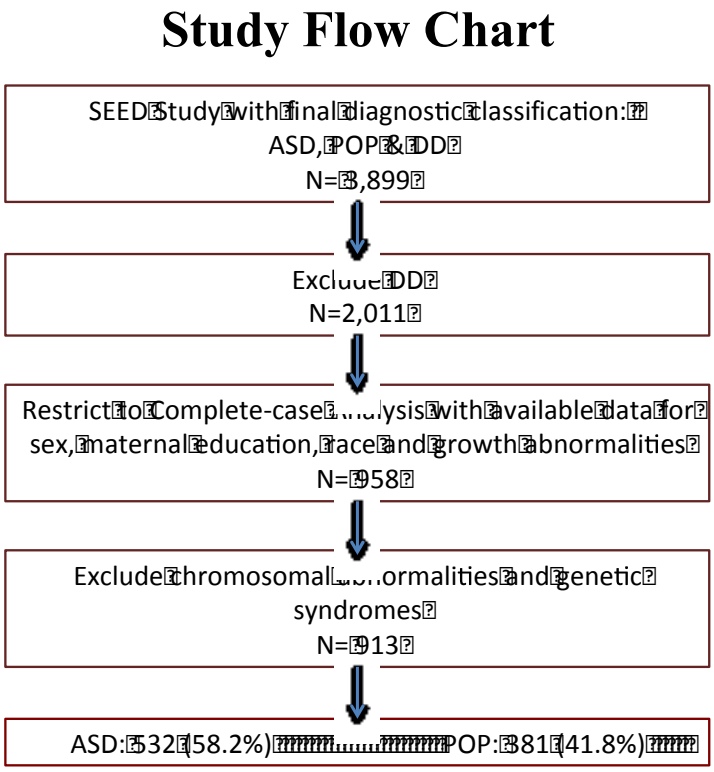


Figure 2: Percentile distribution of a. Head Circumference b. Height c. Weight d. BMI for ASD and POP children

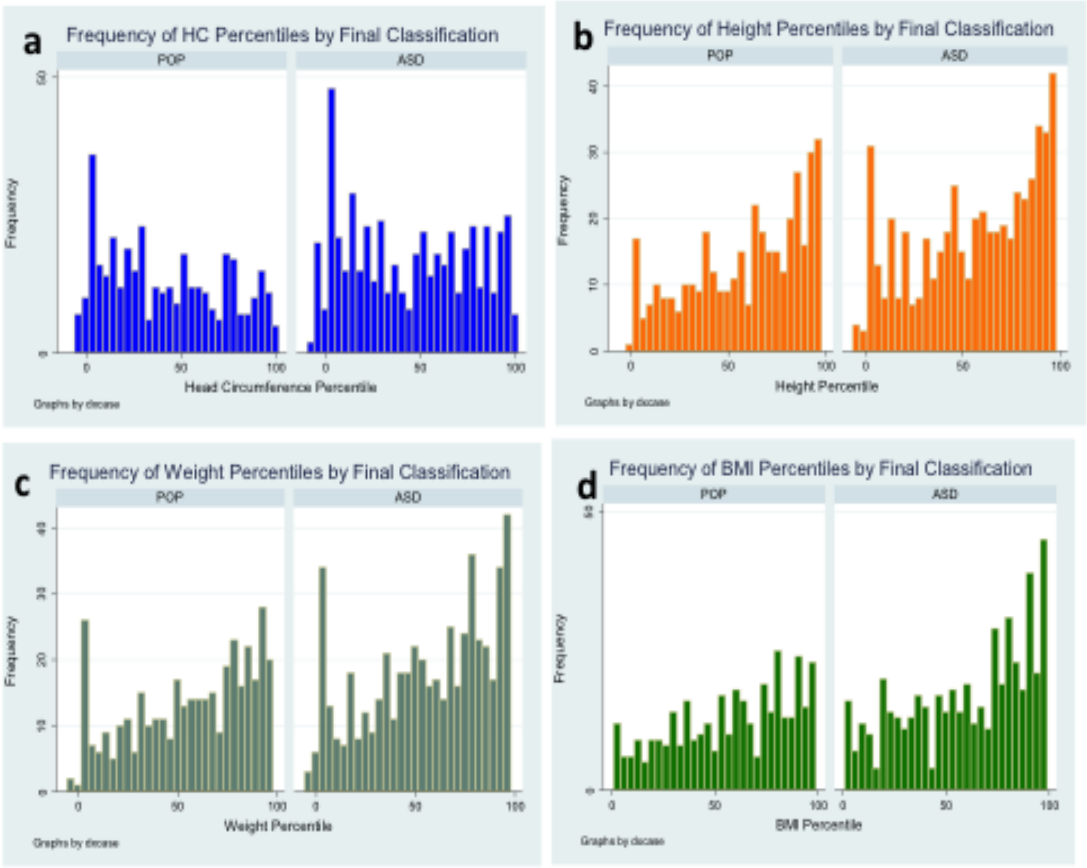
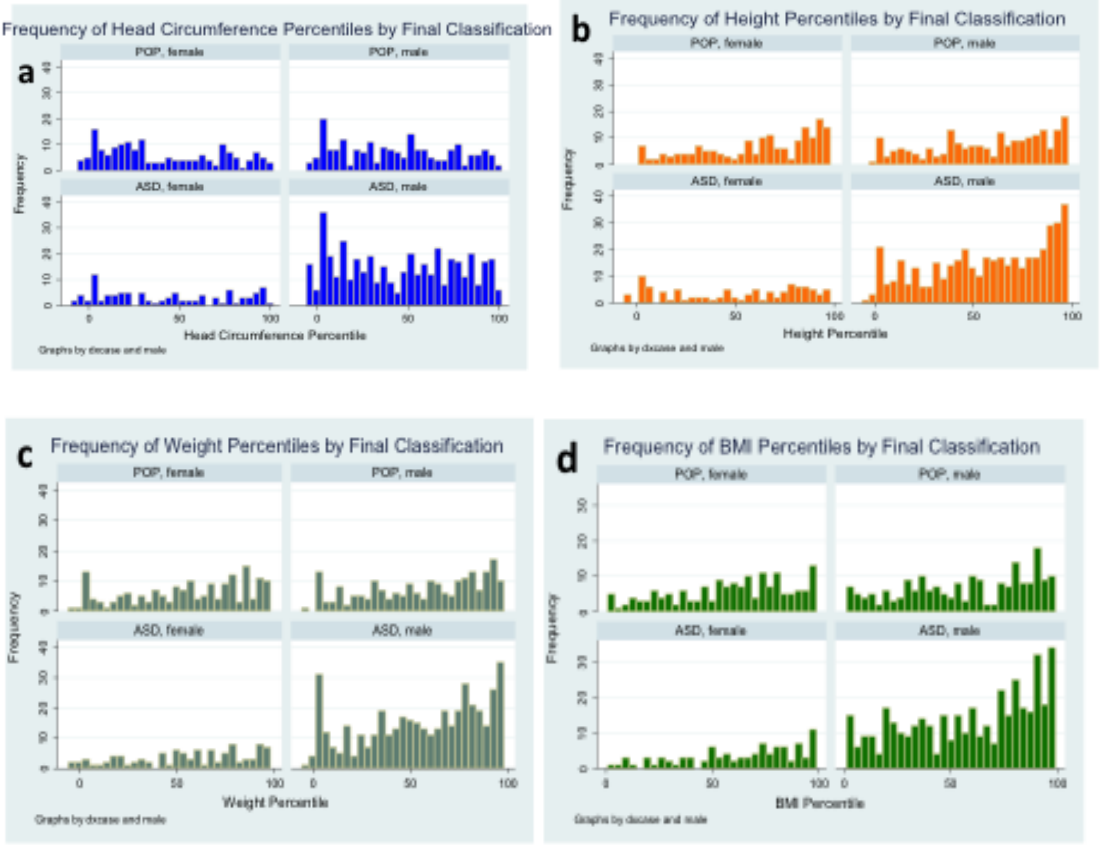


Figure 3: Percentile distribution of a. Head Circumference b. Height c. Weight d. BMI for ASD and POP children stratified by sex



Supplement

Table I. Baseline Characteristics of Excluded Children with Chromosomal Abnormalities and Genetic Syndromes comparing POP vs. ASD

	ASD Status		**Chi-squared p-value**
Exclusion Criteria	**POP /No ASD (%)**	**ASD (%)**	
Chromosomal Abnormalities			**0.045**
Present	0.8	2.5	
Absent	99.2	97.5	
Non-Chromosomal Genetic Syndromes			0.318
Present	2.3	3.4	
Absent	97.7	96.6	

Bolded: p-value <0.05

Chapter 3. Association Between CNV Burden and Dysmorphology In the SEED Study

Norazlin Kamal Nor, MBBS[1], Julie Elizabeth Hoover-Fong, MD, PhD[3],

Christine Ladd-Acosta, PhD[1], Terri H. Beaty, PhD[1],

Margaret Daniele Fallin, PhD[2]

Department of Epidemiology[1] and Mental Health[2],

Johns Hopkins Bloomberg School of Public Health and

Department of Medicine[3], Johns Hopkins University School of Medicine,

Baltimore, MD, USA

Abstract

Introduction:

Autism Spectrum Disorder (ASD) is highly heterogeneous. Use of endophenotypes, or subgrouping of cases, may improve our ability to detect genes related to specific etiologies. Dysmorphism, physical malformations occurring in the embryonic and fetal period, indicates disruption in early development and has been associated with ASD. Investigating genetic susceptibilities in dysmorphism on its own and as a sub-phenotype of ASD may help identify genetic factors related to both. Copy number variants (CNVs) have been identified as a significant contributor to several neurodevelopmental conditions. Here we estimate genome-wide burden of CNVs comparing dysmorphic to non-dysmorphic children drawn from an ASD case-control study, which allows association estimates overall, and among ASD cases with and without dysmorphology.

Methods:

Children from the Study to Explore Early Development, phase 1 (SEED 1) were included in these analyses. SEED is a national, multi-site case-control study of ASD where 2-5 year old children born between 2003 to 2006 from six study states across the United States (including California, Colorado, Georgia, Maryland, North Carolina, and Pennsylvania) were recruited. This analysis includes cases and controls who underwent dysmorphology assessment and had genome-wide genotype data. Using a custom tool for classifying children as dysmorphic resulted in n=45 dysmorphic children (based on 397 physical features) and n=443 non-dysmorphic children. Single nucleotide polymorpohis

(SNP) genotyping was performed using the Illumina Human Omni1-Quad array, and PennCNV was used to call deletion and duplication CNVs. CNV burden over all autosomes was estimated as counts of CNVs per person, as well as the cumulative length of the genome identified as representing a CNV in each person. These were compared between dysmorphic vs. not dysmorphic children. CNVs subsets were also considered: large CNVs (>400 kb), only those overlapping genes, and only those overlapping previously implicated ASD gene regions. Comparisons were stratified by self-reported race/ethnicity, ASD status and sex. Associations between candidate CNV regions previously reported to be associated with ASD and dysmorphism were also estimated to further inform this sub-group of ASD.

Results:

Dysmorphic children in SEED 1 showed less CNV burden than non-dysmorphic children (ratio of cumulative length affected=0.76, p=0.02). This was true for large CNVs (ratio of length=0.49, p=0.01), duplication CNVs (ratio=0.65, p=0.01) and large duplication CNVs (ratio=0.48, p=0.02). Analysis of CNV burden restricted to genic regions showed similar results for overall, large and duplication CNVs. This decreased CNV burden among dysmorphic children remained when restricting the analysis to ASD cases, to non-Hispanic black children, and among males, although none of these associations were statistically significant after correction for multiple testing. None of the candidate ASD CNV regions revealed statistically significant associations with any measure of dysmorphism, but the individual CNVs observed were quite rare and, in the current study sample, these analyses were underpowered.

Conclusions:

Our results show reduced CNV burden among children classified as dysmorphic in SEED 1, even among ASD cases. This may be due to the exclusion of children with chromosomal abnormalities and recognized genetic syndromes from our analysis, to undetected single-gene insults among particants in the dysmorphic group and unrecognized genetic burden in the control group, or to actual protective effects. Further exploration of monogenic versus genome-wide genetic associations in this sample, via exome or full genome sequencing, may be necessary to fully characterize the potential utility of observable dysmorphism as a phenotype for ASD.

Key Terms: dysmorphology, copy number variant burden, autism spectrum disorder

Introduction

Autism spectrum disorder (ASD) is a highly heterogeneous neurodevelopmental disorder with multiple subtypes, and is frequently found with co-morbid conditions ranging from psychiatric to physical abnormalities. Although ASD is understood to have both genetic and environmental risk factors, the genetic etiology is postulated to underlie heterogeneity within ASD (Miles et al., 2011). Clustering all ASD cases based on diagnostic classification may combine many distinct genetically-driven sub-groups and thus undermine efforts to determine true causal factors. Greater understanding of observable autism phenotypes may improve identification of genetic risk factors, as well as allow better prognostication, and potentially allow earlier, targeted, interventions. Dysmorphism, physical malformations occurring in the embryonic and fetal period, has been associated with ASD and is known to be heritable itself. Dysmorphic features may therefore be a useful phenotype of ASD that could help identify a sub-group of ASD associated with particular genetic etiologies.

The term 'dysmorphology' was first proposed by David W. Smith in 1966, and refers to "abnormalities of structural development regardless of severity, timing, or etiology" (Smith, 1966). Examples of dysmorphic features include dysplastic ears and abnormal growth measures such as macrocephaly (enlarged head circumference) or extremely short stature. Dysmorphism at any specific feature is typically defined as being at the extremes of expectation in a general population, with observations markedly higher (or lower) compared to age- and sex-specific population means (Aase, 1990; Merks et al., 2003).

Any single dysmorphic feature occurs in approximately 4% or less of the general population (Merks et al., 2003), although overall dysmorphism affects approximately 3% of all newborns in the United States (Centers for Disease Control and Prevention 2008). Multiple dysmorphic features may represent a general marker for departure from normal developmental processes. In general, multiple dysmorphic features rarely occur without co-existing genetic conditions or teratogenic exposure (Merlob et al. 1985).

Various studies indicate the proportion of children diagnosed with ASD who have physical signs of dysmorphism range between 5-30% (Angkustsiri et al., 2011; Ozgen et al., 2010). Children with ASD are more likely to have major congenital anomalies compared to the general population (Weir et al., 2006). Children with ASD and dysmorphic features also have a greater probability of structural cranial abnormality or known genetic syndromes (Ozgen et al., 2011). In 2011, Angkutsiri et al. described clinical heterogeneity of the physical features of ASD, whereby significantly more children with ASD were classified as dysmorphic compared to control children. Numerous studies have described correlations between morphological abnormalities and ASD (Weir et al., 2006; Miles & Hillman, 2000; Miles et al., 2005; Ozgen et al., 2011; Ozgen et a., 2013). Any child with dysmorphology has a higher likelihood of carrying some detectable genetic aberration. By focusing on this ASD sub-phenotype, it is postulated there may be a greater likelihood of finding an association between specific ASD sub-phenotypes and underlying genetic risk factors.

Rigorous research definitions of dysmorphism remain challenging. Previous studies have used a brief 16 abnormality tool – the 'Minor Anomaly Scale' (Waldrop, 1968) or the Miles dysmorphology classification system, based on 200 features used to categorize 5 classes (Miles and Hillman, 2000) and later three groups: dysmorphic, equivocal and non-dysmorphic (Miles et al., 2005). The Miles algorithm was further fine-tuned into the Miles Autism Dysmorphology Measure (ADM) (Miles et al., 2008), and showed 16% of ASD children were dysmorphic. However, this tool was developed only among children with ASD; no comparison was made with typically developing children. Further, the sample was clinic-based patients and mostly white. For these reasons, the Study to Explore Early Development (SEED) Dysmorphology Group developed a research-reliable quantitative method to characterize and classify dysmorphology among a population-based sample with a mix of ethnicities. They developed a novel, standardized dysmorphology review protocol of over 300 potential dysmorphic features (Shapira et al., 2014). This scale was developed with normal children, as well as children with ASD and generalized developmental delay (DD), involved young children aged 2-5 years of age and represented by three ethnic groups, non-Hispanic white (NHW), non-Hispanic black (NHB), and Hispanic. Unpublished data from the SEED study, source of samples used here, found this tool classified 17% of ASD cases as dysmorphic in each ethnic group, compared to 5% among typically developing control children. This difference remained significant even after accounting for children with chromosomal abnormalities and recognized genetic syndromes. These findings suggest considering dysmorphology as an ASD sub-phenotype may improve risk factor detection in ASD, as genetic risk factors for

children with ASD or population based control children who are dysmorphic could be simultaneously assessed.

Although some chromosomal abnormalities and recognized genetic syndromes have been associated with ASD (Hall, Lightbody & Reiss, 2008; DiGuiseppi et al., 2010), the majority of ASD children do not have any known syndrome. Some of the most consistent genetic findings for ASD to date are inherited and *de novo* copy number variants (CNVs) (Robinson et al., 2014). Studies of CNVs associated with ASD have identified specific CNVs and risk of ASD, as well as an overall association between the genome-wide CNV burden a person carries and ASD risk. In terms of specific CNVs, studies have implicated a number of genes and chromosomal regions, including most prominently regions on chromosomes 7q11.23, 15q11-13, 16p11.2, and 22q11.2 (Pinto et al., 2010; Sanders et al., 2011), which were identified from the Simon Simplex Collection such as 1q21.1 and 3q29 (Pinto et al., 2010; Picinelli et al., 2016), as well as regions encompassing the *SHANK2* (Pinto et al., 2010), *SHANK3* (Gauthier et al., 2009), *NRXN1* (Bucan et al., 2009), *CNTN4* (Roohi et al., 2009) and *CNTNAP2* (Bakkaloglu et al., 2008) genes. These genes and regions are not exclusively associated with risk of ASD but are also associated with different neuropsychiatric phenotypes (i.e. as schizophrenia) or fall under a broad range of neurodevelopmental conditions (Malhotra and Sebat, 2012). In our analysis, we assessed the association between ten loci previously implicated with risk of ASD, including some of the chromosomal regions mentioned above, with dysmorphism in SEED children. Increased genome-wide CNV burden, or more precisely, autosome-wide burden (not including sex chromosomes), has been consistently shown as more common

in ASD cases. Pinto et al. (2010) reported a higher count of rare CNVs among ASD cases, as well as increased counts of deletion CNVs. Others have shown a greater length of the genome affected by CNVs, as well as greater counts, among ASD cases, especially for deletions and for rare CNVs (Vulto-van Silfhout et al., 2013). This has been shown multiple times, with some suggestion that rare *de novo* CNVs are a driver (Sanders et al., 2011) and both deletion and duplications influence risk to ASD (Luo et al., 2012).

Here, we examine whether CNV burden, or CNVs in particular ASD candidate regions, are associated with dysmorphism, considering it as an ASD sub-phenotype. There is a dearth of studies testing for association between CNVs and dysmorphism exclusively. Most studies in this area have examined associations between CNVs and intellectual disability and/or ASD with congenital malformations or dysmorphic features as a supplementary condition, or one that is associated merely with differences in severity. Approaches using alternative phenotypes may enable delineation of possible distinct genetic etiologies. The over-arching objective of this study is to investigate one these features, specifically dysmorphology, and test for potential associations with genetic risk factors as summarized by CNV burden.

Materials and Methods

Study Population

The Study to Explore Early Development (SEED), phase I, is a national multi-site case control study funded by the CDC. Children between ages 2-5 years old were recruited and evaluated at one of six sites: California, Colorado, Georgia, Maryland, North Carolina and Pennsylvania. Potential cases, born between 2003 and 2006 were recruited through partnerships with developmental disability service providers including healthcare and educational systems. Population-based comparison children born in the same years from the same catchment areas were recruited through vital statistics. After phone-based screening and in-person evaluations, children were classified as ASD, other non-ASD developmental disabilities, and children without developmental disabilities from the general population (POP) (Schendel et al., 2012). Eligibility included birth in the study catchment area during the period between September 1st 2003 to August 31st 2006, current residence in the area at the time of first contact, and child living with a knowledgeable caregiver who was able to communicate orally in English or Spanish competently and gave informed consent for participation. The enrolled children also had to be between the ages of 30 and 68 months of age at the completion of the in-person clinical developmental assessment.

Of the 3,899 children recruited into the SEED 1 study, 1,132 underwent genotyping. There were 541 children who had complete dysmorphology classification information

and who were also genotyped. Children with a high proportion of dysmorphic features categorized as not scored (80 or more, out of a possible 397), either because the assessment was not completed or due to poor photograph quality, were filtered out. Our analyses included all of SEED 1 children, except 21 children with non-ASD developmental delay, and 32 children with known chromosomal abnormalities and recognized genetic syndromes, resulting in 488 children in the final sample. The study flow chart describing how the final study population was obtained is shown in **Figure 1**.

Dysmorphology Assessment and Review

Children recruited into the SEED 1 Study were evaluated clinically, and this included a developmental assessment followed by a standardized evaluation for dysmorphology developed by the SEED Dysmorphology Group. This Group included clinical geneticists from each of the six SEED sites who also oversaw and trained personnel (dysmorphology aides) at each site to perform the standardized dysmorphology protocol. Unpublished data from the SEED Dysmorphology Workgroup describes the protocol in greater detail (Shapira, personal communication, SEED Dysmorphology Group unpublished manuscript). Briefly, the 6-part protocol consisted of: 1) in-person anthropometric measurements of the child and all available parents; 2) in-person standardized visual examination of the child, including under a Woods lamp; 3) acquiring a standard series of photographs of the child as well as supplementary photographs of any unusual physical findings; 4) obtaining bilateral hand scans; 5) completing a standardized Dysmorphology Examination Form (DEF) with the observations obtained from the dysmorphology assessment; and 6) carrying out a set of measurements from photographs and hand scans

and documenting those measurements on the DEF, and determining and recording percentiles for all measurements obtained after the in-person evaluation. Dysmorphology aides queried the caregivers about congenital abnormalities and reported genetic syndromes for the child (including birth defects, any previous diagnosis of malformation or developmental syndromes and any previous genetic evaluations or surgeries associated with congenital abnormalities). Quality control measures were instituted for anthropometric measurements and photographs taken of different body parts.

The dysmorphology assessment reviewed a total of 397 potential dysmorphic features per child (37 considered major malformations, and 360 judged as minor), which were grouped into 7 body regions: 1) Ears (90 features); 2) Eyes and eyebrows (62 features); 3) Growth and skin (16 features); 4) Head, hair, face and neck (68 features); 5) Hands and feet (83 features); 6) Mouth, lips, and teeth (26 features); 7) Nose and philtrum (52 features). For each body region or system, one clinical geneticist was assigned to assess all children in the study for dysmorphic features in that region using information from the DEF and all photographs of that region. That clinician's review, and classification of any noted dysmorphism was entered on a standardized Dysmorphology Review Form (DRF). The dysmorphology review included all data in the DEF and all photographs for that body region. Each physical feature was assessed on a Likert scale ranging from 1 to 4: 0=normal or absent; 1=possible or questionable; 2=mild; 3=moderate; and 4=severe; denoting how the feature compared to what was expected in a general population. For features that included a quantitative measure, such as height, percentile ranges were also scored on the 4-point Likert scale. The clinical geneticist responsible for the

dysmorphology review of his or her specified body region was blinded to the child's final classification of ASD or POP. The children were assessed sequentially by race, with the clinical geneticists reviewing all non-Hispanic White children followed by non-Hispanic Black children and Hispanic children.

A feature was classified as 'dysmorphic' if it occurred in ≤5% of the POP group, because these children are considered a sample of the general population. To summarize overall dysmorphism per child, a race/ethnic specific Dysmorphology Score (DS) was calculated as DS = [#Dysmorphic Features/Total features assessed] x 100. Children who had missing data for more than 80 features were excluded from further analysis. The distribution of DS for POP children in each race/ethnicity category was found to fit a log-normal distribution. The expected values based on a log-normal distribution were converted to a corresponding percentile score, and percentile scores >95th percentile were categorized as 'dysmorphic', scores ≤90th percentile were categorized as 'not dysmorphic', and scores >90th percentile and ≤95th percentile were considered 'equivocal'. For analyses presented here, the equivocal group was combined into the not dysmorphic group.

ASD Assessment

For all eligible children, a brief screening interview, the Social Communication Questionnaire (Rutter et al., 2003), was administered to the primary caregiver to identify children who required clinical diagnostic assessment to determine final ASD status. For

SEED, a positive screen was defined as an SCQ score ≥11. Tools for ASD assessment included the Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview-Revised (ADI-R) (Falkmer et al., 2013; Lord et al., 2000). Final classification was assigned using a SEED-specific research algorithm based on ADOS, ADI-R and clinical judgment (Wiggins et al., 2015). Regardless of ascertainment source, any eligible children with a previous ASD diagnosis, who were receiving special education services, and who had a positive screen, were assigned to the ASD workflow. This determined which instruments were administered and the type of diagnostic evaluation the child received during data collection. Based on previous diagnosis and SCQ screening, DD and POP children with negative SCQ screens were assigned to the DD or POP workflow, respectively. If a clinician suspected ASD during the clinical evaluation of a child in the DD or POP workflow, the child would be moved into the ASD workflow (Schendel et al., 2012; DiGiuseppe et al., 2016).

Copy Number Variants (CNVs)

Blood and buccal samples were collected by trained local staff and shipped to the SEED Biosample Repository at Johns Hopkins Bloomberg School of Public Health. These were used to isolate DNA via the QIAsymphony DNA Investigator and QIAsymphony DNA Midi kits (Qiagen) for buccal and blood, respectively. A total of 1,132 SEED 1 cases and controls were genotyped at 1 million single nucleotide polymorphisms (SNPs) on Illumina Human Omni1-Quad array. Genotyping and initial data cleaning was carried out at the Johns Hopkins University SNP Center. Quality control measures at the SNP and sample levels were performed. Samples were excluded if <98% of all markers were

called successfully, if estimated identity by descent (IBD) sharing suggested cryptic relatedness between subjects, if there were sex discrepancies, or if there was excess heterozygosity/homozygosity possibly due to genotyping error. SNPs were excluded for call rates <0.95, minor allele frequency (MAF) <0.01 and if there was evidence of deviation from expected genotype frequencies predicted by Hardy-Weinberg equilibrium ($p<1.0x10^{-8}$ in controls).

CNVs were called using a hidden Markov model implemented in PennCNV (Wang et al., 2007). Hidden copy number state along each autosome was estimated using total signal intensity, allelic intensity ratio, SNP allele frequency, distance between neighboring SNPs, and genomic GC content (Diskin et al., 2008). Quality control (QC) filters were applied at both the CNV and sample levels. CNVs were filtered out if they contained < 10 SNPs, were < 30 kb, or in centromere and telomere regions; samples were excluded if the standard deviation of the log R ratio (LRR) >0.3, the B-allele frequency (BAF) >0.01, or absolute value of a 'wave' factor (due to high GC content over the region) >0.05. The overall data quality pipeline for calling CNVs is shown in **Figure 2**.

Burden Metrics. It is important to remember that CNVs are estimated from SNP data, and they can vary in their beginning and ending positions, as well at the content of the genomic region they encompass. Therefore, we considered summary measures of autosomal CNVs, both CNV counts across all autosomes, as well as their summed lengths in this study. CNV count was based on the total number of unique CNV sites (estimated by PennCNV) for each individual; and cumulative length was determined

summing the length of these unique CNVs in kilobases (kb) per individual. Overall count and length burden metrics included both duplications and deletions, and CNVs occurring anywhere in an autosome (sex chromosomes were omitted). Measures were also calculated separately for duplications and deletions, and for only large CNVs (i.e. those spanning >400kb). Finally, subsets of autosomes were considered: only CNVs overlapping known genes, using hg19 gene boundaries (categorized as "genic CNVs"), and only CNVs overlapping genes reported to be associated with ASD, using the Simons Foundation Autism Research Initiative gene list (termed as "SFARI CNVs"). The UCSC genome database using "TxDb.Hsapiens.UCSC.hg19.knownGene", "annotate", and "org.Hs.eg.db" Bioconductor packages (Goldstein et al., 2016; Carlson et al., 2016) were used to establish known genes and their boundaries. For the "SFARI CNVs", SFARI gene 2.0_ENREF_22 was consulted, and a list of 757 autosomal candidate genes for ASD were used.

CNV Candidate Regions. Malhotra and Sebat (2012) reviewed specific CNV regions associated with ASD, and identified the precise boundaries for each region compared to the catalog on CNVs available on SFARI (Malhotra & Sebat, 2012; https://gene.sfari.org). The largest interval between start and end of any CNV was used to best define the affected autosomal region. We focused on CNVs in 10 specific autosomal regions: chromosomes 1q21.1, 3q29, 7q11.23, 15q11.2, 15q11.2.13.1, 15q13.3, 16p11.2, 16p13.11, 17q12 and 22q11.21.

Statistical Analysis

We characterized the analytic sample using counts and percentages for children with dysmorphism and those without. CNV burden was assessed for both CNV counts and CNV lengths, with effect sizes estimated as the ratio of each measure between dysmorphic and non-dysmorphic SEED 1 children (both ASD cases and POP controls). Mean counts or lengths of CNVs were compared between dysmorphic and non-dysmorphic groups using t-tests. These analyses were carried out overall, for deletions and duplications separately, and then restricting to CNVs to: 1) large CNVs, i.e. those>400 kb, 2) Genic CNVs, i.e. CNVs overlapping with known genes, and 3) SFARI Genic CNVs, i.e. those CNVs overlapping with genes associated with ASD. Analyses were also stratified by self-identified race (Non-Hispanic White, Non-Hispanic Black and Hispanic), ASD classification, and sex. Finally, sensitivity analyses were performed to assess the effect of excluding or including children with chromosomal abnormalities and recognized genetic syndromes, and excluding and including children diagnosed with developmental delay. For ASD candidate region analyses, CNV counts between dysmorphic and non-dysmorphic children were compared via Fisher's exact tests. All computational analyses were performed using R 3.3.2.

Results

Characteristics of the Study Sample

After restricting to SEED 1 participants with genotype and complete dysmorphology data, and further sample filtering for genotyping and CNV calling QC, there were 488 children available for analyses; 45 classified as dysmorphic and 443 not dysmorphic. Baseline characteristics of this sample are shown in **Table 1**. Those classified as dysmorphic were more likely to be male and to have ASD. There was no statistical difference in self-identified race between these two groups. There were higher proportions of children with both chromosomal abnormalities and recognized genetic syndromes in the dysmorphic group as shown in Supplemental Table I, but these were excluded from our analysis.

CNV Burden

There were 10,394 CNVs detected in the whole sample, with a mean CNV count of 20 CNVs per person among dysmorphic children and 21 per person among non-dysmorphic children (Supplemental Table I). The mean combined CNV length in dysmorphic children was 2.02 Mb per person, compared to 2.63 Mb per person in children who were not dysmorphic (Supplemental Table II). Both CNV counts and lengths were generally lower among dysmorphic children compared to non-dysmorphic children in SEED (**Table 2**). Dysmorphic children had 76.6% less genome affected by CNVs than non-

dysmorphic children (p=0.022), and this remained true when considering only CNVs overlapping gene boundaries (74%, p=0.034). The ratio of length affected among dysmorphic versus non-dysmorphic children was even lower when considering only large CNVs (D/ND = 0.491, p=0.015) and only large gene-overlapping CNVs (D/ND = 0.519, p=0.029). When considering only duplications, there were fewer duplications CNVs and less length affected among dysmorphic children (D/ND_{count} = 0.834, p= 0.035; D/ND_{length} = 0.655, p = 0.011). This was consistent when only considering duplications overlapping with genes (D/ND_{count} = 0.766, p =0.035; D/ND_{length} = 0.583, p=0.008). Further restricting the CNVs to only large duplication CNVs also showed a smaller cummulative length among dysmorphic children (D/ND_{length} =0.483, p=0.022) (Supplemental Table II). However, none of these tests remained statistically significant after correcting for multiple testing.

The SEED 1 samples with genotypic data and dysmorphism classification included 297 non-Hispanic White children (NHW), 32 of whom were categorized as dysmorphic. There were also 88 non-Hispanic black (NHB) children, 6 with dysmorphism, and 103 Hispanic children, of whom 8 were dysmorphic. The length-based effect sizes were generally in the same direction across all ethnic groups, with D/ND ratios < 1. However, none of these were statistically significant in the NHW group, despite it representing the largest subset of children (**Table 3**). In the NHB group, D/ND comparisons were nominally significant for all CNVs and for all large CNVs (D/ND_{length} = 0.557, p = 0.027; D/ND_{length} = 0.132, p = 0.008). Similar patterns were observed for genic, and large genic CNV lengths (D/ND_{length} = 0.502, p = 0.034; D/ND_{length} = 0.15, p = 0.016). Among

Hispanic children, deletions, and genic deletions, showed slightly lower genome burden among dysmorphic children, although this was not seen in other groups or the overall sample.

The trend for less CNV burden among SEED 1 children with dysmorphism was consistent when stratifying by ASD status (**Table 4**). There were 267 children with ASD, 34 of them categorized as dysmorphic, while only 8 of 212 POP children were dysmorphic. Results were also similar among 345 males, 38 of whom had some dysmorphism (**Table 5**).

ASD Candidate CNVs Regions

There were few observations of CNVs among ASD candidate regions in dysmorphic children compared to non-dysmorphic children (**Table 6**). In the few regions where CNVs were observed, there was no statistically significant difference in CNV burden between dysmorphic and non-dysmorphic groups, and CNV counts were often more frequent in the non-dysmorphic group.

Discussion

We explored the use of dysmorphism, observable physical abnormalities known to be associated with autism spectrum disorder (ASD), as an ASD sub-phenotype for genetic association studies, in hopes that shared genetic association between dysmorphism and ASD might further illuminate etiological mechanisms for ASD in this sub-group. In our sample of young children from the Study to Explore Early Development, Phase 1 (SEED 1), a national case-control study of autism, we observed trends for decreased CNV burden among children with dysmorphism. Many specific associations with length of the genome affected by CNVs, including overall, large CNVs, CNVs in known genes, and duplications showed nominally statistically significant comparisons, although these did not reach significance after Bonferroni correction for multiple testing. The reduced CNV burden results were consistent when considering only ASD cases or only boys. Results were also generally consistent across ethnic groups, although nominal statistical significance was seen in non-Hispanic black children. None of the CNV burden analyses restricted to previously genes previously associated ASD (based on the SFARI ASD list) were associated with dysmorphism.

Our analyses did not include the few (N=21) SEED 1 children with non-ASD developmental delay (DD) that had available genotyping and dysmorphology data. Sensitivity analyses including those DD children did not change the results (Supplementary Table III).

Importantly, our analyses excluded 8 (13%) of the dysmorphic children and 6 (1%) of the non-dysmorphic children with chromosomal abnormalities and the 10 (16%) dysmorphic and 9 (2%) non-dysmorphic children with non-chromosomal genetic syndromes (Supplementary Table IV). Had these children and the DD children with been included in our analyses, we would have observed excess burden in CNV length among dysmorphic children, across all analyses (Supplementary Table VId). It is likely that the combination of congenital abnormalities and recognized genetic syndromes with dysmorphic features is an indication of some underlying deviation from normal developmental processes resulting from either exposure to teratogens prenatally or other genetic causes. The excluded children thus appear to increase CNV burden when comparing dysmorphic to non-dysmorphic children.

In this study, we endeavored to use the phenotype of dysmorphism, the study of physical malformations present early in life, classified using a newly developed tool established as part of the SEED Study, as sub-phenotype of ASD. As a condition with origins during the embryonic or fetal period, dysmorphism is understood to have both genetic and idiopathic risk factors, with environmental or an interaction between genes and environment likely to represent distinct idiopathic etiologies. In terms of sensitive time periods, dysmorphism and ASD likely share overlapping periods of vulnerability early in life. Dysmorphism is a developmental condition with certain associated genetic susceptibilities.

Across neurodevelopmental conditions such as ASD, some CNVs have emerged as a strong genetic risk factor. However, even the most recurrent CNVs are individually rare. Therefore, global CNV burden as an aggregate may be one way to assess the impact of CNVs on developmental conditions and potentially support the importance gene dosage in leading to these conditions (Walsh et al., 2008; Sebat et al., 2009). CNVs are present not only in those with disease, but also in healthy populations. A comparison with healthy controls is thus important to understand the contribution of CNVs to neurodevelopmental conditions (Rosenfeld & Patel, 2016). Studies of individuals with developmental conditions such as congenital anomalies (Geng et al., 2014; Stark et al., 2015), intellectual disability/developmental delay (ID/DD) (Girirajan et al., 2011; Cooper et al., 2011; Di Grigorio et al., 2017) and ASD (Pinto et al., 2010; Sanders et al., 2011; Leppa et al., 2016) have all supported some role for CNVs in a subset of individuals with these conditions. However, previously published studies investigating the association between CNV burden and dysmorphism specifically are harder to find. Dysmorphism is challenging to quantify, and no specific tool has been accepted as a gold standard. Further, it is considered an accompanying feature to other outcomes of primary clinical relevance. Published studies including dysmorphism often consider it only as a supplementary component of the phenotype, or a marker of increased severity, rather than as a primary phenotype.

