Investigating Clinically Relevant Phenotypes of Neurodevelopmental Copy Number Variants in Children and Adults

A thesis submitted for the degree of
Doctor of Philosophy

Dr. Claire Foley

Supervisors: Professor Louise Gallagher and
Professor Aiden Corvin

Department of Psychiatry
School of Medicine
Trinity College Dublin, the University of Dublin
2021

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work, with contributions from others duly acknowledged within the text.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Signed

Claire Foley

28th May 2020

Summary

Background:

Neurodevelopmental copy number variants (ND CNVs) present with clinically pleiotropic effects, transcending psychiatric diagnostic borders. Gaining clarity on the neurodevelopmental and psychiatric phenotypic effects of ND CNVs in different populations is essential to providing clinically relevant information and services to carriers.

Aims and structure of thesis:

The central aim of this thesis was to investigate the relationship between psychiatric and neurodevelopmental phenotypes and ND CNV carrier status to inform our understanding of ND CNV phenotypes and provide clinically translatable insights. Existing datasets were leveraged for secondary data analysis of large cohorts of previously understudied populations.

Chapter 1 discusses findings in the literature supporting the exploration of psychiatric and neurodevelopmental phenotypes associated with ND CNVs in relevant populations.

Chapter 2 describes an analysis of psychiatric risk outcomes associated with ND CNV carrier status in a cohort of youths with autism spectrum disorder (ASD) and explores the moderating effect of sex on specific outcomes. A follow up analysis is conducted in a sample of ASD-unaffected siblings, to explore identified associations between ND CNV carriers and psychopathology in the absence of ASD.

Chapter 3 presents an investigation of psychiatric risk outcomes associated with ND CNVs in a large clinical cohort of youths and examines for a moderating effect of sex in specific outcomes. The effect of neurodevelopmental disorder comorbidity in the relationship between ND CNVs and psychopathology is explored for relevant outcomes.

Chapter 4 describes an analysis investigating the utility of psychiatric and neurodevelopmental phenotypic features in modelling ND CNV carrier status in individuals with schizophrenia. Data are presented from a discovery and replication cohort.

Finally, Chapter 5 presents a general discussion of the findings.

Results:

The analysis of association between psychiatric phenotypes and ND CNVs in youths with ASD revealed an interaction between sex and ND CNV status on risk of affective problems. Females with an ND CNV were significantly more likely to present with depressive symptoms than males with or

without an ND CNV. ND CNV status was significantly associated with increased risk of affective problems in the sample of typically developing siblings; this effect was driven by female carriers of ND CNVs only.

The analysis of the association between psychiatric symptoms and ND CNVs in a large clinical cohort of youths showed an increased risk of subclinical psychotic symptoms in youths with ND CNVs. This association was also present when individuals with an ASD diagnosis or significant cognitive deficits were excluded, indicating that the association was independent of neurodevelopmental comorbidity in this sample.

In adults with schizophrenia three phenotypic variables (specific learning disorder, developmental delay, and comorbid neurodevelopmental disorder) were significant in modelling positive carrier status for a schizophrenia associated CNV. Replication analysis in a separate cohort confirmed developmental delay and comorbid neurodevelopmental disorder as significant predictors of positive carrier status for schizophrenia associated CNV status.

Conclusion:

The results from this work provide novel, clinically relevant insights into neurodevelopmental and psychiatric associations with ND CNVs in previously understudied populations. An improved understanding of phenotypic associations with ND CNVs can inform genetic testing and counselling and may help with monitoring risks, prevention strategies and early intervention for ND CNV carriers.

Contributions to this Work

All studies were designed in collaboration with my supervisors Professor Louise Gallagher and Professor Aiden Corvin. I was responsible for applying for and processing secondary phenotype datasets, and for the data analysis and interpretation of the work described in this thesis under the supervision of Prof Gallagher and Prof Corvin. Dr Eleisa Heron provided consultation on statistical methods for the studies.

For the study described in chapter 2, genetic data was publicly available and is referenced in the chapter. Phenotype data was obtained for the dataset through the Simons Foundation Autism Research Initiative (SFARI) and is referenced in the chapter.

For the study described in chapter 3, genetic data was collected through the Center for Applied Genomics, Children's Hospital of Philadelphia, and was analysed and provided by Dr Joseph Glessner and Professor Hakon Hakonarson (Center for Applied Genomics, Children's Hospital of Philadelphia). Phenotype data was obtained for the dataset through dbGaP and is referenced in the chapter.

For the study described in chapter 4, students Amy Durand, Razaq Durodoye and Lindsey White contributed to the data entry of the phenotype data from existing research cohorts for the discovery dataset. Ms Christina Mooney, Dr Eric Kelleher, Dr Denise Harold, Dr Derek Morris, Dr Paul Cormican, Dr Carlos Pinto collected and processed genetic and clinical research data from the pre-existing Irish discovery dataset while working in the Neuropsychiatric Genetics Group in TCD. Prof Gary Donohoe, Prof Michael Gill and Prof Aiden Corvin were principal investigators on the projects that originally collected data for the discovery dataset. Prof James Walters, Prof Michael Owen and Prof Michael O'Donovan contributed the data for the replication dataset and provided feedback with regard to the data analysis and interpretation of that dataset. Prof Jonathan Sebat contributed to the core analysis of the genetic data used in the study.

Acknowledgements

I would like to express my profound respect and gratitude to my supervisors Professor Louise Gallagher and Professor Aiden Corvin. The support, guidance, insights and encouragement that you both provided throughout this work have been incredible and I sincerely appreciate all of the time and thought you have both given to this work. Special thanks also to Dr Eleisa Heron- you have been so generous with your time and insightful reflections, and you contributed hugely to my enjoyment of this project. Thank you so much to Professor Gill for giving me the initial opportunity to work on this project.

A massive thank you to everyone in the Neuropsychiatric Genetics department, you are all incredible people and I am so lucky to have met and worked with all of you. Any time I looked for help or advice you went above and beyond to provide assistance, encouragement and insights and I have been blown away by your generosity of spirit and time.

Thank you so much to everyone who contributed their data to the datasets that I worked on. I hope that the work in this thesis benefits you in some way. Thank you to all those who created the datasets I worked on and spent time collecting and working on the data before I got my hands on it, I am so grateful. A huge thank you to everyone named in the contributions section, I am forever indebted to all of you.

Thank you in advance to the examiners who read this thesis for your time, thoughts and expertise.

To my friends from Trinity, UCD, St Kilian's and everywhere else along the way, thank you for being such legends. To all the colleagues and patients I have worked with in different psychiatry posts, thank you for teaching me so much.

I can't thank my family enough. Thank you, for everything. "Hello?"

It's hard to put into words my gratitude to all of you and I hope that I have made it known to you along the way what it has all meant to me. Thank you so much and this work is dedicated to all of you.

Related Publications and Presentations

Journal Articles

Foley, C., Corvin, A., & Nakagome, S. (2017). Genetics of Schizophrenia: Ready to Translate?. Current psychiatry reports, 19(9), 61. https://doi.org/10.1007/s11920-017-0807-5

Foley, C., Heron, E. A., Harold, D., Walters, J., Owen, M., O'Donovan, M., Sebat, J., Kelleher, E., Mooney, C., Durand, A., Pinto, C., Cormican, P., Morris, D., Donohoe, G., Gill, M., Gallagher, L., & Corvin, A. (2020). Identifying schizophrenia patients who carry pathogenic genetic copy number variants using standard clinical assessment: retrospective cohort study. The British journal of psychiatry : the journal of mental science, 216(5), 275–279. https://doi.org/10.1192/bjp.2019.262

Conference Presentations

Claire Foley, Eleisa Heron, James Walters, Louise Gallagher, Aiden Corvin. Clinical Predictors of Pathogenic Copy Number Variants in a Psychosis Population. Oral presentation at Royal Academy of Medicine in Ireland, Psychiatry Section Higher Specialist Training Competition, Dublin, May 2018.