Our results revealed decreased burden among children with dysmorphism in SEED 1. While several other studies have seen associations with length of CNVs burden in dysmorphic children, particularly for large CNVs (Girirajan et al., 2011), the effect

direction has been towards greater length, not less. Previous studies also found increased burden of large CNVs in individuals with intellectual disability (ID) with multiple congenital anomalies (MCA) compared to ID alone, and Cooper et al. (2011) reported a higher prevalence of large CNVs (>400kb) in more severe developmental phenotypes associated with MCA. Individuals with ID, MCA and dysmorphism have been reported to have an increased frequency of *de novo* CNVs, particularly those with more severe phenotypes such as abnormal head circumference (Vulto-van Silfhout et al., 2013). Our results are not consistent with these findings. One possible reason for our finding of smaller mean combined length of CNVs in dysmorphic children may be because we excluded children with the more severe phenotypes of congenital abnormalities and recognized genetic syndromes. Indeed, when they are included, our results are consistent with previous literature. In that literature, children with chromosomal abnormalities and genetic syndromes were included (Cooper et al., 2011; Girirajan et al., 2011; Qiao et al., 2014). Thus, our study is asking a different question – what is the CNV burden comparison among those without any reported congenital abnormalities, major or minor? Another possible explanation for this finding of reduced CNV burden is that children categorized as 'dysmorphic' in SEED may represent a sub-group with a high incidence of single-gene insults or undiagnosed syndromes, or potentially those who have had teratogenic exposure *in-utero* that selectively affected a specific gene or region, but not aggregate measures of CNVs burden. Alternatively, the group of children classified as 'non-dysmorphic', because they did not have enough features across the multitude considered to be declared dysmorphic, may nonetheless have isolated dysmorphic features corresponding to otherwise undiagnosed genetic burden. If this were a generally

large proportion, it could explain our counter-intuitive results. The observation that this reduced burden among dysmorphic children is strong in ASD cases would be consistent with these latter two hypotheses, if the single-gene insults among non-dysmorphic children or the isolated physical abnormalities among the non-dysmorphic children still contributed to ASD risk. Full genome sequencing, or even exome sequencing, could help to resolve this, but such data are not yet available for these SEED samples.

Reduced CNV burden is rarely reported in neurodevelopmental conditions, which has generally been associated with increased CNV burden. Grozeva et al. (2013) reported that in adult patients with bipolar disorder, the rate of very large (>1Mb) and rare CNVs were significantly lower compared to controls, which they postulate may have resulted from increased rates of CNVs in some other phenotype in the controls not accounted for, such as diabetes. It is possible there were other phenotypes not accounted for in our analyses. Another interesting finding from our analysis is that comparing dysmorphic to non-dysmorphic children, there was reduced genome affected by CNVs driven by large CNV duplications. CNV burden involving duplications is not as commonly reported as CNV burden involving deletions, although duplications are more difficult to estimate from SNP data. One a plausible mechanism for how variability in CNV duplications could affect gene expression is if specific regulatory mechanisms (such as STOP codons) were duplicated in the CNV. A study by Martin et al. (2014) found total CNV rare duplications showed a negative correlation with positive symptoms of schizophrenia, and hypothesized this indicated CNV duplication burden may have a small protective effect against symptoms of schizophrenia. Extrapolating their results to our study, in our

analysis large CNV duplications appear to have a negative association with risk of dysmorphism in children, and CNV burden for large duplications may potentially have a small protective effect against risk of dysmorphology. Our results, like Martin et al.'s, did not survive correction for multiple testing. Our study population was small, and this may have impacted statistical power in our analysis.

Our study sample was obtained from SEED 1, a case-control study of ASD, and thus was not truly representative of a population-based sample. This is important to consider in when interpreting our results of risk of dysmorphisms. Our sample was made up 54.7% children with ASD, who represent 75.6% of children found to be dysmorphic. This was by design, given our interest in dysmorphism as an ASD sub-phenotype, but does limit generalizability to dysmorphism *per se*. To improve generalizability of future studies, it would be of interest to perform the analysis in another sample that is more representative of the general population when testing for any role of CNV burden on risk of dysmorphology.

We performed analysis restricted to self-identified race to manage potential confounding due to race/ethnic group. The majority of study participants were Non-Hispanic Whites, with only 6 (6.8%) of Non-Hispanic Blacks and 7 (6.8%) of Hispanics classified as dysmorphic. Previous studies have assessed the impact of race in general on CNV burden, with different populations harboring different average number of CNVs per sample, and admixed populations having a higher number of CNVs (Jakobsson et al., 2008). Analysis on larger and more representative samples would provide clearer

understanding of the role of race in affecting CNV burden for dysmorphology. Of interest, although we used self-reported race instead of genetic ancestry for our analysis, correspondence between self-reported race and genetically predicted ancestry via Principal Components analysis is high (Ladd-Acosta, personal communication).

Our results were also most prominent in boys. Unpublished analysis by the SEED Dysmorphology group on this same analytic population found (after excluding children with developmental delay (DD) from the SEED analyses), there were little differences in dysmorphism prevalence comparing males to females (Shapira, personal communication, SEED Dysmorphology Group unpublished manuscript). Our analysis showed dysmorphic males appear to drive the association between ASD and CNV burden, although these results need to be considered with caution as there were significantly fewer females in the study overall and only 7 (4.8%) of all female children assessed were categorized as dysmorphic.

There were very few overlaps observed when assessing CNV association with dysmorphism in genes previously reported to be associated with ASD. These individual CNVs are rare, and it is possible that our study population was not large enough to detect these overlapping CNVs. As Kaminsky et al. (2011) observed, obtaining adequate evidence for the functional role of rare CNVs in disease causation requires very large sample sizes and large control populations.

Study Limitations and Strengths

Our cross-sectional design, culled from a national case-control study may not be fully representative of the United States in terms of racial make-up (Schendel et al., 2012). Although the study recruitment included a population-based approach, ascertainment bias for selection into the study is still possible with families self-selecting for participation. In terms of limitations associated with the dysmorphology outcome, the young age of the children recruited (2-5 years) restricts the generalizability of our results to only major or minor dysmorphisms identified in early childhood, and we cannot consider morphological changes with age. Also, despite various quality control measures instituted to maintain the quality of the photographs used for dysmorphology assessment, there was variability in missing data for the dysmorphology data gathered attributed to poor photo quality in one recruitment site. However, unpublished data from the SEED study performed by the SEED Dysmorphology subgroup found the missing data had no significant effect on the observed results following sensitivity analyses and multiple imputations.

An important limitation already discussed is undiagnosed chromosomal abnormalities and recognized genetic syndromes in our sample. The information we did have on such conditions was based solely on parental report, and no new genetic testing was available. Thus, it is likely that additional chromosomal abnormalities or genetic syndromes are present and may have influenced our tests for association between CNV burden and dysmorphism.

We were also not able to stratify analyses on *de novo* versus inherited CNVs, yet *de novo* CNVs have been implicated in previous literature (Vulto-van Silfhout et al., 2013). Unfortunately, we did not have parental genotyping information and hence are not able to determine the impact of *de novo* CNVs. In addition, rare CNVs have also been associated with congenital malformations (Serra-Juhe et al., 2012). Although the SEED Study is one of the largest population-based samples of ASD in the US, for analysis of dysmorphism, we only had a subset of the full SEED 1 sample, and are underpowered for detecting rare CNVs.

This study also has a number of strengths. The SEED Study is one of the largest ASD studies with population-based ascertainment. Previous studies have used clinic-based samples, and have often been smaller. Although this study was not fully representative of the United States in terms of racial/ethnic make-up, it included three major groups by design (Schendel et al., 2012), and allowed for the potential effects of race to be considered in the analyses. Finally, the SEED Study involves varied geographical locations across the United States. A major strength is the uniform developmental assessment of all study participants including research-reliable ASD classification, as well as a customized standardized dysmorphology assessment tool with quality assurance across sites. Previous dysmorphology classification protocols were based on clinic patient populations, and on individuals who were primarily white (Miles et al. 2005; Miles et al. 2008).

In summary, we found overall CNV burden, specifically overall large duplication CNV burden, is lower among children with dysmorphology in SEED, particularly among SEED ASD children. Although these results were not anticipated, and the estimated associations were no longer significant after correcting for multiple testing, we think our observations are worthy of further investigation. Autistic sub-phenotypes like dysmorphology may help parse out risk factors, particularly genetic risk factors leading to subsets of ASD. Due to the relatively small sample sizes, these results still need to be considered as preliminary, and future work replicating these findings would be needed to further evaluate genotype-phenotype associations in ASD, a complex and heterogeneous disorder. In addition, genetic analyses incorporating full genome sequencing or at least whole exome sequencing on these subjects should enable improved detection of underlying genetic abnormalities. This information would have helped to more definitively exclude children with undiagnosed chromosomal abnormalities from our study population.

The relevance of CNV burden in children with ASD, particularly those with dysmorphism or multiple congenital anomalies, is supported by the consensus statement issued by the American Academy of Pediatrics recommending the use of microarray analysis in these children, which has a much higher yield than those without (Miller et al., 2010; Shen et al., 2010). Genotype-phenotype studies is an important area for research that could potentially lead to better understanding of the biologic basis of disease and the development of individualized management. Discovering the genetic basis of any condition, particularly in ASD with dysmorphism, will not only allow earlier screening

for and intervention of both ASD and any co-occurring medical conditions, and possibly also improved understanding of prognosis.

This study was given approval by the IRB Committee of Johns Hopkins School of Public Health. Informed consent was obtained from all caregivers before clinical assessment as part of the SEED Study protocol. Diagnoses and assessments were performed at the six SEED sites in the United States.

References

Aase, J. M. (1990). The physical examination in dysmorphology. In Diagnostic dysmorphology (pp. 33-42). New York: Plenum Medical Book Co.

Abrahams, B. S., Arking, D. E., Campbell, D. B., Mefford, H. C., Morrow, E. M., Weiss, L. a, … Packer, A. (2013). SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). Molecular Autism, 4, 36. http://doi.org/10.1186/2040-2392-4-36

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*. Arlington, VA: American Psychiatric Publishing.

Angkustsiri, K., Krakowiak, P., Moghaddam, B., Wardinsky, T., Gardner, J., Kalamkarian, N., … Hansen, R. L. (2011). Minor physical anomalies in children with autism spectrum disorders. *Autism : The International Journal of Research and Practice*, *15*(6), 746–60.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological Medicine*, *25*, 63–77.

Bakkaloglu, B., O'Roak, B. J., Louvi, A., Gupta, A. R., Abelson, J. F., Morgan, T. M., … State, M. W. (2008). Molecular Cytogenetic Analysis and Resequencing of Contactin Associated Protein-Like 2 in Autism Spectrum Disorders. *American Journal of Human Genetics*, *82*(1), 165–173.

Boyd, B. A., Odom, S. L., Humphreys, B. P., & Sam, A. M. (2010). Infants and Toddlers with Autism Spectrum Disorder: Early Identification and Early Intervention. *Journal of Early Intervention*.

Bucan, M., Abrahams, B. S., Wang, K., Glessner, J. T., Herman, E. I., Sonnenblick, L. I., … Hakonarson, H. (2009). Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genetics*, *5*(6).

Carlson, M.R., Pagès, H., Arora, S., Obenchain, V., Morgan, M. (2016). Genomic Annotation Resources in R/Bioconductor. *Methods Mol Biol*.. 1418:67-90.

Centers for Disease Control and Prevention. (2008). Update on overall prevalence of major birth defects--Atlanta, Georgia, 1978-2005. *Morbidity and Mortality Weekly Report*, 57(1), 1-5.

Chakrabarti, S., Fombonne, E. (2001). Pervasive Developmental Disorder in Preschool Children, *Journal of American Medical Association, 285*(24), 3093–3099.

Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S. R., Woodhouse, E., … Bolton, P. (2015). Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. *JAMA Psychiatry*, 1–9.

Constantino JN. Recurrence Rates in Autism Spectrum Disorders. *JAMA*. 2014;312(11):1154-1155.

Cooper, G. M., Coe, B. P., Girirajan, S., Rosenfeld, J. A., Vu, T. H., Baker, C., … Eichler, E. E. (2011). A copy number variation morbidity map of developmental delay. *Nature Genetics*, *43*(9), 838–46.

Database of Genomic Variants. http://projects.tcag.ca/variation/

Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Development and Psychopathology*, *20*, 775–803.

Di Gregorio, E., Riberi, E., Belligni, E. F., Biamino, E., Spielmann, M., Ala, U., … Ferrero, G. B. (2017). CNVs analysis in a cohort of isolated and syndromic DD/ID reveals novel genomic disorders, position effects and candidate disease genes. *Clinical Genetics*, (February), 1–8.

Diguiseppi, C., Hepburn, S., Davis, J. M., Fidler, D. J., Hartway, S., Lee, N. R., … Robinson, C. (2015). Screening for Autism Spectrum Disorders in Children with Down Syndrome: Population Prevalence and Screening Test Characteristics. *Journal of Developmental & Behavioral Pediatrics*, *31*(3), 181–191.

Diskin, S. J., Li, M., Hou, C., Yang, S., Glessner, J., Hakonarson, H., … Wang, K. (2008). Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Research*, *36*(19), 1–12.

Feuk, L., Marshall, C. R., Wintle, R. F., & Scherer, S. W. (2006). Structural variants: changing the landscape of chromosomes and design of disease studies. *Human Molecular Genetics*, *15 Spec No 1*(1), 57–66.

Flor, J., Bellando, J., Lopez, M., & Shul, A. (2017). Developmental functioning and medical co-morbidity profile of children with complex and essential autism. *Autism Research*,10(8):1344-1352.

Folstein, S., Rutter, M. (1977). Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry,*18(4):297-321.

Freitag, C. M., Staal, W., Klauck, S. M., Duketis, E., & Waltes, R. (2010). Genetics of autistic disorders: review and clinical implications. *Eur Child Adolesc Psychiatry*, *19*(3), 169–178.

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., … Buxbaum, J. D. (2014). Most genetic risk for autism resides with common variation. *Nature Genetics*, *46*(8), 881–5.

Gauthier, J., Spiegelman, D., Piton, A., Lafreni??re, R. G., Laurent, S., St-Onge, J., … Rouleau, G. A. (2009). Novel de novo SHANK3 mutation in autistic patients. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *150*(3), 421–424.

Geng, J., Picker, J., Zheng, Z. (2014). Chromosome microarray testing for patients with congenital heart defects reveals novel disease causing loci and high diagnostic yield. *BMC Genomics*, 15: 11279.

Geschwind, D. H. (2011). Genetics of autism spectrum disorders. *Trends in Cognitive Sciences*.

Girirajan, S., Brkanac, Z., Coe, B. P., Baker, C., Vives, L., Vu, T. H., … Eichler, E. E. (2011). Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genetics*, *7*(11).

Girirajan, S., Rosenfeld, J. A., Coe, B. P., Parikh, S., Friedman, N., Goldstein, A., … Eichler, E. E. (2012). Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *The New England Journal of Medicine*, *367*(14), 1321–31.

Girirajan, S., Johnson, R. L., Tassone, F., Balciuniene, J., Katiyar, N., Fox, K., … Selleck, S. B. (2013). Global increases in both common and rare copy number load associated with autism. *Human Molecular Genetics*, *22*(14), 2870–2880.

Goldstein, L. D., Cao, Y., Pau, G., Lawrence, M., Wu, T. D., Seshagiri, S., & Gentleman, R. (2016). Prediction and quantification of splice events from RNA-seq data. *PLoS ONE*, *11*(5), 1–18.

Grayton, H.M., Fernandes, C., Rujescu, D., Collier, D.A.(2012). Copy number variations in neurodevelopmental disorders. *Prog Neurobiol*. 99(1):81-91.

Grozeva, D., Kirov, G., Conrad, D. F., Barnes, C. P., Hurles, M., Owen, M. J., … Craddock, N. (2013). Reduced burden of very large and rare CNVs in bipolar affective disorder. *Bipolar Disorders*, 15(8), 893–898.

Hall, S. S., Lightbody, A. a, & Reiss, A. L. (2008). Compulsive, self-injurious, and autistic behavior in children and adolescents with fragile X syndrome. *American Journal of Mental Retardation: AJMR*, *113*(1), 44–53.

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., Miller, J., Fedele, A., Collins, J., Smith, K., Lotspeich, L., Croen, L.A., Ozonoff, S., Lajonchere,

C., Grether, J.K., Risch, N. Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry* 68(11):1095-102 (2011).

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, 515(7526), 216-221.

Kaminsky, E. B., Kaul, V., Paschall, J., Church, D. M., Bunke, B., Kunig, D., … Martin, C. L. (2011). An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genetics in Medicine*, *13*(9), 777–784.

Lauritsen, M. B., Pedersen, C. B., & Mortensen, P. B. (2005). Effects of familial risk factors and place of birth on the risk of autism: A nationwide register-based study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*, 963–971.

Leppa, V. M. M., Kravitz, S. N. N., Martin, C. L. L., Andrieux, J., Le Caignec, C., Martin-Coignard, D., … Geschwind, D. H. H. (2016). Rare Inherited and De Novo CNVs Reveal Complex Contributions to ASD Risk in Multiplex Families. *American Journal of Human Genetics*, *99*(3), 540–554.

Luo, R., Sanders, S. J., Tian, Y., Voineagu, I., Huang, N., Chu, S. H., … Geschwind, D. H. (2012). Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. *American Journal of Human Genetics*, *91*, 38–55.

Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., … Newschaffer, C. (2017). The Changing Epidemiology of Autism Spectrum Disorders. *Annual Review of Public Health*, 38(1), 81–102.

Martin, A. K., Robinson, G., Reutens, D., & Mowry, B. (2015). Clinical and parental age characteristics of rare copy number variant burden in patients with schizophrenia. American Journal of Medical Genetics, Part B: *Neuropsychiatric Genetics*, 168(5), 374–382.

Mazina, V., Gerdts, J., Trinh, S., Ankenman, K., Ward, T., Dennis, M. Y., … Bernier, R. (2015). Epigenetics of Autism-related Impairment: Copy Number Variation and Maternal Infection*, J Dev Behav Pediatr*, *36*(2), 61–67.

Malhotra, D., Sebat, J. (2012). CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell*. 148(6):1223-41.

Merks, J. H. M., van Karnebeek, C. D. M., Caron, H. N., Hennekam, R. C. M., Karnebeek, C. D. M. Van, Caron, H. N., & Hennekam, R. C. M. (2003). Phenotypic abnormalities: terminology and classification. *American Journal of Medical Genetics. Part A*, *123A*(3), 211–230.

Merlob, P., Papier, C. M., Klingberg, M. A., & Reisner, S. H. (1985). Incidence of congenital malformations in the newborn, particularly minor abnormalities. *Progress in Clinical and Biological Research*, 163C, 51–55.

Miles, J. H., & Hillman, R. E. (2000). Value of a clinical morphology examination in autism. *American Journal of Medical Genetics*, *91*(January 1999), 245–253.

Miles, J. H., Takahashi, T. N., Bagby, S., Sahota, P. K., Vaslow, D. F., Wang, C. H., … Farmer, J. E. (2005). Essential versus complex autism: definition of fundamental prognostic subtypes. *American Journal of Medical Genetics. Part A*, *135*(2), 171–80.

Miles, J. H., Takahashi, T. N., Hong, J., Munden, N., Flournoy, N., Braddock, S. R., et al. (2008). Development and validation of a measure of dysmorphology: useful for autism subgroup classification. *American Journal of Medical Genetics Part A*, 146A(9), 1101-1116.

Miller, D.T., Adam, M.P., Aradhya, S., et al. (2010). Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet*, 86: 749-764.

Morrow, E. M. (2010). Genomic copy number variation in disorders of cognitive development. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(11), 1091–1104.

Ozgen, H. M., Hop, J. W., Hox, J. J., Beemer, F. a, & van Engeland, H. (2010). Minor physical anomalies in autism: a meta-analysis. *Molecular Psychiatry*, *15*(3), 300–7.

Ozgen, H., Hellemann, G. S., Stellato, R. K., Lahuis, B., van Daalen, E., Staal, W. G., … van Engeland, H. (2011). Morphological features in children with autism spectrum disorders: a matched case-control study. *Journal of Autism and Developmental Disorders*, *41*(1), 23–31.

Ozgen, H., Hellemann, G. S., de Jonge, M. V, Beemer, F. a, & van Engeland, H. (2013). Predictive value of morphological features in patients with autism versus normal controls. *Journal of Autism and Developmental Disorders*, *43*(1), 147–55.

Ozonoff, S., Young, G. S., Carter, a., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Stone, W. L. (2011). Recurrence Risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium Study. *Pediatrics*. 128(3):e488-95

Picinelli, C., Lintas, C., Piras, I. S., Gabriele, S., Sacco, R., Brogna, C., & Persico, A. M. (2016). Recurrent 15q11.2 BP1-BP2 microdeletions and microduplications in the etiology of neurodevelopmental disorders. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, 171(8), 1088–1098.

Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., … Betancur, C. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, *466*(7304), 368–372.

Pinto D., Darvishi K., Shi X.H., Rajan D., Rigler D., Fitzgerald T., Lionel A.C., Thiruvahindrapuram B., MacDonald J.R., Mills R., et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nat. Biotechnol*. 2011;29:512–520.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., … Scherer, S. W. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *American Journal of Human Genetics*, *94*(5), 677–694.

Qiao, Y., Mercier, E., Dastan, J., Hurlburt, J., McGillivray, B., Chudley, A. E., … Rajcan-Separovic, E. (2014). Copy number variants (CNVs) analysis in a deeply phenotyped cohort of individuals with intellectual disability (ID). *BMC Medical Genetics*, *15*(Id), 82.

The R Project for Statistical computing: http://www.R-681project.org

Robinson, E. B., Samocha, K. E., Kosmicki, J. a., McGrath, L., Neale, B. M., Perlis, R. H., & Daly, M. J. (2014). Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. *Proceedings of the National Academy of Sciences*, 111(42), 15161–15165.

Robinson, E. B., Neale, B. M., & Hyman, S. E. (2015). Genetic research in autism spectrum disorders. *Current Opinion in Pediatrics*, *27*(6), 685–91.

Roohi, J., Montagna, C., Tegay, D. H., Palmer, L. E., DeVincent, C., Pomeroy, J. C., … Hatchwell, E. (2009). Disruption of contactin 4 in three subjects with autism spectrum disorder. *Journal of Medical Genetics*, *46*(3), 176–82.

Rosenberg, R. E., Law, J. K., Yenokyan, G., McGready, J., Kaufmann, W. E., & Law, P. a. (2009). Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Archives of Pediatrics & Adolescent Medicine*, *163*(10), 907–914.

Rosenfeld, J., & Patel, A. (2016). Chromosomal Microarrays: Understanding Genetics of Neurodevelopmental Disorders and Congenital Anomalies. *Journal of Pediatric Genetics*, *6*(1), 042–050.

Sanders, S., Ercan-sencicek, a G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., & State, M. (2011). Multiple recurrent de novo copy number variations (CNVs), including duplications of the 7q11.23 Williams-Beuren syndrome region, are strongly associated with autism. *Neuron*, *70*(5), 863–885.

Sanders, S. J. (2015). First glimpses of the neurobiology of autism spectrum disorder. *Current Opinion in Genetics and Development*, *33*, 80–92.

Sanders, S. J., He, X., Willsey, a. J., Ercan-Sencicek, a. G., Samocha, K. E., Cicek, a. E., … State, M. W. (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron*, *87*(6), 1215–1233.

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Hultman, C., Larsson, H., Reichenberg, A. (2017). The Heritability of Autism Spectrum Disorder. *Research Letter, JAMA Network.* 318(12):1182-1184.

Schendel, D. E., Diguiseppi, C., Croen, L. A., Fallin, M. D., Reed, P. L., Schieve, L. A., et al. (2012). The Study to Explore Early Development (SEED): a multisite epidemiologic study of autism by the Centers for Autism and Developmental Disabilities Research and Epidemiology (CADDRE) network. *Journal of Autism and Developmental Disorders*, 42(10), 2121-2140.

Sebat, J., Levy, D. L., & McCarthy, S. E. (2009). Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders. *Trends in Genetics*, *25*(12), 528–535.

Serra-Juhé, C., Rodríguez-Santiago, B., Cuscó, I., Vendrell, T., Camats, N., Torán, N., & Pérez-Jurado, L. A. (2012). Contribution of Rare Copy Number Variants to Isolated Human Malformations. *PLoS ONE*, *7*(10).

SFARI Gene Database. 32. https://gene.sfari.org/autdb/CNVHome.do

SFARI Gene Database Copy Number Variants https://gene.sfari.org/database/cnv/

Shapira, K., Tian, L., Aylsworth, A., Elias, E., Hoover-Fong, J., Meeks, N., Souders, M., Tsai, C., Alexander, A., Schieve, L. (2014). Development of a Novel Protocol for Characterizing Dysmorphology to Enhance the Phenotypic Classification of Autism Spectrum Disorders, presented at International Meeting for Autism Research, Atlanta, 2014. Georgia: INSAR.

Shapira, S.K., Lin, H. T., Aylsworth A.S., Elias, E.R., Hoover-Fong, J.E., Meeks, N.J.L., Souders, M.C., Tsai, A.C.H., Zackai, A.H., Alexander, A.A., Yeargin-Allsopp, M.H., Schieve, L.A. (personal communication, SEED Dysmorphology Workgroup unpublished manuscript). A Novel Protocol for Characterizing Dysmorphology to Enhance the Phenotypic Classification of Autism Spectrum Disorder in the Study to Explore Early Development.

Shen, Y., Dies, K.A., Holm, I.A., et al. (2010). Clinical genetic testing for patients with autism spectrum disorders. *Pediatrics*. 125: e727-e735

Sheppard, B., Benke, K.S., Croen, L., Daniels, J., Newschaffer, C., Reynolds, A., Schendel, D., Schieve, L., Ladd-Acosta, C., Fallin, D. (2015). Polygene –by– Prenatal Environment Interaction in Autism Spectrum Disorder using Copy Number Variant Burden, presented at IGES 2015, Baltimore. Maryland: IGES.

Shishido, E., Aleksic, B., & Ozaki, N. (2014). Copy-number variation in the pathogenesis of autism spectrum disorder. *Psychiatry and Clinical Neurosciences*.

Shoukier, M., Klein, N., Auber, B., Wickert, J., Schroeder, J., Zoll, B., … Zirn, B. (2013). Array CGH in patients with developmental delay or intellectual disability: Are there phenotypic clues to pathogenic copy number variants? *Clinical Genetics*, *83*(1), 53–65.

Smith, D. W. (1966). Dysmorphology (teratology). *J Pediatr.* 69(6):1150-69.

Stark, Z., Behrsin, J., Burgess, T. , et al. (2015). SNP microarray abnormalities in a cohort of 28 infants with congenital diaphragmatic hernia. (2015) *Am J Med Genet A*, 167A(10): 2319-26.

Torres, F., Barbosa, M., & Maciel, P. (2016). Recurrent copy number variations as risk factors for neurodevelopmental disorders: critical overview and analysis of clinical implications. *J Med Genet*, *53*(1), 73–90.

Vulto-van Silfhout, A. T., Hehir-Kwa, J. Y., van Bon, B. W. M., Schuurs-Hoeijmakers, J. H. M., Meader, S., Hellebrekers, C. J. M., … de Vries, B. B. a. (2013). Clinical Significance of De Novo and Inherited Copy-Number Variation. *Human Mutation*, 34, 1679–1687.

Waldrop, M. F., Pederson, F. A., & Bell, R. Q. (1968). Minor physical anomalies and behavior in preschool children. *Child Development*, 39(2), 391-400.

Walsh, C.A., Morrow, E.M., Rubenstein, J.L. (2008). Autism and brain development. *Cell.* 135: 396-400

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F. a, … Bucan, M. (2007). PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research*, 17, 1665–1674.

Wier, M. L., Yoshida, C. K., Odouli, R., Grether, J. K., & Croen, L. a. (2006). Congenital anomalies associated with autism spectrum disorders. *Developmental Medicine and Child Neurology*, *48*, 500–507.

Wiggins, L. D., Levy, S. E., Daniels, J., Schieve, L., Croen, L. A., DiGuiseppi, C., … Schendel, D. (2015). Autism Spectrum Disorder Symptoms Among Children

Enrolled in the Study to Explore Early Development (SEED). *Journal of Autism and Developmental Disorders*, *45*(10), 3183–3194.

Winchester, L., Yau, C., Ragoussis, J. (2009). Comparing CNV detection methods for SNP arrays. *Brief. Funct. Genomic Proteomic*. 8:353–366.

Zhang, X., Du, R., Li, S., Zhang, F., Jin, L., & Wang, H. (2014). Evaluation of copy number variation detection for a SNP array platform. *BMC Bioinformatics*, *15*(1), 50.

Tables

Table 1: Characteristics of Children Classified as Dysmorphic and Not Dysmorphic in the SEED 1 Study

Baseline Characteristics Total N= 488		**Dysmorphic N= 45 (9.2%)**	**Not Dysmorphic N= 443 (90.8%)**	**p-value**
Diagnosis	ASD	34 (75.6%)	233 (52.6%)	<0.001
	Possible Case	3 (6.7%)	6 (1.4%)	
	POP	8 (17.8%)	204 (46.0%)	
Sex	Male	38 (84.4%)	307 (69.3%)	0.033
	Female	7 (15.6%)	136 (30.7%)	
Race	Non-Hispanic White	32 (71.1%)	265 (59.8%)	0.335
	Non-Hispanic Black	6 (13.3%)	82 (18.5%)	
	Hispanic	7 (15.6%)	96 (21.7%)	

Table 2: CNV Burden for Counts and Lengths in Dysmorphic vs. Non-Dysmorphic Children

N=488		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (D)= 45 N (ND)= 443	**CNV Characteristics**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**
All CNVs	Overall	0.934	0.381	0.766	**0.022**
	Overall Genic	0.933	0.425	0.740	**0.034**
	Overall SFARI	1.017	0.911	1.014	0.935
All Large CNVs	Overall large	0.640	0.052	0.491	**0.015**
	Overall large Genic	0.654	0.083	0.519	**0.029**
	Overall large SFARI	1.737	0.471	1.080	0.909
Deletions	Overall Deletion	1.023	0.821	0.967	0.755
	Deletion Genic	1.080	0.384	1.046	0.709
	Deletion SFARI	1.113	0.618	1.029	0.904
Duplications	Overall Dup.	0.834	**0.035**	0.655	**0.011**
	Dup. Genic	0.766	**0.035**	0.583	**0.008**
	Dup. SFARI	0.883	0.574	0.989	0.971

Bolded: significant t-test $p<0.05$
D= Dysmorphic
ND= Not Dysmorphic

Table 3: CNV Burden for Lengths in Dysmorphic vs. Non-Dysmorphic Children stratified by Race/Ethnicity

		Average CNV Length (kb) All		Average CNV Length (kb) NHW		Average CNV Length (kb) NHB		Average CNV Length (kb) Hispanic	
CNVs	**CNV Characteristics**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**
All CNVs	Overall	0.766	**0.022**	0.825	0.210	0.557	**0.027**	0.757	0.129
	Overall Genic	0.740	**0.034**	0.842	0.351	0.502	**0.034**	0.588	**0.026**
	Overall SFARI	1.014	0.935	0.818	0.366	1.740	0.225	1.260	0.543
All Large CNVs	Overall large	0.491	**0.015**	0.584	0.150	0.132	**0.008**	0.617	0.456
	Overall large Genic	0.519	**0.029**	0.616	0.207	0.150	**0.016**	0.785	0.727
	Overall large SFARI	1.080	0.909	0.311	0.174	--	--	22.29	0.189
Del.	Overall Deletion	0.967	0.755	1.014	0.909	1.161	0.640	0.635	**0.012**
	Deletion Genic	1.046	0.709	1.153	0.346	1.018	0.947	0.630	**0.033**
	Deletion SFARI	1.029	0.904	0.847	0.552	2.334	0.181	0.609	0.382
Dup.	Overall Dup.	0.655	**0.011**	0.721	0.134	0.313	**0.002**	0.863	0.666
	Dup. Genic	0.583	**0.008**	0.677	0.148	0.337	**0.018**	0.552	0.159
	Dup. SFARI	0.989	0.971	0.775	0.503	0.520	0.310	2.520	0.116
Sample size		N(All)= 488 N(D)= 45 N(ND)= 443		N(NHW)= 297 $N(D_{NHW})$= 32 $N(ND_{NHW})$= 265		N(NHB)= 88 $N(D_{NHB})$= 6 $N(ND_{NHB})$= 82		N(Hisp.)= 103 $N(D_H)$= 7 $N(ND_H)$= 96	

Bolded: t-test $p<0.05$
D: Dysmorphic ND: Not Dysmorphic
NHW: Non-Hispanic White NHB: Non-Hispanic Black Hisp.: Hispanic

Table 4: CNV Burden for CNV Lengths in Dysmorphic vs. Non-Dysmorphic Children Stratified by ASD Status

		Average CNV Length (kb) Entire		Average CNV Length (kb) ASD		Average CNV Length (kb) POP	
CNVs	**CNV Characteristics**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**
All CNVs	Overall	0.766	**0.022**	0.693	**0.027**	0.839	0.402
	Overall Genic	0.740	**0.034**	0.676	0.055	0.749	0.224
	Overall SFARI	1.014	0.935	1.039	0.855	1.097	0.820
All Large CNVs	Overall large	0.491	**0.015**	0.359	**0.012**	0.752	0.698
	Overall large Genic	0.519	**0.029**	0.404	**0.031**	0.638	0.453
	Overall large SFARI	1.080	0.909	1.147	0.849	--	--
Deletions	Overall Deletion	0.967	0.755	1.014	0.911	0.752	0.059
	Deletion Genic	1.046	0.709	1.149	0.336	0.664	0.079
	Deletion SFARI	1.029	0.904	0.928	0.787	1.509	0.518
Dup.	Overall Dup.	0.655	**0.011**	0.530	**0.009**	0.893	0.735
	Dup. Genic	0.583	**0.008**	0.469	**0.012**	0.801	0.557
	Dup. SFARI	0.989	0.971	1.275	0.505	0.575	0.244
Total and Group totals		N(All)= 488 N(D)= 45 N(ND)= 443		N(ASD)= 267 $N(D_{ASD})$= 34 $N(ND_{ASD})$= 233		N(POP)= 212 $N(D_{POP})$= 8 $N(ND_{POP})$= 204	

Bolded: t-test $p<0.05$
ASD: Children with Autistic Spectrum Disorder
POP: Typically developing children

Table 5: CNV Burden in Dysmorphic vs. Non-Dysmorphic Children Stratified by Sex

		Average CNV Length (kb) All		Average CNV Length (kb) Male		Average CNV Length (kb) Female	
CNVs	**CNV Characteristics**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**
All CNVs	Overall	0.766	**0.022**	0.771	**0.047**	0.787	0.253
	Overall Genic	0.740	**0.034**	0.768	0.097	0.707	0.193
	Overall SFARI	1.014	0.935	0.985	0.945	1.140	0.643
All Large CNVs	Overall large	0.491	**0.015**	0.496	**0.028**	0.540	0.277
	Overall large Genic	0.519	**0.029**	0.524	0.054	0.605	0.358
	Overall large SFARI	1.080	0.909	1.145	0.849	0	0.093
Deletions	Overall Deletion	0.967	0.755	1.000	0.999	0.825	0.356
	Deletion Genic	1.046	0.709	1.094	0.514	0.874	0.579
	Deletion SFARI	1.029	0.904	0.959	0.885	1.402	0.401
Dup.	Overall Dup.	0.655	**0.011**	0.643	**0.014**	0.767	0.453
	Dup. Genic	0.583	**0.008**	0.595	**0.019**	0.628	0.310
	Dup. SFARI	0.989	0.971	1.029	0.927	0.676	0.581
Sample Size		N(All)= 488 N(D)= 45 N(ND)= 443		N(Male)= 345 $N(D_{Male})$= 38 $N(ND_{Male})$= 307		N(Female)= 143 $N(D_{Female})$= 7 $N(ND_{Female})$= 136	

Bolded: t-test $p<0.05$

Table 6: CNV Associations with Dysmorphism at ASD CNV Candidate regions

Region	All and large CNVs			Deletions and large deletions CNVs			Duplications and large duplications CNVs		
	D (N=45) Count (%)	ND (N=443) Count (%)	p-value	D (N=45) Count (%)	ND (N=443) Count (%)	p-value	D (N=45) Count (%)	ND (N=443) Count (%)	p-value
1q21.1A	0 (0%)	11 (2.5%)	0.61	0 (0%)	3 (0.6%)	1.00	0 (0%)	8 (1.8%)	1.00
1q21.1L	0 (0%)	4 (0.9%)	1.00	0 (0%)	0 (0%)	--	0 (0%)	4 (0.9%)	1.00
3q29A	0 (0%)	4 (0.9%)	1.00	0 (0%)	0 (0%)	--	0 (0%)	4 (0.9%)	1.00
3q29L	0 (0%)	0 (0%)	--	0 (0%)	0 (0%)	--	0 (0%)	0 (0%)	--
7q11.23A	1 (2.2%)	9 (1.9%)	1.00	0 (0%)	2 (0.4%)	1.00	1 (2.2%)	7 (1.5%)	0.54
7q11.23L	0 (0%)	0 (0%)	--	0 (0%)	0 (0%)	--	0 (0%)	0 (0%)	--
15q11.2A	4 (8.9%)	58 (13.0%)	0.63	2 (4.4%)	35 (7.9%)	0.56	2 (4.4%)	23 (5.2%)	1.00
15q11.2L	0 (0%)	2 (0.4%)	1.00	0 (0%)	0 (0%)	--	0 (0%)	2 (0.4%)	1.00
15q11.2.13.1A	4 (8.9%)	55 (12.4%)	0.80	2 (4.4%)	33 (7.4%)	0.75	2 (4.4%)	22 (4.9%)	1.00
15q11.2.13.1L	0 (0%)	0 (0%)	--	0 (0%)	0 (0%)	--	0 (0%)	0 (0%)	--
15q13.3A	1 (2.2%)	7 (1.5%)	0.54	0 (0%)	0 (0%)	--	1 (2.2%)	7 (1.5%)	0.54
15q13.3L	1 (2.2%)	4 (0.9%)	0.38	0 (0%)	0 (0%)	--	1 (2.2%)	4 (0.9%)	0.38
16p11.2A	0 (0%)	2 (0.4%)	1.00	0 (0%)	1 (0.2%)	1.00	0 (0%)	1 (0.2%)	1.00
16p11.2L	0 (0%)	2 (0.4%)	1.00	0 (0%)	1 (0.2%)	1.00	0 (0%)	1 (0.2%)	1.00
16p13.11A	4 (8.9%)	55 (12.4%)	0.63	1 (2.2%)	10 (2.2%)	1.00	3 (6.7%)	45 (10.1%)	0.60
16p13.11L	0 (0%)	2 (0.4%)	1.00	0 (0%)	0 (0%)	--	0 (0%)	2 (0.2%)	1.00
17q12A	8 (17.8%)	65 (14.6%)	0.51	1 (2.2%)	4 (0.9%)	0.38	7 (15.5%)	61 (13.7%)	0.65
17q12L	0 (0%)	2 (0.4%)	1.00	0 (0%)	0 (0%)	--	0 (0%)	2 (0.2%)	1.00
22q11.21A	2 (4.4%)	18 (4.0%)	0.53	1 (2.2%)	2 (0.4%)	0.25	1 (2.2%)	16 (3.6%)	1.00
22q11.21L	1 (2.2%)	1 (0.2%)	0.17	0 (0%)	0 (0%)	--	1 (2.2%)	1 (0.2%)	1.00

A=All
L=Large
D= Dysmorphic
ND= Not Dysmorphic

Figures

Figure 1: Flow Chart of Study Population Selection

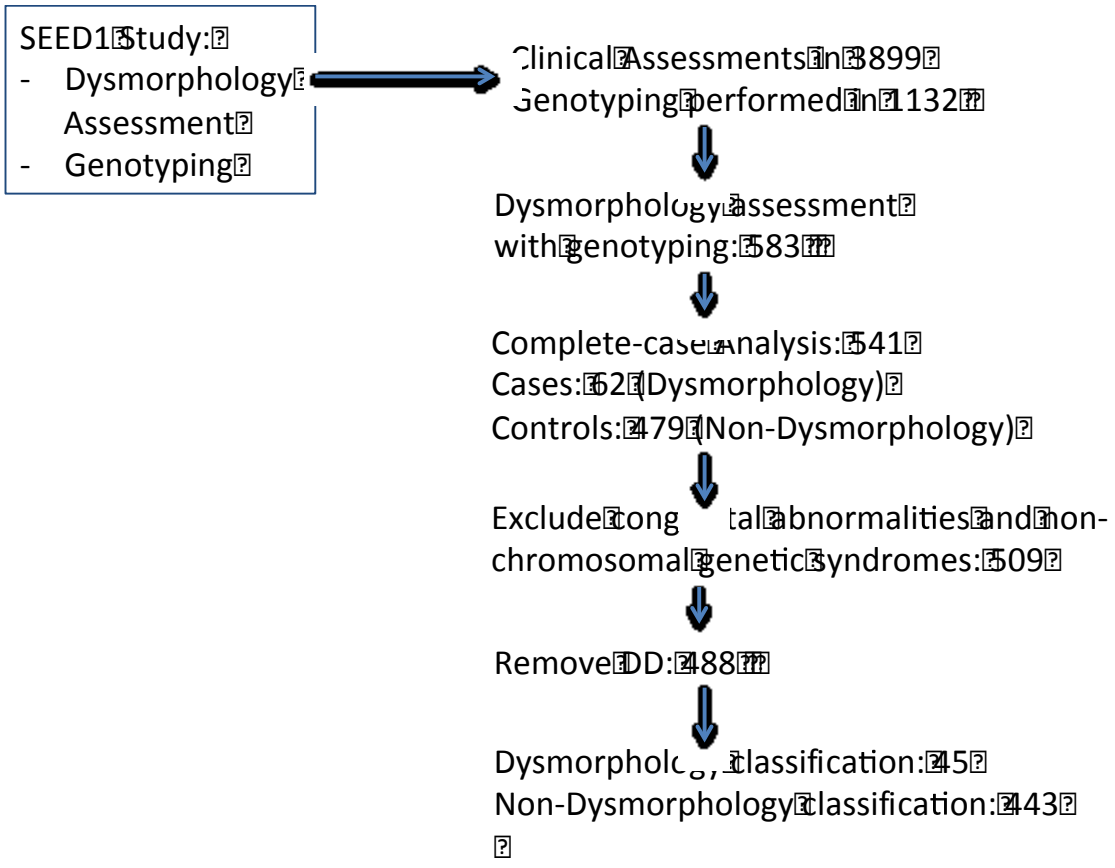
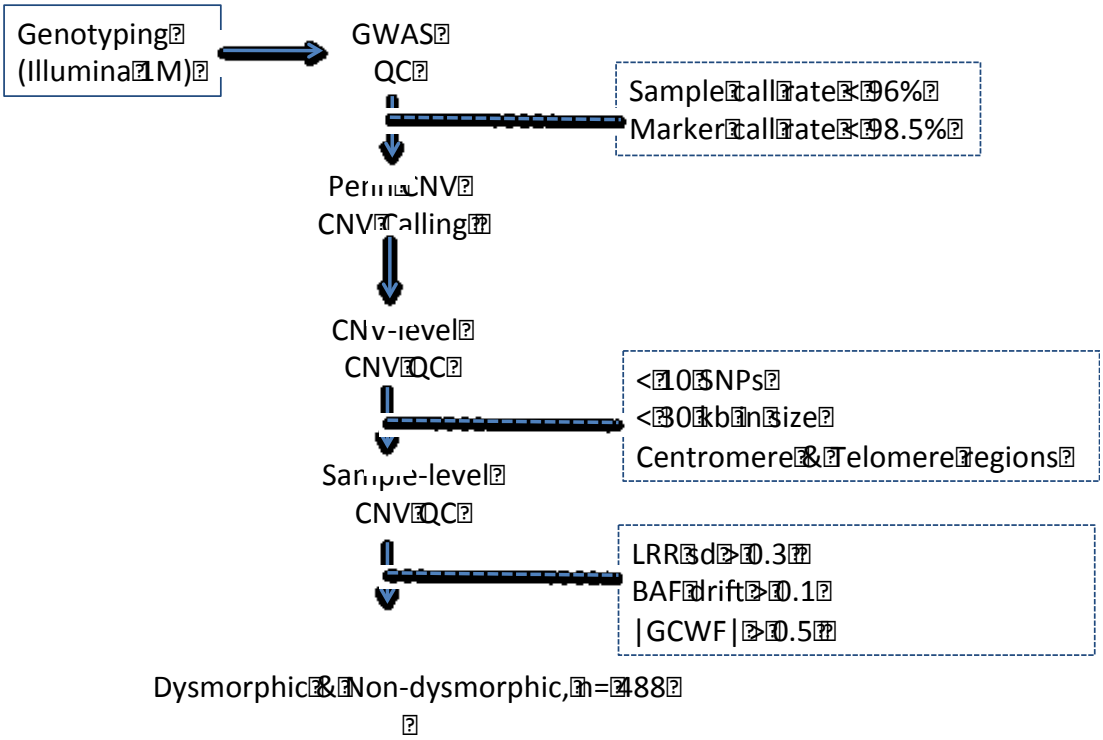


Figure 2: CNV SEED Quality Control Pipeline



Supplement

S. Table I: CNV Count/Rate by Dysmorphic Status

All CNVs	**CNV Count / Rate**				
CNV characteristics N=488	**Total CNVs**	**Dysmorphic Rate, N= 45**	**Non-Dysmorphic Rate, N= 443**	**D/ ND Ratio**	**P-value**
Overall	10,394	20.022	21.428	0.934	0.381
Overall Genic	5,082	9.777	10.478	0.933	0.425
Overall SFARI	619	1.288	1.266	1.017	0.911
Overall large	360	0.488	0.763	0.640	0.052
Overall large Genic	305	0.422	0.645	0.654	0.083
Overall large SFARI	20	0.066	0.038	1.737	0.471
Deletion Total	5,545	11.600	11.338	1.023	0.821
Deletion Genic	2,730	6.000	5.553	1.080	0.384
Deletion SFARI	364	0.822	0.738	1.113	0.618
Deletions large	50	0.066	0.106	0.628	0.337
Deletions large Genic	39	0.066	0.081	0.820	0.717
Deletions large SFARI	5	0	0.011	0	--
Overall Duplication	4849	8.422	10.090	0.834	**0.035**
Duplication Genic	2352	3.777	4.925	0.766	**0.035**
Duplication SFARI	255	0.466	0.528	0.883	0.574
Duplication large	310	0.422	0.656	0.642	0.083
Duplication large Genic	266	0.355	0.564	0.630	0.088
Duplication large SFARI	15	0.066	0.027	2.461	0.311

Bolded: t-test $p<0.05$

S. Table II: Average CNV Length by Dysmorphic Status

ALL	**Ave. CNV length (kb)**			
CNV characteristics N=488	**Length in D n=45**	**Length in ND n=443**	**D/ND Ratio**	**P-value**
Overall	2,018.41	2,632.16	0.76	**0.02**
Overall Genic	672.15	907.90	0.74	**0.03**
Overall SFARI	145.25	143.20	1.01	0.93
Overall large	430.1	875.19	0.49	**0.01**
Overall large Genic	189.731	365.60	0.52	**0.03**
Overall large SFARI	8.29	7.67	1.08	0.91
Deletion Total	908.19	938.60	0.96	0.75
Deletion Genic	321.77	307.57	1.04	0.71
Deletion SFARI	92.42	89.82	1.03	0.90
Deletion large	54.21	98.12	0.55	0.22
Deletion large Genic	36.12	40.94	0.88	0.83
Deletion large SFARI	0	2.53	0	--
Overall Duplication	1,110.22	1,693.55	0.65	**0.01**
Duplication Genic	350.38	600.32	0.58	**0.01**
Duplication SFARI	52.82	53.37	0.99	0.97
Duplication large	375.91	777.07	0.48	**0.02**
Duplication large Genic	153.60	324.66	0.47	**0.02**
Duplication large SFARI	8.29	5.13	1.61	0.54

Bolded: t-test $p<0.05$
D= Dysmorphic
ND= Not Dysmorphic

S. Table III: CNV Burden for Lengths in Dysmorphic vs. Non-dysmorphic: Analysis With DD Children (+DD) and Without (-DD)

N (+DD)= 509 **N (DD)=21**		**Average CNV Length (kb) -DD**		**Average CNV Length (kb) +DD**	
N (D): 45 **N (ND): 509**	**D/ND Ratio**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**
All CNVs	Overall	0.766	**0.022**	0.778	**0.029**
	Overall Genic	0.740	**0.034**	0.755	**0.046**
	Overall SFARI	1.014	0.935	1.027	0.870
All Large CNVs	Overall large	0.491	**0.015**	0.509	**0.020**
	Overall large Genic	0.519	**0.029**	0.539	**0.039**
	Overall large SFARI	1.080	0.909	1.119	0.869
Deletions	Overall Deletion	0.967	0.755	0.970	0.773
	Deletion Genic	1.046	0.709	1.053	0.665
	Deletion SFARI	1.029	0.904	1.042	0.860
Duplications	Duplication Overall	0.655	**0.011**	0.670	**0.015**
	Duplication Genic	0.583	**0.008**	0.599	**0.012**
	Duplication SFARI	0.989	0.971	1.002	0.993

Bolded: t-test $p<0.05$
D= Dysmorphic
ND= Not Dysmorphic
DD= Developmental delay
+DD= Analysis With DD Children (+DD)
-DD= Analysis Without DD Children (-DD)
Note:
CNV burden for lengths in dysmorphic vs. non-dysmorphic without DD children (-DD) is the same as the CNV burden for lengths in dysmorphic vs. non-dysmorphic in Table 2.

S. Table IV: Proportion of Children with Congenital Abnormalities and Non-Genetic Syndromes in Dysmorphic vs. Non-dysmorphic Children for SEED 1

Co-existing Abnormalities or Syndromes N(All)[#] = 541 N(DD) = 21	**Dysmorphic, N=62**		**Non-dysmorphic, N=479**		
	Yes	**No**	**Yes**	**No**	**p-value**
Chromosomal Abnormalities	8 (13%)	54 (87%)	6 (1%)	473 (99%)	<0.05 (χ^2=29.56)
Non-chromosomal Genetic Syndromes	10 (16%)	52 (84%)	9 (2%)	470 (98%)	<0.05 (χ^2=32.89)

DD= Developmental delay
N(All)[#]= 541 includes all children with chromosomal abnormalities, non-chromosomal genetic syndromes and children with developmental delay

S. Table V: CNV Burden and CNV Lengths in Dysmorphic Children vs. Not Dysmorphic Children: Analyses without (-CAGS) and with (+CAGS) Children with Chromosomal Abnormality and Non-Chromosomal Genetic Syndromes

		Average CNV Length (kb) -CAGS		**Average CNV Length (kb) +CAGS**	
CNVs	**CNV Characteristics**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**
All CNVs	Overall	0.766	**0.022**	1.722	0.125
	Overall Genic	0.740	**0.034**	1.716	0.150
	Overall SFARI	1.014	0.935	1.499	0.102
All Large CNVs	Overall large	0.491	**0.015**	3.022	0.103
	Overall large Genic	0.519	**0.029**	2.756	0.122
	Overall large SFARI	1.080	0.909	7.154	**0.041**
Deletions	Overall Deletion	0.967	0.755	1.329	0.228
	Deletion Genic	1.046	0.709	1.286	0.234
	Deletion SFARI	1.029	0.904	1.300	0.314
Duplications	Duplic. Overall	0.655	**0.011**	1.934	0.159
	Duplic. Genic	0.583	**0.008**	1.932	0.183
	Duplic. SFARI	0.989	0.971	1.822	0.106
Sample size		N(All)= 488 N(D)= 45 N(ND)= 443		N(All)= 516 N(D)= 60 N(ND)= 456	

Bolded: t-test $p<0.05$
CAGS= Chromosomal Abnormality and Non-Chromosomal Genetic Syndromes
D= Dysmorphic
ND= Non Dysmorphic

S. Table VI: Overall CNV Burden for Length in Dysmorphic vs. Non-dysmorphic: Analyses Comparing Combinations of DD and CAGS Children

Analysis of Overall CNV Burden for Length for D vs. ND using combinations		**a. Average CNV Length (kb) D vs. ND, -DD, -CAGS (N=488)**		**b. Average CNV Length (kb) D vs. ND, +DD, -CAGS (N=509)**		**c. Average CNV Length (kb) D vs. ND, -DD, +CAGS (N=516)**	
CNVs	**CNV Characteristics**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**
All CNVs	Overall	0.766	**0.022**	0.779	**0.029**	1.722	0.125
	Overall Genic	0.740	**0.034**	0.755	**0.046**	1.716	0.150
	Overall SFARI	1.014	0.935	1.027	0.876	1.499	0.102

Analysis of Overall CNV Burden for Length using different combinations		**d. Average CNV Length (kb) D vs. ND, +DD, +CAGS (N=541)**		**e. Average CNV Length (kb) Dysmorphic with CAGS vs. ND (no CAGS) +DD (N=481)**		**f. Average CNV Length (kb) Non-dysmorphic with CAGS vs. ND (no CAGS) +DD (N=479)**	
CNVs	**CNV Characteristics**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**	**D/ND Ratio**	**p-value**
All CNVs	Overall	2.068	**0.041**	5.837	**0.009**	2.513	0.215
	Overall Genic	2.063	0.052	5.886	**0.013**	2.532	0.192
	Overall SFARI	1.687	**0.036**	3.631	**0.016**	2.037	0.118

Bolded: t-test $p<0.05$

D= Dysmorphic ND= Not Dysmorphic DD= Developmental delay

+DD= Analysis With DD Children (+DD) -DD= Analysis Without DD Children (-DD)

CAGS= Chromosomal Abnormality and Non-Chromosomal Genetic Syndromes

+CAGS= Analysis With CAG Children (+CAG)

-CAGS= Analysis Without CAG Children (-CAG)

Analysis 'a' is the same analysis as for Table 2

Analysis 'a' & 'b' is the same as S. Table III; analysis 'a' & 'c' is the same as S.Table V

S. Table VII: List of Chromosomal Abnormalities and Non-Chromosomal Genetic Syndromes Excluded from Analysis

Chromosomal Abnormalities

POP	ASD
47, XXY Klinefelter Syndrome	15q11.2q13 duplication
Mosaic Down Syndrome	1q44 deletion
Trisomy 21 (Down Syndrome)	Trisomy 21 (Down Syndrome)
	Chromosome 17p13.2
	Deletion 15q13.2q1
	Mosaic 45X/46XY
	Partial Monosomy 21
	Unbalanced Translocation
	Williams Syndrome

Non-Chromosomal Genetic Syndromes

POP	ASD
Glucose-6-Phosphate deficiency	Alopecia
Hypothyroidism	Cytomegalovirus (CMV)
Neurofibromatosis	Ehlers-Danlos Syndrome
Retinoblastoma	Fragile X Syndrome
Rheumatoid Arthritis	Hypothyroidism
	Mitochondrial Disorders
	Menke Syndrome
	Proteus Syndrome
	Rubinstein-Taybi Syndrome
	Septo-optic Dysplasia
	Sturge-Weber Syndrome

Chapter 4. Association Between CNV Burden and Abnormal Physical Growth in the SEED Study

Norazlin Kamal Nor, MBBS[1], Julie Elizabeth Hoover-Fong, MD, PhD[3],

Christine Ladd-Acosta, PhD[1], Terri H. Beaty, PhD[1],

Margaret Daniele Fallin, PhD[2]

Department of Epidemiology[1] and Mental Health[2],

Johns Hopkins Bloomberg School of Public Health and

Department of Medicine[3], Johns Hopkins University School of Medicine,

Baltimore, MD, USA

Abstract

Introduction:

Autism Spectrum Disorder (ASD) is highly heterogeneous, with strong evidence of some genetic control. Use of endophenotypes, or sub-grouping of cases, may improve our ability to detect genes related to specific sub-groups of ASD. Abnormal growth for sex and age, including large or small head circumference (macro- or microcephaly), abnormally tall or short stature, overweight or underweight, or high/low BMI, is also highly heritable, and may be a related feature among some children with ASD. Investigating genetic susceptibilities for growth abnormalities and as a sub-phenotype of ASD may help identify genetic factors related to both. Copy number variants (CNVs) have been identified as a significant contributor to several neurodevelopmental conditions (including ASD) and to growth. Here we estimate burden of CNVs comparing children with abnormal to normal growth in an ASD case-control study. We also examined whether particular CNVs in regions previously associated with ASD are also associated with abnormal growth in this ASD case-control study.

Methods:

Participants included 840 children aged 2-5 years (born between 2003 and 2006) from six sites across the United States enrolled into Phase 1 of the Study to Explore Early Development (SEED 1). This analysis includes ASD cases and controls who had genome-wide genotype data from the Illumina Omni1 array and also had anthropometric measurements to determine physical growth, standardized to age and sex. CNVs were called using the PennCNV algorithm. CNV burden over all autosomes was estimated as

counts of CNVs per person, as well as the cumulative length of the genome affected by CNVs per person. Burden was compared between SEED 1 children classified with a growth abnormality (large: >90th percentile or small: <10th percentile) and children in the normal range, for head circumference (HC), height, weight, and BMI. CNV types, duplications versus deletions, and CNV subsets were also considered: large CNVs (>400 kb), only those overlapping known genes, and only those overlapping previously implicated ASD genes. All analyses were stratified by sex. Associations between candidate CNV regions previously reported to be associated with ASD and growth abnormalities were tested separately to assess for possible common genetic links between ASD and growth abnormalities.

Results:

The burden ratio for tall versus normal stature children was <1 for both CNV counts and lengths ($Tall/Norm_{counts} = 0.909$, $p = 0.029$; $Tall/Norm_{length} = 0.833$, $p = 0.033$), although this was not statistically significant after Bonferroni correction for multiple tests. Similar patterns were observed when restricting to genic CNVs and CNVs encompassing recognized genes associated with ASD, as identified by SFARI ($T/N_{genic} = 0.83$, $p=0.06$; $T/N_{SFARI} = 0.79$, $p=0.044$), and among deletions overlapping SFARI genes ($T/N_{SFARI} = 0.63$, $p=0.005$). Results for tall stature among girls were similar, with burden ratios < 1. Deletions, and genic deletions, showed nominally statistically significant differences ($T/S_{del} = 0.76$, $p = 0.010$; $T/S_{delgenic} = 0.78$, $p = 0.037$). Some signals were seen for short stature, although less of the genome was affected by deletion CNVs ($S/N=0.78$, $p=0.008$)

and large deletion CNVs (S/N=0.23, p=0.001), when both were restricted to genic regions.

For HC, children with macrocephaly had less their genomes affected when considering CNVs that overlapped with SFARI regions: large SFARI CNVs (Mac/N=0.21, p=0.032), duplication SFARI CNVs (Mac/N=0.58, p=0.049) and large duplication SFARI CNVs (Mac/N=0.01, p=0.008). Decreased CNV burden with macrocephaly was consistent in boys, although the estimated effect sizes were in the opposite direction for macrocephaly and microcephaly in girls, although sex-stratified results for HC were not statistically significant.

Compared to children with normal BMI, children with high BMI had less of their genome affected by large duplication SFARI CNVs (H/N=0.21, p=0.049). Girls, specifically, showed less CNV burden from deletions with high BMI (H/N ranged from 0.1 to 0.86 when analyzing deletion CNV, with corresponding p-values ranging from 0.001 to 0.18), and also showed some evidence for decreased CNV burden among overweight girls, particularly for deletion CNVs.

CNVs, particularly duplications, in the ASD candidate region 1q21.1 showed positive association with macrocephaly, especially in girls. ASD candidate region 15q11.2 was positively associated with short stature, again only in girls. Finally, 15q11.2 and 15q11.2.13.1 showed negative associations with microcephaly, among both boys and girls, but only achieved nominal statistical significance among boys. None of these

remained significant after correction for multiple testing. No other regions or growth features showed nominally significant comparisons.

Conclusions:

We observed nominally significant decreased CNV burden among children with tall stature in the SEED 1 sample, consistently for boys and girls. However, we also observed marked differences between sexes for overweight and high BMI, where only girls showed evidence for differential (decreased) CNV burden. Macrocephaly may also be associated with decreased CNV burden in these samples, although estimated effects were in the same direction for boys, but in the opposing direction among girls. Associations with specific CNVs in previously identified ASD candidate regions also showed sex-specific results. Our study shows CNVs may contribute to genetic risk of abnormal growth, and there appears to be some potential for common pathways involving CNVs for ASD and abnormal physical growth differing by sex. Establishing potential genotype-phenotype associations between ASD and growth abnormalities may improve risk factor identification in the subset of ASD individuals with growth abnormalities.

Key Terms: macrocephaly, microcephaly, tall stature, short stature, overweight, underweight, high BMI, low BMI, copy number variant burden, autism spectrum disorder

Introduction

Autism spectrum disorder (ASD) is a complex and highly heterogeneous neurodevelopmental disorder with multiple subtypes, and is frequently found with co-morbid conditions ranging from psychiatric to physical abnormalities. Although ASD is understood to have both genetic and environmental risk factors, no single genetic etiology has been identified and it is likely quite heterogeneous (Miles et al., 2011). Clustering all ASD cases based solely on diagnostic classification may combine many distinct genetically-driven sub-phenotypes, and thus undermine efforts to determine true causal factors. Better understanding of observable autism sub-phenotypes may improve identification of genetic risk factors, as well as allow better prognostication, and potentially allow earlier, targeted interventions. Abnormal growth, particularly in early childhood, has been associated with ASD and is known to be heritable. Consideration of abnormal growth, in the context of ASD, may be a useful sub-phenotype that could help identify a sub-group of ASD associated with genetic etiologies. In this study, we aim to investigate association between copy number variant (CNV) burden and abnormalities of physical growth in ASD and compare cases to typically developing (control) children. We also tested for association between growth abnormalities and ASD by analyzing for association between CNVs previously implicated with ASD and growth abnormalities. By investigating growth abnormality as a sub-phenotype of ASD, we may be able to parse out distinct genetic etiologies.

ASD is known to have multiple etiologies including genetic, environmental and epigenetic risk factors. Although all affected individuals share the core features of ASD,

ASD sub-phenotypes are frequently observed. These co-occurring conditions in ASD include medical, psychiatric, and behavioral conditions, e.g. cognitive impairment, gastrointestinal disturbance, and abnormalities of physical growth. As a complex disease with high heterogeneity, lumping all autism cases based strictly on diagnostic classification may hamper our ability to identify autism risk factors. Approaches utilizing sub-phenotypes of ASD may permit delineation of distinct genetic sub-groups. The over-arching objective of this study is to investigate growth abnormalities that may identify a sub-phenotype of ASD, and test for potential associations with genetic risk factors.

The exact etiologies and pathways leading to ASD are not well established, with both genetic and environmental risk factors reported, in addition to likely interactions between genes and between genes and the environment (Geschwind, 2011). Although some chromosomal abnormalities and recognized syndromes have been associated with ASD (Hall, Lightbody & Reiss, 2008; DiGuiseppi et al., 2010), the majority of ASD children do not have any known syndrome. Genetic discovery in ASD has found associations with both inherited and *de novo* (newly occurring in a child resulting from a germline change in a parent) mutations (Robinson et al., 2015). These include common and rare single nucleotide variants (SNVs) as well as copy number variants (CNVs). Some of the most consistent ASD genetic findings to date are CNVs (Robinson et al., 2014). Studies of CNVs in ASD have shown associations between specific CNVs and risk of ASD, as well as an overall association between the genome-wide CNV burden a person carries and risk of ASD. In terms of specific CNVs, studies have implicated a number of genes and chromosomal regions, including most prominently regions on chromosomes 7q11.23,

15q11-13, 16p11.2, and 22q11.2 (Pinto et al., 2010; Sanders et al., 2011), loci identified from the Simon Simplex Collection such as 1q21.1 and 3q29 (Pinto et al., 2010; Picinelli et al., 2016), as well as genomic regions encompassing the *SHANK2* (Pinto et al., 2010), *SHANK3* (Gauthier et al., 2009), *NRXN1* (Bucan et al., 2009), *CNTN4* (Roohi et al., 2009) and *CNTNAP2* (Bakkaloglu et al., 2008) genes. These genes and regions are not exclusively associated with ASD, but are associated with different neuropsychiatric phenotypes such as schizophrenia or fall under the more general term 'neurodevelopmental conditions' (Malhotra and Sebat, 2012). In our analysis, we tested for the association between ten regions previously reported to be associated with ASD, including some of the regions mentioned above, with growth abnormalities in SEED children. Increased genome-wide CNV burden, or more precisely, autosome-wide CNV burden (i.e. the sex chromosomes X and Y were excluded), has been consistently shown to be associated with ASD using several measures. Pinto et al. (2010) reported a higher count of rare CNVs among ASD cases, as well as increased deletion CNVs. Others have shown a greater length of the genome affected by CNVs, as well as greater counts, among ASD cases, especially for deletions and for rare CNVs (Vulto-van Silfhout et al., 2013). This has been shown multiple times, with some suggestion that rare *de novo* CNVs are drive evidence of association (Sanders et al., 2011) and that both deletion and duplications are relevant (Luo et al., 2012).

Various studies indicate the proportion of children diagnosed with ASD who have physical signs of growth anomalies range between 5-30% (Angkustsiri et al., 2011; Ozgen et al., 2010; Weir et al., 2006; Miles & Hillman, 2000; Miles et al., 2005; Ozgen et

al., 2011; Ozgen et al., 2013). Any child with dysmorphology, abnormal physical features, has a higher likelihood of carrying detectable genetic aberrations. Abnormal growth is a specific set of dysmorphic features, typically including abnormal head circumference, height, or weight for a child's sex and age. Children with ASD have been recognized to have abnormalities in several modalities of physical growth. Amongst these abnormalities is accelerated overgrowth of the head early in development, leading to macrocephaly in young children with ASD (Lainhart et al., 1997; Miles & Hillman, 2000; Courchesne et al., 2001; Redcay & Courchesne, 2005; Courchesne et al., 2003). Approximately 15-20% of children with autism have been reported to have macrocephaly in various studies, with some variability (Fombonne et al., 1999; Dementieva et al., 2005; Lainhart et al., 2006). The child's age influences when growth abnormalities are seen. Macrocephaly is often not observed at birth, but by the first year of life it begins to be more frequently recognized (Courchesne, Carper & Akshoomoff, 2003; Hazlett et al., 2005, Fukumoto et al., 2008, Mraz et al., 2007).

In addition to abnormal head growth, children with ASD have also been reported to have abnormal growth for both weight and height. Recent studies on
ASD children show an association between autism and obesity or being overweight (Curtin et al., 2005; Curtin et al., 2010; Rimmer et al., 2010; Evans et al., 2012; Broder-Fingert, 2014; Zuckerman & Fombonne, 2015; Must et al., 2016). ASD children have also been reported to have tall stature and generalized overgrowth, especially in boys (van Daalen et al., 2007; Xiong et al., 2009; Chawarska et al., 2011).

Multiple factors likely affect different aspects of physical growth. Growth in children is influenced by genetic, environmental and gene-by-environment interactions. Taking stature as an example, height is a complex phenotype influenced by many genetic factors (Marouli et al., 2017; Lettre, 2009). Genome-wide association studies have identified hundreds of common genetic variants influencing adult height, but together these only explain a small proportion of the estimated genetic variation. Variation in height due to genetic markers may reflect combined effects of genes (both common and rare variants), gene-by-gene or gene-by-environment interactions (Hirschhorn & Lettre, 2009; Lango et al., 2010). Genetic control of height may vary across the distribution of height in percentiles. While height is largely attributable to the combined effects of multiple genes, extreme height abnormalities (e.g. short and tall stature), may be controlled by single rare variants which exert large effects (Hirschhorn & Lettre, 2009; Lango et al., 2010; Hemani et al., 2013). These rare variants may be individual SNPs or CNVs. Some studies report up to 10% of children with idiopathic short stature carry pathogenic CNVs (Canton et al., 2014; Zahnleiter et al., 2013; van Duyvenvoorde et al., 2014). Rare, genic CNVs have also been implicated in short stature (Dauber et al., 2011; Zahnleiter et al., 2013). Zahnleiter et al. (2013) found both CNV deletions and duplications in individuals with extremely short stature, while Dauber et al. (2011) reported CNV deletions were associated with short stature in children presenting with neurodevelopmental disorders. Dauber et al. (2011) performed a genome-wide CNV burden analysis, and showed children with short stature had higher combined CNV burden, with both longer lengths and higher counts of CNVs. They did not find any association between CNVs and tall stature. Specific associations between other growth abnormalities, e.g. obesity, have also

been found in some CNV regions (Jarick et al., 2011). Genome-wide CNV burden may be one of the genetic factors contributing to genetic variation not only in stature, but also in other modalities of physical growth.

Some studies have reported specific CNV regions to be associated with growth abnormalities and ASD, most commonly macrocephaly (Klein et al., 2013). Associations with novel CNVs located at 6q23.2 and 10q24.32 have been observed with macrocephaly (Conti et al., 2012). A phenomenon described as 'mirror phenotypes' has also been described for CNVs at regions 16p11.2 and 1q21.1, where the growth phenotype observed (abnormally small or abnormally large) depends on whether the CNV is a deletion or duplication. CNVs at location 16p11.2 show opposite effects on BMI and head circumference based on the variant present: CNV deletions were associated with ASD, obesity and macrocephaly while duplications were associated with ASD, schizophrenia, underweight and microcephaly (Shinawi et al., 2010; Jacquemont et al., 2011; Qureshi et al., 2014; Stein, 2015). The 1q21.1 CNV deletion presents with microcephaly, while the duplication CNVs showed association with macrocephaly (Rosenfeld et al., 2012; Bernier et al., 2016). It has been suggested that gene dosage leads to differential gene expression may be the biological mechanism by which CNVs can potentially affect phenotypes in opposite directions (McCarroll et al., 2006; Di Gregorio et al., 2017).

In this study, we examined CNV burden over all autosomes comparing children with abnormalities in head circumference, height, weight and BMI. Characterization of CNV

burden for specific growth patterns, and assessment of this signal in ASD children versus typically developing children may support the co-occurrence of abnormal growth and ASD as a specific sub-phenotype.

Materials and Methods

Study Population

The Study to Explore Early Development (SEED), phase I, is a national multi-site case control study funded by the CDC. Children between ages 2-5 years old were recruited and evaluated at one of six states: California, Colorado, Georgia, Maryland, North Carolina and Pennsylvania. Potential cases, born between 2003 and 2006 were recruited through partnerships with developmental disability service providers, including healthcare and educational systems. A population-based sample of control children born in the same years from the same catchment areas was recruited through vital statistics. After phone-based screening and in-person evaluations, children were classified as ASD, other non-ASD developmental disabilities, and children without developmental disabilities from the general population (POP) (Schendel et al., 2012). Eligibility included birth in the study catchment area during the period 9/1/2003-8/31/2006, current residence in the area at the time of first contact, and child living with a knowledgeable caregiver who was able to communicate orally in English or Spanish competently and gave informed consent for participation. The enrolled children also were between the ages of 30 and 68 months of age at the completion of the clinical developmental assessment.

Growth Measures

SEED participants underwent anthropometric measurements and had standardized photographs taken of specific regions of the body. The specific growth measures of interest here were collected using a standardized procedure by trained clinic staff using

standardized supplies including a tape measure for head circumference, the stadiometer and the weighing scale. For head circumference, a non-stretchable, plasticized measuring tape was used to measure the head circumference for maximum circumference of the head. The tape was placed just above the eyebrows, above the ears and around the most protuberant part of the back of the head (occiput), pulled snugly to compress hair and read to the nearest 0.1cm, after which the measurement was recorded on the Dysmorphology Review Form (DRF) form and repeated, until repeated measurements were within 0.2cm. For height measurement, an accurate and appropriate stadiometer was used: a vertical board with an attached metric rule and a horizontal headpiece that could be brought into direct contact with the most superior (top) part of the head, and read to the closest 0.1cm. Height measurement was performed for all children with hair accessories removed and without shoes. The child was measured standing with heels, buttocks, shoulders and head touching a flat upright surface. The arms were held on the side, with shoulders relaxed and legs straight, and heels close together. The child was asked to look straight ahead and the perpendicular headpiece lowered to the crown of the head snugly with compression of the hair. The measurer's eyes were parallel with the headpiece. The measurement was repeated, with agreement to within 1cm, and recorded on a growth chart appropriate for the child's age and sex. The raw measurement and percentile growth from the percentile chart was then transferred to the DRF form. The stadiometer position was standardized, so there were no attachments to the wall and no underlying carpet. The stadiometer was also calibrated monthly. For the measurement of weight, a safe and accurate scale with a wide enough platform to support the child being weighed was used. The scale was required to be calibrated with standard weights, and

could be zeroed, and not be positioned on a carpeted surface. The child stood on the weighing platform without assistance and wearing only light undergarments or gown. The reading was recorded, and repeated until agreement within 0.1kg.

For this analysis, raw values of head circumference, height, and weight were converted to z-scores and percentiles using the 2000 CDC Head Circumference-for-Age Growth Charts for girls and boys ages 2–20, Stature for-Age Growth Charts for girls and boys ages 2–20, Weight-for-Age Growth Charts for girls and boys ages 2–20 and BMI- for-Age Growth Charts for girls and boys ages 2–20.

Children were categorized as abnormal if they were beyond the upper and lower 10^{th} percentiles for any growth feature: above 90% designated macrocephaly, tall stature, overweight, or high BMI; and below 10% designated microcephaly, short stature, underweight or low BMI. Normal growth children were designated as those with measures between 10^{th} and 90^{th} percentiles, based on age and sex-specific references.

ASD Assessment

For all eligible children, a brief screening interview, the Social Communication Questionnaire (SCQ; Rutter et al. (2003)), was administered to the primary caregiver to identify children who required clinical diagnostic assessment to determine final ASD status. For SEED, a positive screen was defined as an SCQ score ≥11. Tools for ASD assessment included the Autism Diagnostic Observation Schedule (ADOS) and Autism

Diagnostic Interview-Revised (ADI-R) (Falkmer et al., 2013; Lord et al., 2000). Final classification was assigned using a SEED-specific research algorithm based on ADOS, ADI-R and clinical judgment (Wiggins et al., 2015). Regardless of ascertainment source, any eligible children with a previous ASD diagnosis, who were receiving special education services, and who had a positive screen, were assigned to the ASD workflow. This determined which instruments were administered and the type of diagnostic evaluation the child received during data collection. Based on previous diagnosis and SCQ screening, DD and POP children with negative SCQ screens were assigned to the DD or POP workflow, respectively. If a clinician suspects ASD during the clinical evaluation of a child in the DD or POP workflow, the child would be moved into the ASD workflow (Schendel et al., 2012; DiGiuseppe et al., 2016).