Claire Foley, Eleisa Heron, James Walters, Louise Gallagher, Aiden Corvin. Identification of Phenotypic Predictors of Pathogenic Copy Number Variants in Psychosis Populations. Poster presentation at World Congress of Psychiatric Genetics, Orlando, Florida, October 2017.

Claire Foley, Eleisa Heron, Louise Gallagher, Aiden Corvin. Identification of Phenotypic Predictors of Pathogenic Copy Number Variants in a Psychosis Population. Poster presentation at European College of Neuropsychopharmacology Congress, Paris, France, September 2017.

Claire Foley, Eleisa Heron, Louise Gallagher, Aiden Corvin. Can we identify the subgroup of schizophrenia patients most likely to benefit from clinical genetic testing? Poster presentation at College of Psychiatrists of Ireland, Spring Conference, 2017

List of Tables

List of Figures

List of Abbreviations

ACI	Autism Comorbidity Interview
ADHD	Attention Deficit Hyperactivity Disorder
ADI-R	Autism Diagnostic Interview-Revised
ADOS	Autism Diagnostic Observation Schedule
EU-AIMS	European Autism Interventions - A Multicentre Study for Developing New Medications
ALSPAC	Avon Longitudinal Study of Parents and Children
ASD	Autism Spectrum Disorder
AUC	Area Under the Curve
AUROC	Area Under the Receiver Operating Characteristic Curve
BP	Breakpoint
CBCL	Child Behaviour Checklist
CD	Conduct Disorder
CNDD	Comorbid Neurodevelopmental Disorder
CHOP	Children's Hospital of Philadelphia
CI	Confidence Interval
CIDI	Composite International Diagnostic Interview
CMA	Chromosomal Microarray Analysis
CNB	Computerized Neurocognitive Battery
CNV	Copy Number Variant
COST	European Cooperation in Science and Technology
DAS	Differential Ability Scale
DECIPHER	DatabasE of genomiC varIation and Phenotype in Humans using Ensembl Resources

DNA	Deoxyribonucleic acid
DSM	Diagnostic and Statistical Manual of Mental Disorders
ECHO	Cardiff University Experiences of Children With Copy Number Variants Study
EHR	Electronic health record
EMR	Electronic Medical Records
FDR	false discovery rate
GABA	Gamma aminobutyric acid
GOASSESS	Computerised structured interview based on Kiddie-Schedule for Affective Disorders and Schizophrenia
GP	General Practitioner
GWA	Genome Wide Association
GWAS	Genome Wide Association Study
HDD	History of developmental delay
HLD	History of learning difficulties
ICD	International Classification of Diseases
ID	Intellectual Disability
IMAGINE ID	Intellectual disability and mental health: assessing the genomic impact on neurodevelopment study
IQ	Intelligence Quotient
IQR	Interquartile range
K-SADS	Kiddie-Schedule for Affective Disorders and Schizophrenia
LEAP	Longitudinal European Autism Project
LOF	Loss of Function
MAF	Minor allele frequency
MINDDS	Maximising Impact of research in Neurodevelopmental Disorders

NAHR	Nonallelic homologous recombination
NDD	Neurodevelopmental Disorder
NIMH	National Institute of Mental Health
NPV	Negative predictive value
NRXN	Neurexin
NVIQ	Nonverbal Intelligence Quotient
ODD	Oppositional Defiant Disorder
OR	Odds Ratio
PFC	Prefrontal Cortex
PGC	Psychiatric Genomics Consortium
PNC	Philadelphia Neurodevelopmental Cohort
PPV	Positive Predictive Value
PRIME	Prevention through Risk Identification, Management, and Education (PRIME) screen questionnaire for subclinical psychotic symptoms
PRS	Polygenic Risk Score
PS-R	PRIME screen revised
ROC	Receiver operating characteristic curve
SCAN	Schedules for Clinical Assessment in Neuropsychiatry
SCID	Structured Clinical Interview for DSM-IV-TR Axis I Disorders
SCZ	Schizophrenia
SD	Standard deviation
SFARI	Simons Foundation Autism Research Initiative
SI	Suicidal Ideation
SIPS	Structured Interview for Prodromal Symptoms

SLD	Specific learning disorder
SNP	Single nucleotide polymorphism
SOPS	Scale of Prodromal Symptoms
SSC	Simons Simplex Collection
TRF	Teachers Report Form
UCSC	University of California Santa Cruz
VABS	Vineland Adaptive Behaviour Scale
VCFS	Velo-Cardio-Facial Syndrome
VOUS	Variant of Unknown Significance
WASI	Wechsler Abbreviated Scale of Intelligence
WHO	World Health Organisation
WISC	Wechsler Intelligence Scale for Children
WRAT	Wide Range Achievement Test
YSR	Youth Self Report

Contents

Chapter 1: Introduction

Parts of this chapter have been adapted from the following article:

Foley, C., Corvin, A., & Nakagome, S. (2017). Genetics of schizophrenia: ready to translate?. Current psychiatry reports, 19(9), 61.

1.1 Background

Psychiatric and neurodevelopmental disorders (NDDs) account for nearly a third of disability burden worldwide in adults (Anderson et al., 2011; Organization, 2008; Sullivan et al., 2012a) and also contribute substantially to global disability burden in youths (Baranne et al., 2018; Erskine et al., 2015). The functional limitations imposed by psychiatric disorders and NDDs have implications for the well-being of affected individuals, their families, and society. Individuals with psychiatric and developmental impairment require increased supports of health, social, educational and occupational services (Arim et al., 2017; Boyle et al., 2011; Jensen et al., 2011; Ochoa et al., 2003).

Schizophrenia, autism spectrum disorder (ASD) and intellectual disability (ID) are NDDs that each present with considerable clinical heterogeneity and psychiatric and neurodevelopmental comorbidity. They have complex aetiologies with substantial genetic heterogeneity (Bass et al., 2018; Weiner et al., 2017; Woodbury-Smith et al., 2018) .

Genetic overlap in the form of shared genetic risk factors has also been identified between these disorders (Anttila et al., 2018; Clarke et al., 2016; Lee et al., 2019; Smeland et al., 2019). A major breakthrough in psychiatric genetics in recent times has been the discovery that structural genomic variants termed copy number variants (CNVs) contribute substantially to risk of ASD, ID and schizophrenia in subsets of individuals with the disorders and that many of the same CNVs are shared as risk factors across the three disorders (Coe et al., 2014; Sanders et al., 2015). There is evidence that these neurodevelopmental CNVs (ND CNVs) may contribute to clinical heterogeneity within disorders (Bishop et al., 2017; Lowther et al., 2017; Marshall et al., 2017) and there is now accumulating evidence of association between ND CNVs and a broader range of psychiatric and cognitive phenotypes (Chawner et al., 2019; Hanson et al., 2015; Kendall et al., 2019; Stefansson et al., 2014), indicating a complex pattern of risk conferred by these CNVs. Identification of the psychiatric and neurodevelopmental phenotypes associated with ND CNVs can contribute timely clinical information that can directly impact individuals who carry these CNVs and may be helpful in understanding the complex and overlapping clinical presentations of NDDs and psychiatric disorders.

1.2 Neurodevelopmental disorders

1.2.1 A Note on Terminology

There is considerable variation in the literature with regard to the disorders that the term "neurodevelopmental disorder" refers to. The term is non-specific and can refer to a wide range of presentations but is generally accepted to refer to conditions characterised by some form of disruption of brain development (Thapar et al., 2017). Childhood onset disorders such as ASD and ID are widely

referred to as neurodevelopmental disorders (American Psychiatric Association, 2013). There is substantial evidence that schizophrenia also has a neurodevelopmental basis (Dazzan et al., 2002; MacCabe, 2008; Malaspina et al., 2001; Owen et al., 2017) and schizophrenia is often referred to as an NDD (Chawner et al., 2019; Grayton et al., 2012; Kendall et al., 2020; Merikangas et al., 2009a; Rund, 2018; Thygesen et al., 2018; Wolfe et al., 2017). When referring to ASD, ID and schizophrenia collectively in this thesis, I use the term "neurodevelopmental disorder".