Copy Number Variants (CNVs)

Blood and buccal samples were collected by trained local staff and shipped to the SEED Biosample Repository at Johns Hopkins Bloomberg School of Public Health. These were used to isolate DNA via the QIAsymphony DNA Investigator and QIAsymphony DNA Midi kits (Qiagen) for buccal and blood, respectively. A total of 1,132 SEED 1 cases and controls were genotyped at 1 million single nucleotide polymorphisms (SNPs) on Illumina Human Omni1-Quad array. Genotyping and initial data cleaning was carried out at the Johns Hopkins University SNP Center. Quality control measures at the SNP and sample levels were performed. Samples were excluded if <98% of all markers were called successfully, if estimated identity by descent (IBD) sharing suggested cryptic

relatedness between subjects, if there were sex discrepancies, or if there was excess heterozygosity/homozygosity. SNPs were excluded for call rates <0.95, minor allele frequency (MAF) <0.01 and if there was evidence of deviation from expected genotype frequencies predicted by Hardy-Weinberg equilibrium ($p<1.0x10^{-8}$ in controls).

CNVs were called using a hidden Markov model implemented in PennCNV (Wang et al., 2007). Hidden copy number state along each chromosome was estimated using total signal intensity, allelic intensity ratio, SNP allele frequency, distance between neighboring SNPs, and genomic GC content (Diskin et al., 2008). Quality control (QC) filters were applied at both the CNV and sample levels. CNVs were filtered out if they contained < 10 SNPs, were < 30 kb, or were in centromere and telomere regions; samples were excluded if the standard deviation of the log R ratio (LRR) >0.3, the B-allele frequency (BAF) >0.01, or absolute value of a 'wave' factor (due to high GC content over the region) >0.05. The overall data quality pipeline is shown in Figure 2.

Burden Metrics. We considered both CNV counts across all autosomes, as well as their summed lengths. CNV count was based on the total number of unique CNV sites for each individual; length was determined by summing the length of these unique CNVs in kilobases (kb) per individual. Overall count and length burden metrics included both duplications and deletions, and CNVs occurring anywhere in an autosome. Measures were also calculated separately for duplications and deletions, and for only large (>400kb) CNVs. Finally, subsets of autosomal CNVs were considered: only CNVs overlapping known genes, using hg19 gene boundaries (categorized as "genic CNVs"),

and only CNVs overlapping genes associated with ASD, using the Simons Foundation Autism Research Initiative gene list (categorized as "SFARI CNVs"). The UCSC genome database using "TxDb.Hsapiens.UCSC.hg19.knownGene", "annotate", and "org.Hs.eg.db" Bioconductor packages (Goldstein et al., 2016; Carlson, 2016) were used to establish known genes and their boundaries. For the "SFARI CNVs", SFARI gene 2.0_ENREF_22 was consulted, and a list of 757 autosomal candidate genes for ASD was used.

CNV Candidate Regions. Malhotra and Sebat (2012) reviewed specific CNV regions associated with ASD, and identified the precise boundaries for each region compared to the catalog on CNVs available from SFARI (Malhotra & Sebat, 2012; https://gene.sfari.org). The largest interval between start and end of any CNV was used to best define the affected chromosomal region. We focused on 10 ASD-associated CNVs for regional analysis: chromosomes 1q21.1, 3q29, 7q11.23, 15q11.2, 15q11.2.13.1, 15q13.3, 16p11.2, 16p13.11, 17q12 and 22q11.21.

Statistical Analysis

CNV burden analyses were carried out for eight growth abnormalities: macrocephaly, microcephaly, tall and short stature, overweight and underweight plus high and low BMI, comparing children with each abnormality to children with typical growth in that domain. CNV burden was assessed for both CNV counts and CNV lengths, with effect sizes estimated as the ratio of each measure between the children with abnormal growth feature

(e.g. macrocephaly) to those with normal growth. Mean counts or cumulative lengths were compared between growth groups using t-tests. These analyses were carried out overall, for deletions and duplications separately, and then restricting to: CNVs >400 kb (large CNVs), CNVs overlapping with genes (Genic CNVs), and CNVs overlapping with ASD-associated genes (SFARI Genic CNVs). Analyses were also stratified by sex. Finally, sensitivity analyses were performed to assess the effect of excluding or including children with chromosomal abnormalities and recognized genetic syndromes, and excluding and including children diagnosed with developmental delay. We used STATA (MP12.1) graphics to compare effect size between the whole analytic population and Non-Hispanic White children only using scatter plots for each growth abnormality. For candidate region analyses, CNV counts between abnormal and typical growth phenotypes for each growth feature were compared via Fisher's exact tests for CNVs over all autosomes and for deletion and duplication CNVs, as well as for large CNVs. Sex-stratified analyses were also performed. All computational analyses were performed using R 3.3.2.

Results

Characteristics of the Study Sample

There were 3,899 children recruited into the SEED 1 study. Of these, genotyping was performed on 1,132, and of these 1,016 had complete data on growth. Our analyses included all SEED 1 children, except 147 children with non-ASD developmental delay and 29 with known chromosomal abnormalities and non-genetic syndromes, resulting in

840 samples, 341 (38.8%) children with ASD and possible ASD, and 499 (59.4%) typically developing children. The study flow chart describing how the final study population was obtained, and frequencies of each growth abnormality is shown in **Figure 1**. The growth abnormality with the largest number of individuals and highest proportion was microcephaly, with 165 (19.6%) children in the study categorized with microcephaly. The growth abnormality with the fewest individual and lowest percentage was macrocephaly, with 52 (6.2%) children. For growth abnormalities on the lower end of the spectrum, with growth measures $\leq$ the 10th percentile for growth, growth abnormalities range from 7.5% to 9% (underweight and low BMI respectively) of the whole sample for each growth abnormality, and for growth abnormalities on the upper end of the spectrum, with growth measures at or above the 90th percentile of growth, growth abnormalities range between 12.6% to 14% (overweight and high BMI respectively) of the total sample.

The sex, race, and ASD status frequencies among each type of growth abnormality are shown in **Table 1**. There was a greater preponderance of males, with 549 (65.4%) males and 291 (34.6%) females in the whole sample. The majority (60.4%) of the children were Non-Hispanic White, with 17.8% Non-Hispanic Black and 21.8% Hispanic. Children with macrocephaly and children with high BMI had a higher proportion of ASD cases than children with normal head circumference. Children with high BMI were also more frequently Hispanic than normal-BMI children.

CNV Burden in Growth Abnormalities

CNV burden is reported as the ratio of counts or cumulative CNV lengths per child among each abnormal growth group compared to control children with normal growth for that feature. Complete results for each growth abnormality as well as results stratified by sex are reported in Supplementary Tables I through XVI. For the study sample, there were a total of 18,226 individual CNVs calculated, with a mean CNV count of 21.69 per person. The summary of the effect sizes and t-test results for CNV burden by count and lengths for each of the eight growth abnormalities considered are shown in **Table 2**. The burden ratio for tall versus normal stature children was <1 for both counts and lengths (Tall/Norm counts = 0.909, p = 0.029; Tall/Norm length = 0.833, p = 0.033), although this was not statistically significant after Bonferroni correction for multiple tests.

Effect sizes and nominal p-values for each type of CNV test performed (CNVs overall, by deletions and duplications, and restricting to large CNVs, all genic CNVs and SFARI CNVs), across all eight growth abnormalities, are shown in **Table 3**, to allow comparison across both growth and CNV types. Across all growth abnormalities, the ratio of mean CNV length for those with a growth abnormality to those without for overall CNVs range from 0.83 (in tall stature vs. normal stature) to 1.62 (in macrocephaly vs. normal head size). Consistent with the overall CNV length differences between tall children and normal height children described in **Table 2**, similar patterns were observed when restricting to genic and SFARI genic CNVs (T/N_{genic} = 0.83, p=0.06; T/N_{SFARI} = 0.79, p=0.044), and among deletions overlapping SFARI genes (T/N_{SFARI} = 0.63, p=0.005) (**Table 3**). For CNVs overall, children with normal stature had a 1.10-fold increase in

total CNV count per individual, compared to children with tall stature (Supplementary Table V).

Some signals were seen for short stature, with less of the genome affected by deletion CNVs (S/N=0.78, p=0.008) and large deletion CNVs (S/N=0.23, p=0.001), when restricted to genic regions. For head circumference, children with macrocephaly had less of the genome affected when considering CNVs overlappin SFARI regions: large SFARI CNVs (Mac/N=0.21, p=0.032), duplication SFARI CNVs (Mac/N=0.58, p=0.049) and large duplication SFARI CNVs (Mac/N=0.01, p=0.008). Compared to children with normal BMI, children with high BMI had less of their genome affected by large duplication SFARI CNVs (highBMI/N=0.21, p=0.049). In tall and short stature, the CNV subtype involved were deletion CNVs, while in macrocephaly and high BMI, the CNV subtypes were duplication CNVs. For the main analysis, there were no significant findings for CNV burden comparing overweight or underweight children with normal weight children, for both CNV counts and lengths. None of these nominally significant differences met Bonferroni criteria for multiple testing (set at $\alpha<0.0005$).

Results for the same analyses, stratified by sex, are shown in **Tables 4 and 5**. In males, burden among macrocephaly and microcephaly boys appeared to be lower than typical boys, although results were not statistically significant. Similar results to non-stratified analyses were observed for tall stature. In particular, male children with tall stature had fewer SFARI-overlapping CNVs, and SFARI-overlapping deletion CNVs, cumulative lengths (T/N=0.74, p=0.047; T/Ndel=0.57, p=0.012). Other effect sizes were similar, but

did not reach nominal statistical significance. Males with short stature also showed less CNV length burden across overall and subtype analyses. The strongest statistical significance was for deletion CNVs in genes, large deletion CNVs in genes, and large duplication CNVs in genes (S/N delgenic = 0.72, p = 0.0002; S/Nlargedelgenic = 0.13, p = 0.0002; S/Nlargedupgenic = 0.48, p = 0.04). After correcting for multiple testing at $p<0.0005$, the only findings still significant were the association between CNV burden and short stature in males, for deletion CNVs in known genes and large deletion CNVs in genes.

Among females, macrocephaly and microcephaly had burden ratio estimates >1 for most analyses, although none were statistically significant. This was in the opposite direction from male results, although in both strata, these estimates were not significant. Results for tall stature among girls were similar to unstratified analyses, with burden ratios < 1. Deletions, and genic deletions, showed nominally statistically significant differences (T/Sdel = 0.76, p = 0.010; T/Sdelgenic = 0.78, p = 0.037). Girls also showed some evidence for decreased CNV burden among overweight children, particularly for deletion CNVs. This corresponded with a smaller deletion burden among girls with high BMI (H/N ranged from 0.1 to 0.86 among deletion analyses, with corresponding p-values ranging from 0.001 to 0.18). However, none of the significant results survived correction for multiple testing at $p<0.0005$.

Considering **Tables 3 – 5** together, there are consistent associations across sex, such as decreased burden among tall stature, and sex-specific associations, particularly decreased burden among overweight and high BMI among girls.

ASD Candidate CNV Regions

To test for a potentially shared genetic risk for growth abnormalities and ASD, we tested for association between each of the eight growth abnormalities and CNVs in regions previously associated with ASD. Counts of all CNVs, deletion and duplication CNVs and large CNVs for CNV regions reported to be associated with growth abnormalities are shown in Supplementary Tables XVII-XXIV. **Tables 6 and 7** summarize results for any comparisons with nominal $p<0.05$ in overall, and sex-stratified analyses, respectively. Region 1q21.1 showed increased numbers of CNVs among children with macrocephaly, particularly duplication CNVs, but this appears to be specific to girls (**Table 7**). Region 15q11.2 and 15q11.2.13.1 showed decreased numbers of CNVs among children with microcephaly, and 15q11.2 had increased counts among children with short stature. Decreased frequencies among the microcephaly group were observed among both boys and girls, but only achieved nominal statistical significance among boys. The increased counts at 15q11.2 related to short stature appear to be driven by girls (**Table 7**). None of these remain significant after correction for multiple testing. No other regions or growth features showed nominally significant comparisons in the unstratified analyses.

Discussion

We explored the use of observed growth abnormalities as sub-phenotype of ASD for genetic association studies, in hopes that shared genetic association between growth abnormalities and ASD could reveal possible shared mechanisms for ASD and growth. In our sample of young children from the Study to Explore Early Development, Phase 1 (SEED 1), a national case-control study of autism, we observed nominally significant decreased CNV burden among children with tall stature. This pattern was consistent across boys and girls. However, we also observed marked differences between sexes for overweight and high BMI, where only girls showed evidence for differential (decreased) CNV burden. Macrocephaly may also be associated with decreased CNV burden in these samples, although effect estimates were in the same direction for boys, but in opposing direction among girls. Notably, CNV subset analysis restricting to consideration of only those overlapping candidate genes for ASD (SFARI – genic) were generally implicated in the burden associations, showing decreased burden among children with particular growth abnormalities. Consistent with this specificity to ASD-associated regions, CNV associations in 10 specific regions previously identified as ASD candidate regions were also significant, but also showed sex-specific results. None of these nominally significant associations survive correction for multiple testing.

The most striking finding across our nominally significant results is the reduced, rather than increased, genome CNV burden among children with abnormal growth. This is surprising, as previous studies of CNV burden and growth, or neurodevelopmental conditions, have reported increased CNV burden. One possible reason is our exclusion of

children with known chromosomal abnormalities, recognized genetic syndromes and non-ASD developmental delay. This was intentional, and should allow clearer understanding of risk of ASD in a generally idiopathic group, but excluding the cases with chromosomal abnormalities, recognized genetic syndromes and non-ASD developmental delay likely excluded children with more severe ASD phenotypes. Previous studies showing excess burden did not exclude these types of children (Dauber et al., 2011). As a sensitivity analysis, we ran the same analyses including the 147 non-ASD developmental delay (DD) children who were also genotyped in SEED and had growth data. The decreased burden signal goes away when DD children are included, and in some cases, increased burden was observed (Figure 3). As an example, with DD children in the analysis, there was increased genome burden with short stature. This is more consistent with Dauber et al.'s (2011) findings of increased CNV burden in children presenting with clinical indications. Our design, nested within an ASD case-control study, with a high proportion of ASD children, may have also influenced our results. The expected direction, assuming some genetic overlap between ASD and growth abnormalities, and previous findings of increased CNV burden among ASD cases, would be opposite of our observation of smaller CNV burden in ASD males. Decreased burden has occasionally been observed for neurodevelopmental disorders. Martin et al. (2014) found a negative association between CNV burden for total rare duplication CNVs and positive symptoms of schizophrenia. They hypothesized this represented CNV duplication burden may have a small protective effect against positive symptoms of schizophrenia. As far as we are aware, no studies on CNV burden and growth have reported this type of negative association.

Previous research on CNV burden and growth abnormalities has focused on single features of growth, such as short stature (Dauber et al., 2011) or obesity (Jarick et al., 2011), rather than considering multiple features simultaneously as done here. We set out to examine how CNV burden as a genetic risk factor, may be associated with each growth abnormality, while also being able to look across abnormalities in the same children. This helped us identify potential patterns of CNV subtype associations across growth abnormalities, like duplications versus deletions, and restrictions related to genic and ASD genic regions versus general. Both deletions and duplications were found to be associated with abnormal growth in our sample. Height abnormalities tended to be associated with deletions, while in macrocephaly and BMI were associated with duplications. Previous studies have reported pathogenic CNVs are often deletions (Serra-Juhe et al., 2017). However, it has been discovered that CNV duplications can also lead to aberrant gene expression; an example of this is the shift of reading frame that encodes a stop codon leading to null mutation seen in SRGAP3 gene duplication in childhood schizophrenia (Wilson et al., 2011).

Our burden and candidate region analyses showed different effects between males and females. Tall stature has been previously reported in ASD, particularly in boys (Curtin et al., 2005; van Daalen et al., 2007; Chawarska et al., 2011), and potentially the decreased CNV burden in our sample may contribute to this genetic risk. For females, less of the genome was affected by CNVs with macrocephaly, tall stature, overweight, underweight and high BMI. These CNVs were primarily deletions, and often showed stronger

evidence with restricted to SFARI genes, which are canidates for ASD. Macrocephaly and high BMI have both been described in ASD children, including in girls with ASD (Courchesne et al., 2003; Hazlett et al., 2005; Fukumoto et al., 2008; Mraz et al., 2007; Ahearn et al., 2001; Evans et al., 2012; Chen et al., 2009; Curtin et al., 2010), and CNVs may be a potential genetic risk factor.

When considering specific ASD candidate regions, CNV associations further appeared to be sex-specific. Females may be driving the association between duplication CNVs in the 15q11.2 region with short stature, and association for CNVs in region 1q21.1 with macrocephaly. In both cases, there were greater CNV counts in the abnormal growth girls compared to normal girls, but not as strongly differential in boys. The association between CNV region 15q11.2 and short stature in females has not described in published literature. The region encompasses *TUBGCP5, CYFIP1, NIPA2,* and *NIPA1* genes, which have been implicated in axonal growth and neural connectivity, with duplication carriers having variable phenotype (Burnside et al., 2011; Picinelli et al., 2016). We also observed significant overlap between ASD-associated CNVs in regions 15q11.2 and 15q11.2.13.1 with microcephaly, but only in males. The association was in the negative direction, with less CNV counts seen among microcephalic boys.

These findings generally support a hypothesis of differential genetic risks in males and females for both ASD and growth. Sex dimorphism in growth abnormalities has been described in ASD (Wells et al., 2007; Suren et al., 2013; Werling & Geschwind, 2013; Campbell et al., 2014; Lai et al., 2015; Lai, Baron-Cohen & Buxbaum, 2015; Werling,

2016) and these CNV burden findings suggest the mechanism underlying these observations may be differ by sex.

Our cross-sectional analysis sample was drawn from a case-control study of ASD, with over-representation in the ASD case group compared to the population. This may have influenced our results, particularly those showing burden specific to CNVs that overlap ASD-associated genes. However, most associations suggest decreased burden among cases, which is not intuitively consistent with ASD sample enrichment for boys. Thus, other explanations, such as excluding ASD cases with reported genetic anomalies, are the most likely clue to interpretation. One advantage of our design is the ability to test for the association between CNVs influencing risk to ASD and growth abnormalities. For example, complementary growth phenotypes have been seen for CNVs at 16p11.2, a region with recognized CNVs for ASD. Individuals with 16p11.2 deletions are more likely to have ASD, be overweight and have macrocephaly and those with 16p11.2 duplications have increased risk of ASD and schizophrenia, underweight and microcephaly (Shinawi et al., 2010; Jacquemont et al., 2011; Qureshi et al., 2014; Maillard et al., 2014; Stein, 2015). Other regions such as 7q11.23 and 1q21.1 have been implicated with ASD and abnormal head size (Merla et al., 2010; Sanders et al., 2011). In our study, we did not observe any significant growth associations for CNVs on 16p11.2.

We did observe a 12-fold increase in CNV deletions on 1q21.1 among children with microcephaly. CNVs may influence gene expression not only through coding sequence disruption, but also through gene dosage effects (either loss or gain of genes in the

region) (McCarroll et al., 2006). The opposite head size phenotypes associated with different CNVs on 1q21.1 was previously reported by Rosenfeld et al. (2012). Potential biological mechanisms for abnormalities in head circumference resulting from CNVs have also been reported. Brunetti-Perri et al. (2008) postulated the *HYDIN* gene might play a causal role because it has been implicated in regulating the cerebral cortex size, while Rosenfeld et al. (2012) suggested that *PIAS3*, a regulator of hematopoietic growth factor signaling and *LIX1L*, which has been associated with limb development, might be part of the biological pathway leading to abnormal head size. Sequences in the 1q21.1 region encoding DUF1220 protein domains with variable gene dosage in the *NBPF* gene have been implicated with brain size (Davis et al., 2014). This group further suggested a potential biologic mechanism underlying abnormal brain size results from variation in DUF1220 domain could lead to disruption of mitotic cell regulation and neuronal migration.

We also found significant association between ASD-associated genes in 15q11.2 and 15q11.2.13.1 regions with microcephaly, and the 15q11.2 region and short stature. The region 15q11.2 has been linked with specific learning disabilities and abnormalities of brain region size, and reported to have complementary phenotype effects based on pattern of CNV loss or gain (Ulfarsson et al., 2017). The association between region 15q11.2 and head circumference (as well as stature) suggest possible pleiotropy in this CNV region, which could influence multiple traits, like ASD and different growth abnormalities. These shared genetic susceptibilities may point towards a common biological pathway.

Study Limitations and Strengths

Our cross-sectional design, drawn from a national case-control study may not be fully representative of the United States in terms of racial/ethnic make-up (Schendel et al., 2012). Although the study recruitment included a population-based approach, ascertainment bias for selection into the study is still possible. In terms of limitations associated with the growth abnormality outcome, the young age of the children limits generalizability, as growth trajectories are still in process. This is further limited by the cross-sectional nature of our measurement. To truly understand physical growth, repeated measurement over time is preferred.

Despite being one of the largest autism case-control samples in the US, after restricting to complete cases analyses, our sample size, particularly the number of children categorized with each growth abnormality, was modest. The smallest subsample with growth abnormalities was for macrocephaly, with only 52 children. This limits power and the ability to adjust for a large number of potential confounders. In addition, the study population size was too small to consider more extreme thresholds for defining growth abnormality. For example, the SEED dysmorphology assessment for each growth feature, define children with height at or greater than 3 s.d. from the mean as 'tall', which would have resulted in even smaller numbers.

An important limitation already discussed is that there may have been undiagnosed chromosomal abnormalities and recognized genetic syndromes in our sample. The

information we did have on such conditions was based on parental report, not direct genetic evaluation. Thus, it is likely additional chromosomal abnormalities or genetic syndromes were present and may have influenced our analyses.

We were also not able to stratify analyses on *de novo* versus inherited CNVs, yet *de novo* CNVs have been directly implicated in previous literature on growth abnormalities (Dauber et al., 2011; Zahnleiter et al., 2013). Unfortunately, we did not have parental genotyping information and could not determine the impact of *de novo* CNVs. In addition, rare CNVs have also been associated with congenital malformations (Serra-Juhe et al., 2012). Although the SEED Study is one of the largest population-based samples of ASD in the US, for the purpose of studying growth abnormalities, we only had a subset of the full SEED 1 sample, and were underpowered.

This study also has several strengths. The SEED Study is one of the largest ASD studies with population-based ascertainment. Previous studies have used clinic-based samples, and had smaller sample sizes. The SEED Study involves a varied geographical location across the United States, to best represent the diverse experiences in different regions. A major strength is the uniform developmental assessment of all study participants including research-reliable ASD classification and rigorous, standardized anthropometric measures across multiple domains of growth.

In summary, we found overall CNV burden is lower among children with a variety of growth abnormalities in SEED, but there differences by sex. Although the results were

not as we originally anticipated, and none of the nomially significant associations survive correction for multiple testing, we think these observations are worthy of further investigation. Growth abnormalities affect a substantial number of children, and understanding their etiology is an important component of understanding the health of children. Growth abnormalities may also represent a distinct sub-phenotype for ASD, which currently affects 1 in 68 children in the United States (CDC, 2016). Elucidating ASD phenotypes and sub-phenotypes (including growth abnormalities) may aid toward greater understanding of the multiple facets of ASD and a deeper understanding of the biology causing ASD. Greater appreciation of the association between ASD and abnormal growth, and its genetic influences, may also have clinical applications by leading to improved detection of ASD in children with growth abnormalities and allowing for earlier intervention, as well as recognition of growth abnormalities in ASD that could then be better addressed and managed.

Due to the relatively small sample sizes, these results must be considered preliminary, and future work replicating these findings need further evaluation for genotype-phenotype associations in ASD.

The IRB Committee of Johns Hopkins School of Public Health approved this study. Informed consent was obtained from all caregivers before clinical assessment as part of the SEED Study protocol. Diagnoses and assessments were performed at the six SEED sites in the United States.

References

Abrahams, B. S., Arking, D. E., Campbell, D. B., Mefford, H. C., Morrow, E. M., Weiss, L. a, … Packer, A. (2013). SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Molecular Autism*, 4, 36.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*. Arlington, VA: American Psychiatric Publishing.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological Medicine*, *25*, 63–77.

Bakkaloglu, B., O'Roak, B. J., Louvi, A., Gupta, A. R., Abelson, J. F., Morgan, T. M., … State, M. W. (2008). Molecular Cytogenetic Analysis and Resequencing of Contactin Associated Protein-Like 2 in Autism Spectrum Disorders. *American Journal of Human Genetics*, *82*(1), 165–173.

Bernier, R., Steinman, K. J., Reilly, B., Wallace, A. S., Sherr, E. H., Pojman, N., … Chung, W. K. (2015). Clinical phenotype of the recurrent 1q21.1 copy-number variant. *Genetics in Medicine*, *18*(April), 1–9.

Broder-Fingert, S., Brazauskas, K., Lindgren, K., Iannuzzi, D., & Van Cleave, J. (2014). Prevalence of overweight and obesity in a large clinical sample of children with autism. *Academic Pediatrics*, 14(4), 408–414.

Brunetti-Pierri, N., Berg, J. S., Scaglia, F., Belmont, J., Bacino, C. A., Sahoo, T., … Sally, S. (2008). Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat. Genet. 40*(12), 1466–1471.

Bucan, M., Abrahams, B. S., Wang, K., Glessner, J. T., Herman, E. I., Sonnenblick, L. I., … Hakonarson, H. (2009). Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genetics*, *5*(6).

Buescher, A. V. S., Cidav, Z., Knapp, M., & Mandell, D. S. (2014). Costs of Autism Spectrum Disorders in the United Kingdom and the United States. *JAMA Pediatrics*, *168*(8), 721.

Burnside, R. D., Pasion, R., Mikhail, F. M., Carroll, A. J., Robin, N. H., Youngs, E. L., … Butler, M. G. (2011). Microdeletion/microduplication of proximal 15q11.2 between BP1 and BP2: A susceptibility region for neurological dysfunction including developmental and language delay. *Human Genetics*, *130*(4), 517–528.

Canton, A. P. M., Costa, S. S., Rodrigues, T. C., Bertola, D. R., Malaquias, A. C., Correa, F. A., … Jorge, A. A. L. (2014). Genome-wide screening of copy number variants in children born small for gestational age reveals several candidate genes involved in growth pathways. *European Journal of Endocrinology*, *171*(2), 253–262.

Carlson, M.R., Pagès, H., Arora, S., Obenchain, V., Morgan, M. (2016). Genomic Annotation Resources in R/Bioconductor. *Methods Mol Biol*.. 1418:67-90.

Campbell, D. J., Chang, J., & Chawarska, K. (2014). Early generalized overgrowth in autism spectrum disorder: Prevalence rates, gender effects, and clinical outcomes. *Journal of the American Acad of Child and Adol Psych*, 53(10), 1063–1073.

The Centers for Disease Control and Prevention (CDC). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. Surveillance Summaries. (2016) 65(3); 1–23 http://www.cdc.gov/mmwr/volumes/65/ss/ss6503a1.htm

Chakrabarti, S., Fombonne , E. (2001). Pervasive Developmental Disorder in Preschool Children, *Journal of American Medical Association, 285*(24), 3093–3099.

Chawarska, K., Campbell, D., Chen, L., Shic, F., Klin, a., & Chang, J. (2011). Early Generalized Overgrowth in Boys With Autism. *Archives of Gen Psychiatry*, 68(10), 1021–1031.

Chen, A.Y., Kim, S.E., Houtrow, A.J., Newacheck, P.W. (2010) Prevalence of obesity among children with chronic conditions. *Obesity* (Silver Spring).18(1):210-3.

Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S. R., Woodhouse, E., … Bolton, P. (2015). Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. *JAMA Psychiatry*, 1–9.

Conti, S., Condò, M., Posar, A., Mari, F., Resta, N., Renieri, A., Neri, I., Patrizi, A., Parmeggiani, A. (2012). Phosphatase and tensin homolog (PTEN) gene mutations and autism: literature review and a case report of a patient with Cowden syndrome, autistic disorder, and epilepsy. *J Child Neurol*. 27(3):392-7.

Courchesne, E., Karns, C. M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D., … Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study . *Neurology*, 57(2), 245–254.

Courchesne, E., Carper, R., & Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. *JAMA*, 290(3), 337–44.

Curtin, C., Bandini, L.G., Perrin, E.C., Tybor, D.J., Must, A. (2005). Prevalence of overweight in children and adolescents with attention deficit hyperactivity disorder and autism spectrum disorders: a chart review. *BMC Pediatr*., 5:48.

Curtin, C., Anderson, S. E., Must, A., & Bandini, L. G. (2010). The prevalence of obesity in children with autism: a secondary data analysis using nationally representative data from the National Survey of Children's Health. *BMC Pediatr.*, 10, 11.

Dauber, A., Yu, Y., Turchin, M. C., Chiang, C. W., Meng, Y. A., Demerath, E. W., … Hirschhorn, J. N. (2011). Genome-wide association of copy-number variation reveals an association between short stature and the presence of low-frequency genomic deletions. *American Journal of Human Genetics*, *89*(6), 751–759.

Davis, J. M., Searles, V. B., Anderson, N., Keeney, J., Dumas, L., & Sikela, J. M. (2014). DUF1220 Dosage Is Linearly Associated with Increasing Severity of the Three Primary Symptoms of Autism. *PLoS Genetics*, *10*(3), 1–5.

Dementieva, Y.A., Vance, D.D., Donnelly, S.L., Elston, L.A., Wolpert, C.M., Ravan, S.A., DeLong, G.R., Abramson, R.K., Wright, H.H., Cuccaro, M.L. (2005). Accelerated head growth in early development of individuals with autism. *Pediatr Neurol.,*32(2):102-8

Di Gregorio, E., Riberi, E., Belligni, E. F., Biamino, E., Spielmann, M., Ala, U., … Ferrero, G. B. (2017). CNVs analysis in a cohort of isolated and syndromic DD/ID reveals novel genomic disorders, position effects and candidate disease genes. *Clinical Genetics*, (February), 1–8.

Diskin, S. J., Li, M., Hou, C., Yang, S., Glessner, J., Hakonarson, H., … Wang, K. (2008). Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Research*, *36*(19), 1–12.

Evans, E. W., Must, A., Anderson, S. E., Curtin, C., Scampini, R., Maslin, M., & Bandini, L. (2012). Dietary patterns and body mass index in children with autism and typically developing children. *Research in Autism Spectrum Disorders*, 6(1), 399–405.

Feuk, L., Marshall, C. R., Wintle, R. F., & Scherer, S. W. (2006). Structural variants: changing the landscape of chromosomes and design of disease studies. *Human Molecular Genetics*, *15 Spec No 1*(1), 57–66.

Folstein, S., Rutter, M. (1977). Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry,*18(4):297-321.

Fombonne, E., Rogé, B., Claverie, J., Courty, S., Frémolle, J. (1999). Microcephaly and macrocephaly in autism. *J Autism Dev Disord*. 29(2):113-9.

Fukumoto, A., Hashimoto, T., Ito, H., Nishimura, M., Tsuda, Y., Miyazaki, M., … Kagami, S. (2008). Growth of head circumference in autistic infants during the first year of life. *Journal of Autism and Developmental Disorders*, 38(3), 411–418.

Geschwind, D. H. (2011). Genetics of autism spectrum disorders. *Trends in Cognitive Sciences*.

Goldstein, L. D., Cao, Y., Pau, G., Lawrence, M., Wu, T. D., Seshagiri, S., & Gentleman, R. (2016). Prediction and quantification of splice events from RNA-seq data. *PLoS ONE*, *11*(5), 1–18.

Grayton, H.M., Fernandes, C., Rujescu, D., Collier, D.A.(2012). Copy number variations in neurodevelopmental disorders. *Prog Neurobiol*. 99(1):81-91.

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., Miller, J., Fedele, A., Collins, J., Smith, K., Lotspeich, L., Croen, L.A., Ozonoff, S., Lajonchere, C., Grether, J.K., Risch, N. Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry* 68(11):1095-102 (2011).

Hazlett, H., Poe, M., Gerig, G., & al, et. (2005). Magnetic resonance imaging and head circumference study of brain size in autism: Birth through age 2 years. *Archives of Gen Psyciatry*, 62(12), 1366–1376.

Hemani, G., Yang, J., Vinkhuyzen, A., Powell, J. E., Willemsen, G., Hottenga, J. J., … Visscher, P. M. (2013). Inference of the genetic architecture underlying bmi and height with the use of 20,240 sibling pairs. *American Journal of Human Genetics*, *93*(5), 865–875.

Hirschhorn, J. N., & Lettre, G. (2009). Progress in genome-wide association studies of human height. *Hormone Research*, *71*(SUPPL. 2), 5–13.

Hochstenbach, R., Buizer-Voskamp, J. E., Vorstman, J. A. S., & Ophoff, R. A. (2011). Genome arrays for the detection of copy number variations in idiopathic mental retardation, idiopathic generalized epilepsy and neuropsychiatric disorders: Lessons for diagnostic workflow and research. *Cytogenetic and Genome Research*, 135(3–4), 174–202.

Jacquemont, S., Reymond, A., Zufferey, F., Harewood, L., Walters, R. G., Kutalik, Z., … Froguel, P. (2011). Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature*, *478*(7367), 97–102.

Jarick, I., Vogel, C. I. G., Scherag, S., Schäfer, H., Hebebrand, J., Hinney, A., & Scherag, A. (2011). Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis. *Human Molecular Genetics*, *20*(4), 840–852.

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, 515(7526), 216-221.

Kang, S. J., Chiang, C. W. K., Palmer, C. D., Tayo, B. O., Lettre, G., Butler, J. L., … Hirschhorn, J. N. (2010). Genome-wide association of anthropometric traits in African- and African-derived populations. *Human Molecular Genetics*, 19(13), 2725–2738.

Klein, S., Sharifi-Hannauer, P., & Martinez-Agosto, J. (2013). Macrocephaly as a clinical indicator of genetic subtypes in autism. *Autism Research*, 6(January), 51–56.

Kuczmarski, R.J., Ogden, C.L., Guo, S.S., et al. (2002). 2000 CDC growth charts for the United States: methods and development. *Vital Health Stat* 11. 246:1

Lai, M.-C., Baron-Cohen, S., & Buxbaum, J. D. (2015). Understanding autism in the light of sex/gender. *Molecular Autism*, 6(1), 24.

Lai, M. -C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *J Am Acad Child Adolesc Psychiatry*, 54(1), 11–24.

Lainhart, J.E., Piven, J., Wzorek, M., Landa, R., Santangelo, S.L., Coon, H., Folstein, S.E.(1997). Macrocephaly in children and adults with autism. *J Am Acad Child Adolesc Psychiatry*. 36(2):282-90.

Lainhart, J. E., Bigler, E. D., Bocian, M., Coon, H., Dinh, E., Dawson, G., Deutsch, C. K., Dunn, M., Estes, A., Tager-Flusberg, H., Folstein, S., Hepburn, S., Hyman, S., McMahon, W., Minshew, N., Munson, J., Osann, K., Ozonoff, S., Rodier, P., Rogers, S., Sigman, M., Spence, M. A., Stodgell, C. J. and Volkmar, F. (2006), Head circumference and height in autism: A study by the collaborative program of excellence in autism. *Am. J. Med. Genet*., 140A: 2257–2274.

Lango Allen, H., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., … Hirschhorn, J. N. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, *467*(7317), 832–838. 0

Lavelle, T.A., Weinstein, M.C., Newhouse, J.P., Munir, K., Kuhlthau, K.A., Prosser, L.A. Economic Burden of Childhood Autism Spectrum Disorders. *Pediatrics* 133(3):e520-9 (2014)

Lauritsen, M. B., Pedersen, C. B., & Mortensen, P. B. (2005). Effects of familial risk factors and place of birth on the risk of autism: A nationwide register-based study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*, 963–971.

Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., … Newschaffer, C. (2017). The Changing Epidemiology of Autism Spectrum Disorders. *Annual Review of Public Health*, 38(1), 81–102.

Li, X., Tan, L., Liu, X., Lei, S., Yang, T., Chen, X., … Deng Hongwen, H. (2010). A genome wide association study between copy number variation (CNV) and human height in Chinese population. *Journal of Genetics and Genomics*, 37(12), 779–785.

Lettre, G. (2009). Genetic regulation of adult stature. http://doi.org/10.1097/MOP.0b013e32832c6dce

Luo, R., Sanders, S. J., Tian, Y., Voineagu, I., Huang, N., Chu, S. H., … Geschwind, D. H. (2012). Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. *American Journal of Human Genetics*, *91*, 38–55.

Maillard, A. M., Hippolyte, L., Rodriguez-Herreros, B., Chawner, S. J. R. A., Dremmel, D., Agüera, Z., … Jacquemont, S. (2015). 16P11.2 Locus Modulates Response To Satiety Before the Onset of Obesity. *International Journal of Obesity (2005)*, (August), 1–7.

Malhotra, D., Sebat J. (2012). CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell*. 148(6):1223-41.

Marouli, E., Graff, M., Medina-Gomez, C., Lo, K. S., Wood, A. R., Kjaer, T. R., … Lettre, G. (2017). Rare and low-frequency coding variants alter human adult height. *Nature*, 542(7640), 186–190.

Martin, A. K., Robinson, G., Reutens, D., & Mowry, B. (2015). Clinical and parental age characteristics of rare copy number variant burden in patients with schizophrenia. American Journal of Medical Genetics, Part B: *Neuropsychiatric Genetics*, 168(5), 374–382.

McCarroll, S.,Hadnott, T., Perry, G., Sabeti, P., Zody, M., Barrett, J., Dallaire, S., Gabriel, B., Lee, C., Daly, M., D. Altshuler. (2006). Common deletion polymorphisms in the human genome. *Nat. Genet*., 38: 86–92

Merla, G., Brunetti-Pierri, N., Micale, L., & Fusco, C. (2010). Copy number variants at Williams-Beuren syndrome 7q11.23 region. *Human Genetics*, 128(1), 3–26.

Miles, J. H., & Hillman, R. E. (2000). Value of a clinical morphology examination in autism. *American Journal of Medical Genetics*, *91*(January 1999), 245–253.

Miles, J. (2011). Autism spectrum disorders- A genetics review *Genetics In Medicine,* 13:4.

Moreno-De-Luca, D., Mulle, J. G., Kaminsky, E. B., Sanders, S. J., Myers, S. M., Adam, M. P., … Ledbetter, D. H. (2010). Deletion 17q12 is a recurrent copy number variant that confers high risk of autism and schizophrenia. *American Journal of Human Genetics*, *87*(5), 618–630.

Mraz, K. D., Green, J., Dumont-mathieu, T., Makin, S., & Fein, D. (2007). Correlates of Head Circumference Growth, 7–9.

Must, A., Eliasziw, M., Phillips, S. M., Curtin, C., Kral, T. V. E., Segal, M., … Bandini, L. G. (2016). The Effect of Age on the Prevalence of Obesity among US Youth with Autism Spectrum Disorder. *Childhood Obesity* (Print), X(X), 1–11.

Ogden, C. L., Carroll, M. D., & Flegal, K. M. (2010). Prevalence of high body mass index in U.S. children and adolescents. *Journal of the American Medical Association*, *303*(3), 242–249.

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2014). Prevalence of Childhood and Adult Obesity in the United States, 2011-2012. *Jama*, *311*(8), 806.

Ozonoff, S., Young, G. S., Carter, a., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Stone, W. L. (2011). Recurrence Risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium Study. *Pediatrics*.

Picinelli, C., Lintas, C., Piras, I. S., Gabriele, S., Sacco, R., Brogna, C., & Persico, A. M. (2016). Recurrent 15q11.2 BP1-BP2 microdeletions and microduplications in the etiology of neurodevelopmental disorders. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *171*(8), 1088–1098.

Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., … Betancur, C. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, *466*(7304), 368–372.

Pinto D., Darvishi K., Shi X.H., Rajan D., Rigler D., Fitzgerald T., Lionel A.C., Thiruvahindrapuram B., MacDonald J.R., Mills R., et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nat. Biotechnol.,* 2011;29:512–520.

Qureshi, A. Y., Mueller, S., Snyder, A. Z., Mukherjee, P., Berman, J. I., Roberts, T. P. L., … Buckner, R. L. (2014). Opposing Brain Differences in 16p11.2 Deletion and Duplication Carriers. *Journal of Neuroscience*, *34*(34), 11199–11211.

Rimmer, J.H., Yamaki, K., Lowry, B.M., Wang, E., Vogel, L.C. (2010). Obesity and obesity-related secondary conditions in adolescents with intellectual/developmental disabilities. *J Intellect Disabil Res.,* 54(9):787-94.

Redcay, E., Courchesne, E. (2005). When Is the Brain Enlarged in Autism? A Meta-Analysis of All Brain Size Reports. *Biological Psychiatry* , Volume 58 , Issue 1 , 1 – 9

Robinson, E. B., Neale, B. M., & Hyman, S. E. (2015). Genetic research in autism spectrum disorders. *Current Opinion in Pediatrics*, *27*(6), 685–91.

Roohi, J., Montagna, C., Tegay, D. H., Palmer, L. E., DeVincent, C., Pomeroy, J. C., … Hatchwell, E. (2009). Disruption of contactin 4 in three subjects with autism spectrum disorder. *Journal of Medical Genetics*, *46*(3), 176–82.

Rosenberg, R. E., Law, J. K., Yenokyan, G., McGready, J., Kaufmann, W. E., & Law, P. a. (2009). Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Archives of Pediatrics & Adolescent Medicine*, *163*(10), 907–914.

Rosenfeld, J. A., Traylor, R. N., Schaefer, G. B., McPherson, E. W., Ballif, B. C., Klopocki, E., Aylsworth, A. S. (2012). Proximal microdeletions and microduplications of 1q21.1 contribute to variable abnormal phenotypes. *European Journal of Human Genetics : EJHG*, *20*(7), 754–61.

Sanders, S., Ercan-sencicek, a G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., & State, M. (2011). Multiple recurrent de novo copy number variations (CNVs), including duplications of the 7q11.23 Williams-Beuren syndrome region, are strongly associated with autism. *Neuron*, *70*(5), 863–885.

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Hultman, C., Larsson, H., Reichenberg, A. (2017). The Heritability of Autism Spectrum Disorder. *Research Letter, JAMA Network.* 318(12):1182-1184.

Serra-Juhé, C., Martos-Moreno, G., Bou de Pieri, F., Flores, R., González, J. R., Rodríguez-Santiago, B., … Pérez-Jurado, L. A. (2017). Novel genes involved in severe early-onset obesity revealed by rare copy number and sequence variants. *PLoS Genetics*, *13*(5), 1–19.

Schendel, D. E., Diguiseppi, C., Croen, L. a, Fallin, M. D., Reed, P. L., Schieve, L. a, … Yeargin-Allsopp, M. (2012). The Study to Explore Early Development (SEED): a multisite epidemiologic study of autism by the Centers for Autism and Developmental Disabilities Research and Epidemiology (CADDRE) network. *Journal of Autism and Developmental Disorders*, 42(10), 2121–40.

Schüle, B., Albalwi, M., Northrop, E., Francis, D. I., Rowell, M., Slater, H. R., … Francke, U. (2005). Molecular breakpoint cloning and gene expression studies of a novel translocation t(4;15)(q27;q11.2) associated with Prader-Willi syndrome. *BMC Medical Genetics*, 6(1), 18.

SFARI Gene Database. 32. https://gene.sfari.org/autdb/CNVHome.do

Shapira, K., Tian, L., Aylsworth, A., Elias, E., Hoover-Fong, J., Meeks, N., Souders, M., Tsai, C., Alexander, A., Schieve, L. (2014). Development of a Novel Protocol for Characterizing Dysmorphology to Enhance the Phenotypic Classification of Autism Spectrum Disorders, presented at International Meeting for Autism Research, Atlanta, 2014. Georgia: INSAR.

Shapira, S.K., Lin, H. T., Aylsworth A.S., Elias, E.R., Hoover-Fong, J.E., Meeks, N.J.L., Souders, M.C., Tsai, A.C.H., Zackai, A.H., Alexander, A.A., Yeargin-Allsopp, M.H., Schieve, L.A. (personal communication, SEED Dysmorphology Workgroup unpublished manuscript). A Novel Protocol for Characterizing Dysmorphology to Enhance the Phenotypic Classification of Autism Spectrum Disorder in the Study to Explore Early Development.

Sheppard, B., Benke, K.S., Croen, L., Daniels, J., Newschaffer, C., Reynolds, A., Schendel, D., Schieve, L., Ladd-Acosta, C., Fallin, D. (2015). Polygene –by– Prenatal Environment Interaction in Autism Spectrum Disorder using Copy Number Variant Burden, presenred at IGES, Baltimore, 2015. Maryland: IGES.

Shinawi, M., Liu, P., Kang, S.-H. L., Shen, J., Belmont, J. W., Scott, D. A., Lupski, J. R. (2010). Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *Journal of Medical Genetics*, *47*(5), 332–341.

Shishido, E., Aleksic, B., & Ozaki, N. (2014). Copy-number variation in the pathogenesis of autism spectrum disorder. *Psychiatry and Clinical Neurosciences*. 68(2):85-95.

Shoukier, M., Klein, N., Auber, B., Wickert, J., Schroeder, J., Zoll, B., Zirn, B. (2013). Array CGH in patients with developmental delay or intellectual disability: Are there phenotypic clues to pathogenic copy number variants? *Clinical Genetics*, *83*(1), 53–65.

Stein, J. L. (2015). Copy number variation and brain structure: lessons learned from chromosome 16p11.2. *Genome Medicine*, *7*(1), 13.

Suren, P., Stoltenberg, C., Bresnahan, M., Hirtz, D., Lie, K. K., Lipkin, W. I., … Hornig, M. (2013). Early growth patterns in children with autism. *Epidemiology* (Cambridge, Mass.), 24(5), 660–670.

Ulfarsson, M. O., Walters, G. B., Gustafsson, O., Steinberg, S., Silva, A., Doyle, O. M., … Stefansson, K. (2017). 15q11.2 CNV affects cognitive, structural and functional correlates of dyslexia and dyscalculia. *Translational Psychiatry*, *7*(4), e1109.

van Daalen, E., Swinkels, S. H. N., Dietz, C., van Engeland, H., & Buitelaar, J. K. (2007). Body Length and Head Growth in the First Year of Life in Autism. Pediatric *Neurology*, 37(5), 324–330.

van Duyvenvoorde, H. A., Lui, J. C., Kant, S. G., Oostdijk, W., Gijsbers, A. C. J., Hoffer, M. J. V, Wit, J. M. (2014). Copy number variants in patients with short stature. *European Journal of Human Genetics : EJHG*, *22*(5), 602–9.

Vulto-van Silfhout, A. T., Hehir-Kwa, J. Y., van Bon, B. W. M., Schuurs-Hoeijmakers, J. H. M., Meader, S., Hellebrekers, C. J. M., … de Vries, B. B. a. (2013). Clinical Significance of De Novo and Inherited Copy-Number Variation. *Human Mutation*, 34, 1679–1687.

Wells, J. (2007). Sexual dimorphism of body composition. Best Practice & Research *Clinical Endocrinology & Metabolism*. 21(3), 415-430.

Werling, D. M. (2016). The role of sex-differential biology in risk for autism spectrum disorder. *Biology of Sex Differences*, 7, 58.

Werling, D. M., & Geschwind, D. H. (2013). Sex differences in autism spectrum disorders. *Current Opinion in Neurology*, 26(2), 146–153.

WHO Multicentre Growth Reference Study Group (2006a). Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. *Acta Paediatrica*, Suppl 450:57–66.

Wilson, N. K. a, Lee, Y., Long, R., Hermetz, K., Rudd, M. K., Miller, R., … Addington, A. M. (2011). A Novel Microduplication in the Neurodevelopmental Gene SRGAP3 That Segregates with Psychotic Illness in the Family of a COS Proband. *Case Reports in Genetics*, *2011*, 585893.

Winchester L., Yau C., Ragoussis J. Comparing CNV detection methods for SNP arrays. Brief. Funct. Genomic Proteomic. 2009;8:353–366.

Xiong, N., Ji, C., Li, Y., He, Z., Bo, H., & Zhao, Y. (2009). The physical status of children with autism in China. *Research in Developmental Disabilities*, 30(1), 70–76.

Zahnleiter, D., Uebe, S., Ekici, A. B., Hoyer, J., Wiesener, A., Wieczorek, D., … Thiel, C. T. (2013). Rare Copy Number Variants Are a Common Cause of Short Stature. *PLoS Genetics*, *9*(3), 1–11.

Zhang, X., Du, R., Li, S., Zhang, F., Jin, L., & Wang, H. (2014). Evaluation of copy number variation detection for a SNP array platform. *BMC Bioinformatics*, *15*(1), 50.

Zuckerman, K. E., & Fombonne, E. (2015). Obesity and Autism. *Pediatrics*, 136(6), 1051–1061.

Tables

Table 1: Characteristics of Children With Growth Abnormalities In the SEED 1 Study

N=840		Normal Growth (>10th, <90th)	Growth Abnormality - high (≥90th %)		Growth Abnormality - low (≤10th %)	
HEAD CIRCUMFERENCE		**Percentile, N=623**	**Macrocephaly, N=52**	**p-value***	**Microcephaly, N=165**	**p-value***
Diagnosis	ASD	37.1%	51.9%	**0.032**	41.2%	0.38
	P. Case	1.4%	3.8%		2.4%	
	POP	61.5%	44.2%		56.4%	
Sex	Male	66.6%	55.8%	0.113	63.6%	0.473
	Female	33.4%	44.2%		36.4%	
Race	NHW	63%	62.1%	0.148	48.8%	0.058
	NHB	15.7%	27.6%		23.2%	
	Hisp	21.3%	10.3%		28.0%	
STATURE		**N=656**	**Tall, N=116**		**Short, N=68**	
Diagnosis	ASD	38.0%	40.5%	0.686	44.1%	0.461
	P. Case	1.8%	0.9%		2.9%	
	POP	60.2%	58.6%		52.9%	
Sex	Male	67.2%	58.6%	0.071	58.8%	0.162
	Female	32.8%	41.4%		41.2%	
Race	NHW	60%	59.4%	0.782	65.7%	0.776
	NHB	18.3%	16.6%		17.1%	
	Hisp	21.7%	25%		17.1%	
WEIGHT		**N=671**	**Overwt,N=106**		**Underwt,N=63**	
Diagnosis	ASD	38.2%	44.3%	0.198	36.5%	0.789
	P. Case	1.9%	0		3.2%	
	POP	59.9%	55.7%		60.3%	
Sex	Male	65.3%	66%	0.880	65.1%	0.97
	Female	34.7%	34%		34.9%	
Race	NHW	62.2%	52.9%	0.355	55.2%	0.75
	NHB	17.1%	20.6%		20.7%	
	Hisp	20.7%	26.5%		24.1%	
BODY MASS INDEX		**N=647**	**High, N=118**		**Low, N=75**	
Diagnosis	ASD	36.9%	51.7%	**0.005**	34.7%	0.876
	P. Case	2.0%	0		2.7%	
	POP	61.1%	48.3%		62.7%	
Sex	Male	64.8%	67.8%	0.524	66.7%	0.743
	Female	35.2%	32.2%		33.3%	
Race	NHW	61.7%	48.7%	**0.040**	75.9%	0.182
	NHB	17.9%	17.9%		17.2%	
	Hisp	20.5%	33.3%		6.9%	

*Compared to normal growth children

ASD=Autism Spectrum Disorder P. Case=Possible Case Bolded: p<0.05

NHW=Non-Hispanic White NHB=Non-Hispanic Black Hisp=Hispanic

Table 2: Association Analyses of Overall Genome-wide CNV Burden in Growth Abnormalities

Macrocephaly and Microcephaly

CNV Burden	Normal HC	Macro-cephaly	Mac/ N ratio	p-value	Micro-cephaly	Mic/ N ratio	p-value
CNV Rate	21.79	23.96	1.099	0.459	20.63	0.946	0.224
Average Length (kb)	2,581	4,183	1.620	0.227	2,491	0.965	0.741

Tall and Short Stature

CNV Burden	Normal Height	Tall	Tall/ N ratio	p-value	Short	Short/N ratio	p-value
CNV Rate	22.14	20.14	**0.909**	0.029	20.07	0.906	0.089
Average Length (kb)	2,718	2,264	**0.833**	0.033	2,803	1.031	0.879

Overweight and Underweight

CNV Burden	Normal Weight	Over-weight	Over-weight/ N ratio	p-value	Under-weight	Under-weight/ N ratio	p-value
CNV Rate	21.56	23.30	1.080	0.366	20.38	0.945	0.413
Average Length (kb)	2,624	2,811	1.071	0.621	2,819	1.074	0.740

High BMI and Low BMI

CNV Burden	Normal BMI	High BMI	High BMI/ N ratio	p-value	Low BMI	Low BMI/ N ratio	p-value
CNV Rate	21.65	22.56	1.042	0.589	20.73	0.957	0.446
Average Length (kb)	2,704	2,629	0.972	0.831	2,357	0.871	0.219

CNV rate: CNV count per individual
Average Length: Average CNV length per individual in kilobase pairs (kb)
N: Normal growth
Bolded: t-test $p < 0.05$

Table 3: Summary of CNV Burden by Length (kb) for Growth Abnormalities in SEED

CNV		Macro-cephaly		Micro-cephaly		Tall		Short		Over-weight		Under-weight		High BMI		Low BMI	
		Mac /N	P	Mic/ N	P	T/ N	P	S/ N	P	O/ N	P	U/ N	P	H/ N	P	L/ N	P
Over-all	O	1.62	0.22	0.96	0.74	**0.83**	0.03	1.03	0.87	1.07	0.62	1.07	0.74	0.97	0.83	0.87	0.22
	OG	1.81	0.20	0.92	0.47	0.83	0.06	0.95	0.82	1.08	0.59	1.03	0.87	0.99	0.84	0.86	0.25
	OS	0.67	0.07	0.97	0.84	**0.79**	0.04	0.91	0.63	1.08	0.72	1.14	0.48	0.97	0.83	1.14	0.48
Large (>400 kb)	L	2.68	0.22	1.03	0.90	0.69	0.09	1.24	0.67	0.99	0.99	1.27	0.67	0.79	0.40	0.64	0.11
	LG	2.89	0.21	0.93	0.80	0.69	0.09	1.01	0.97	1.04	0.87	1.17	0.74	0.84	0.53	0.72	0.24
	LS	**0.20**	0.03	2.38	0.27	0.81	0.74	2.55	0.41	0.67	0.51	1.92	0.62	0.38	0.11	1.61	0.66
Dele-tion	De	1.03	0.84	0.97	0.76	0.86	0.05	0.89	0.23	1.12	0.45	0.96	0.75	1.03	0.78	0.94	0.54
	DeG	0.97	0.90	0.91	0.29	0.87	0.11	**0.78**	.008	1.10	0.46	0.89	0.31	1.00	0.97	0.93	0.51
	DeS	0.72	0.23	1.07	0.74	**0.63**	.005	0.77	0.15	1.01	0.97	1.09	0.65	1.06	0.84	1.19	0.51
Del. Large	LDe	0.74	0.48	1.22	0.61	0.87	0.76	0.75	0.57	0.93	0.88	0.52	0.06	0.73	0.42	1.17	0.81
	LDeG	0.65	0.35	0.95	0.89	0.71	0.37	**0.23**	.001	0.84	0.67	0.68	0.33	0.75	0.46	1.24	0.70
	LDeS	0.94	0.95	0.79	0.25	0.90	0.93	0.49	0.52	0.76	0.77	0	0.08	1.36	0.76	9.23	0.37
Dup.	Du	1.98	0.21	0.95	0.79	0.81	0.09	1.11	0.72	1.04	0.81	1.13	0.68	0.93	0.68	0.83	0.22
	DuG	2.35	0.18	0.92	0.65	0.80	0.14	1.05	0.86	1.07	0.73	1.11	0.74	0.95	0.81	0.83	0.28
	DuS	**0.58**	0.04	0.79	0.25	1.10	0.52	1.18	0.64	1.21	0.26	1.24	0.55	1.15	0.45	1.04	0.84
Dup. Large	LDu	2.98	0.21	1.01	0.97	0.67	0.10	1.31	0.63	1.00	0.98	1.38	0.60	0.80	0.47	0.57	0.05
	LDuG	3.32	0.19	0.93	0.82	0.69	0.14	1.15	0.78	1.07	0.73	1.25	0.68	0.85	0.61	0.64	0.13
	LDuS	**0.01**	.008	1.93	0.45	0.77	0.72	3.51	0.36	1.21	0.26	2.80	0.50	**0.21**	0.04	0.29	0.08

O=Overall CNVs OG= Overall Genic CNVs OS=Overall SFARI CNVs
L=Large CNVs LG=Large Genic CNVs LS=Large SFARI CNVs
De=Deletion CNVs DeG=Deletion Genic CNVs DeS=Deletion SFARI CNVs
Du=Duplication CNVs DuG=Duplication Genic CNVs DuS=Duplication SFARI
P: p-value for t-tests comparing growth abnormality to normal growth
Green-colored cell and bolded effect size: $p < 0.05$.

Table 4: CNV Burden by Length (kb) for Growth Abnormalities Stratified by Sex: Male

CNV		Macro-cephaly		Micro-cephaly		Tall		Short		Over-weight		Under-weight		High BMI		Low BMI	
		Mac/N	P	Mic/N	P	T/N	P	S/N	P	O/N	P	U/N	P	H/N	P	L/N	P
Overall	O	0.99	0.97	0.88	0.20	0.84	0.09	0.81	0.13	1.13	0.52	0.94	0.62	1.02	0.87	0.93	0.60
	OG	1.02	0.93	0.83	0.08	0.83	0.12	**0.70**	**0.02**	1.16	0.44	0.92	0.57	1.03	0.84	0.96	0.77
	OS	0.79	0.45	0.86	0.42	**0.74**	**0.04**	0.80	0.26	1.21	0.54	1.10	0.63	1.27	0.40	1.19	0.46
Large (>400 kb)	L	0.84	0.65	0.79	0.30	0.77	0.29	0.60	0.08	1.14	0.70	0.71	0.23	0.96	0.91	0.67	0.19
	LG	0.86	0.68	0.74	0.19	0.72	0.18	**0.43**	**0.01**	1.24	0.54	0.72	0.25	1.02	0.95	0.82	0.52
	LS	0.30	0.15	0.72	0.68	1.17	0.83	0	--	1.09	0.90	**0.15**	**0.04**	0.59	0.45	0.39	0.21
Deletion	De	1.10	0.74	0.91	0.33	0.92	0.49	0.83	0.11	1.28	0.25	1.01	0.91	1.12	0.55	0.91	0.33
	DeG	1.01	0.97	0.85	0.09	0.94	0.59	**0.72***	**0.0002**	1.28	0.19	0.96	0.78	1.09	0.62	0.91	0.37
	DeS	0.89	0.73	0.93	0.79	**0.57**	**0.01**	0.77	0.27	1.23	0.66	1.15	0.72	1.34	0.49	1.30	0.44
Del. Large	LDe	0.57	0.33	0.80	0.57	1.21	0.73	0.67	0.57	1.38	0.55	0.72	0.46	0.94	0.91	0.59	0.24
	LDeG	0.39	0.06	0.71	0.35	0.93	0.89	**0.13***	**0.0002**	1.36	0.52	0.94	0.91	1.06	0.89	0.89	0.82
	LDeS	1.80	0.67	0	--	4.51	0.44	0	--	3.12	0.41	0	--	2.61	0.46	0	--
Dup.	Du	0.92	0.75	0.87	0.31	0.79	0.10	0.80	0.25	1.04	0.86	0.89	0.50	0.97	0.89	0.94	0.75
	DuG	1.02	0.92	0.81	0.20	0.77	0.12	0.70	0.11	1.10	0.71	0.90	0.58	1.00	0.98	0.98	0.94
	DuS	0.63	0.21	0.75	0.17	1.07	0.73	0.87	0.68	1.18	0.44	1.00	0.99	1.17	0.52	1.00	0.99
Dup. Large	LDu	0.88	0.74	0.79	0.36	0.71	0.21	0.59	0.10	1.10	0.79	0.70	0.27	0.96	0.92	0.68	0.28
	LDuG	0.94	0.89	0.75	0.27	0.68	0.17	**0.48**	**0.04**	1.22	0.62	0.68	0.24	1.01	0.97	0.81	0.55
	LDuS	0	--	0.86	0.86	0.76	0.75	0	--	0.80	0.80	0.17	0.07	0.32	0.20	0.45	0.31

O=Overall CNVs OG= Overall Genic CNVs OS=Overall SFARI CNVs
L=Large CNVs LG=Large Genic CNVs LS=Large SFARI CNVs
De=Deletion DeG=Deletion Genic CNVs DeS=Deletion SFARI CNVs
Du=Duplication DuG=Duplication Genic CNVs DuS=Duplication SFARI
P: p-value for t-tests comparing growth abnormality to no growth abnormality
Blue-colored cell and bolded effect size: $p < 0.05$
*Statistically significant result after correcting for multiple testing at $p=0.05/96=0.0005$

Table 5: CNV Burden by Length (kb) for Growth Abnormalities Stratified by Sex: Female

CNV		Macro-cephaly		Micro-cephaly		Tall		Short		Over-weight		Under-weight		High BMI		Low BMI	
		Mac /N	P	Mic/ N	P	T/ N	P	S/ N	P	O/ N	P	U/ N	P	H/ N	P	L/ N	P
Over-all	O	2.55	0.20	1.13	0.61	0.81	0.20	1.33	0.48	0.96	0.82	1.31	0.59	0.87	0.42	0.76	0.18
	OG	2.98	0.20	1.11	0.65	0.82	0.30	1.28	0.52	0.93	0.74	1.22	0.69	0.86	0.46	0.71	0.16
	OS	**0.52**	0.03	1.15	0.63	0.86	0.41	1.07	0.85	0.84	0.36	1.23	0.62	0.77	0.22	1.05	0.87
Large (>400 kb)	L	6.32	0.19	1.68	0.47	0.59	0.22	2.13	0.40	0.73	0.47	2.25	0.47	0.53	0.14	0.59	0.35
	LG	6.81	0.18	1.43	0.57	0.66	0.34	1.79	0.46	0.71	0.43	1.91	0.50	0.56	0.18	0.56	0.32
	LS	0.04	0.05	7.47	0.16	0.31	0.31	6.19	0.26	0	--	4.67	0.44	0	--	3.70	0.47
Dele-tion	De	0.96	0.80	1.09	0.56	**0.76**	0.01	0.97	0.89	0.82	0.11	0.86	0.19	0.86	0.18	0.99	0.96
	DeG	0.95	0.79	1.03	0.83	**0.78**	0.03	0.88	0.47	**0.78**	0.04	0.78	0.08	0.82	0.08	0.95	0.87
	DeS	0.52	0.13	1.33	0.46	0.71	0.15	0.78	0.37	0.61	0.05	0.99	0.97	**0.57**	0.02	1.01	0.97
Del. Large	LDe	1.10	0.88	2.28	0.31	0.43	0.19	0.84	0.81	**0.21**	0.01	**0.19**	0.01	**0.23**	.003	2.44	0.52
	LDeG	1.10	0.90	1.51	0.58	0.39	0.13	0.36	0.15	0	--	0.28	0.06	**0.10**	.001	1.94	0.58
	LDeS	0	0.31	10.0	0.33	0	0.19	0.51	0.59	0	--	0	--	0	--	19.4	0.35
Dup.	Du	3.62	0.18	1.15	0.58	0.84	0.46	1.54	0.46	1.04	0.86	1.58	0.53	0.87	0.58	0.64	0.07
	DuG	4.29	0.18	1.16	0.68	0.83	0.53	1.50	0.46	1.02	0.94	1.45	0.59	0.88	0.66	0.59	0.08
	DuS	0.51	0.12	0.86	0.71	1.16	0.53	1.64	0.45	1.28	0.41	1.71	0.50	1.14	0.68	1.12	0.72
Dup. Large	LDu	7.15	0.19	1.59	0.58	0.61	0.31	2.33	0.39	0.82	0.67	2.59	0.43	0.56	0.22	0.39	0.08
	LDuG	8.00	0.18	1.42	0.64	0.70	0.48	2.05	0.40	0.84	0.72	2.21	0.45	0.31	0.97	0.38	0.09
	LDuS	0.07	0.10	6.11	0.33	1.23	0.87	23.0	0.22	0	--	12.0	0	--	0.20	0	--

O=Overall CNVs OG= Overall Genic CNVs OS=Overall SFARI CNVs
L=Large CNVs LG=Large Genic CNVs LS=Large SFARI CNVs
De=Deletion CNVs DeG=Deletion Genic CNVs DeS=Deletion SFARI CNVs
Du=Duplication CNVs DuG=Duplication Genic CNVs DuS=Duplication SFARI
P: p-value for t-tests comparing growth abnormality to no growth abnormality
Red-colored cell and bolded effect size: $p < 0.05$
No statistically significant result after correcting for multiple testing at 0.05/96=0.0005

Table 6: Summary of Associations (nominal P<0.05) between CNVs in ASD-associated regions and Growth Abnormalities

Growth Abnormality	ASD Candidate CNV region	Fisher's exact p value	Frequency (%) In Abnormal	Frequency (%) In Normal	Ratio of Abnor-mal / Normal
Macrocephaly	1q21.1 All CNVs	**0.028**	4 (7.7)	12 (1.9)	4.05
n(Mac)=52 n(Normal)=623	1q21.1 Large CNVs	**0.012**	3 (5.7)	4 (0.6)	9.50
	1q21.1 Dup	**0.022**	4 (7.7)	11 (1.7)	4.50
	1q21.1 Large Dup	**0.012**	3 (5.7)	4 (0.6)	9.50
Microcephaly	1q21.1 Del	**0.015**	2 (1.2)	1 (0.1)	12.00
n(Mic)=165 n(Normal)=623	15q11.2 All CNVs	**0.023**	12 (7.2)	86 (13.8)	0.52
	15q11.2.13.1 All CNVs	**0.002**	9 (5.4)	85 (13.6)	0.39
	15q11.2.13.1 Del	**0.043**	6 (3.6)	51 (8.1)	0.44
Tall Stature	-	None < 0.05			
Short Stature n(SS)=68 n(Normal)=656	15q11.2 Large Dup	**0.046**	2 (2.9)	2 (0.3)	9.67
Overweight	-	None < 0.05			
Underweight	-	None < 0.05			
High BMI	-	None < 0.05			
Low BMI	-	None < 0.05			

Bolded: p< 0.05

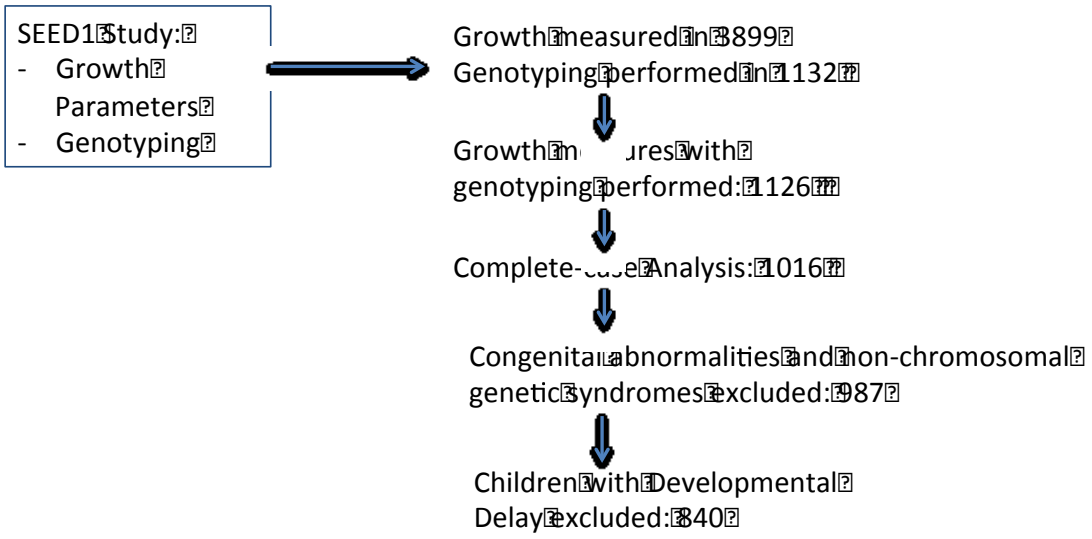
Table 7: Summary of Associations between CNVs in ASD-associated Regions and Growth Abnormalities Stratified by Sex

Growth Abnormality	CNV Region	CNV Sub-type	Males				Females			
			Fisher's exact p value	Freq. (%) In Abnormal (a)	Freq. (%) In Normal (b)	Ratio of a/b	Fisher's exact p value	Freq. (%) In Abnormal (c)	Freq. (%) In Normal (d)	Ratio of c/d
Macrocephaly	1q21.1	All	0.560	1(3)	11(1.7)	2.0	**0.003**	3(12.5)	1(0.5)	25
		Large	0.287	1(3)	4(0.6)	3.7	--	2(8.3)	0(0)	--
		Dup	0.528	1(3.4)	10(2.4)	1.4	**0.003**	3(12.5)	1(0.5)	25
		L.Dup	0.287	1(3.4)	4(0.9)	3.7	--	2(8.3)	0(0)	--
Microcephaly	1q21.1	Del	**0.363**	1(0.9)	1(0.2)	4.5	--	1(0.5)	0(0)	--
	15q11.2	All	**0.040**	6(5.7)	55(13.2)	0.4	0.518	6(10)	29(13.9)	0.7
	15q11.2.13.1	All	**0.015**	5(4.7)	55(13.2)	0.3	0.180	4(6.6)	28(13.4)	0.5
		Del	0.063	2(1.9)	30(7.2)	0.2	0.307	1(1.6)	9(4.3)	0.3
Short Stature	15q11.2	All	0.106	1(2.5)	51(11.5)	0.2	**0.034**	8(28.5)	25(11.6)	2.4
		Large	--	0(0)	2(0.4)	0	**0.035**	2(7.1)	1(0.4)	2.5
		Dup	--	0(0)	22(4.9)	0	**0.048**	4(14.2)	9(4.1)	3.4
		L.Dup	--	0(0)	2(0.4)	0	--	2(7.1)	0(0)	--

L. Dup=Large Duplication
Bolded: $p < 0.05$

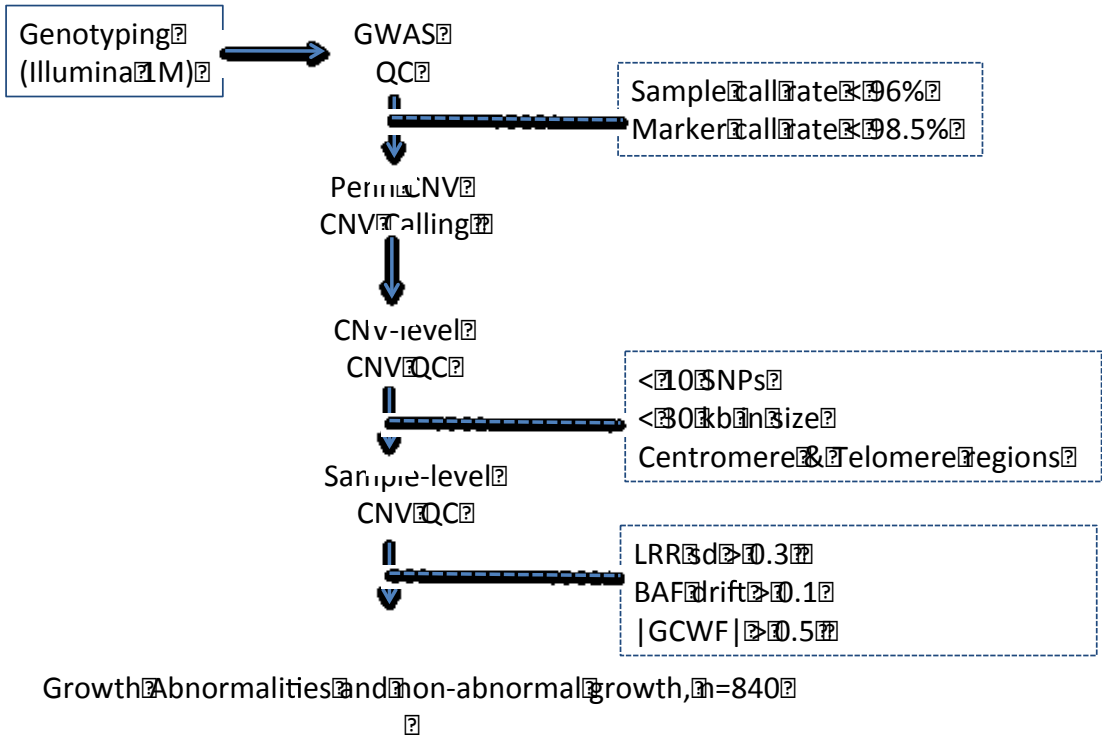
Figures

Figure 1: Flow Chart for Sample Inclusion in Analysis



≤10th percentile for growth	>10th, <90th perc.	≥90th percentile for growth
Microcephaly: 165 (19.6%)	623 (74.2%)	Macrocephaly: 52 (6.2%)
Short Stature: 68 (8.1%)	656 (78.1%)	Tall Stature: 116 (13.8%)
Underweight: 63 (7.5%)	671 (79.9%)	Overweight: 106 (12.6%)
Low BMI: 75 (9.0%)	647 (77.0%)	High BMI: 118 (14.0%)

Figure 2: CNV SEED Quality Control Pipeline (Sheppard et al., 2015)



CNV= Copy Number Variants
SNP= Single Nucleotide Polymorphism
QC= Quality Control
LRR= Log R Ratio
BAF= B Allele Frequency
GCWF= Guanine-Cytosine base pair wave factor

Figure 3. CNV Burden Results Comparison With and Without DD

Without DD

CNV	Growth Abnormality	Macro-cephaly	Micro-cephaly	Tall Stature	Short Stature	Over-weight	Under-weight
		M/NM	I/NI	T/NT	S/NS	O/NO	U/NU
Overall	Overall	1.62	0.96	**0.83**	1.03	1.07	1.07
	Overall Genic	1.81	0.92	0.83	0.95	1.08	1.03
	Overall SFARI	0.67	0.97	**0.79**	0.91	1.08	1.14
Large (>400kb)	Large	2.68	1.03	0.69	1.24	0.99	1.27
	Large Genic	2.89	0.93	0.69	1.01	1.04	1.17
	Large SFARI	**0.20**	2.38	0.81	2.55	0.67	1.92
Deletion	All Del.	1.03	0.97	0.86	0.89	1.12	0.96
	Del. Genic	0.97	0.91	0.87	**0.78**	1.10	0.89
	Del. SFARI	0.72	1.07	**0.63**	0.77	1.01	1.09
Duplicati on	All Dup.	1.98	0.95	0.81	1.11	1.04	1.13
	Dup. Genic	2.35	0.92	0.80	1.05	1.07	1.11
	Dup. SFARI	**0.58**	0.79	1.10	1.18	1.21	1.24

[Table 3]

With DD

CNV	Growth Abnormality	Macro-cephaly	Micro-cephaly	Tall Stature	Short Stature	Over-weight	Under-weight
		M/NM	I/NI	T/NT	S/NS	O/NO	U/NU
Overall	Overall	1.43	**1.35**	**0.82**	**1.81**	1.08	1.48
	Overall Genic	1.57	1.30	0.83	**1.75**	1.09	1.47
	Overall SFARI	**0.64**	1.12	**0.78**	1.19	1.04	1.29
Large (>400kb)	Large	2.19	**2.26**	**0.66**	**3.56**	1.08	2.41
	Large Genic	2.34	**1.92**	0.69	**2.93**	1.10	2.20
	Large SFARI	**0.14**	**4.63**	0.91	**7.58**	0.96	3.66
Deletion	All Del.	0.93	1.02	**0.83**	1.28	1.06	1.41
	Del. Genic	0.90	0.98	**0.85**	1.27	1.04	1.41
	Del. SFARI	0.64	1.04	**0.64**	0.79	0.93	1.18
Duplica-tion	All Dup.	1.77	**1.57**	0.81	**2.13**	1.09	1.53
	Dup. Genic	2.02	1.51	0.81	2.03	1.04	1.51
	Dup. SFARI	**0.63**	1.26	1.06	1.95	0.93	1.48

DD= Developmental Delay
Values in cells: Effect Size (ratio of mean CNV in growth abnormality to mean CNV in normal growth)
Green cells: significant p-values for t-test comparing means for CNV length in analysis excluding children with developmental delay
Orange cells: significant p-values for t-test comparing means for CNV length in analysis including children with developmental delay

Supplement

S. Table 1: CNV Burden for Counts and Lengths Comparing Children With Macrocephaly and Children with Normal Head Circumference

MACROCEPHALY N= 675		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (Macrocephaly): 52 N (Normal HC): 623	**CNV Characteristics**	**Mac/N Ratio**	**p-value**	**Mac/N Ratio**	**p-value**
All CNVs	Overall	1.099	0.458	1.620	0.227
	Overall Genic	1.129	0.376	1.812	0.209
	Overall SFARI	0.922	0.702	0.675	0.070
All Large CNVs	Overall large	1.635	0.269	2.686	0.227
	Overall large Genic	1.554	0.327	2.890	0.214
	Overall large SFARI	1.497	0.661	0.209	**0.032**
Deletions	Overall Deletion	1.045	0.763	1.036	0.843
	Deletion Genic	1.025	0.840	0.979	0.906
	Deletion SFARI	1.037	0.891	0.729	0.238
Deletions Large	Deletion Large	1.012	0.979	0.739	0.483
	Deletion Large Genic	0.773	0.568	0.649	0.350
	Deletion Large SFARI	4.792	0.434	0.940	0.956
Duplications	Dup. Overall	1.165	0.294	1.985	0.212
	Dup. Genic	1.257	0.246	2.357	0.186
	Dup. SFARI	0.755	0.179	0.586	**0.049**
Duplications Large	Dup. Large Overall	1.759	0.240	2.981	0.216
	Dup. Large Genic	1.711	0.274	3.321	0.197
	Dup. Large SFARI	0.630	0.587	0.017	**0.008**

Bolded: significant at $p<0.05$. No statistically significant result after multiple testing corrections.