There is also no standardised definition of the term neurodevelopmental CNVs and no definitive list of such CNVs. In the main discussion of this thesis, the term ND CNVs will be used as it is generally used in published literature- referring to CNVs that have shown robust association with ASD, ID and/or schizophrenia (Chawner et al., 2019; Kendall et al., 2019; Kendall et al., 2016; Martin et al., 2019; Stefansson et al., 2014). Where the term is used in reference to an analysis undertaken in the thesis, a specific list of CNVs is specified and presented in the methods section of the relevant chapter.

1.3 Phenotypic heterogeneity, overlap and comorbidity in neurodevelopmental disorders

ASD, ID and schizophrenia show substantial clinical heterogeneity in their presentations. Phenotypic overlap occurs between the disorders and comorbidity between these conditions and other psychiatric disorders is common.

1.3.1 Autism Spectrum Disorder

Autism spectrum disorder is characterised by persistent deficits in social communication and social interaction across multiple contexts and restricted, repetitive patterns of behaviour, interests, or activities with associated clinically significant impairments in social, occupational, or other important areas of functioning. There is marked heterogeneity in the clinical presentation of ASD with variation between individuals in severity of core symptoms, level of intellectual function, adaptive behaviour and functional communication ability (Charman et al., 2017).

The estimated prevalence rates of comorbid psychiatric disorders in people with ASD have varied from ~55-94% (Lugo-Marín et al., 2019; Mukaddes et al., 2010; Simonoff et al., 2008). These rates vary significantly on the basis of differences in methodology, measures and sample selection (age ranges, cognitive levels, sex distributions) and ascertainment. Individuals with ASD are more likely to have psychiatric symptomatology than neurotypical populations (Abdallah et al., 2011; Lugo-Marín et al., 2019; Skokauskas et al., 2012a).

There is wide variability in the rates of reported diagnosis of schizophrenia in autism. Low prevalence rates of schizophrenia have been reported in some studies (<0.4%) (Levy et al., 2010; Leyfer et al., 2006), while others have reported much higher prevalence rates- up to 13% (Joshi et al., 2013). The marked variation may be due to differences in studies in terms of sample size, sampling methods, subjects' demographics and diagnostic criteria utilised (Skokauskas et al., 2010).

Epidemiological studies of rates of comorbid ID in autism spectrum disorders also vary widely depending on sampling methods, demographics of participants and assessment instruments used. It was previously thought that a majority of individuals with ASD (up to ~75%) presented with comorbid ID (Cohen et al., 1997; Volkmar et al., 2004), however, more recent evidence suggests that only ~50% of individuals with ASD present with ID (Chakrabarti et al., 2005; Charman et al., 2011).

Systematic review has identified rates of suicidality to range between 11–66% in individuals with ASD (adult and child) samples (Hedley et al., 2018). Rates of suicidal ideation in adults with ASD have been reported to be significantly higher than both general adult populations and patients with psychosis (Cassidy et al., 2014; Cassidy et al., 2017), and there is an increased risk of premature death by suicide in people with ASD compared to the general population (Hedley et al., 2018; Hirvikoski et al., 2016; Kirby et al., 2019). There is accumulating evidence of increased prevalence of suicidal ideation and behaviour in children and adolescents with ASD compared with typically developing populations (Aitken et al., 2016; Chen et al., 2017; Hunsche et al., 2020).

Identifying psychiatric comorbidities and differentiating comorbid psychiatric symptoms from core features of ASD can be clinically challenging. Impairments in communication are a core feature of ASD (Leyfer et al., 2006). Children and adolescents with ASD also have impairments in complex information processing, theory of mind and executive functioning (Leyfer et al., 2006). These impairments culminate in significant difficulties for many children and adolescents with ASD being able to process and describe their experiences, emotions and mental states (Leyfer et al., 2006; Mazzone et al., 2012), presenting challenges for clinicians in identifying psychiatric symptomatology experienced by the young person.

A further challenge is differentiating psychiatric symptomatology from the core features of ASD. Symptoms such as repetitive behaviour, rituals, social withdrawal, avoidance of social situations, poor motivation are frequently seen in young people with psychiatric disorders such as obsessive compulsive disorder, social anxiety and depression (Cholemkery et al., 2014). For young people with ASD presenting with these symptoms, it may be unclear whether they are presenting with co-occurring separate disorders or whether the symptoms are a manifestation of core ASD symptoms.

1.3.2 Intellectual Disability

Intellectual disability is defined as "a condition of arrested or incomplete development of the mind, which is especially characterized by impairment of skills manifested during the developmental period, which contribute to the overall level of intelligence" (Maulik et al., 2011; Organization, 1988). Meta-analysis of population-based studies from across the globe indicate that the prevalence of intellectual disability is around 1% (Linehan, 2019; Maulik et al., 2011). Intellectual disability is generally classified as mild, moderate, severe or profound (Harris, 2006; Maulik et al., 2011). The clinical manifestations, and cognitive and adaptive abilities vary widely between individuals with ID.

Reported rates of psychiatric comorbidity in children and adolescents with ID range between 30-50% (Einfeld et al., 2011; Emerson et al., 2007). The relative risk of psychiatric disorder ranges from 2.8-4.5 in children and adolescents with ID compared with those without (Einfeld et al., 2011). Comorbid rates of psychopathology are also high in adults with ID (30-40%) (Cooper et al., 2007; Morgan et al., 2008) and they are at higher risk of psychiatric disorders compared with individuals without ID (Cooper et al., 2001).

Features of ASD present more commonly in individuals with ID than those without. Reported prevalence rates of ASD in children with ID vary between 4-28% (Bryson et al., 2008; Gillberg et al., 1986; Nordin et al., 1996; Tonnsen et al., 2016). A large population-based study estimated ASD to present in 18% of children with ID (Tonnsen et al., 2016)

Schizophrenia is also over-represented in people with ID. The prevalence of psychosis spectrum disorders in individuals with ID is estimated at approximately 3.5-5% (Aman et al., 2016; Morgan et al., 2008; Sheehan et al., 2015). A meta-analysis indicated that prevalence rates of schizophrenia decreased as level of severity of intellectual disability increased- prevalence rates in mild, moderate and severe ID were estimated at 5.55%, 4.21% and 0.89% respectively (Aman et al., 2016).

1.3.3 Schizophrenia

Schizophrenia is a heterogeneous clinical syndrome characterised by pronounced disturbances of thought, emotion, perception and behaviour. Schizophrenia affects ~0.5-1% of populations worldwide (Fischer et al., 2017; McGrath et al., 2008) and the WHO classifies the disorder within the top 10 leading causes of disability globally (Murray et al., 1996). Schizophrenia is associated with significant morbidity and premature mortality (by 10-20 years) (Tiihonen et al., 2009).

A systematic review by Kincaid et al. reported rates of between 9.6-61% for autistic symptoms presenting in outpatient and inpatient populations with schizophrenia, with rates of diagnosed ASD ranging from <1% to 52% (Kincaid et al., 2017).

As a group, individuals with schizophrenia have been identified premorbidly to have mean IQ scores approximately one-half of a standard deviation below mean scores of healthy comparison subjects (Woodberry et al., 2008). Risk of schizophrenia is also reported to be elevated in individuals with ID, with estimates of ~3-5% (Cooper et al., 2007; Hemmings, 2006; Morgan et al., 2008; Turner, 1989), compared with the lifetime risk in the general population of ~1%. Psychiatric comorbidities including anxiety, depression and substance abuse disorders frequently present with schizophrenia (Buckley et al., 2009).

1.4 Copy number variants

The clinical heterogeneity observed in NDDs and frequency of phenotypic overlap between NDDs suggest that complex aetiological processes underlie their presentations and that there may be shared aetiological risk factors between disorders. The identification of CNVs as important risk factors contributing to the pathogenesis of NDDs and possibly to other psychiatric disorders is an important development in psychiatric genetics.