Mac= Macrocephaly N=Normal Head Circumference HC=Head Circumference

S. Table II: CNV Burden for Lengths in Macrocephaly Stratified by Sex

MACROCEPHALY Stratified by Sex		**Average CNV Length (kb) All**		**Average CNV Length (kb) Male**		**Average CNV Length (kb) Female**	
Copy Number Variants (CNVs)		Mac/N Ratio	p-value	Mac/N Ratio	p-value	Mac/N Ratio	p-value
All CNVs	Overall	1.620	0.227	0.991	0.972	2.555	0.200
	Overall Genic	1.812	0.209	1.022	0.934	2.987	0.200
	Overall SFARI	0.675	0.070	0.798	0.450	0.521	**0.034**
All Large CNVs	Overall large	2.686	0.227	0.842	0.654	6.327	0.191
	Overall large Genic	2.890	0.214	0.861	0.688	6.808	0.186
	Overall large SFARI	0.209	**0.032**	0.302	0.149	0.045	0.058
Deletions	Overall Deletion	1.036	0.843	1.103	0.739	0.961	0.802
	Deletion Genic	0.979	0.906	1.010	0.969	0.955	0.799
	Deletion SFARI	0.729	0.238	0.892	0.737	0.523	0.135
Deletions Large	Deletion Large	0.739	0.483	0.569	0.335	1.101	0.888
	Deletion Large Genic	0.649	0.350	0.396	0.068	1.103	0.904
	Deletion Large SFARI	0.940	0.956	1.805	0.673	0	--
Duplica-tions	Duplic. Overall	1.985	0.212	0.923	0.750	3.622	0.181
	Duplic. Genic	2.357	0.186	1.028	0.919	4.299	0.185
	Duplic. SFARI	0.586	**0.049**	0.638	0.215	0.517	0.119
Duplica-tions Large	Duplic. Large Overall	2.981	0.216	0.882	0.746	7.151	0.191
	Duplic. Large Genic	3.321	0.197	0.948	0.890	8.006	0.184
	Duplic. Large SFARI	0.017	**0.008**	0	**--**	0.069	0.099
Total and group totals Bolded: significant p<0.05 Mac=Macrocephaly N=Normal Head Circ.		N(All)=675 N(Mac)= 52 N(N)= 623		N(Male)=444 $N(Mac_{Male})$= 29 $N(N_{Male})$= 415		N(Female)=231 $N(Mac_{Female})$= 23 $N(N_{Female})$= 208	

S. Table III: CNV Burden for Counts and Lengths Comparing Children With Microcephaly and Children with Normal Head Circumference

MICROCEPHALY N= 788		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (Microcephaly): 165 N (Normal HC): 623	**CNV Characteristics**	**Mic/N Ratio**	**p-value**	**Mic/N Ratio**	**p-value**
All CNVs	Overall	0.946	0.224	0.965	0.741
	Overall Genic	0.929	0.117	0.922	0.474
	Overall SFARI	0.898	0.249	0.970	0.849
All Large CNVs	Overall large	0.817	0.126	1.038	0.901
	Overall large Genic	0.775	0.067	0.934	0.803
	Overall large SFARI	1.416	0.488	2.387	0.271
Deletions	Overall Deletion	0.957	0.531	0.974	0.760
	Deletion Genic	0.933	0.262	0.914	0.299
	Deletion SFARI	0.966	0.799	1.074	0.741
Deletions Large	Deletion Large	0.957	0.867	1.224	0.614
	Deletion Large Genic	0.852	0.562	0.954	0.897
	Deletion Large SFARI	1.510	0.658	4.134	0.415
Duplications	Dup. Overall	0.933	0.172	0.959	0.798
	Dup. Genic	0.925	0.229	0.927	0.657
	Dup. SFARI	0.800	0.084	0.798	0.253
Duplications Large	Dup. Large Overall	0.789	0.116	1.009	0.977
	Dup. Large Genic	0.760	0.082	0.931	0.819
	Dup. Large SFARI	1.391	0.576	1.929	0.454

Bolded: significant at $p<0.05$. No statistically significant result after multiple testing corrections.

Mic= Microcephaly N=Normal Head Circumference HC=Head Circumference

S. Table IV: CNV Burden for Lengths in Microcephaly Stratified by Sex

MICROCEPHALY Stratified by Sex		**Average CNV Length (kb) All**		**Average CNV Length (kb) Male**		**Average CNV Length (kb) Female**	
Copy Number Variants (CNVs)		Mic/N Ratio	p-value	Mic/N Ratio	p-value	Mic/N Ratio	p-value
All CNVs	Overall	0.965	0.741	0.885	0.207	1.130	0.610
	Overall Genic	0.922	0.474	0.829	0.088	1.112	0.657
	Overall SFARI	0.970	0.849	0.868	0.425	1.151	0.630
All Large CNVs	Overall large	1.038	0.901	0.794	0.302	1.688	0.477
	Overall large Genic	0.934	0.803	0.744	0.195	1.438	0.572
	Overall large SFARI	2.387	0.271	0.723	0.679	7.469	0.167
Deletions	Overall Deletion	0.974	0.760	0.909	0.335	1.098	0.561
	Deletion Genic	0.914	0.299	0.852	0.091	1.035	0.838
	Deletion SFARI	1.074	0.741	0.935	0.791	1.333	0.464
Deletions Large	Deletion Large	1.224	0.614	0.801	0.573	2.279	0.318
	Deletion Large Genic	0.954	0.897	0.713	0.354	1.516	0.582
	Deletion Large SFARI	4.134	0.415	0	--	10.054	0.335
Duplica-tions	Duplic. Overall	0.959	0.798	0.871	0.309	1.152	0.707
	Duplic. Genic	0.927	0.657	0.816	0.203	1.163	0.683
	Duplic. SFARI	0.798	0.253	0.754	0.172	0.867	0.716
Duplica-tions Large	Duplic. Large Overall	1.009	0.977	0.793	0.359	1.595	0.587
	Duplic. Large Genic	0.931	0.819	0.750	0.277	1.421	0.641
	Duplic. Large SFARI	1.929	0.454	0.868	0.868	6.115	0.331
Total and group totals Bolded: significant $p<0.05$ Mic=Macrocephaly N=Normal Head Circ.		N(All)=788 N(Mic)= 165 N(N)= 623		N(Male)=520 $N(Mic_{Male})$= 105 $N(N_{Male})$= 415		N(Female)=268 $N(Mic_{Female})$= 60 $N(N_{Female})$= 208	

S. Table V: CNV Burden for Counts and Lengths Comparing Children With Tall Stature and Children with Normal Height

TALL STATURE N= 772		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (Tall Stature): 116 N (Normal Height): 656	**CNV Characteristics**	**TS/N Ratio**	**p-value**	**TS/N Ratio**	**p-value**
All CNVs	Overall	0.909	**0.029**	0.832	**0.033**
	Overall Genic	0.937	0.195	0.831	0.069
	Overall SFARI	0.798	**0.012**	0.792	**0.044**
All Large CNVs	Overall large	0.751	0.070	0.699	0.095
	Overall large Genic	0.778	0.121	0.699	0.099
	Overall large SFARI	0.837	0.728	0.819	0.744
Deletions	Overall Deletion	0.878	**0.020**	0.861	0.054
	Deletion Genic	0.932	0.198	0.875	0.119
	Deletion SFARI	0.685	**0.004**	0.631	**0.005**
Deletions Large	Deletion Large	0.545	0.056	0.877	0.759
	Deletion Large Genic	0.646	0.202	0.710	0.374
	Deletion Large SFARI	0.807	0.833	0.907	0.932
Duplications	Dup. Overall	0.948	0.389	0.815	0.099
	Dup. Genic	0.942	0.463	0.806	0.148
	Dup. SFARI	0.977	0.862	1.109	0.529
Duplications Large	Dup. Large Overall	0.795	0.195	0.674	0.100
	Dup. Large Genic	0.805	0.241	0.697	0.145
	Dup. Large SFARI	0.848	0.781	0.778	0.728

Bolded: significant at $p<0.05$. No statistically significant result after multiple testing corrections.

TS= Tall Stature N=Normal Height

S. Table VI: CNV Burden for Lengths in Tall Stature Stratified by Sex

TALL STATURE Stratified by Sex		**Average CNV Length (kb) All**		**Average CNV Length (kb) Male**		**Average CNV Length (kb) Female**	
Copy Number Variants (CNVs)		TS/N Ratio	p-value	TS/N Ratio	p-value	TS/N Ratio	p-value
All CNVs	Overall	0.832	**0.033**	0.845	0.091	0.813	0.203
	Overall Genic	0.831	0.069	0.834	0.123	0.819	0.308
	Overall SFARI	0.792	**0.044**	0.745	**0.047**	0.861	0.410
All Large CNVs	Overall large	0.699	0.095	0.775	0.294	0.589	0.228
	Overall large Genic	0.699	0.099	0.721	0.182	0.661	0.343
	Overall large SFARI	0.819	0.744	1.176	0.831	0.313	0.317
Deletions	Overall Deletion	0.861	0.054	0.928	0.492	0.766	**0.010**
	Deletion Genic	0.875	0.119	0.939	0.598	0.786	**0.037**
	Deletion SFARI	0.631	**0.005**	0.573	**0.012**	0.711	0.158
Deletions Large	Deletion Large	0.877	0.759	1.211	0.738	0.431	0.191
	Deletion Large Genic	0.710	0.374	0.936	0.896	0.399	0.132
	Deletion Large SFARI	0.907	0.932	4.514	0.442	0	--
Duplications	Duplic. Overall	0.815	0.099	0.794	0.100	0.840	0.465
	Duplic. Genic	0.806	0.148	0.773	0.124	0.837	0.532
	Duplic. SFARI	1.109	0.529	1.078	0.738	1.162	0.538
Duplications Large	Duplic. Large Overall	0.674	0.100	0.716	0.211	0.613	0.316
	Duplic. Large Genic	0.697	0.145	0.683	0.172	0.708	0.482
	Duplic. Large SFARI	0.778	0.728	0.768	0.757	1.238	0.870
Total and group totals Bolded: significant $p<0.05$ *: significant at $p<0.0005$ TS= Tall Stature, N=Normal Ht.		N(All)=772 N(TS)=116 N(N)=656		N(Male)=509 $N(TS_{Male})=68$ $N(N_{Male})=441$		N(Female)=263 $N(TS_{Female})=48$ $N(N_{Female})=215$	

S. Table VII: CNV Burden for Counts and Lengths Comparing Children Short Stature and Children with Normal Height

SHORT STATURE N= 724		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (Short Stature): 68 N (Normal Ht): 656	**CNV Characteristics**	**SS/N Ratio**	**p-value**	**SS/N Ratio**	**p-value**
All CNVs	Overall	0.906	0.089	1.031	0.879
	Overall Genic	0.931	0.313	0.956	0.829
	Overall SFARI	0.886	0.332	0.914	0.636
All Large CNVs	Overall large	0.773	0.159	1.248	0.670
	Overall large Genic	0.734	0.117	1.015	0.973
	Overall large SFARI	1.786	0.411	2.552	0.413
Deletions	Overall Deletion	0.917	0.229	0.896	0.236
	Deletion Genic	0.950	0.496	0.785	**0.008**
	Deletion SFARI	0.911	0.540	0.778	0.159
Deletions Large	Deletion Large	0.465	**0.036**	0.756	0.574
	Deletion Large Genic	0.275	**0.001**	0.233	**0.001**
	Deletion Large SFARI	1.378	0.794	0.494	0.520
Duplications	Duplic. Overall	0.894	0.175	1.112	0.725
	Duplic. Genic	0.909	0.388	1.054	0.863
	Duplic. SFARI	0.847	0.386	1.180	0.645
Duplications Large	Duplic. Large Overall	0.838	0.395	1.317	0.633
	Duplic. Large Genic	0.830	0.399	1.155	0.779
	Duplic. Large SFARI	1.929	0.438	3.511	0.364

Bolded: significant at $p<0.05$. No statistically significant result after multiple testing corrections.

SS= Short Stature N= Normal Height

S. Table VIII: CNV Burden for Lengths in Short Stature Stratified by Sex

TALL STATURE Stratified by Sex		**Average CNV Length (kb) All**		**Average CNV Length (kb) Male**		**Average CNV Length (kb) Female**	
Copy Number Variants (CNVs)		SS/N Ratio	p-value	SS/N Ratio	p-value	SS/N Ratio	p-value
All CNVs	Overall	1.031	0.879	0.814	0.129	1.332	0.479
	Overall Genic	0.956	0.829	0.708	**0.024**	1.288	0.523
	Overall SFARI	0.914	0.636	0.808	0.264	1.069	0.851
All Large CNVs	Overall large	1.248	0.670	0.600	0.083	2.138	0.406
	Overall large Genic	1.015	0.973	0.433	**0.008**	1.799	0.459
	Overall large SFARI	2.552	0.413	0	--	6.199	0.259
Deletions	Overall Deletion	0.896	0.236	0.838	0.112	0.978	0.891
	Deletion Genic	0.785	**0.008**	0.721	**.0002***	0.883	0.478
	Deletion SFARI	0.778	0.159	0.775	0.273	0.782	0.375
Deletions Large	Deletion Large	0.756	0.574	0.677	0.574	0.842	0.815
	Deletion Large Genic	0.233	**.0002***	0.137	**.0002***	0.359	0.152
	Deletion Large SFARI	0.494	0.520	0	--	0.511	0.599
Duplications	Duplic. Overall	1.112	0.725	0.800	0.255	1.539	0.459
	Duplic. Genic	1.054	0.863	0.700	0.118	1.506	0.461
	Duplic. SFARI	1.180	0.645	0.870	0.681	1.645	0.458
Duplications Large	Duplic. Large Overall	1.317	0.633	0.590	0.105	2.331	0.398
	Duplic. Large Genic	1.155	0.779	0.485	**0.038**	2.057	0.407
	Duplic. Large SFARI	3.511	0.364	0	--	22.99	0.225
Total and group totals Bolded: significant p<0.05 *: significant at p<0.0005 SS= Short Stature, N=Normal Ht.		N(All)=724 N(SS)=68 N(N)=656		N(Male)=40 $N(SS_{Male})$=441 $N(N_{Male})$=481		N(Female)=28 $N(SS_{Female})$=215 $N(N_{Female})$=243	

S. Table IX: CNV Burden for Counts and Lengths Comparing Overweight With Normal-Weight Children

OVERWEIGHT N= 777		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (Overweight): 106 N (Normal Wt): 671	**CNV Characteristics**	**O/N Ratio**	**p-value**	**O/N Ratio**	**p-value**
All CNVs	Overall	1.080	0.366	1.071	0.621
	Overall Genic	1.088	0.308	1.083	0.594
	Overall SFARI	1.096	0.571	1.083	0.729
All Large CNVs	Overall large	1.103	0.667	0.997	0.994
	Overall large Genic	1.146	0.591	1.045	0.874
	Overall large SFARI	1.130	0.835	0.673	0.515
Deletions	Overall Deletion	1.087	0.510	1.120	0.458
	Deletion Genic	1.074	0.500	1.105	0.466
	Deletion SFARI	1.063	0.801	1.011	0.973
Deletions Large	Deletion Large	0.712	0.302	0.937	0.880
	Deletion Large Genic	0.661	0.265	0.842	0.673
	Deletion Large SFARI	1.808	0.548	0.761	0.777
Duplications	Dup. Overall	1.072	0.356	1.042	0.813
	Dup. Genic	1.104	0.317	1.070	0.731
	Dup. SFARI	1.147	0.322	1.214	0.267
Duplications Large	Dup. Large Overall	1.185	0.509	1.007	0.981
	Dup. Large Genic	1.246	0.436	1.081	0.801
	Dup. Large SFARI	0.904	0.893	0.632	0.558

Bolded: significant at $p<0.05$. No statistically significant result after multiple testing corrections.

O= Overweight N= Normal weight

S. Table X: CNV Burden for Lengths in Overweight Stratified by Sex

OVERWEIGHT Stratified by Sex		**Average CNV Length (kb) All**		**Average CNV Length (kb) Male**		**Average CNV Length (kb) Female**	
Copy Number Variants (CNVs)		O/N Ratio	p-value	O/N Ratio	p-value	O/N Ratio	p-value
All CNVs	Overall	1.071	0.621	1.130	0.519	0.961	0.822
	Overall Genic	1.083	0.594	1.167	0.442	0.935	0.740
	Overall SFARI	1.083	0.729	1.216	0.548	0.840	0.360
All Large CNVs	Overall large	0.997	0.994	1.142	0.707	0.736	0.473
	Overall large Genic	1.045	0.874	1.241	0.546	0.712	0.434
	Overall large SFARI	0.673	0.515	1.090	0.904	0	--
Deletions	Overall Deletion	1.120	0.458	1.279	0.251	0.827	0.116
	Deletion Genic	1.105	0.466	1.279	0.192	0.780	**0.040**
	Deletion SFARI	1.011	0.973	1.236	0.665	0.615	0.052
Deletions Large	Deletion Large	0.937	0.880	1.387	0.551	0.213	**0.008**
	Deletion Large Genic	0.842	0.673	1.368	0.527	0	--
	Deletion Large SFARI	0.761	0.777	3.124	0.415	0	--
Duplications	Duplic. Overall	1.042	0.813	1.043	0.859	1.041	0.866
	Duplic. Genic	1.070	0.731	1.101	0.716	1.019	0.946
	Duplic. SFARI	1.214	0.267	1.179	0.443	1.282	0.418
Duplications Large	Duplic. Large Overall	1.007	0.981	1.108	0.797	0.821	0.674
	Duplic. Large Genic	1.081	0.801	1.219	0.622	0.844	0.719
	Duplic. Large SFARI	0.632	0.558	0.802	0.801	0	--
Total and group totals Bolded: significant $p<0.05$ *: significant at $p<0.0005$ O= Overweight, N=Normal Wt.		N(All)=777 N(O)=106 N(N)=671		N(Male)=508 $N(O_{Male})=70$ $N(N_{Male})=438$		N(Female)=269 $N(O_{Female})=36$ $N(N_{Female})=233$	

S. Table XI: CNV Burden for Counts and Lengths Comparing Underweight With Normal-Weight Children

UNDERWEIGHT N= 734		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (Underweight): 63 N (Normal Wt): 671	**CNV Characteristics**	**U/N Ratio**	**p-value**	**U/N Ratio**	**p-value**
All CNVs	Overall	0.945	0.413	1.074	0.740
	Overall Genic	0.943	0.408	1.034	0.876
	Overall SFARI	1.006	0.964	1.148	0.488
All Large CNVs	Overall large	0.835	0.297	1.275	0.672
	Overall large Genic	0.842	0.356	1.171	0.747
	Overall large SFARI	1.141	0.871	1.924	0.620
Deletions	Overall Deletion	0.974	0.802	0.964	0.751
	Deletion Genic	0.944	0.507	0.899	0.313
	Deletion SFARI	1.112	0.589	1.093	0.657
Deletions Large	Deletion Large	0.798	0.546	0.520	0.066
	Deletion Large Genic	0.953	0.906	0.687	0.335
	Deletion Large SFARI	0	--	0	--
Duplications	Dup. Overall	0.910	0.258	1.139	0.686
	Dup. Genic	0.941	0.566	1.111	0.742
	Dup. SFARI	0.844	0.414	1.249	0.554
Duplications Large	Dup. Large Overall	0.843	0.372	1.387	0.604
	Dup. Large Genic	0.819	0.345	1.257	0.680
	Dup. Large SFARI	1.521	0.652	2.805	0.505

Bolded: significant at $p<0.05$. No statistically significant result after multiple testing correction.

U= Underweight N= Normal Weight

S. Table XII: CNV Burden for Lengths in Underweight Stratified by Sex

UNDERWEIGHT Stratified by Sex		**Average CNV Length (kb) All**		**Average CNV Length (kb) Male**		**Average CNV Length (kb) Female**	
Copy Number Variants (CNVs)		U/N Ratio	p-value	U/N Ratio	p-value	U/N Ratio	p-value
All CNVs	Overall	1.074	0.740	0.939	0.619	1.315	0.596
	Overall Genic	1.034	0.876	0.924	0.569	1.222	0.693
	Overall SFARI	1.148	0.488	1.099	0.629	1.234	0.623
All Large CNVs	Overall large	1.275	0.672	0.711	0.229	2.257	0.474
	Overall large Genic	1.171	0.747	0.722	0.253	1.909	0.506
	Overall large SFARI	1.924	0.620	0.149	**0.039**	4.678	0.442
Deletions	Overall Deletion	0.964	0.751	1.018	0.911	0.866	0.191
	Deletion Genic	0.899	0.313	0.962	0.788	0.786	0.080
	Deletion SFARI	1.093	0.657	1.153	0.594	0.990	0.973
Deletions Large	Deletion Large	0.520	0.066	0.727	0.465	0.199	**0.014**
	Deletion Large Genic	0.687	0.335	0.948	0.913	0.284	0.060
	Deletion Large SFARI	0	--	0	--	0	--
Duplications	Duplic. Overall	1.139	0.686	0.891	0.506	1.583	0.532
	Duplic. Genic	1.111	0.742	0.901	0.581	1.458	0.596
	Duplic. SFARI	1.249	0.554	1.002	0.993	1.713	0.502
Duplications Large	Duplic. Large Overall	1.387	0.604	0.708	0.278	2.592	0.437
	Duplic. Large Genic	1.257	0.680	0.683	0.242	2.211	0.456
	Duplic. Large SFARI	2.805	0.505	0.171	0.072	12.05	0.369
Total and group totals Bolded: significant $p<0.05$ *: significant at $p<0.0005$ U= Underweight N=Normal Wt		N(All)=734 N(U)=63 N(N)=671		N(Male)=479 $N(U_{Male})=41$ $N(N_{Male})=438$		N(Female)=255 $N(U_{Female})=22$ $N(N_{Female})=233$	

S. Table XIII: CNV Burden for Counts and Lengths Comparing Children With High BMI and Children with Normal BMI

HIGH BMI N= 765		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (High BMI): 118 N (Normal BMI): 647	**CNV Characteristics**	**H/N Ratio**	**p-value**	**H/N Ratio**	**p-value**
All CNVs	Overall	1.041	0.589	0.972	0.831
	Overall Genic	1.025	0.740	0.997	0.847
	Overall SFARI	1.040	0.787	0.972	0.831
All Large CNVs	Overall large	0.984	0.943	0.799	0.407
	Overall large Genic	1.013	0.954	0.844	0.533
	Overall large SFARI	0.548	0.231	0.382	0.109
Deletions	Overall Deletion	1.031	0.778	1.038	0.786
	Deletion Genic	1.035	0.714	1.003	0.976
	Deletion SFARI	1.025	0.906	1.066	0.841
Deletions Large	Deletion Large	0.541	**0.035**	0.731	0.424
	Deletion Large Genic	0.572	0.081	0.757	0.463
	Deletion Large SFARI	1.827	0.546	1.365	0.764
Duplications	Dup. Overall	1.054	0.460	0.934	0.686
	Dup. Genic	1.013	0.888	0.955	0.810
	Dup. SFARI	1.063	0.648	1.157	0.453
Duplications Large	Dup. Large Overall	1.081	0.746	0.808	0.476
	Dup. Large Genic	1.107	0.702	0.858	0.611
	Dup. Large SFARI	0.228	**0.018**	0.211	**0.049**

Bolded: significant at $p<0.05$. No statistically significant result after multiple testing corrections.

H= High BMI N= Normal BMI

S. Table XIV: CNV Burden for Lengths in High BMI Stratified by Sex

HIGH BMI Stratified by Sex		**Average CNV Length (kb) All**		**Average CNV Length (kb) Male**		**Average CNV Length (kb) Female**	
Copy Number Variants (CNVs)		H/N Ratio	p-value	H/N Ratio	p-value	H/N Ratio	p-value
All CNVs	Overall	1.099	0.660	1.028	0.875	0.872	0.420
	Overall Genic	0.997	0.847	1.037	0.845	0.864	0.459
	Overall SFARI	0.972	0.831	1.276	0.403	0.772	0.227
All Large CNVs	Overall large	0.799	0.407	0.963	0.912	0.529	0.148
	Overall large Genic	0.844	0.533	1.022	0.949	0.564	0.188
	Overall large SFARI	0.382	0.109	0.599	0.454	0	--
Deletions	Overall Deletion	1.038	0.786	1.120	0.551	0.866	0.182
	Deletion Genic	1.003	0.976	1.089	0.621	0.821	0.085
	Deletion SFARI	1.066	0.841	1.339	0.496	0.575	**0.024**
Deletions Large	Deletion Large	0.731	0.424	0.947	0.911	0.229	**0.003**
	Deletion Large Genic	0.757	0.463	1.064	0.890	0.102	**0.001**
	Deletion Large SFARI	1.365	0.764	2.615	0.465	0	--
Duplications	Duplic. Overall	0.934	0.686	0.972	0.898	0.875	0.584
	Duplic. Genic	0.955	0.810	1.006	0.979	0.885	0.662
	Duplic. SFARI	1.157	0.453	1.169	0.522	1.140	0.688
Duplications Large	Duplic. Large Overall	0.808	0.476	0.965	0.926	0.561	0.225
	Duplic. Large Genic	0.858	0.611	1.014	0.970	0.626	0.316
	Duplic. Large SFARI	0.211	**0.049**	0.321	0.201	0	--
Total and group totals Bolded: significant p<0.05 *: significant at p<0.0005 H=High BMI, N=Normal BMI		N(All)=765 N(H)=118 N(N)=647		N(Male)=499 $N(H_{Male})$=80 $N(N_{Male})$=419		N(Female)=266 $N(H_{Female})$=38 $N(N_{Female})$=228	

S. Table XV: CNV Burden for Counts and Lengths Comparing Children With Low BMI and Children with Normal BMI

LOW BMI N= 722		**CNV Counts/ Rate**		**Average CNV Length (kb)**	
N (Low BMI): 75 N (Normal BMI): 647	**CNV Characteristics**	**L/N Ratio**	**p-value**	**L/N Ratio**	**p-value**
All CNVs	Overall	0.957	0.446	0.871	0.219
	Overall Genic	0.944	0.358	0.867	0.253
	Overall SFARI	1.054	0.638	1.142	0.485
All Large CNVs	Overall large	0.822	0.363	0.641	0.118
	Overall large Genic	0.876	0.556	0.723	0.242
	Overall large SFARI	0.862	0.797	1.616	0.663
Deletions	Overall Deletion	0.896	0.096	0.939	0.543
	Deletion Genic	0.922	0.184	0.930	0.515
	Deletion SFARI	0.987	0.938	1.197	0.513
Deletions Large	Deletion Large	0.639	0.193	1.175	0.813
	Deletion Large Genic	0.772	0.490	1.241	0.706
	Deletion Large SFARI	1.437	0.773	9.233	0.376
Duplications	Dup. Overall	1.031	0.680	0.833	0.220
	Dup. Genic	0.971	0.766	0.832	0.282
	Dup. SFARI	1.158	0.425	1.043	0.841
Duplications Large	Dup. Large Overall	0.862	0.544	0.571	0.055
	Dup. Large Genic	0.898	0.671	0.641	0.133
	Dup. Large SFARI	0.718	0.614	0.291	0.082

Bolded: significant at $p<0.05$. No statistically significant result after multiple testing corrections.

L= Low BMI N= Normal BMI

S. Table XVI: CNV Burden for Lengths in Low BMI Stratified by Sex

LOW BMI Stratified by Sex		**Average CNV Length (kb) All**		**Average CNV Length (kb) Male**		**Average CNV Length (kb) Female**	
Copy Number Variants (CNVs)		L/N Ratio	p-value	L/N Ratio	p-value	L/N Ratio	p-value
All CNVs	Overall	0.871	0.219	0.932	0.601	0.764	0.188
	Overall Genic	0.867	0.253	0.960	0.773	0.710	0.164
	Overall SFARI	1.142	0.485	1.195	0.460	1.052	0.871
All Large CNVs	Overall large	0.641	0.118	0.676	0.199	0.591	0.358
	Overall large Genic	0.723	0.242	0.826	0.523	0.568	0.324
	Overall large SFARI	1.616	0.663	0.395	0.213	3.708	0.475
Deletions	Overall Deletion	0.939	0.543	0.913	0.334	0.989	0.965
	Deletion Genic	0.930	0.515	0.916	0.369	0.957	0.874
	Deletion SFARI	1.197	0.513	1.307	0.443	1.015	0.974
Deletions Large	Deletion Large	1.175	0.813	0.590	0.240	2.439	0.526
	Deletion Large Genic	1.241	0.706	0.893	0.824	1.944	0.579
	Deletion Large SFARI	9.233	0.376	0	--	19.47	0.352
Duplications	Duplic. Overall	0.833	0.220	0.944	0.755	0.644	0.078
	Duplic. Genic	0.832	0.282	0.986	0.946	0.589	0.084
	Duplic. SFARI	1.043	0.841	1.002	0.993	1.123	0.724
Duplications Large	Duplic. Large Overall	0.571	0.055	0.688	0.280	0.391	0.083
	Duplic. Large Genic	0.641	0.133	0.815	0.554	0.385	0.094
	Duplic. Large SFARI	0.291	0.082	0.450	0.313	0	--
Total and group totals Bolded: significant p<0.05 *: significant at p<0.0005 L=Low BMI, N=Normal BMI		N(All)=722 N(L)=75 N(N)=647		N(Male)=469 $N(L_{Male})$=50 $N(N_{Male})$=419		N(Female)=253 $N(L_{Female})$=25 $N(N_{Female})$=228	

S. Table XVII: Association between CNVs in ASD-associated regions and Macrocephaly

Region	All and All large CNVs			Deletion and Deletion large CNVs			Duplication and Duplication large CNVs		
	Mac, N=52	N, N=623	p-value	Mac, N=52	N, N=623	p-value	Mac, N=52	N, N=623	p-value
1q21.1A	4(7.7%)	12(1.9%)	**0.028**	0(0%)	1(0.1%)	1	4(7.7%)	11(1.7%)	**0.022**
1q21.1L	3(5.7%)	4(0.6%)	**0.012**	0(0%)	0(0%)	--	3(5.7%)	4(0.6%)	**0.012**
3q29A	1(1.9%)	5(0.8%)	0.383	0(0%)	0(0%)	--	1(1.9%)	5(0.8%)	0.383
3q29L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
7q11.23A	0(0%)	6(0.9%)	1	0(0%)	0(0%)	--	0(0%)	6(0.9%)	1
7q11.23L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q11.2A	6(11.5%)	86(13.8%)	0.833	5(9.6%)	51(8.1%)	0.608	1(1.9%)	35(5.6%)	0.513
15q11.2L	0(0%)	2(0.3%)	1	0(0%)	0(0%)	--	0(0%)	2(0.3%)	1
15q11.2.13.1A	6(11.5%)	85(13.6%)	0.833	5(9.6%)	51(8.1%)	0.608	1(1.9%)	34(5.4%)	0.509
15q11.2.13.1L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q13.3A	0(0%)	11(1.7%)	1	0(0%)	0(0%)	--	0(0%)	11(1.7%)	1
15q13.3L	0(0%)	8(1.2%)	1	0(0%)	0(0%)	--	0(0%)	8(1.2%)	1
16p11.2A	0(0%)	3(0.5%)	1	0(0%)	1(0.1%)	1	0(0%)	2(0.3%)	1
16p11.2L	0(0%)	2(0.3%)	1	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1
16p13.11A	6(11.5%)	73(11.7%)	1	1(1.9%)	13(2%)	1	5(9.6%)	60(9.6%)	1
16p13.11L	0(0%)	3(0.5%)	1	0(0%)	1(0.1%)	1	0(0%)	2(0.3%)	1
17q12A	4(7.7%)	107(17.1%)	0.081	0(0%)	7(1.1%)	1	4(7.7%)	100(16%)	0.158
17q12L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
22q11.21A	1(1.9%)	29(4.6%)	0.721	0(0%)	5(0.8%)	1	1(1.9%)	24	0.712
22q11.21L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--

A= All
L= Large
Mac=Macrocephaly
N=Normal Head Circumference
Bolded: $p<0.05$

S. Table XVIII: Association between CNVs in ASD-associated regions and Microcephaly

Region	All and All large CNVs			Deletion and Deletion large CNVs			Duplication and Duplication large CNVs		
	Mic, N=165	N, N=623	p-value	Mic, N=165	N, N=623	p-value	Mic, N=165	N, N=623	p-value
1q21.1A	2(1.2%)	12(1.9%)	0.745	2(1.2%)	1(0.1%)	**0.015**	0(0%)	11 (1.7%)	0.132
1q21.1L	0(0%)	4(0.6%)	0.585	0(0%)	0(0%)	--	0(0%)	4(0.6%)	0.585
3q29A	0(0%)	5(0.8%)	0.589	0(0%)	0(0%)	--	0(0%)	5(0.8%)	0.589
3q29L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
7q11.23A	5(3.0%)	6(0.9%)	0.058	2(1.2%)	0(0%)	**0.043**	3(1.8%)	6(0.9%)	0.405
7q11.23L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q11.2A	12(7.2%)	86 (13.8%)	**0.023**	7(4.2%)	51(8.1%)	0.094	5(3.0%)	35(5.6%)	0.231
15q11.2L	3(1.8%)	2(0.3%)	0.064	1(0.6%)	0(0%)	0.209	2(1.2%)	2(0.3%)	0.195
15q11.2.13.1A	9(5.4%)	85(13.6%)	**0.002**	6(3.6%)	51(8.1%)	**0.043**	3(1.8%)	34(5.4%)	0.060
15q11.2.13.1L	1(0.6%)	0(0%)	0.209	1(0.6%)	0(0%)	0.209	0(0%)	0(0%)	--
15q13.3A	3(1.8%)	11(1.7%)	1	0(0%)	0(0%)	--	3(1.8%)	11(1.7%)	1
15q13.3L	1(0.6%)	8(1.2%)	0.693	0(0%)	0(0%)	--	1(0.6%)	8(1.2%)	0.693
16p11.2A	1(0.6%)	3(0.5%)	1	0(0%)	1(0.1%)	1	1(0.6%)	2(0.3%)	0.506
16p11.2L	0(0%)	2(0.3%)	1	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1
16p13.11A	17(10.3%)	73(11.7%)	0.681	4(2.4%)	13(2%)	0.765	13(7.8%)	60(9.6%)	0.548
16p13.11L	0(0%)	3(0.5%)	1	0(0%)	1(0.1%)	1	0(0%)	2(0.3%)	1
17q12A	24(14.5%)	107(17.1%)	0.481	0(0%)	7(1.1%)	0.355	24(14.5%)	100(16%)	0.718
17q12L	2(1.2%)	0(0%)	**0.043**	0(0%)	0(0%)	--	2(1.2%)	0(0%)	**0.043**
22q11.21A	9(5.4%)	29(4.6%)	0.683	0(0%)	5(0.8%)	0.589	9(5.4%)	24	0.381
22q11.21L	2(1.2%)	0(0%)	**0.043**	0(0%)	0(0%)	--	2(1.2%)	0(0%)	**0.043**