CNVs are stretches of DNA larger than 1,000 base pairs (bp) for which copy number differences have been observed in comparing two or more genomes. These structural variants can be copy number gains (duplications) or losses (deletions) (Scherer et al., 2007). The mechanisms resulting in copy number variation are still uncertain. One important identified mechanism responsible for CNV formation is nonallelic homologous recombination (NAHR). This mechanism was identified in line with the observation that copy number variation frequently occurs in close proximity or within duplicated sequences of DNA. During meiosis, parental chromosomes generally align, using similar stretches of DNA as indicators guiding where to pair and begin recombination. In regions of duplicated sequences, recombination mechanisms can begin crossover events in erroneously identified regions. As a result, copies of the duplicated sequence can be gained or lost. These recombination variations tend to occur in regions containing long sequences of repeats, with these "recombination hotspots" showing higher rates of chromosomal rearrangement and resulting in recurrent but rare copy number variants (Eichler, 2008; Lupski, 1998; Sharp et al., 2005).

Current estimates indicate that just under 10% of the human genome is subject to copy number variation (~7.5% losses, ~3.9% gains) (Zarrei et al., 2015). CNVs may be inherited or arise *de novo*. Evidence suggests that over 99% of CNVs are inherited, with the remainder generated *de novo* (McCarroll et al., 2008; van Ommen, 2005). CNVs can disrupt gene sequence and/or alter gene dosage directly altering phenotype, or they can change expression patterns of other genes by altering regulatory regions (Freeman et al., 2006; Lupski et al., 1992; McCarroll et al., 2006; Nowakowska, 2017). Effects of copy number variation range from adaptive traits to pathological presentations to lethal effects in utero (Zarrei et al., 2015). Copy number variants have been associated with a wide

range of morbidities including congenital cardiac disease (Edwards et al., 2016), human immunodeficiency virus (HIV) (Gonzalez et al., 2005), autoimmune dysfunction (Aitman et al., 2006; Willcocks et al., 2008), malignancies (Shao et al., 2019; Shlien et al., 2009) and are important pathogenic risk factors in neurodevelopmental and neuropsychiatric disorders (Cooper et al., 2011; Kirov et al., 2009a; Merikangas et al., 2009a; Rees et al., 2016a; Sanders et al., 2015; Torres et al., 2016).

1.4.1 Copy number variants in neurodevelopmental disorders

1.4.1.1 Autism Spectrum Disorder

The aetiology of ASD is likely to involve both genetic and environmental contributors. To date, environmental risk factors for ASD have not been extensively explored or characterised (Modabbernia et al., 2017), with a small number of identified risk factors such as exposure to sodium valproate in the prenatal period having a likelihood of contribution to the risk of developing ASD (Christensen et al., 2013; Ramaswami et al., 2018). The significant genetic contribution to ASD is indicated by high heritability estimates for the disorder. Early twin studies estimated ASD heritability to be as high as 90% (Tick et al., 2016). Some population-based studies have produced lower heritability estimates at 50-60% (Sandin et al., 2014), but most estimates are in the region of 70-80% (Colvert et al., 2015; Hallmayer et al., 2011; Ramaswami et al., 2018; Sandin et al., 2017). Concordance rates estimates for monozygotic twins, dizygotic twins and siblings vary significantly at 30-99%, 0-65%, and 3-30% respectively (Hallmayer et al., 2011; Nordenbæk et al., 2014; Ramaswami et al., 2018; Tick et al., 2016).

Both common and rare genetic variants have been identified to contribute to the risk of developing autism (Anney et al., 2017; Weiner et al., 2017; Yuen et al., 2017) and genetic susceptibility varies between individuals (Leblond et al., 2019). Common variants have been estimated to contribute a large proportion of genetic risk for ASD, accounting for 40-60% of the overall liability (Gaugler et al., 2014; Ramaswami et al., 2018). Rare variants are likely to explain a lower proportion of collective heritability in ASD than common variants but have larger individual effects. Between 10-25% of individuals with ASD carry a single genetic mutation (chromosomal rearrangements, CNVs, small insertions and deletions (indels), single nucleotide variants) that is likely to be the pathological factor accounting for development of the disorder (Bourgeron, 2015). Copy number variants are the most prevalent type of rare variants contributing to ASD (Geschwind, 2011).

1.4.1.1.1 Copy number variation

Higher rates of copy number variation have been identified in patients with ASD as compared with controls. The frequency of clinically detectable CNVs in individuals with ASD is estimated at 8-21%, with ~10% of detected CNVs identified as likely pathogenic (Schaefer et al., 2013; Sebat et al.,

2007; Vicari et al., 2019). *De novo* CNVs present in 4-10% of individuals with ASD (Gilman et al., 2011; Marshall et al., 2008; Sebat et al., 2007; Vicari et al., 2019), a higher rate compared with unaffected siblings (~1-2%) (Bourgeron, 2016; Gilman et al., 2011). An enrichment for rare inherited CNVs has also been identified in ASD probands compared with unaffected siblings (Krumm et al., 2015; Leppa et al., 2016). Individuals with ASD also appear to carry a higher burden of gene-disruptive CNVs (Krumm et al., 2015; Turner et al., 2017).

Pinto et al. assessed the genome-wide characteristics of rare CNVs in individuals with ASD. An increased burden of rare genic CNVs (OR 1.19, $p = 0.012$) was identified in ASD probands. The effect was noted to be more significant in the case of loci that were previously associated with ASD and/or ID where 7.6% of individuals with ASD had rare CNVs preferentially affecting ASD/ID genes compared to 4.5% in controls (Pinto et al., 2010). In 2014, Pinto et al. expanded this study, confirming the excess of genic CNVs in those with ASD as compared with controls and the increased risk of ASD in those carrying exonic pathogenic CNVs which overlapped with known ASD and ID loci (Pinto et al., 2014).

An analysis of rare CNV in the Simons Simplex Collection (the SSC- a collection of phenotype and genotype data on families with one ASD affected child) identified 83 rare *de novo* CNVs in probands and unaffected siblings; these rare *de novo* CNVs were significantly more common in the probands (5.8% had at least one rare *de novo* CNV) compared with the siblings (1.7% had at least one rare *de novo* CNV). The proportion in probands increased further when examining only genic rare, *de novo* CNVs. The rare *de novo* CNVs in probands tended to be larger than in siblings and included a greater number of genes (Sanders et al., 2011). A later analysis combining CNV data from the SSC with data from the Autism Genome Project, strongly implicated six ASD CNV risk loci (1q21.1, 3q29,7q11.23, 16p11.2, 15q11.2–13, and 22q11.2); findings from this analysis strongly implicated *de novo* mutations in ASD risk (Sanders et al., 2015).

In summary, copy number variation appears to play a significant role in ASD pathogenesis. There is evidence for a higher burden of copy number variation in ASD affected individuals compared with controls, particularly in the context of CNVs directly affecting genes(Levy et al., 2011; Pinto et al., 2014; Pinto et al., 2010; Sanders et al., 2011; Sanders et al., 2015). Recurrent CNVs at a number of specific CNV loci have been robustly associated with ASD (Sanders et al., 2015). Recurrent ASD-associated CNVs may be inherited (Krumm et al., 2015; Pinto et al., 2014) or present *de novo* (Pinto et al., 2014; Sanders et al., 2015); however, *de novo* mutations appear to be more strongly implicated in ASD risk (Sanders et al., 2015). Individually these CNVs are rare, each accounting for less than 1% of ASD cases (Luo et al., 2012).

1.4.1.2 Intellectual Disability

Intellectual disability may result from a wide range of aetiological insults including environmental factors (e.g. prenatal insults: alcohol exposure, iodine deficiency, infections; perinatal insults: perinatal anoxia/asphyxia; postnatal insults: infant cerebral infection, traumatic brain injury) and genetic factors (Boat et al., 2015). For the majority of individuals with ID, the cause of the disorder is unknown (Bass et al., 2018). As for other NDDs, genetic studies are challenged by marked clinical and genetic heterogeneity.

There is a dearth of research specifically investigating the heritability of ID, however high heritability has been identified in family and population studies of intelligence (Vissers et al., 2016). Heritability of intelligence is significantly affected by age. In early childhood, there is evidence of a strong environmental effect on intelligence, with heritability estimates at less than 50%. In adulthood, heritability of intelligence appears to be higher at 60-80%, with a lesser effect for environmental factors (McGue et al., 1993; Pesta et al., 2020).