A= All
L= Large
Mic=Microcephaly
N=Normal Head Circumference
Bolded: $p<0.05$

S. Table XIX: Association between CNVs in ASD-associated regions and Tall Stature

Region	All and All large CNVs			Deletion and Deletion large CNVs			Duplication and Duplication large CNVs		
	TS, N=116	N, N=656	p-value	TS, N=116	N, N=656	p-value	TS, N=116	N, N=656	p-value
1q21.1A	1(0.8%)	16(2.4%)	0.492	1(0.8%)	2(0.3%)	0.386	0(0%)	14(2.1%)	0.245
1q21.1L	0(0%)	7(1.0%)	0.602	0(0%)	0(0%)	--	0(0%)	7(1.0%)	0.602
3q29A	1(0.8%)	4(0.6%)	0.558	0(0%)	0(0%)	--	1(0.8%)	4(0.6%)	0.558
3q29L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
7q11.23A	3(2.5%)	8(1.2%)	0.221	0(0%)	2(0.3%)	1	3(2.5%)	6(0.9%)	0.140
7q11.23L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q11.2A	17(14.6%)	78(11.9%)	0.442	11(9.5%)	47(7.1%)	0.443	6(5.1%)	31(4.7%)	0.814
15q11.2L	0(0%)	3(0.4%)	1	0(0%)	1(0.1%)	1	0(0%)	2(0.3%)	1
15q11.2.13.1A	16(12.7%)	78(11.9%)	0.540	11(9.5%)	47(7.1%)	0.443	5(3.9%)	31(4.7%)	1
15q11.2.13.1L	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--
15q13.3A	2(1.7%)	10(1.5%)	0.698	0(0%)	0(0%)	--	2(1.7%)	10(1.5%)	0.698
15q13.3L	2(1.7%)	6(0.9%)	0.343	0(0%)	0(0%)	--	2(1.7%)	6(0.9%)	0.343
16p11.2A	0(0%)	4(0.6%)	1	0(0%)	1(0.1%)	1	0(0%)	3(0.4%)	1
16p11.2L	0(0%)	2(0.3%)	1	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1
16p13.11A	10(8.6%)	79(12.0%)	0.345	3(2.5%)	15(2.2%)	1	7(6.0%)	64(9.7%)	0.226
16p13.11L	0(0%)	3(0.4%)	1	0(0%)	1(0.1%)	1	0(0%)	2(0.3%)	1
17q12A	20(17.2%)	103(15.7%)	0.680	1(0.8%)	4(0.6%)	0.558	19(16.3%)	99(15.1%)	0.779
17q12L	0(0%)	2(0.3%)	1	0(0%)	0(0%)	--	0(0%)	2(0.3%)	1
22q11.21A	7(6.0%)	32	0.644	1(0.8%)	4(0.6%)	0.558	6(5.1%)	28(4.2%)	0.625
22q11.21L	0(0%)	2(0.3%)	1	0(0%)	0(0%)	--	0(0%)	2(0.3%)	1

A= All
L= Large
TS=Tall Stature
N=Normal Height
Bolded: $p<0.05$

S. Table XX: Association between CNVs in ASD-associated regions and Short Stature

Region	All and All large CNVs			Deletion and Deletion large CNVs			Duplication and Duplication large CNVs		
	SS, N= 68	N, N= 656	p-value	SS, N= 68	N, N=656	p-value	SS, N= 68	N, N=656	p-value
1q21.1A	1(1.4%)	16(2.4%)	1	0(0%)	2(0.3%)	1	1(1.4%)	14 (2.1%)	1
1q21.1L	0(0%)	7(1.0%)	1	0(0%)	0(0%)	--	0(0%)	7(1.0%)	1
3q29A	1(1.4%)	4(0.6%)	0.390	0(0%)	0(0%)	--	1(1.4%)	4(0.6%)	0.390
3q29L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
7q11.23A	0(0%)	8(1.2%)	1	0(0%)	2(0.3%)	1	0(0%)	6(0.9%)	1
7q11.23L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q11.2A	9(13.2%)	78(11.9%)	0.697	5(7.3%)	47(7.1%)	1	4(5.8%)	31(4.7%)	0.561
15q11.2L	2(2.9%)	3(0.4%)	0.072	0(0%)	1(0.1%)	1	2(2.9%)	2(0.3%)	**0.046**
15q11.2.13.1A	6(8.8%)	78(11.9%)	0.553	4(5.8%)	47(7.1%)	1	2(2.9%)	31(4.7%)	0.760
15q11.2.13.1L	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--
15q13.3A	2(2.9%)	10(1.5%)	0.313	0(0%)	0(0%)	--	2(2.9%)	10(1.5%)	0.313
15q13.3L	1(1.4%)	6(0.9%)	0.500	0(0%)	0(0%)	--	1(1.4%)	6(0.9%)	0.500
16p11.2A	0(0%)	4(0.6%)	1	0(0%)	1(0.1%)	1	0(0%)	3(0.4%)	1
16p11.2L	0(0%)	2(0.3%)	1	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1
16p13.11A	7(10.3%)	79(12.0%)	0.843	0(0%)	15(2.2%)	0.383	7 (10.3%)	64(9.7%)	0.828
16p13.11L	0(0%)	3(0.4%)	1	0(0%)	1(0.1%)	1	0(0%)	2(0.3%)	1
17q12A	12(17.6%)	103(15.7%)	0.727	2(2.9%)	4(0.6%)	0.101	10(14.7%)	99(15.1%)	1
17q12L	0(0%)	2(0.3%)	1	0(0%)	0(0%)	--	0(0%)	2(0.3%)	1
22q11.21A	0(0%)	32	0.063	0(0%)	4(0.6%)	1	0(0%)	28(4.2%)	0.099
22q11.21L	0(0%)	2(0.3%)	1	0(0%)	0(0%)	--	0(0%)	2(0.3%)	1

A= All
L= Large
SS=Short Stature
N=Normal Height
Bolded: $p<0.05$

S. Table XXI: Association between CNVs in ASD-associated regions and Overweight

Region	All and All large CNVs			Deletion and Deletion large CNVs			Duplication and Duplication large CNVs		
	O, N=106	N, N= 671	p-value	O, N=106	N, N=671	p-value	O, N= 106	N, N=671	p-value
1q21.1A	1(0.9%)	17(2.5%)	0.492	0(0%)	3(0.4%)	1	1(0.9%)	14(2.1%)	0.707
1q21.1L	1(0.9%)	6(0.9%)	1	0(0%)	0(0%)	--	1(0.9%)	6(0.9%)	1
3q29A	1(0.9%)	5(0.7%)	0.586	0(0%)	0(0%)	--	1(0.9%)	5(0.7%)	0.586
3q29L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
7q11.23A	2(1.9%)	8(1.2%)	0.634	0(0%)	2(0.3%)	1	2(1.9%)	6(0.9%)	0.299
7q11.23L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q11.2A	11(10.3%)	84(12.5%)	0.633	6(5.6%)	52(7.7%)	0.553	5(4.7%)	32(4.7%)	1
15q11.2L	1(0.9%)	3(0.4%)	0.444	0(0%)	1(0.1%)	1	1(0.9%)	2(0.3%)	0.356
15q11.2.13.1A	10 (9.4%)	83(12.3%)	0.518	6(5.6%)	52(7.7%)	0.553	4(3.7%)	31(4.6%)	1
15q11.2.13.1L	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--
15q13.3A	1(0.9%)	12(1.7%)	1	0(0%)	0(0%)	--	1(0.9%)	12(1.7%)	1
15q13.3L	1(0.9%)	8(1.2%)	1	0(0%)	0(0%)	--	1(0.9%)	8(1.2%)	1
16p11.2A	1(0.9%)	2(0.3%)	0.356	1(0.9%)	0(0%)	0.136	0(0%)	2(0.3%)	1
16p11.2L	1(0.9%)	0(0%)	0.136	1(0.9%)	0(0%)	0.136	0(0%)	0(0%)	--
16p13.11A	11(10.3%)	0(0%)	0.746	0(0%)	17(2.5%)	0.149	11(10.3%)	62(9.2%)	0.720
16p13.11L	1(0.9%)	2(0.3%)	0.356	0(0%)	1(0.1%)	1	1(0.9%)	1(0.1%)	0.254
17q12A	23(21.7%)	102(15.2%)	0.116	2(1.9%)	5(0.7%)	0.245	21(19.8%)	97(14.4%)	0.188
17q12L	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--	0(0%)	1(0.1%)	1
22q11.21A	4(3.7%)	35(5.2%)	0.639	1(0.9%)	4(0.6%)	0.520	3(2.8%)	31(4.6%)	0.608
22q11.21L	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--	0(0%)	1(0.1%)	1

A= All
L= Large
O=Overweight
N= Normal weight

S. Table XXII: Association between CNVs in ASD-associated regions and Underweight

Region	All and All large CNVs			Deletion and Deletion large CNVs			Duplication and Duplication large CNVs		
	U, N= 63	N, N=671	p-value	U, N= 63	N, N=671	p-value	U, N= 63	N, N=671	p-value
1q21.1A	0(0%)	17(2.5%)	0.386	0(0%)	3(0.4%)	1	0(0%)	14(2.1%)	0.623
1q21.1L	0(0%)	6(0.9%)	1	0(0%)	0(0%)	--	0(0%)	6(0.9%)	1
3q29A	0(0%)	5(0.7%)	1	0(0%)	0(0%)	--	0(0%)	5(0.7%)	1
3q29L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
7q11.23A	1(1.5%)	8(1.2%)	0.556	0(0%)	2(0.3%)	1	1(1.5%)	6(0.9%)	0.468
7q11.23L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q11.2A	9(14.2%)	84(12.5%)	0.691	5(7.9%)	52(7.7%)	1	4(6.3%)	32(4.7%)	0.539
15q11.2L	1(1.5%)	3(0.4%)	0.302	0(0%)	1(0.1%)	1	1(1.5%)	2(0.3%)	0.236
15q11.2.13.1A	7(11.1%)	83(12.3%)	1	4(6.3%)	52(7.7%)	1	3(4.7%)	31(4.6%)	1
15q11.2.13.1L	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--
15q13.3A	1(1.5%)	12(1.7%)	1	0(0%)	0(0%)	--	1(1.5%)	12(1.7%)	1
15q13.3L	0(0%)	8(1.2%)	1	0(0%)	0(0%)	--	0(0%)	8(1.2%)	1
16p11.2A	1(1.5%)	2(0.3%)	0.236	0(0%)	0(0%)	--	1(1.5%)	2(0.3%)	0.236
16p11.2L	1(1.5%)	0(0%)	0.085	0(0%)	0(0%)	--	1(1.5%)	0(0%)	0.085
16p13.11A	6(9.5%)	0(0%)	1	1(1.5%)	17(2.5%)	1	5(7.9%)	62(9.2%)	1
16p13.11L	0(0%)	2(0.3%)	1	0(0%)	1(0.1%)	1	0(0%)	1(0.1%)	1
17q12A	10(15.8%)	102(15.2%)	0.855	0(0%)	5(0.7%)	1	10(15.8%)	97(14.4%)	0.711
17q12L	1(1.5%)	1(0.1%)	0.164	0(0%)	0(0%)	--	1(1.5%)	1(0.1%)	0.164
22q11.21A	0(0%)	35(5.2%)	0.063	0(0%)	4(0.6%)	1	0(0%)	31(4.6%)	0.100
22q11.21L	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--	0(0%)	1(0.1%)	1

A= All
L= Large
UW=Underweight
N= Normal weight

S. Table XXIII: Association between CNVs in ASD-associated regions and High BMI

Region	All and All large CNVs			Deletion and Deletion large CNVs			Duplication and Duplication large CNVs		
	HB, N=118	N, N=647	p-value	HB, N=118	N, N=647	p-value	HB, N=118	N, N=647	p-value
1q21.1A	1(0.8%)	16(2.4%)	0.494	0(0%)	3(0.4%)	1	1(0.8%)	13 (2.0%)	0.707
1q21.1L	1(0.8%)	6(0.9%)	1	0(0%)	0(0%)	--	1(0.8%)	6(0.9%)	1
3q29A	1(0.8%)	5(0.7%)	1	0(0%)	0(0%)	--	1(0.8%)	5(0.7%)	1
3q29L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
7q11.23A	3(2.5%)	8(1.2%)	0.389	0(0%)	2(0.3%)	1	3(2.5%)	6(0.9%)	0.149
7q11.23L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q11.2A	14(11.8%)	82(12.6%)	0.881	9(7.6%)	47(7.2%)	0.848	5(4.2%)	35(5.4%)	0.821
15q11.2L	1(0.8%)	3(0.4%)	0.489	0(0%)	0(0%)	--	1(0.8%)	3(0.4%)	0.489
15q11.2.13.1A	14(11.8%)	79(12.2%)	1	10(8.4%)	46(7.1%)	0.567	4(3.3%)	33(5.1%)	0.639
15q11.2.13.1L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q13.3A	0(0%)	11(1.7%)	0.230	0(0%)	0(0%)	--	0(0%)	11(1.7%)	0.230
15q13.3L	0(0%)	8(1.2%)	0.616	0(0%)	0(0%)	--	0(0%)	8(1.2%)	0.616
16p11.2A	1(0.8%)	2(0.3%)	0.395	1(0.8%)	0(0%)	0.154	0(0%)	2(0.3%)	1
16p11.2L	1(0.8%)	0(0%)	0.154	1(0.8%)	0(0%)	0.154	0(0%)	0(0%)	--
16p13.11A	10(8.4%)	75(11.6%)	0.425	0(0%)	16(2.5%)	0.151	10(8.4%)	59(9.1%)	1
16p13.11L	0(0%)	3(0.4%)	1	0(0%)	1(0.1%)	1	0(0%)	2(0.3%)	1
17q12A	24(20.3%)	97(14.9%)	0.169	1(0.8%)	6(0.9%)	1	23(19.5%)	91(14.0%)	0.158
17q12L	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--	0(0%)	1(0.1%)	1
22q11.21A	7(5.9%)	29(4.4%)	0.479	1(0.8%)	4(0.6%)	0.568	6(5.1%)	25(3.8%)	0.609
22q11.21L	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--	0(0%)	1(0.1%)	1

A= All
L= Large
HB=High BMI
N=Normal BMI

S. Table XXIV: Association between CNVs in ASD-associated regions and Low BMI

Region	All and All large CNVs			Deletion and Deletion large CNVs			Duplication and Duplication large CNVs		
	LB, N= 75	N, N= 647	p-value	LB, N= 75	N, N=647	p-value	LB, N= 75	N, N=647	p-value
1q21.1A	1(1.3%)	16(2.4%)	1	0(0%)	3(0.4%)	1	1(1.3%)	13 (2.0%)	1
1q21.1L	0(0%)	6(0.9%)	1	0(0%)	0(0%)	--	0(0%)	6(0.9%)	1
3q29A	0(0%)	5(0.7%)	1	0(0%)	0(0%)	--	0(0%)	5(0.7%)	1
3q29L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
7q11.23A	0(0%)	8(1.2%)	1	0(0%)	2(0.3%)	1	0(0%)	6(0.9%)	1
7q11.23L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
15q11.2A	8 (10.6%)	82 (12.6%)	0.714	7(9.3%)	47(7.2%)	0.488	1(1.3%)	35(5.4%)	0.163
15q11.2L	1(1.3%)	3(0.4%)	0.355	1(1.3%)	0(0%)	0.104	0(0%)	3(0.4%)	1
15q11.2.13.1A	7(9.3%)	79(12.2%)	0.574	6(8.0%)	46(7.1%)	0.812	1(1.3%)	33(5.1%)	0.243
15q11.2.13.1L	1(1.3%)	0(0%)	0.104	1(1.3%)	0(0%)	0.104	0(0%)	0(0%)	--
15q13.3A	3(4.0%)	11(1.7%)	0.170	1(1.3%)	0(0%)	0.104	3(4.0%)	11(1.7%)	0.170
15q13.3L	1(1.3%)	8(1.2%)	1	0(0%)	0(0%)	--	1(1.3%)	8(1.2%)	1
16p11.2A	1(1.3%)	2(0.3%)	0.280	0(0%)	0(0%)	--	1(1.3%)	2(0.3%)	0.280
16p11.2L	1(1.3%)	0(0%)	0.104	0(0%)	0(0%)	--	1(1.3%)	0(0%)	0.104
16p13.11A	11(14.6%)	75(11.6%)	0.450	2(2.6%)	16 (2.5%)	0.709	9(12%)	59(9.1%)	0.404
16p13.11L	0(0%)	3(0.4%)	1	0(0%)	1(0.1%)	1	0(0%)	2(0.3%)	1
17q12A	14(18.6%)	97(14.9%)	0.399	1(1.3%)	6(0.9%)	0.537	14(18.6%)	91(14.0%)	0.299
17q12L	1(1.3%)	1(0.1%)	0.197	0(0%)	0(0%)	--	1(1.3%)	1(0.1%)	0.197
22q11.21A	3(4.0%)	29(4.4%)	1	0(0%)	4(0.6%)	1	3(4.0%)	25(3.8%)	1
22q11.21L	0(0%)	1(0.1%)	1	0(0%)	0(0%)	--	0(0%)	1(0.1%)	1

A= All
L= Large
LB=Low BMI
N=Normal BMI

S. Table XXV: Summary of Overall Association Between ASD-associated CNVs and Growth Abnormalities by CNV Region

Region of Overlap	Total overlapping CNV counts across growth abnormalities per CNV region			Growth abnormality with significant Fisher's exact p by All CNVs comparing frequency of overlap in abnormal to normal growth		Growth abnormality with significant Fisher's exact p by Deletion CNVs comparing frequency of overlap in abnormal to normal growth		Growth abnormality with significant Fisher's exact p by Duplication CNVs comparing frequency of overlap in abnormal to normal growth	
	All	**Del**	**Dup**	**All**	**Large**	**All**	**Large**	**All**	**Large**
1q21.1	11	3(27%)	8(73%)	Macro-cephaly	Macro-cephaly	Micro-cephaly		Macro-cephaly	Macro-cephaly
3q29	5	0(0%)	5(100%)						
7q11.23	14	2(7%)	12(93%)			Micro-cephaly			
15q11.2	86	55(64%)	31(36%)	Micro-cephaly					Short Stature
15q11.2.13.1	75	52(69%)	23(31%)	Micro-cephaly		Micro-cephaly			
15q13.3	12	1(8%)	11(92%)						
16p11.2	5	2(40%)	3(60%)						
16p13.11	78	11(14%)	67(86%)						
17q12	131	7(5%)	124(95%)		Micro-cephaly				Micro-cephaly
22q11.21	31	3(10%)	28(90%)		Micro-cephaly				Micro-cephaly

Del: Deletion
Dup: Duplication

S. Table XXVI: Summary of Overall Association Between ASD-associated CNVs and Growth Abnormalities by Growth Abnormality Stratified by Sex

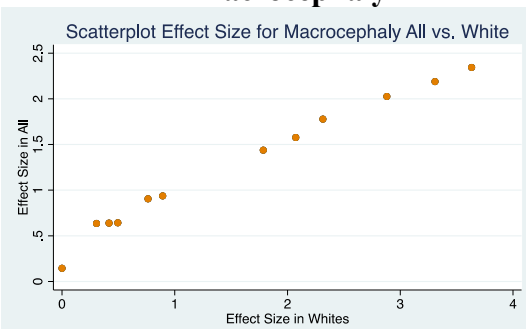
Region	**All and All large CNVs**			**Deletion and Deletion large CNVs**			**Duplication and Duplication large CNVs**		
Growth	Abnormal	Normal	p-value	Abnormal	Normal	p-value	Abnormal	Normal	p-value
Macrocephaly									
ALL, N	52	623		52	623		52	623	
1q21.1A	4(7.7%)	12(1.9%)	**0.028**	0(0%)	1(0.1%)	1	4(7.7%)	11(1.7%)	**0.022**
1q21.1L	3(5.7%)	4(0.6%)	**0.012**	0(0%)	0(0%)	--	3(5.7%)	4(0.6%)	**0.012**
MALE, N	29	415							
1q21.1A	1(3%)	11(1.7%)	0.560	0(0%)	1(0.1%)	1	1(3.4%)	10(2.4%)	0.528
1q21.1L	1(3%)	4(0.6%)	0.287	0(0%)	0(0%)	--	1(3.4%0	4(0.9%)	0.287
FEMALE, N	24	208							
1q21.1A	3(12.5%)	1(0.5%)	**0.003**	0(0%)	0(0%)	--	3(12.5%)	1(0.5%)	**0.003**
1q21.1L	2(8.3%)	0(0%)	**0.010**	0(0%)	0(0%)	--	2(8.3%)	0(0%)	**0.010**
Microcephaly									
ALL	165	623		165	623		165	623	
1q21.1A	2(1.2%)	12(1.9%)	0.745	2(1.2%)	1(0.1%)	**0.015**	0(0%)	11(1.7%)	0.132
1q21.1L	0(0%)	4(0.6%)	0.585	0(0%)	0(0%)	--	0(0%)	4(0.6%)	0.585
MALE, N	105	415		105	415		105	415	
1q21.1A	1(0.9%)	11(2.6%)	0.474	1(0.9%)	1(0.2%)	0.363	0(0%)	10(2.4%)	0.224
1q21.1L	0(0%)	4(0.9%)	0.587	0(0%)	0(0%)	--	0(0%)	4(0.9%)	0.587
FEMALE, N	60	208		60	208		60	208	
1q21.1A	1(1.6%)	1(0.5%)	0.398	1(%)	0(0%)	0.224	0(0%)	1(0.5%)	1
1q21.1L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
ALL	165	623		165	623		165	623	
15q11.2A	12(7.2%)	86(13.8%)	**0.023**	7(4.2%)	51(8.1%)	0.094	5(3.0%)	35(5.6%)	0.231
15q11.2L	3(1.8%)	2(0.3%)	0.064	1(0.6%)	0(0%)	0.209	2(1.2%)	2(0.3%)	0.195
MALE, N	105	415		105	415		105	415	
15q11.2A	6(5.7%)	55(13.2%)	**0.040**	4(3.8%)	30(7.2%)	0.270	2(1.9%)	25(6%)	0.136
15q11.2L	0(0%)	2(0.5%)	1	0(0%)	0(0%)	--	0(0%)	2(0.5%)	1
FEMALE, N	60	208		60	208		60	208	
15q11.2A	6(10%)	29(13.9%)	0.518	3(5%)	19(9.1%)	0.426	3(5%)	10(4.8%)	1
15q11.2L	3(5%)	0(0%)	**0.010**	1(1.6%)	0(0%)	0.224	2(3.3%)	0(0%)	**0.049**
ALL	165	623		165	623		165	623	
15q11.2.13.1A	9(5.4%)	85(13.6%)	**0.002**	6(3.6%)	51(8.1%)	**0.043**	3(1.8%)	34(5.4%)	0.060
15q11.2.13.1L	1(0.6%)	0(0%)	0.209	1(0.6%)	0(0%)	0.209	0(0%)	0(0%)	--
MALE, N	105	415		105	415		105	415	
15q11.2.13.1A	5(4.7%)	55(13.2%)	**0.015**	2(1.9%)	30(7.2%)	0.063	3(2.8%)	25(6.0%)	0.235
15q11.2.13.1L	0(0%)	0(0%)	--	0(0%)	0(0%)	--	0(0%)	0(0%)	--
FEMALE, N	60	208		60	208		60	208	
15q11.2.13.1A	4(6.6%)	28(13.4%)	0.180	3(5%)	21(10.1%)	0.307	1(1.6%)	9(4.3%)	0.465
15q11.2.13.1L	1(1.6%)	0(0%)	0.224	1(1.6%)	0(0%)	0.224	0(0%)	0(0%)	--
Short Stature									
ALL	68	656		68	656		68	656	
15q11.2A	9(13.2%)	78(11.9%)	0.697	5(7.3%)	47(7.1%)	1	4(5.8%)	31(4.7%)	0.561
15q11.2L	2(2.9%)	3(0.4%)	0.072	0(0%)	1(0.1%)	1	2(2.9%)	2(0.3%)	**0.046**
MALE, N	40	441		40	441		40	441	
15q11.2A	1(2.5%)	51(11.5%)	0.106	1(2.5%)	29(6.5%)	0.497	0(0%)	22(4.9%)	0.241
15q11.2L	0(0%)	2(0.4%)	1	0(0%)	0(0%)	--	0(0%)	2(0.4%)	1
FEMALE, N	28	215		28	215		28	215	
15q11.2A	8(28.5%)	25(11.6%)	**0.034**	4(14.2%)	16(7.4%)	0.262	4(14.2%)	9(4.1%)	**0.048**
15q11.2L	2(7.1%)	1(0.4%)	**0.035**	0(0%)	1(0.4%)	1	2(7.1%)	0(0%)	**0.012**

S. Figure I: Scatter Plots of Effect Size in All vs. White for Growth Abnormalities

We created scatter plots of effect size for CNV burden lengths comparing results from the whole analytic sample to results among Non-Hispanic Whites only (S. Figure 1 below). Macrocephaly is the only growth abnormality that shows strong correlation between overall and Non-Hispanic White results. This suggests that Non-Hispanic Whites contribute significantly to the association between CNV burden and macrocephaly. For the other growth abnormalities we assessed, it appears that Non-Hispanic Whites do not drive the association between CNV burden and abnormal growth. Studies assessing genome-wide CNV analyses in different races have shown that there are ethnic differences of CNVs for anthropometric measurements including height and BMI in African and Asian ancestry populations compared to populations with European ancestry (Kang et al., 2010; Li et al., 2010). The potential confounding effect of race needs to be considered in evaluating CNV burden associations for growth abnormalities, although the effect likely varies for different growth abnormalities and in different race/ ethnic groups.

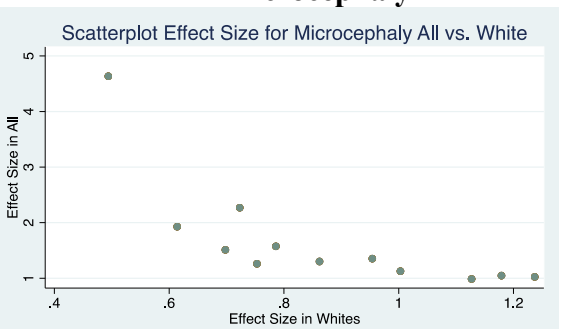
S. Figure I: Scatter Plots of Effect Size in All vs. White for Growth Abnormalities

Macrocephaly



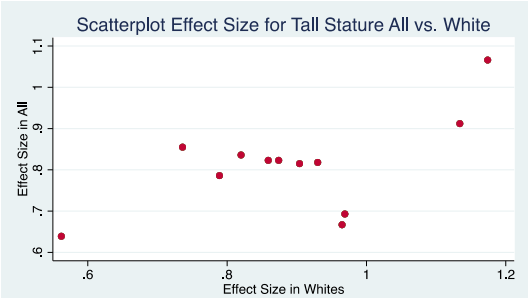
% of NHW with Macrocephaly: 62.1%
% of NHW with No Macrocephaly: 63%

Microcephaly



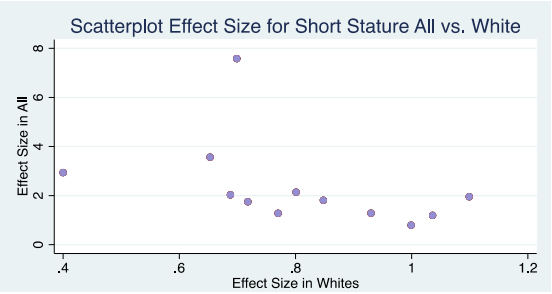
% of NHW with Microcephaly: 48.8%
% of NHW with No Microcephaly: 63%

Tall Stature



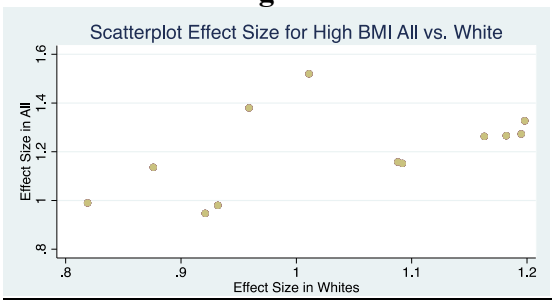
% of NHW with Tall Stature: 54.0%
% of NHW with No Tall Stature: 58%

Short Stature



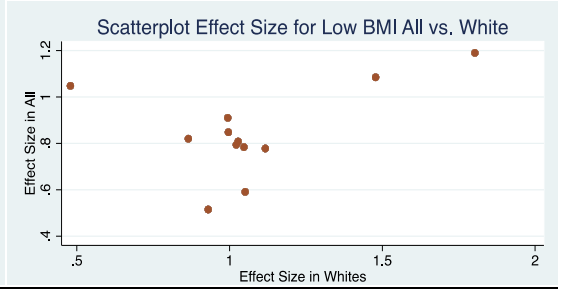
% of NHW with Short Stature: 64.3%
% of NHW with No Short Stature: 58%

High BMI



% of NHW with High BMI: 48.2%
% of NHW with No High BMI: 59.4%

Low BMI



% of NHW with Low BMI: 66.7%
% of NHW with No Low BMI: 59.4%

Further description of S.Figure I provided in the text for appendices on pages that follow.

Text for Appendices

CHAPTER 5: CONCLUSION

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder that can manifest with varying degrees of severity. In addition, ASD often co-occurs with medical, behavioral, or other psychiatric disorders, and characterization of ASD by subgroups as defined by these co-occuring conditions is increasingly recognized (Levy et al., 2010; Muskens, Velders, & Staal, 2017). As ASD may well have multiple etiologies, classifying ASD into one homogenous category could easily limit our ability to identify true causal factors, while also precluding development of tailored approaches to prognostic and therapeutic strategies. Characterizing co-occurring phenotype as ASD sub-groups or sub-phenotypes could improve our understanding of the nuances of autism biology, and potentially unveil associations with risk factors related to distinct etiologies.

In this dissertation, the phenotypes of dysmorphology and specifically abnormal growth as potential ASD sub-phenotypes were examined to help with etiologic research, particularly in genetic subgrouping. Previous work in the SEED 1 study revealed association between ASD and dysmorphism (Shapira et al., 2014). Our assessment of growth abnormalities, a specific subset of dysmorphologic features, in this sample showed females with ASD had short stature and a combination of short stature, microcephaly and normal weight compared to typically developing females. When considering genetic risk via two cumulative measures of CNV burden, we found cumulatie CNV burden was negatively associated with dysmorphism among SEED 1 cases and controls. We also observed decreased CNV burden among those CNVs encompassed previously reported candidate genes for ASD among children with tall stature and macrocephaly. Importantly, these associations also varied by sex. When

considering specific CNVs already known to be associated with ASD, we observed association between ASD and short stature at CNV region 15p11.2 that varied by sex, with significant association only observed in females.

Our results not only highlight the potential utility of dysmorphism co-occurring with ASD, they also highlight the importance of sex differences in ASD phenotypes. Although the high male:female sex ratio in ASD is well established, it was only more recently recognized that ASD can present differently in females compared to males (Rutter et al., 2003; Lai et al., 2015; Ecker et al., 2017). Females with ASD tend to present with more severe phenotypes, and higher likelihood of co-existing conditions such as intellectual disability (Werling & Geschwind, 2013). The 'Female Protective Effect' in ASD was put forth to explain the phenomenon, whereby females can tolerate the presence of more autism risk variants than boys, and would need a larger genetic burden or possibly environmental risk factors before presenting as affected with ASD (Sanders et al., 2011). Research in ASD genetic risk factors has found a higher overall burden, specifically of *de novo* CNV mutations, in females with ASD (Iossifov et al., 2014; Sanders et al., 2011; Sanders et al., 2015; De Rubeis et al., 2014). Given these findings, we decided to further explore sex differences in ASD by stratifying genotype-phenotype associations by sex.

We chose to utilize specific ASD phenotypes and genotypes in our study to explore areas these relative new areas in ASD research. Dymorphism in ASD is not yet fully understood, and there is no specific pathognomic dysmorphic feature associated with ASD. Unpublished work from the SEED population showed a higher likelihood of

dysmorphism in ASD children compared to typically developing children. In addition, there are indications emerging of specific dysmorphic features that are more likely to be associated with ASD, for example short stature and high BMI. Prior to this, research on genotype-phenotype associations for dysmorphology was not performed specifically for dysmorphism, but instead used dysmorphism as an additional condition to intellectual impairment or ASD itself. In addition, unlike in most previous studies, we were able to exclude children with developmental delay and chromosomal abnormalities and recognized genetic syndromes, enabling us to examine dysmophism more specifically. This work should thus allow for a more precise genotype-phenotype association between CNV and dysmorphology to be evaluated. Using various growth abnormalities as phenotypes, we checked across three growth modalities to examine differences in genotype-phenotype associations. Previous research in this field has generally tended to assess only individual modalities of growth.

Review of Specific Aims

In Specific Aim 1, we characterized growth abnormalities in a sample of ASD children compared to typically developing children from Phase 1 of the SEED Study. In addition, we assessed the combination of height, weight and head circumference simultaneously, which we term trivariate growth phenotype, to examine differences in growth symmetry in children with ASD compared to typically developing children. We found sex-specific differences in growth abnormalities among young children with ASD regardless of abnormality definition used. ASD girls had greater odds of short stature compared to

typically developing girls which achieved statistical signifcance, and a higher prevalence of the combined growth phenotype involving microcephaly and short stature, but normal weight.

In Specific Aim 2, we investigated the genotype-phenotype association between CNVs and dysmorphology by estimating the CNV burden measures (both combined counts and cumulative length of CNVs on all autosomes) comparing dysmorphic to non-dysmorphic children and the association between ASD candidate CNV regions and dysmorphism. We found genome-wide CNV burden was negatively associated with dysmorphology, even among ASD cases, and these associations varied by sex. Decreased CNV burden in dysmorphic SEED 1 children were observed for large duplication CNVs and those restricted to genic regions. None of the CNVs in ASD candidate genes revealed statistically significant associations with dysmorphism, but these CNVs observed were quite rare and, in the current study sample, these analyses were underpowered.

In Specific Aim 3, we assessed the genotype-phenotype association between CNVs and growth abnormalities by estimating the CNV burden comparing children with growth abnormality to normal children, and testing for association between ASD candidate CNV regions and abnormal growth. We observed decreased CNV burden associated with each growth modality, and found CNV burden was restricted to ASD candidate genes which were negatively associated with tall stature and macrocephaly, and again these associations varied by sex. Associations with specific CNVs in ASD candidate regions showed sex-specific results. We found there is potential shared genetic risk for ASD and

short stature at CNV region 15p11.2 that varied by sex, as the association was only significant in females.

These findings demonstrate the importance of taking into consideration the influence of sex in characterizing the ASD sub-phenotypes of growth abnormalities and dysmorphology, as well as the variability of genetic risk factors for ASD. To improve discovery of genotype-phenotype associations in ASD, it is imperative to include females in all ASD research. We found a negative association between genome-wide CNV burden and dysmorphology in the SEED 1 Study, an unexpected finding that may be due to: (1) the exclusion of children with chromosomal abnormalities and recognized genetic syndromes, (2) undetected single-gene insults among control samples, or (3) actual protective effects especially among females. We also found genome-wide CNV burden restricted to ASD candidate genes appear to be associated with specific growth abnormalities that vary by sex, suggesting a potential common pathway involving CNVs for ASD and short stature only in females. This potential shared genetic risk factor for ASD and growth abnormality may lead to possible future clinical application in the diagnosis and management of ASD but should be tailored to the child's sex.

Study Limitations

One key limitation is our reliance on a cross-sectional design; children were only assessed at one point in time close to study recruitment. This is of particular importance for Aims 1 and 3, where categorized growth abnormalities were used. In this study, children were only assessed once between the ages of 2 and 5 years. This likely does not impact the dysmorphology classification, which is expected to remain fairly consistent in early childhood. However, for specific growth abnormality determination, this is a notable limitation. Growth in children is a developmental process with sensitive periods and changing trajectories of growth based on age. To truly understand physical growth, sequential measurements over time is preferable, as a single reading may not be representative of the child's growth and cannot establish the child's growth trajectory. In ASD children in particular, growth trajectories and deviations from this is time-sensitive, with abnormal growth changes usually occurring from 6 months of life onwards and starting to decline in severity after approximately 2 years of age.

Thus, the age at which growth measurement was performed is integral to not just the results but also their interpretation. The absence of overgrowth in autistic boys may be associated with this limitation in our study design. We utilized growth measurements assessed at one point in time to determine growth abnormality, and we acknowledge this limits interpretability.