Genetic factors contributing to ID may differ between cases dependant on level of severity. Reichenberger et al. have suggested that most milder intellectual disabilities represent a low extreme in the normal distribution of IQ, whereas severe ID represents a distinct disorder with aetiologies that likely differ from milder forms of ID. This suggestion was based on evidence from their analysis indicating that siblings of individuals with mild IDs presented with lower IQs compared with general populations, whereas siblings of individuals with severe ID were not found to differ from the general population (Reichenberg et al., 2016).

Common and rare variants have been identified to contribute to genetic causes of ID (Bass et al., 2018). Genetic mutations associated with intellectual disability present with a wide range of phenotypes (Bass et al., 2018).

There has been some recent debate in the field of neurodevelopmental genetics regarding the relationship between ASD and ID and whether the two can be meaningfully separated in terms of genetic risk factors. Some large cohort studies have described findings supporting “autism-predominant” neurodevelopmental disorder genes (Satterstrom et al., 2020; Stessman et al., 2017), whereas others have argued strongly against this classification on both theoretical and practical grounds (Myers et al., 2020). Further research will be required to gain clarity on this controversial issue.

1.4.1.2.1 Copy number variation

In a study of 29,085 individuals with ID, developmental delay and/or ASD, Coe et al. identified a significantly higher rate of rare CNVs (frequency of <1%) in cases compared with controls. The excess of CNVs was primarily driven by large, rare deletions (Coe et al., 2014). Pathogenic CNVs

are identified in up to 15% of children with developmental delay referred for genetic testing (Bass et al., 2018; Miller et al., 2010b). *De novo* copy number variants present in approximately 10% of individuals with intellectual disability (de Vries et al., 2005; Vissers et al., 2016). Many rare, recurrent (~70) copy number variants have significant associations with developmental delay (Coe et al., 2014). Even in the absence of meeting full criteria for an intellectual disability, neurodevelopmentally associated CNVs have associations with decreases in cognitive ability and educational attainment (Bass et al., 2018; Kendall et al., 2016; Männik et al., 2015; Stefansson et al., 2014).

1.4.1.3 Schizophrenia

A complex interplay between genetic and environmental risk factors likely contribute to the aetiology of schizophrenia (Zwicker et al., 2018). Twin studies estimating concordance and heritability for schizophrenia report concordance rates between monozygotic twins at 33-65% and between dizygotic twins at 0-28%, with heritability estimates of approximately 79-85% (Cardno et al., 2000; Hilker et al., 2018; Sullivan et al., 2003). For most affected individuals, schizophrenia has a polygenic architecture in which hundreds or even thousands of variants collectively contribute to risk (Ripke et al., 2014). Single nucleotide polymorphism (common variant) heritability across the genome is largely uniform and it is likely that more than 71% of 1Mb genomic regions contain at least one risk variant (Loh et al., 2015). The polygenic risk score (PRS) method is given as the weighted sum of risk alleles with the weights specified by association coefficients (Chatterjee et al., 2016; Pharoah et al., 2002). This method was applied to a dataset of 22,177 schizophrenia cases and 27,629 controls, and estimated 27.4% of the heritability to be explained by SNPs with minor allele frequency >2% (Loh et al., 2015). This suggests that common variants explain a significant fraction, but not all of the heritability of schizophrenia. Although rare variants likely contribute a lower proportion of collective heritability in schizophrenia than common variants, they present with much larger individual effects.

1.4.1.3.1 Copy number variation

Global increases in rare deletions and duplications (>100kb) have been identified in patients with schizophrenia compared with controls (Walsh et al., 2008). Large consortia studies and a centralised analysis of 21,094 cases and 22,227 controls by the Psychiatric Genomics Consortium (PGC) confirmed the finding of large, rare CNV association with schizophrenia and have shown that the effect is driven by CNVs that overlap genes (International Schizophrenia Consortium, 2008; Marshall et al., 2017; Stefansson et al., 2008). These studies have also identified specific loci that were significantly associated with schizophrenia. Confirming the known association with 22q11.2 deletions, early consortia studies also reported novel associations with deletions at 1q21.1, 15q11.2, and 15q13.3. Subsequent studies implicated at least ten other regions with either risk or protective

loci (reviewed in (Kirov, 2015)). In 2017, the PGC analysis provided genome-wide significant evidence for eight of these (1q21.1, 2p16.3 (NRXN1), 3q29, 7q11.2, 15q13.3, distal 16p11.2, proximal 16p11.2 and 22q11.2) and suggestive support for eight more (Marshall et al., 2017).

Collectively, the list of known/probable schizophrenia CNV risk loci are carried by <2.5% of patients (Kirov, 2015). These CNVs increase risk for schizophrenia considerably (OR 2-30); they are of moderate penetrance for the disorder (penetrance 2-18%) (Kirov et al., 2014; Malhotra et al., 2012; Rees et al., 2014a; Rujescu et al., 2009b; Tansey et al., 2016). Most represent new events probably explained by a higher rate of *de novo* CNVs reported in schizophrenia compared to controls (Kirov et al., 2012).

1.5 Genetic overlap in neurodevelopmental disorders

Genetic overlap between schizophrenia, ASD and ID has been implicated in findings of common and rare variant analyses. An analysis of nearly 900,000 individuals by the Brainstorm Consortium demonstrated that a range of psychiatric disorders (particularly ADHD, schizophrenia, major depressive disorder and bipolar disorder) share common variant risk (Anttila et al., 2018). A cross-disorder genome wide association study (GWAS) meta-analysis comprised of 725,000 cases and controls across eight neuropsychiatric disorders modelled genetic correlations between the disorders and identified three groups of disorders based on shared common genetic risk. Schizophrenia clustered into a group with major depressive disorder and bipolar affective disorder; ASD clustered into a group with ADHD (Lee et al., 2019). Extensive common variant overlap between schizophrenia and intelligence has been identified in a genome-wide analysis (Smeland et al., 2019) and common genetic risk for ASD has been associated with general cognitive ability in a general population (Clarke et al., 2016). In a large ASD GWAS, Grove et al. investigated common variant overlap between ASD and other neuropsychiatric and neurodevelopmental phenotypes, identifying significant genetic correlations between ASD and schizophrenia, depression, ADHD and measures of cognitive ability (particularly educational attainment) (Grove et al., 2019).

Exome sequencing studies of schizophrenia trios have identified an enrichment of *de novo* non-synonymous mutations in affected individuals and the genes with *de novo* mutations have been found to overlap with genes implicated in autism, ID and developmental delay (McCarthy et al., 2014) (Genovese et al., 2016; Singh et al., 2016).

1.5.1 Overlap of CNVs in neurodevelopmental disorders

DiGeorge syndrome, arising from the 22q11.2 deletion has long been known as a genetic risk factor shared between NDDs, presenting with increased risk for schizophrenia (Bassett et al., 1998; Chow

et al., 1999; Karayiorgou et al., 1995), ASD (Fine et al., 2005; Niklasson et al., 2001) and ID (Bassett et al., 2005; Goldberg et al., 1993) as well as other clinical characteristics (Bassett et al., 2011). It is now known that at least eight pathogenic CNVs are shared as risk factors between these NDDs. The eight CNV loci identified by the PGC analysis as significantly associated with schizophrenia (Marshall et al., 2017) are each also associated with significant increased risk for ASD or ID (Coe et al., 2014; Sanders et al., 2015). There are many other CNVs associated with autism spectrum disorder and intellectual disability syndromes (Bucan et al., 2009; Cooper et al., 2011; Glessner et al., 2009; Hehir-Kwa et al., 2011; Qiao et al., 2014; Rees et al., 2016a; Sanders et al., 2015)

Neurodevelopmental CNVs (ND CNVs) have reduced penetrance for NDDs meaning that a proportion of individuals with these variants present with no evidence of NDDs. They also have variable expressivity, meaning that the severity of associated phenotypes varies between individuals (Heil et al., 2012). ND CNVs present in general population and control samples but are enriched in NDD populations (Kirov et al., 2014). The penetrance of ND CNVs for early developmental disorders is significantly higher (ranging from 8-88% depending on which specific ND CNV (Kirov et al., 2014)) than the risk of developing schizophrenia (penetrance ranging from 2-18%) (Kirov et al., 2014). The full range of pleiotropy associated with ND CNVs is under ongoing investigation.

1.6 Psychiatric and cognitive phenotypes associated with ND CNV

1.6.1 Population-based retrospective cohort studies

Cognitive deficits in carriers of ND CNVs have been identified in a number of population studies of adults with ND CNVs. In an Icelandic-based population cohort of >100,000 individuals, Stefansson et al. found that carriers of ND CNVs had a global assessment of functioning score that was 0.7 standard deviations below non-carrier population controls. ND CNV carriers performed at a cognitive level between performance levels of schizophrenia patients and population controls (who did not carry ND CNVs) (Stefansson et al., 2014). Kendall et al. also identified reduced cognitive performance in carriers of ND CNVs in a cohort of >100,000 genotyped subjects in a large general UK-based adult population dataset. The cognitive deficits in the CNV carriers were not as severe as cognitive deficits seen in individuals with schizophrenia in the cohort (Kendall et al., 2016). Mannik et al. have identified an association between large, rare CNVs and lower educational attainment in a population cohort in Estonia and replicated the finding in cohorts of adults from Italy and the United States and adolescents from the United Kingdom (Männik et al., 2015).

An association between ND CNVs and depression has been established in two adult population samples. In the Icelandic population sample referred to above, Stefansson et al. found that ND CNV carriers were at increased risk of depression and suicidal ideation (odds ratio = 2.86, P = 0.0017, and

odds ratio = 2.20, P = 0.011, respectively), but ND CNV carriers did not differ significantly from non-carriers in terms of anxiety and substance misuse prevalence (Stefansson et al., 2014). Kendall et al. assessed >400,000 individuals from the UK Biobank and also found an increased risk of depression in ND CNV carriers (OR 1.34, $p = 1.38 \times 10{-}7$) (Kendall et al., 2019). Interestingly, individuals with major neurodevelopmental disorders were excluded from these analyses, indicating that the associations between ND CNVs and depression may be independent of neurodevelopmental comorbidity.

Childhood population samples have also identified associations between ND CNVs and cognitive and neuropsychiatric problems. Martin et al. analysed data from a population sample of 12,982 children and identified an association between large ND CNVs and increased risk of an outcome variable termed neurodevelopmental problems comprised of ADHD, ASD, motor problems, learning difficulties and tic problems. ND CNVs were not associated with risk of anxiety or depression in this study (Martin et al., 2019). In another childhood population cohort of 6,807 individuals, Guyatt et al. explored the relationship between rare CNV burden and neuropsychiatric presentations. An association was identified between CNVs and a continuous measure of ASD and IQ but not anxiety or depression diagnoses or quantitative psychiatric traits of ADHD or psychotic experiences (Guyatt et al., 2018).

1.6.2 Case-control studies

A number of research groups have collected deeply phenotyped data on carriers of specific ND CNVs and utilising case-control study designs have shown a significant increased risk of cognitive deficits and psychiatric phenotypes in ND CNV carriers. Hanson et al. showed that carriers of the 16p11.2 deletion were at substantially higher risk of presenting with psychiatric and/ or neurodevelopmental disorders compared with controls; 93% of deletion carriers had at least one of these disorders compared with 21% of non-carrier controls. IQ scores were shown to vary widely in the deletion carrier group, but on average their IQ scores were 26 points lower than the controls (Hanson et al., 2015). Higher frequencies of psychiatric disorders were also found by Niarchou et al. in 16p11.2 deletion carriers and in duplication carriers compared with familial controls (OR 8.9, $p < 0.001$ and OR 5.3, $p = 0.01$ respectively). Deletion and duplication carriers had a higher frequency ADHD compared with controls and deletion carriers additionally had higher rates of ASD compared with controls. There were no differences in the prevalence of anxiety disorders, oppositional disorder/conduct disorder, and psychotic symptoms between duplication or deletion carriers and controls in this analysis (Niarchou et al., 2019).

Niarchou et al. assessed the prevalence and characteristics of psychiatric disorders and cognitive impairments in children with the 22q11.2 deletion syndrome (22q11.2DS), compared with 39 sibling controls. More than half (54%) of children with 22q11.2DS met diagnostic criteria for one or more

DSM-IV-TR psychiatric disorder, compared with 10% of the siblings, with the children with 22q11.2DS presenting with higher rates of ADHD, ODD and anxiety disorders. More children with 22q11.2DS (26%) met the cut-off for probable ASD diagnosis compared to their siblings (5%). (Niarchou et al., 2014).

An increased risk of psychopathology in ND CNV carriers may be expected, given that they are at higher risk of NDDs and that NDDs themselves confer an increased risk of psychiatric and behavioural disorders occurring co-morbidly (Einfeld et al., 2011; Simonoff et al., 2008). However, a number of these studies have presented data that indicate that the increased risk of psychopathology in those with ND CNVs is not solely mediated by intellectual impairment or ASD symptomatology. Hanson et al. demonstrated that there was a higher than expected number of psychiatric diagnoses in the 16p11.2 deletion carriers even when controlling for non-verbal IQ (NVIQ) and ASD diagnosis (Hanson et al., 2015). Niarchou et al. undertook a mediation analysis to examine the effect of IQ on the psychopathology association with 22q11.2DS and identified no significant relationship between psychopathology and IQ (Niarchou et al., 2014). Royston et al. conducted a meta-analysis assessing prevalence rates of anxiety disorders in individuals with Williams syndrome (microdeletion 7q11.23) and compared rates of individuals with ID of mixed aetiology (Carrasco et al., 2005). Individuals with Williams syndrome were four times more likely to have an anxiety disorder compared with individuals with ID of mixed aetiology, indicating that the 7q11.23 may confer an additional risk for anxiety disorders (Reardon et al., 2015; Royston et al., 2017)**.** These studies suggest that ND CNVs may confer psychiatric risk that is separate to NDD comorbidity.

These studies have contributed to the understanding of psychiatric and cognitive phenotypes associated with specific ND CNVs. When only one genotype is assessed per study, it is difficult to know the extent to which phenotypic findings for different CNVs can be compared across studies; there is variation between studies in terms of sample sizes, ascertainment methods and phenotypic batteries. Chawner et al. used deep phenotype data to assess the effect of a group of ND CNVs on behavioural and cognitive outcomes in a cohort of children with a number of different ND CNVs (loci included: 1q21.1, 2p16.3, 9q34.3, 15q11.2, 15q13.3, 16p11.2, 22q11.2) compared with control siblings. They also compared risk between different genotypes within the study. The CNV carriers had a 14-times higher risk of presenting with one or more psychiatric disorder. The disorders of highest risk were ADHD (OR 6.9, $p = 2.09x10^{-06}$), oppositional defiant disorder (OR 3.6, p = 0.012), anxiety disorders (OR 2.9, p = 0.0146), and autism spectrum disorder traits (OR 44.1, $p = 2.50\times 10^{-09}$). The results remained significant after controlling for IQ (Chawner et al., 2019). In ND CNV carriers, they also identified evidence of differences in phenotypic profiles between different genotypes (Chawner et al., 2019).

1.6.3 Subphenotypes in neurodevelopmental disorders associated with CNVs

A number of studies have sought to examine the role of CNVs in identifying specific patterns in the phenotypic heterogeneity observed within NDDs such as ASD. In an innovative study, Merikangas et al. identified that among cases diagnosed with ASD, presence of CNVs affecting ASD or intellectual disability genes were associated with specific communication and language deficits (Merikangas et al., 2015a). Bishop et al. also identified phenotypic characteristic associated with *de novo* CNV carriers among cases diagnosed with ASD. Individuals with *de novo* mutations were more likely to present with motor developmental delay (walking at later age) but were less impaired on some measures of ASD symptomatology (Bishop et al., 2017).