The relatively small sizes of subsamples for both dysmorphism and abnormal growth limited our ability to adjust for potential confounders beyond race/ethnic group and sex. The sample size was too small to use the criteria for growth abnormality defined by clinically-informed algorithm implemented in the SEED study (Dysmorphology Review Form) for CNV analyses, as this would have resulted in even smaller numbers of individuals categorized with growth abnormalities. We were also unable to distinguish between inherited and *de novo* CNVs. Published literature has shown the importance of *de novo* CNVs in children with dysmorphism (Vulto-van Silfhout et al., 2013) and growth abnormalities (Canton et al., 2014; Zahnleiter et al., 2013; van Duyvenvoorde et al., 2014). Rare CNVs have also been associated with congenital malformations (Serra-Juhe et al., 2012) and growth abnormalities (Dauber et al., 2011; Zahnleiter et al., 2013). Unfortunately, we did not have parental genotyping information and hence were not able to identify *de novo* CNVs. In addition, although the SEED Study itself is one of the largest population-based samples of ASD in the US, for the purpose of dysmorphology and growth abnormalities, this study is still underpowered to detect rare variants. Although the CNV burden measures used in this study reflect both common and rare CNVs, the limited analytic size likely hampered our ability to detect associations for rare variants. Finally, the size of the analytic sample also influenced our ability to obtain statistically significant results after correction for multiple testing. Although a large proportion of the results did not survive correction for multiple testing, these results may still point towards potential factors could be of clinical significance, and may eventually reach statistical significance if replicated using larger sample sizes.

Working within an ASD case-control study presented challenges as well. We had a large percentage of children with ASD in the sample (approximately 40-55%), yet we were examining relationships between CNV burden and dysmorphism or growth phenotypes, not just ASD. The interpretation of CNV burden for both dysmorphism and growth abnormalities were made within this context. In addition, the analytic sample obtained from SEED 1 had a multi-ethnic composition but is not fully representative of the United States in terms of racial make-up, as it oversampled certain minority ethnic groups (Schendel et al., 2012).

Another potential imitation is our use of a single CNV calling algorithm, the PennCNV algorithm. Comparisons of various CNV detection algorithms for using SNP data have shown that there is no gold standard for detection of CNVs (Winchester, Yau & Ragoussis, 2009). Previous studies recommend using more than one calling algorithm to improve specificity, as different CNV calling tools may lead to inconsistent results (Pinto et al., 2011; Winchester, Yau & Ragoussis, 2009). However, some researchers support the use of the PennCNV algorithm even on its own as this program has been observed to outperform other packages in sensitivity and specificity of CNV calling (Zhang et al., 2014).

We must also consider the limits of our dysmorphology evaluation, particularly the process of determining dysmorphic features using clinical photographs of children's body parts where there are missing data. This occurred despite various rigorous quality control measures to maintain the photographic quality of the photographs used for

dysmorphology assessment. However, unpublished sensitivity analyses and multiple imputations from the same analytic sample showed missing data had no significant effect on the observed results (Shapira, personal communication, SEED Dysmorphology Group unpublished manuscript). We also attempted to account for the high proportion of children with ASD in the analytic population (55%) by stratifying the results by ASD diagnosis, and found ASD diagnosis contributed to the association observed for the analysis using dysmorphology as a sub-phenotype.

Study Strengths

This study nonetheless has several important strengths. The first is the SEED Study itself. SEED 1 is the largest ASD case-control study with population-based ascertainment in the United States (Schendel et al., 2012). Multiple studies on dysmorphology and growth abnormalities have been on clinic-based samples without population-based controls. This was especially useful for Aim 1, when we evaluated growth abnormalities in children with ASD and had local controls in the analytic sample. The SEED Study also includes children from varied geographical locations and race/ethnic groups across the United States, representing diverse experiences in different regions and accounting for geographical factors that might affect the associations assessed.

Inclusion of ASD children with typically developing children in the SEED Study enriches the analytic sample with children who are likely to have a higher genetic burden, especially for ASD-related genes, and improve the likelihood of finding genetic

associations between our phenotypes of growth abnormalities and dysmorphology with ASD. Using the phenotypes of dysmorphology and growth abnormalities, which have both been linked with underlying genetic defects, also increases the likelihood of finding associations with genetic risk factors. Morphological abnormalities should reflect genetic abnormalities, and we did find significant results for CNVs in candidate genes associated with ASD and dysmorphology and growth abnormalities.

Another advantage of the SEED study is the number of females in the analytic sample, especially females with ASD. ASD is more prevalent in males, with the current ratio of male-to-female being 4-5 to 1 (Fombonne, 2009; Loomes et al., 2017). Several studies on children with ASD for growth abnormalities did not obtain adequate females with ASD to allow for stratification by sex, resulting in ASD females either being excluded from the study, or stratified analysis yielding no meaningful information on risk among females (Lai et al., 2017). Inclusion of autistic girls in studies on ASD is increasingly recognized as important, and may give insights about potential risk factors.

The SEED study had the advantage of dedicated medical geneticists focused on rigorous research assessment of dysmorphology. They developed a new tool to quantify dysmorphology as part of the SEED Study that use information from cases and population-based controls, blinding of the raters, and standardization of growth measurements across sites via Standard Operating Procedures (SOPs), as well as the implementation of repeated measurements to improve accuracy. Use of this new tool allowed quantification of dysmorphic features and a more robust measure of

dysmorphology. This also included multiple aspects of growth modalities. This allowed analysis of uni-dimensional and bi-dimensional growth abnormalities, as well as the trivariate growth phenotype. Previous studies often concentrated on just one aspect of growth abnormality, for example, macrocephaly, in children with ASD. We performed analyses of growth abnormalities in all modalities including BMI and the trivariate growth phenotype to give a comprehensive view of cross-sectional growth abnormalities in multiple dimensions. The trivariate growth phenotype is not an aspect of growth that has been considered in detail for ASD, and in fact, there is a dearth of literature considering growth asymmetry. Thus, we believe this is a novel way of considering growth in children with ASD, and may give a better perspective of overall physical growth in young children with autism spectrum.

Public Health Implications

ASD is a developmental condition of major public health import. The global prevalence of ASD has increased by almost forty-fold in the past five decades from a prevalence of one in 2,500 children (Gillberg & Wing, 1999) in the 1990s to one in 68 children in a recent report by the CDC (CDC, 2016). It is a source of enormous emotional burden to the families involved, a strain on resources for local communities and a substantial economic burden to society. It has been estimated that the lifetime individual ASD-related costs range between $1.5 to $2.5 million US dollars (Buescher et al., 2014). Addressing this condition of considerable public health significance is therefore timely and urgent.

The positive findings from this study indicate areas where screening, service provision and management of children with ASD could be enhanced, if our results are replicated. Improved recognition of growth abnormalities in ASD children, and use of genotyping as a tool to increase awareness and ASD detection in children with growth abnormalities and dysmorphology would be of significant clinical impact.

Based on our findings, establishing potential genotype-phenotype associations between ASD and growth abnormalities may improve risk factor identification in the subset of ASD individuals with abnormal growth. Recognizing the association of growth abnormalities with ASD could be utilized to increase awareness in providers of care for

ASD children to monitor growth and implement intervention for factors that could be ameliorated, for example dietary management.

The association between dysmorphism and ASD could potentially be developed into a diagnostic clinical tool for early intervention through laboratory-based as well as clinical methods (utilizing measures of dysmorphism). Understanding dysmorphic phenotypes may lead to more focused and potentially earlier provision of intervention in phenotypes such as abnormal growth and potentially to establishing better outcomes (Dawson, 2008; Boyd et al., 2010).

This study shows including autistic females in research, even if it may be limited by smaller samples, is important to identify possible sex-specific risk factors. Sex gives a unique perspective to understand the underlying etiologies in autism, and should become as core principle in autism research to further explore the heterogeneity of this neurodevelopmental condition (Rutter et al., 2003; Lai et al., 2015; Ecker et al., 2017). Understanding sexual dimorphism in ASD could also potentially lead to new and targeted treatment avenues.

Future Directions

Recent advances in research suggest that ASD presents differently in males and females. Researchers are seeking to understand the biological differences between ASD in males and females, as well as re-evaluating the effectiveness of diagnostic tools and treatments for ASD females. Our study shows outcomes that differ by sex, emphasizing the importance of including females in ASD research. Our finding of short stature limited to ASD females has rarely been reported, and thus a crucial future direction is replicating this finding in other studies. In addition, for assessment of ASD phenotypes, including ASD females would allow for stratified analysis for sex-specific risk factors.

Studying these sub-phenotypes of ASD may improve our biological understanding of this complex and heterogeneous disorder, particularly if different risk factor constellations are reflected in different phenotype presentations. This may be particularly helpful in parsing out different genetic risks. In addition, characterization of specific autistic phenotypes may aid in early diagnosis and prediction of outcomes (Walsh, 2011). For future work, one potential avenue for research on genotype-phenotype association is narrowing the phenotype from general dysmorphism to specific dysmorphic features not restricted to growth abnormalities, for example ear abnormalities.

Longitudinal measures of growth would be more representative of true growth trajectories and improve accuracy of categorizing growth abnormalities. For future work, it would be useful to expand this study to a longitudinal analysis, to further understand

developmental trajectories for somatic growth. It would also be beneficial to be able to incorporate covariates of both paternal and maternal growth measures, even just adult head circumference, height and weight. Adjusting for the anthropometric measures in both parents may allow greater insight into the genetic factors related to abnormal growth in ASD. Another avenue for future research is considering other covariates such as medication use (psychotropic medication) for specific ASD growth abnormality phenotypes such as obesity/overweight.

An important limitation in this study is the possibility there may be undiagnosed chromosomal abnormalities and genetic syndromes among our study sample. The information we utilized was based on parental report, and it is probable that additional chromosomal abnormalities or genetic syndromes may be present, and may have influenced associations detected here. This is an important issue that could be addressed for future work in this field. Genetic analyses incorporating full genome sequencing or at least whole exome sequencing on these subjects would enable improved delineation of underlying genetic risk factors in ASD, and in the case of CNV burden analysis, allow for exclusion of children with chromosomal abnormalities or other causes of increased genetic risk detected through sequencing.

Finally, the young age of the children recruited (2-5 years) for SEED restricts the generalizability of our results to only those in early childhood, as morphological changes occurring with age cannot be considered. Future work to expand the analysis on older

children could address this issue and allow assessment of developmental trajectories to be performed.

References

American Psychiatric Association's Diagnostic and Statistical Manual, Fifth Edition (DSM-5) http://www.dsm5.org

Buescher, A. V. S., Cidav, Z., Knapp, M., & Mandell, D. S. (2014). Costs of Autism Spectrum Disorders in the United Kingdom and the United States. *JAMA Pediatrics*, *168*(8), 721.

Canton, A. P. M., Costa, S. S., Rodrigues, T. C., Bertola, D. R., Malaquias, A. C., Correa, F. A., … Jorge, A. A. L. (2014). Genome-wide screening of copy number variants in children born small for gestational age reveals several candidate genes involved in growth pathways. *European Journal of Endocrinology*, *171*(2), 253–262.

The Centers for Disease Control and Prevention (CDC). (2016). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. Surveillance Summaries. (2016) 65(3); 1–23 http://www.cdc.gov/mmwr/volumes/65/ss/ss6503a1.htm

Dauber, A., Yu, Y., Turchin, M. C., Chiang, C. W., Meng, Y. A., Demerath, E. W., … Hirschhorn, J. N. (2011). Genome-wide association of copy-number variation reveals an association between short stature and the presence of low-frequency genomic deletions. *American Journal of Human Genetics*, *89*(6), 751–759.

De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Ercument Cicek, a., … Buxbaum, J. D. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*, 515(7526):209-15.

Dissanayake, C., Bui, Q. M., Huggins, R., & Loesch, D. Z. (2006). Growth in stature and head circumference in high-functioning autism and Asperger disorder during the first 3 years of life. *Development and Psychopathology*, 18(2), 381–93.

Ecker, C., Andrews, D. S., Gudbrandsen, C. M., Marquand, A. F., Ginestet, C. E., Daly, E. M., … Murphy, D. G. M. (2017). Association Between the Probability of Autism Spectrum Disorder and Normative Sex-Related Phenotypic Diversity in Brain Structure. *JAMA Psychiatry*, 74(4), 329.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatr Res*., 65: 591-598

Gillberg,C., Wing L. (1999). Autism: not an extremely rare disorder. *Acta Psychiatr Scand,* 1999;99:399–406

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, 515(7526), 216-221.

Lai, M. C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *Journal of the American Acad of Child and Adol Psych*, 54(1), 11–24.

Lai, M.-C., Lerch, J. P., Floris, D. L., Ruigrok, A. N.V., Pohl, A., Lombardo, M. V. and Baron-Cohen, S. (2017), Imaging sex/gender and autism in the brain: Etiological implications. *J Neuroscience Res*, 95: 380–397.

Malhotra, D., Sebat J. (2012). CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell*. 148(6):1223-41.

Merikangas, K.R., Calkins, M.E., Burstein, M., He, J.P., Chiavacci, R., Lateef, T., Ruparel, K., Gur, R.C., Lehner, T., Hakonarson, H., Gur, R.E. (2015) Comorbidity of physical and mental disorders in the neurodevelopmental genomics cohort study. *Pediatrics*, 135(4):e927–e938.

Miles, J. (2011). Autism spectrum disorders- A genetics review *Genetics In Medicine,* 13:4.

Mraz, K. D., Green, J., Dumont-mathieu, T., Makin, S., & Fein, D. (2007). Correlates of Head Circumference Growth, 7–9.

Muskens, J. B., Velders, F. P., & Staal, W. G. (2017). Medical comorbidities in children and adolescents with autism spectrum disorders and attention deficit hyperactivity disorders: a systematic review. *Eur Child & Adol Psych*.

Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., … Betancur, C. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, *466*(7304), 368–372.

Pinto D., Darvishi K., Shi X.H., Rajan D., Rigler D., Fitzgerald T., Lionel A.C., Thiruvahindrapuram B., MacDonald J.R., Mills R., et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nat. Biotechnol*. 2011;29:512–520.

Rutter, M., Caspi, A., Moffitt, T.E. (2003). Using sex differences in psychopathology to study causal mechanisms: unifying issues and research strategies. *J Child Psychol Psychiatry*. 44(8):1092–115.

Sanders, S., Ercan-sencicek, a G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D.,

& State, M. (2011). Multiple recurrent de novo copy number variations (CNVs), including duplications of the 7q11.23 Williams-Beuren syndrome region, are strongly associated with autism. *Neuron*, *70*(5), 863–885.

Sanders, S. J., He, X., Willsey, a. J., Ercan-Sencicek, a. G., Samocha, K. E., Cicek, a. E., … State, M. W. (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron*, 87(6), 1215–1233.

Schendel, D. E., Diguiseppi, C., Croen, L. a, Fallin, M. D., Reed, P. L., Schieve, L. a, … Yeargin-Allsopp, M. (2012). The Study to Explore Early Development (SEED): a multisite epidemiologic study of autism by the Centers for Autism and Developmental Disabilities Research and Epidemiology (CADDRE) network. *Journal of Autism and Developmental Disorders*, *42*(10), 2121–40.

Schieve, L.A., Gonzalez, V., Boulet, S.L., Visser, S.N., Rice, C.E., Van Naarden, B.K., Boyle, C.A. (2012) Concurrent medical conditions and health care use and needs among children with learning and behavioral developmental disabilities, National Health Interview Survey, 2006–2010. *Res Dev Disabil* 33(2):467–476.

Shapira, K., Tian, L., Aylsworth, A., Elias, E., Hoover-Fong, J., Meeks, N., Souders, M., Tsai, C., Alexander, A., Schieve, L. (2014). Development of a Novel Protocol for Characterizing Dysmorphology to Enhance the Phenotypic Classification of Autism Spectrum Disorders, presented at International Meeting for Autism Research, Atlanta, 2014. Georgia: INSAR.

Serra-Juhé, C., Martos-Moreno, G., Bou de Pieri, F., Flores, R., González, J. R., Rodríguez-Santiago, B., … Pérez-Jurado, L. A. (2017). Novel genes involved in severe early-onset obesity revealed by rare copy number and sequence variants. *PLoS Genetics*, *13*(5), 1–19.

Torrey, E. F., Dhavale, D., Lawlor, J. P., & Yolken, R. H. (2004). Autism and head circumference in the first year of life. Biological Psychiatry, 56(11), 892–894.

van Duyvenvoorde, H. A., Lui, J. C., Kant, S. G., Oostdijk, W., Gijsbers, A. C. J., Hoffer, M. J. V, Wit, J. M. (2014). Copy number variants in patients with short stature. *European Journal of Human Genetics : EJHG*, *22*(5), 602–9.

Vulto-van Silfhout, A. T., Hehir-Kwa, J. Y., van Bon, B. W. M., Schuurs-Hoeijmakers, J. H. M., Meader, S., Hellebrekers, C. J. M., … de Vries, B. B. a. (2013). Clinical Significance of De Novo and Inherited Copy-Number Variation. *Human Mutation*, 34, 1679–1687.

Winchester L., Yau C., Ragoussis J. Comparing CNV detection methods for SNP arrays. *Brief. Funct. Genomic Proteomic*. 2009;8:353–366.

Zahnleiter, D., Uebe, S., Ekici, A. B., Hoyer, J., Wiesener, A., Wieczorek, D., … Thiel,

C. T. (2013). Rare Copy Number Variants Are a Common Cause of Short Stature. *PLoS Genetics*, *9*(3), 1–11.

Zhang, X., Du, R., Li, S., Zhang, F., Jin, L., & Wang, H. (2014). Evaluation of copy number variation detection for a SNP array platform. *BMC Bioinformatics*, *15*(1), 50.

Zwaigenbaum, L., Young, G. S., Stone, W. L., Dobkins, K., Ozonoff, S., Brian, J., … Messinger, D. (2014). Early head growth in infants at risk of autism: A baby siblings research consortium study. Journal of the *American Acad of Child and Adol Psych*, 53(10), 1053–1062.

CURRICULUM VITAE

BIOGRAPHICAL

Contact Address:

53 Jalan PJU 1A/ 29

Ara Damansara

Petaling Jaya

Selangor 47301

MALAYSIA

Cellphone : +60-12-210-7263

E-mail: norazlinabouzid@yahoo.com, nkamaln1@jhu.edu,

norazlinabouzid3672@gmail.com

EDUCATION AND TRAINING

9/2013- 2018

Doctoral candidate in Genetic Epidemiology

Dept of Epidemiology

Johns Hopkins Bloomberg School of Public Health

Baltimore, USA

1/2000 – 8/2002

M.R.C.P.C.H. (UK) Membership of the Royal College of Paediatrics and Child Health

Royal College of Paediatrics and Child Health, United Kingdom

9/1992 – 7/1998

M.B.B.S. Bachelor of Medicine, Bachelor of Surgery

King's College, University of London (formerly known as United Medical and Dental Schools of Guys and St Thomas' Hospitals), United Kingdom

Distinctions awarded for MBBS Parts 4,5,7

9/1992 – 8/1995

B.Sc. in Neuroscience

King's College, University of London (formerly known as United Medical and Dental Schools of Guys and St Thomas' Hospitals), United Kingdom

Intercalated degree, second-class upper (2:1)

9/1990 – 6/1992

A-Levels

Roedean School, United Kingdom

3A(Maths, Biology,Chemistry), 1B(Physics)

CERTIFICATION AND LICENSURE

Annual Practising Certificate in Medicine from the MalaysianMedical Council (2004-present)

General Medical Council certificate of practice, United Kingdom (1998-2002)

Certification (Paediatrics) : Membership of the Royal College of Paediatrics and Child Health (2002)

ACADEMIC, ADMINISTRATIVE AND CLINICAL APPOINTMENTS

ACADEMIC AND ADMINISTRATIVE EMPLOYMENT HISTORY

1/2004-present

Paediatric Specialist, Department of Paediatrics, National University of Malaysia, Kuala Lumpur, Malaysia

1/2007-8/2013

Coordinator of Suspected Child Abuse and Neglect (SCAN) cases, National University of Malaysia, Kuala Lumpur, Malaysia

1/2012-8/2013

Co-coordinator for Year Four Medical Students Personal and Professional Development module

CLINICAL EMPLOYMENT AND EXPERIENCE

1/2004-present

Lecturer, Department of Paediatrics, National University of Malaysia, Kuala Lumpur, Malaysia

1/2008-present

Commenced Developmental Paediatrics experience and training at the Child Development Center, National University of Malaysia, Kuala Lumpur, Malaysia

3/2002- 9/2002

Paediatric Resident Medical Officer, St Mary's Hospital NHS Trust, London, United Kingdom

8/2001-2/2002

Paediatric Senior House Officer, St Mary's HospitalNHS Trust, London, United Kingdom

8/2000-7/2001

Paediatric Senior House Officer, Guy's HospitalNHS Trust, London, United Kingdom

8/1999- 7/2000

Paediatric Senior House Officer, Pembury HospitalNHS Trust, Pembury, United Kingdom

2/1999- 7/1999

Surgical House Officer, Eastbourne HospitalNHS Trust, Eastbourne, United Kingdom

8/1998- 2/1999

Medical House Officer, St Thomas' HospNHS Trust, London, United Kingdom

RESEARCH INTERESTS AND PROJECTS

Research Interests:

Autism Spectrum Disorders

Developmental Psychopathology

Child Maltreatment and Neglect

Research Projects: As Primary investigator:

- (2012) Level of knowledge regarding child safety at home among parents attending primary medical centre at UKMM(UniversitiKebangsaan Malaysia Medical Center)
- (2011) Knowledge, attitude and practices of paediatric medical officers regarding child abuse and neglect

As Joint investigator:

- (2012) School absenteeism among paediatric clinic patients in UKMMC
- (2012) Infant Feeding Practices in Wilayah Persekutuan Kuala Lumpur
- (2011) Teenage pregnancies: Exploring the issues in the light of prevention

- (2011) Screening for Autism Spectrum Disorder: Use of the Modified Checklist for Autism in Toddlers (M-CHAT)-Malay Version
- (2011) Use of Zolendronic Acid in the Treatment of Thalassemia Induced Osteoporosis
- (2011) Use of Parent Administered Picture Activity Cards to Facilitate Self Care Activities in Children with Autism
- (2011) Can Music Distraction Therapy Replace EMLA Cream for Pain Management in Infants Undergoing venepuncture?
- (2010) Retrospective Review of Factors Affecting Compliance to Methylphenidate amongst Children with Development Disorders at the Child Development Centre, UKMMC

HONORS AND AWARDS

Invited as a reviewer for an article by Pediatrics International (official journal of the Japan Pediatric Society) in February 2013

Exploratory Research Grant awarded by Ministry of Higher Education of Malaysia for research on teenage pregnancies in Malaysia (2012)

EACD Bursary awarded by the European Academy of Childhood Disabilities(EACD) to present poster at the 23rd EACD Annual General Meeting in Rome,Italy (2011)

EACD Bursary awarded by the European Academy of Childhood Disabilities(EACD) to present poster at the 22nd EACD Annual General Meeting in Brussels,Belgium (2010)

CFAK Senior Scholar Award awarded by the American Association of Child and Adolescent Award Psychiatrists and Allied Professionals to present Clinical Perspectives session "Crossing Borders to Protect Children of the World: International Perspectives on Child Abuse and Maltreatment" with international consortium (2010)

Awarded Distinctions in Medical School for Part 4 MBBS in Clinical Pharmacology and Therapeutics, Part 5 MBBS in Obstetrics and Gynaecology, Part 7 MBBS in Clinical Practice by the King's College, University of London (1996, 1997, 1998)

Awarded full scholarship by the Malaysian government for undergraduate and medical school educationin the United Kingdom (1990-1998)

MEMBERSHIPS IN PROFESSIONAL SOCIETIES AND ORGANIZATIONS

International Association of Child and Adolescent Psychiatrists and Allied Health (IACAPAP) (2010-2014)

International Society for the Prevention of Child Abuse and Neglect (IPSCAN) (2010-2014)

European Academy of Childhood Disability (EACD) (2010-2014)

Chapter of Child Neurology and Developmental Paediatrics -Malaysian Society of Neurosciences (CCNDP-MSN) (2010-present)

Malaysian Metabolic Society (2010-present)

DEPARTMENTAL AND UNIVERSITY COMMITTEES

National University of Malaysia

Member, National University of Malaysia Medical Library Committee (2006-present)

Member, One Stop Crisis Centre Committee, National University of Malaysia Medical Centre (01/2009-08/2013)

Committee, Personal Professional Development Module for Medical Students, National University of Malaysia Medical Centre (2012)

TEACHING

Undergraduate

Direct supervision: 30 medical students per year

Lecture/seminar-based teaching: 300 medical students per year,

120 allied professional students per year (Speech pathologists, Dietitians, Pharmacists, Social workers, Occupational therapists)

Postgraduate

Direct supervision: 10 Pediatric Mastersstudents per year

Lecture/seminar-based teaching: 30 Pediatrics Mastersstudents per year

Dissertation/Research project supervision

Main supervisor in 2 Master dissertations (2011, 2012)

Co-supervisor in 5 Master dissertations (2010, 2011, 2012)

Main supervisor in 1 medical undergraduate research project (2012)

Co-supervisor in 1 medical undergraduate research project (2012)

Courses

2009-2013

Learning disabilities and developmental delay (Paediatric Graduate Course)

2009-2013

Child Maltreatment (Family Medicine Graduate Course)

2009-2013

Child Maltreatment (Social Work Undergraduate Course)

Teaching Assistant (as Graduate student)

Fundamentals of Epidemiology Fall Term, 2014

COMMITTEE, ORGANIZATIONAL AND VOLUNTARY SERVICES

Permata Autisme committee member (2013)

Child Protection Committee for Federal Territories of Kuala Lumpur (2008-2014)

Treasurer of the Chapter of Child Neurology and Developmental Paediatrics -Malaysian Society of Neurosciences (CCNDP-MSN) (2010-2012)

Facilitator and presenter for Child Safety Child Friendly campaign of the National University of Malaysia Medical Center (2012)

Panel member for General Autism Forum, state of Terengganu, Malaysia (2012)

Panel member for forum on "General paediatric problems" for general public organized by Oakleaf Park housing association, Kuala Lumpur (2012)

Interview slot on "Autism" for 'Dari Fail Doktor' television program for Malaysian terrestrial TV, RTM TV2 (2012)

Panel member for Safe Motherhood Congress and Abandoned Babies Forum, Malaysian Maternal and Neonatal Health (2011)

Organizing committee for the 7th Malaysian Indonesia Brunei medical sciences conference (20011)

Interview on autism for 'Meet the Expert' program onterrestrial Malaysian radio (Radio IKIM) (2011)

Voluntary administration of vaccines under 'free vaccination' programs for welfare homes by Malaysian Paediatric Association, sponsored by pharmaceutical company – vaccination for children in three welfare homes (Ti Ratana, RumahSayang) (2011)

Lecture series for school teachers: (2011) *Learning Difficulties* at Sekolah Rendah Al-Amin primary school

(2010) *Autism and ADHD* at Sekolah Menengah Cheras secondary school

Main student organizer of 'Operation Zilina', Ophthalmic Aid to Eastern Europe- a project to transport and donate ophthalmic equipment from St Thomas' Hospital, United Kindom to Prague, Czech Republic, Zilina, Slovakia and Krakow, Poland (1994) Student Coordinator of European Medical Students' Association(EMSA) in University of London. Organised an exchange program with medical students in Moscow and taught 'Medical English' to Russian medical students in Moscow, Russia (1995). Also helped to organize and run EMSA summits in Zagreb, Croatia(1996) and Crete, Greece(1997)

SEMINARS AND INVITED PRESENTATIONS

PRESENTATIONS AT MEETINGS

1. Kamal Nor, N. (2012, September). *Developmental paediatric management of a child planned for cochlear implant*. Oral presentation at thePaediatric Neurology Updates, CCNDP-MSN meeting, Kota Bharu, Kelantan, Malaysia.
2. Kamal Nor, N., Tan,S.M.K., MohdIshak, N., Omar, K., Siraj, H. ,Wahab, S., Tohid, H., MohdTohit, N. (2012, July). *Teenage pregnancies in Malaysia: Exploring issues in ,the light of prevention.* Poster presented at the 20th World Congress of the

International Association of Child and Adolescent Psychiatrists and Allied Health, Paris, France.

3. Kamal Nor, N., Osman, R., Syed Zulkifli S.Z. (2012, May). *Level of knowledge of child abuse amongst paediatric medical doctors in Malaysia*. Poster presented at the 24th Annual meeting of the European Academy of Childhood Disability, Istanbul, Turkey.
4. Kamal Nor, N., Chandran, V.(2011, July). *Management of challenging behaviours in children with developmental disabilities*. Oral presentation at the PaediatricNeuro Updates, CCCNDP-MSN Meeting, Johor Baru, Malaysia.
5. Kamal Nor, N., Raja Lope,R.J., Tan, S.M.K.,Ismail, J., Chan, L.F., Loh, S. F.,Chandran, V.(2011, June).*Parental knowledge and perception of autism spectrum disorder in Malaysia.* Poster presented at the 23rd European Academy of childhood disability 2011 Annual General Meeting, Rome, Italy.Awarded EACD Bursary
6. Kamal Nor, N. (2011, Oct). *Assessment of child abuse in Malaysia.* Oral presentation, part of the 'Symposium on international perspectives of child abuse' presented at the American Academy of Child And Adolescent Psychiatry 57th Annual General Meeting, New York, USA. Awarded CFAK Senior Scholar Travel Grant
7. Kamal Nor, N., Tan, S.N.M.(2010, September). *Group therapy to reduce Bohsia.* Oral presentation at the 18th ISPCAN International Congress Honolulu, Hawaii, USA .

8. Kamal Nor, N., Raja Lope, R.J., Ismail, J.(2010, May). *Assessment of developmental scores of children with suspected child abuse and neglect in a university hospital in Malaysia.* Poster presentation at the22nd European Academy of Childhood Disability Annual General Meeting , Brussels, Belgium. Awarded EACD Bursary
9. Kamal Nor, N., Yang, W.W.. (2009, October). *The PPUKM experience of statutory rape cases.* Poster presentation at the 9th World Congress International Association for Adolescent Health and 31st Annual Congress of Malaysian Paediatric Association. Kuala Lumpur, Malaysia.
10. Wahab, S., Zakuan, E.K., Tan, S.M.K.,Loh, S. F., Kamal Nor, N. (2009, October). *When males become victims.*Poster presentation at the 9th World Congress International Association for Adolescent Health and 31st Annual Congress of Malaysian Paediatric Association. Kuala Lumpur, Malaysia.
11. Chan, L.F., Tan, S.M.K., Loh, S.F., KamalNor, N. (2009, October).*Sexual abuse by a traditional faith healer.* Poster presentation at the 9th World Congress International Association for Adolescent Health and 31st Annual Congress of Malaysian Paediatric Association. Kuala Lumpur, Malaysia.
12. Kamal Nor, N., Rohana, A.J., Shareena, I. (2006, Aug). *Bleeding Meckel`s diverticulum in a premature neonate.* Oral case presentation at the 1st Asean Congress of Paediatric Surgery and 28th Annual Congress of Malaysian Paediatric Association , Kuala Lumpur, Malaysia.

INVITED LECTURES

1. Kamal Nor, N. (2013, July). *Child Abuse and Neglect*. Invited speaker to symposium: Making our children safe. Family Medicine Scientific Conference 2013, Kuantan, Pahang, Malaysia.
2. Kamal Nor, N. (2013, June). *Supporting the Learning Process in Children*. Organised by PIBG of Al-Amin School, Kuala Lumpur, Malaysia.
3. Kamal Nor, N. (2012, September). *Living with autism*. General Forum on Autism organized by the state hospital of Terengganu, Terengganu, Malaysia.
4. Kamal Nor, N. (2012, June).*Assessment of the child with learning difficulties in a general paediatric clinic.* 5th National Paediatric Research Conference, Kuala Lumpur, Malaysia.
5. Kamal Nor, N. (2012, April).*Recognising stress in your child.* 'Understanding Children's hearts and minds: Emotional functioning and learning difficulties' forum organized by Vijayaratnam Foundation, Petaling Jaya, Malaysia.
6. Kamal Nor, N. (2012, March). *Enhancing Child Development*.'I Love Me' Health Conference, Kuala Lumpur
7. Kamal Nor, N. (2012, February). *Child maltreatment: Services available in Malaysia.* Women and Child Health Seminar, Department of Family Health, UniversitiKebangsaan Malaysia Medical Center, Kuala Lumpur, Malaysia.
8. Kamal Nor, N. (2011, December). *Child maltreatment*. Management of trauma cases workshop, Department of Medical Social Welfare, UniversitiKebangsaan Malaysia Medical Center, Kuala Lumpur, Malaysia.
9. Kamal Nor, N. (2011, June).*Autistic or not?*Symposium at the 33rd Annual Congress of the MalaysianPaediatric Association, Kota Bharu, Malaysia.

10. Kamal Nor, N. (2011, March). *The Role of the Suspected Child Abuse and Neglect (SCAN) team in UKMMC.* Seminar on Forensic Gynaecology, Dept of Obstetrics and Gynaecology, UniversitiKebangsaan Malaysia Medical Center, Kuala Lumpur, Malaysia.
11. Kamal Nor, N. (2010, Oct).*Physician-directed Developmental Screening.*Pre-Congress meeting of the 32nd Annual Congress of the Malaysian PaediatricAssociation , Kuala Lumpur, Malaysia.
12. Kamal Nor, N. (2010, July). *Abandoned babies in Malaysia.* MAMANEH(Malaysian Maternal and Neonatal Health) Annual General Meeting, Kuala Lumpur, Malaysia.

BIBLIOGRAPHY

ORIGINAL ARTICLES

1. Fong CY, Aye AM, Peyman M, Kamal N, Visvaraja S, Tajunisah I, Ong LC. (2013). Neonatal Herpes Simplex Virus Type-1 Central Nervous System Disease with Acute Retinal Necrosis. *Pediatr Infect Dis J. 2013 Dec 30*

2. Nik Ruzyanei Nik Jaafar, Mohammad Daud Tuti Iryani, Wan Ismail Wan Salwina, Abdul Rahman Fairuz Nazri, Nor Azlin Kamal, Reddy Jaya Prakash & Shamsul Azhar Shah. (2013) Externalizing and internalizing syndromes in

relation to school truancy among adolescents in high-risk urban schools. *Asia-Pacific Psychiatry.* 5(S1):27-34

3. N Kamal Nor, S Syed Zakaria, R Osman. (2012). Level of knowledge of child abuse in Malaysia amongst paediatric medical doctors. *Developmental Medicine & Child Neurology.* 54(s3):51

4. N. Kamal Nor, S. Tan, L.F. Chan, N. Mohd Ishak, S. Wahab, F. Loh Sit, N. Mohd Tohit, Z. Kalil, N.A.A.B. Wahab, H. Siraj. (2012). Teenage pregnancies: Exploring the issues in the light of prevention. *Neuropsychiatrie de l Enfance et de l Adolescence*. 60(5S):152

5. Chan, L.F., Tan, S.M.K.,Ang, J.K., Kamal Nor, N., Sharip, S. (2012). A Case of Sexual Abuse by a Traditional Faith Healer: Are There Potential Preventions? *Journal of Child Sexual Abuse.* 21:6, 613-620

6. Hussain, I., Kamal Nor, N., Sarvananthan, R., Ortiz, M.H. (2011). Keeping developmental on schedule: the role of screening tools. *Journal of Paediatrics, Obstetrics &Gynaecology.* 2011 Jan/Feb

7. Alias, H., MdZin, R., Abdul.Halim, A.R., Sharaf, I., Eguchi, M., ALatiff, Z., Kamal Nor,N., Rahman, J. and Hirokazu.(2011). Successful treatment of very large congenital infantile fibrosarcoma. *Pediatrics International.* 53(5):768-770

8. Zarina, A.L., Kamal Nor, N., Hamidah, A., Aziz,D.A.,SyedZulkifli, S.Z.(2010). Spectrum of infections in splenectomisedthalassaemia patients. *Medical Journal of Malaysia.* 65(4):284-286

9. Wahab, S., Zakuan, E.K., Tan, S.M.K.,Loh, S.F., Kamal Nor, N. (2010).When males rbecome victims - A case report.*ASEAN Journal of Psychiatry.*Vol. 11 (2) July – December 2010

10. Zarina, A. L., KamalNor,N., Jeevanan, J., Hamidah, A., Goh, B.S., Syed Zulkifli, S.Z., and Rahman J. (2010). Vincristine-induced vocal cord palsy case report and review of the literature. *Journal of Pediatric Hematology Oncology.* 32(5):407-410

11. Rohana, J., Kamal Nor, N., Sithasanan, N., Boey, C.C.M,,Shareena, I., Boo, N.Y., Thambidorai. (2008). Bleeding Meckel's diverticulum in a premature neonate *Journal of Neonatal-Perinatal Medicine.*1:119-21

BOOK CHAPTERS AND OTHER PUBLICATIONS

1. Tan, S.M.K., Kamal Nor, N., Loh, S.F., Wahab,S., Marimuthu, S., Chan, L.F. (2012).The developmental impact of early maltreatment and exposure to violence.

IACAPAP textbook of Child and Adolescent Mental health, Perinatal B.3. Ebook at http://iacapap.org/iacapap-textbook-of-child-and-adolescent-mental-health

2. Kamal Nor, N. (2010). Abandoned babies on the rise. *MAMANEH (Malaysian Maternal and Neonatal Health) bulletin article*