1.6.4 Sex differences in ND CNV phenotypes

Neurodevelopmental disorders are reported as presenting more frequently in males compared with females. The prevalence ratio of males to females with ASD is commonly reported at ~4:1 (Fombonne, 2009). Female-to-male prevalence ratio of ID has been identified in adults as 0.7-0.9 and in children and adolescents as 0.4-1.0 (Maulik et al., 2011; Maulik et al., 2013). Incidence of schizophrenia is higher in males than females with a reported male to female incidence ratio of ~1.4:1 (Aleman et al., 2003; Falkenburg et al., 2014; McGrath et al., 2008). These epidemiological differences suggest that there may be sex differences in NDD aetiologies.

Differences have been identified in CNV burden between sexes in NDDs. ASD affected females have been found to present with a higher burden of *de novo* CNVs than affected males (Levy et al., 2011; Sanders et al., 2011; Sanders et al., 2015). More genes appear to map within *dn*CNVs in female probands compared with male probands (Jacquemont et al., 2014; Sanders et al., 2011; Sanders et al., 2015). Females with schizophrenia have demonstrated a higher burden of schizophrenia associated CNVs compared with males (OR 1.38, $P = 0.0055$) (Han et al., 2016).

Differences in rare CNV phenotypes have also been observed between sexes. *De novo* CNVs or loss of function mutations (LOFs) have demonstrated more significant impact on cognition in females than males with ASD: non-verbal IQ was decreased by 10 points more in female carriers of *de novo* CNVs or LOFs compared with male carriers in a study by Sanders et al. (Sanders et al., 2015). In a childhood population sample, Martin et al. observed that females diagnosed with anxiety or depression were nearly four times more likely than males to have large CNVs (Martin et al., 2019). In an adult population study of ND CNV association with depression, Kendall et al. undertook an exploratory analysis which identified weak evidence for a higher rate of depression among female ND CNV carriers compared with male carriers (Kendall et al., 2019).

1.6.5 Summary of psychiatric and cognitive phenotypic associations with ND CNVs

Population studies have identified cognitive deficits in adult ND CNV carriers which present even in the absence of major NDDs such as schizophrenia, ASD and ID (Stefansson et al., 2014) (Kendall et al., 2016). Depression has also been significantly associated with ND CNV status in large adult population studies (Kendall et al., 2019). In childhood population samples, there is evidence of association between ND CNVs and reduced cognitive function (Guyatt et al., 2018) and neurodevelopmental problems such as ASD, ADHD, motor disorders, learning difficulties, tic problems (Martin et al., 2019), but associations between ND CNVs and other psychiatric disorders have not to date been identified in such studies (Guyatt et al., 2018; Martin et al., 2019). Analyses in childhood retrospective cohorts have been undertaken in relatively small samples and there have been some limitations in studies with regard to psychiatric outcomes assessed (e.g. some psychiatric outcomes examined in aggregate precluding estimation of association for individual outcomes).

Case-control studies indicate that ND CNV carriers are at higher risk of developing psychiatric symptomatology (Chawner et al., 2019; Hanson et al., 2015; Niarchou et al., 2014; Royston et al., 2017). There is evidence from population studies and case-control studies that the increased risk of psychiatric disorder may be independent of ASD co-morbidity (Hanson et al., 2015; Kendall et al., 2019) or cognitive level (Hanson et al., 2015; Niarchou et al., 2014; Royston et al., 2017). There is some evidence of sex differences in the phenotypic effects of ND CNVs (Kendall et al., 2019; Martin et al., 2019).

1.7 Clinical Utility of ND CNV diagnosis in neurodevelopmental disorders

1.7.1 Autism spectrum disorder and intellectual disability

Clinical genetic testing is a standard of care for the investigation of unexplained neurodevelopmental disorders ASD and ID in paediatric populations. Chromosomal microarray analysis (CMA)[1] is recommended as a first-tier investigation for the detection of pathogenic CNVs in both disorders (Battaglia et al., 2013; Schaefer et al., 2013). The diagnostic yield of CMA is ~15% in the investigation of unexplained ASD (Carter et al., 2013; Schaefer et al., 2013; Velinov, 2019) and ID or developmental delay (Miller et al., 2010a).

The benefits of diagnosing ND CNVs for individuals with neurodevelopmental disorders include a sense of empowerment through the acquisition of information and understanding around a child's presentation, better social, medical and educational service provision to the child and family, specific screening for medical, psychiatric or developmental risks associated with a genetic syndrome,

[1] CMA includes both array comparative genomic hybridization (aCGH) and single nucleotide polymorphism arrays

specific recurrence risk counselling and improved access to support and research networks (Moeschler et al., 2014; Schaefer et al., 2013).

1.7.2 Schizophrenia

Genetic testing is not currently utilised in standard clinical practice for investigation and management of individuals with schizophrenia. The identification of well-established pathogenic CNVs that confer high risk for schizophrenia has led to increasing interest in introducing genetic testing to routine psychiatric management of the disorder (Baker et al., 2014; Miles et al., 2008a).

Copy number variants with known schizophrenia risk are found in ~2.5% of individuals with the disorder (Kirov, 2015); clinically significant CNVs have been reported in up to 8% of individuals with the disorder (Costain et al., 2013). Pick up rates for CNVs in the general schizophrenia population are likely to be low. In autism and intellectual disability, 'syndromal' cases where additional phenotypes and dysmorphic features are also present, are more likely to have a genomic aetiology (Miles et al., 2008a). Recent studies have suggested that identifying schizophrenia patients with co-morbid ID is likely to be helpful in identifying subsets of individuals with genomic disorders. Thygesen and colleagues reported an approximately three-fold higher rate of pathogenic CNVs in patients with psychosis and intellectual disability compared to rates in the general schizophrenia population (Thygesen et al., 2018). Lowther et al. examined the genome-wide burden of pathogenic CNVs in a schizophrenia cohort (n=546) and demonstrated a significantly higher burden of pathogenic CNVs (OR 5.01, $p = 0.0001$) in patients with schizophrenia and low IQ (IQ< 85) compared with those with average IQ ($IQ \geq 85$). Based on their findings, the authors concluded that individuals with schizophrenia and low IQ should be prioritised for clinical microarray testing in clinical and research contexts (Lowther et al., 2017).

Psychiatrists have endorsed positive attitudes towards the incorporation of genetics into psychiatric clinical practice (Hoop et al., 2008), but have indicated a perceived lack of competence in their capacity to deliver this kind of information to patients and families (Blacker et al., 2005). Genetic counselling has been identified as a need by patients with schizophrenia and has been shown to be helpful in improving patients' understanding of empiric recurrence risk of their disorder (Costain et al., 2012b).

1.7.3 Prenatal identification of ND CNVs

CMA testing is recommended as a gold standard investigation in the prenatal setting when foetuses are identified as having a congenital malformation (Brabbing-Goldstein et al., 2018; Shaffer et al., 2012; Shkedi-Rafid et al., 2016). Although the use of CMA in this context is widely accepted, there is debate on its use in investigating structurally normal foetuses. CMA testing is an option in some settings for patients who undergo invasive testing for various reasons including elective testing

(Brabbing-Goldstein et al., 2018; Dugoff et al., 2016; Obstetricians et al., 2013); the frequency of pathogenic CNVs is approximately 1% in this group (Miny et al., 2013). The limited phenotypic data at a population level, incomplete penetrance and variable expressivity of ND CNVs have been identified as considerable challenges to their clinical interpretation in the prenatal setting (Brabbing-Goldstein et al., 2018). The identification of ND CNVs in the prenatal setting can impact on parental choices around pregnancy, including termination of pregnancy (Brabbing-Goldstein et al., 2018).

1.7.4 Clinical challenges in ND CNV interpretation

Penetrance for various phenotypic features of ND CNVs varies significantly between individuals, posing challenges to clinicians counselling parents of children or individuals with these CNVs. For some neurodevelopmentally associated CNVs, such as the 22q11.2 deletion or those implicated in Prader-Willi syndrome/Angelman syndrome, the phenotypes are well established and clinical guidelines inform the optimal management and investigation of individuals with the variant (Bassett et al., 2011; McCandless, 2010). The evidence bases for clinical and phenotypic presentations of other recurrent ND CNVs are not as well-established. International databases have been established to support interpretation of variants (Feenstra et al., 2006; Firth et al., 2009; Kaminsky et al., 2011) and phenotyping and population studies are contributing valuable information to the understanding of ND CNV clinical profiles (Al Shehhi et al., 2019; Dolcetti et al., 2013; Hanson et al., 2015).

Variants of unknown significance (VOUS) and incidental findings may be identified on CMA testing. VOUS are genetic variants which many have associations with the disorder of interest or other developmental disorders but do not have enough clinical evidence to be categorized as either a pathogenic or a benign variant. The ambiguity of VOUS have been suggested to cause stress when identified on CMA; however there is also evidence to suggest that identification of VOUS are important to parents of children with developmental disorders, with genetic counselling contributing to positive outcomes (Jez et al., 2015). Incidental findings are secondary findings unrelated to the indication for ordering the genetic test but that may be of medical value for patient care. Such findings present ethical concerns for genetic testing (Roche et al., 2015). Many experts suggest that patients and research participants should make the choice of whether to attain information on incidental findings for themselves, however complex issues regarding what constitutes a fully informed decision continue to provide challenges in this context (Appelbaum et al., 2017). Recommendations have been made to support appropriate reporting of incidental findings (Green et al., 2013).

There is no one approach to that will align with the specific needs of every individual with an ND CNV and their family. Careful clinical consideration weighing up the potential risks and benefits of genetic testing on a case by case basis is necessary.

1.8 Thesis rationale

We are beginning to comprehend the range of neurodevelopmental features and psychopathologies associated with ND CNVs. The phenotypic associations with ND CNVs transcend the borders of conventional psychiatric diagnosis and contribute to the clinical heterogeneity observed within disorders. Getting a clear picture of ND CNV phenotypes in different populations is essential to informing genetic testing for ND CNVs and to being able to provide relevant clinical information to those affected by ND CNVs. A comprehensive and refined understanding of phenotypes associated with ND CNVs may lead to insights into the aetiologies of psychiatric disorders and NDDs and may in the longer-term help to improve diagnostic, treatment and management strategies.

For children presenting with unexplained ASD or developmental delay, chromosomal microarray (CMA) testing is now a standard diagnostic investigation (Battaglia et al., 2013; Schaefer et al., 2013); clinically significant CNVs are identified in up to 15% of individuals with these conditions. The benefits associated with identification of CNVs in children with ASD are well documented (Moeschler et al., 2014; Schaefer et al., 2013). There is a rapidly developing evidence base indicating that ND CNVs may be associated with a wide spectrum of psychiatric risk in carriers, however there is currently insufficient data to be able to provide reliable information on psychiatric risk profiles associated with the majority of ND CNVs in children with ASD. Understanding psychiatric risk profiles in the subset of individuals with ASD who carry ND CNVs will be important in providing informed genetic counselling for these individuals and their families and may also provide insights for surveillance and management of risk for ND CNV carriers.

A better understanding of psychiatric risk profiles associated with ND CNVs in children who do not present with a major NDD is also essential. ND CNVs have incomplete penetrance for NDDs and many carriers may not have major childhood NDDs. There is limited data available in the current literature on ND CNV associations with psychiatric phenotypes in large cohorts of youths. Identifying whether increased risk of major classes of childhood psychopathologies are associated with ND CNVs will be vital for child and adolescent carriers of the ND CNVs and their families.

Evidence suggests that individuals with schizophrenia are at increased risk of carrying ND CNV, however, unlike children with NDDs they are not routinely tested. A major challenge to the routine integration of CNV screening for schizophrenia patients is that confirmed variants are carried by <2.5% of patients (Kirov, 2015). The identification of clinical symptoms or demographic features that differentiate schizophrenia patients that carry schizophrenia associated (SCZ-associated) CNVs would be helpful in clarifying who might benefit most from testing. Recent studies have suggested that identifying schizophrenia patients with co-morbid ID or multiple congenital malformations or dysmorphic features is likely to be helpful in identifying subsets of individuals with genomic disorders (Christian G. Bouwkamp et al., 2017; Lowther et al., 2017; Thygesen et al., 2018). The utility of other developmental indices in identifying such subsets of patients is less well explored.

The central aim of this thesis was to investigate the relationship between psychiatric and neurodevelopmental phenotypes and ND CNV carrier status using secondary data analysis of existing datasets, to inform our understanding of ND CNV phenotypes and provide clinically translatable insights.

1.8.1 Aims and Hypothesis

Analysis 1: Investigation of psychiatric phenotypes associated with neurodevelopmental copy number variants in a cohort of youths with ASD

In this analysis, the first aim was to establish whether ND CNVs are associated with increased risk for a range of psychiatric phenotypes in a large cohort of youths with ASD. It was hypothesised that youths with ND CNVs would have higher rates of psychopathologies compared with youths without ND CNVs. Based on recent evidence suggesting sex differences in depression and anxiety phenotypic associations with ND CNVs (Kendall et al., 2019; Martin et al., 2019), it was hypothesised that there would be differences in the rates of these disorders in male and female carriers of ND CNVs in the sample.

A follow up analysis was conducted in a sample of ASD-unaffected siblings, to explore whether the identified associations between ND CNV carriers and psychopathology in individuals with ASD also presented in the absence of ASD. It was hypothesised that increased risk of psychopathology associated with ND CNV status in the ASD sample would also present in the ASD-unaffected siblings.

The data for this analysis were obtained from the Simons Simplex Collection, a cohort comprised of autism simplex families (n=2,644), including ASD proband and unaffected sibling genotype and phenotype data (Fischbach et al., 2010).

Analysis 2: Investigation of psychiatric phenotypes associated with neurodevelopmental copy number variants in a large population-based clinical cohort of youths

In this analysis, the first aim was to investigate the relationship between ND CNVs and four major outcomes reflective of significant psychopathology (internalising disorders, externalising disorders, subclinical psychotic symptoms and suicidal ideation) in a large population-based clinical cohort of youths. It was hypothesised that youths with ND CNVs would have higher rates of psychopathology compared with those without ND CNVs.

The second aim of the analysis was to assess sex differences in the effects of ND CNVs on internalising disorders. It was hypothesised that ND CNV effects on internalising psychiatric disorders would differ between males and females. As for the first analysis, this hypothesis was based

on recent evidence suggesting sex differences in depression and anxiety phenotypic associations with ND CNVs (Kendall et al., 2019; Martin et al., 2019).

The third aim was to identify whether NDD comorbidity was a major determinant of increased risk of psychopathology associated with ND CNVs. Increased risk of NDDs ASD and ID are well-established in ND CNV carriers (Coe et al., 2014; Sanders et al., 2015). ASD and ID are associated with increased risk of psychiatric disorders compared with neurotypical populations (Abdallah et al., 2011; Einfeld et al., 2011; Skokauskas et al., 2012a). Therefore, increased risk of psychopathology associated with ND CNVs may occur as a result of increased rates of NDDs in ND CNV carriers. However, a number of studies have indicated that increased risk of psychopathology associated with ND CNV status is not solely attributable to neurodevelopmental comorbidity (Hanson et al., 2015; Kendall et al., 2019; Niarchou et al., 2014; Stefansson et al., 2014). In this analysis, it was hypothesised that increased risk of psychopathology associated with ND CNVs would be conferred separately to ASD or ID comorbidity.

These hypotheses were tested using genetic and phenotype data from a publicly available large cohort of youths (n=8,205), the Philadelphia Neurodevelopmental Cohort (PNC) (Calkins et al., 2015).

Analysis 3: Identifying schizophrenia patients who carry pathogenic genetic copy number variants using standard clinical assessment: a retrospective cohort study

In this analysis, the aim was to determine whether clinically identifiable phenotypic features were predictive of SCZ-associated CNV carrier status in a large schizophrenia cohort.

Based on the known overlap with other neurodevelopmental disorders and previously reported phenotype studies (Ahn et al., 2014; Costain et al., 2014; Derks et al., 2013; Kirov et al., 2014; Philip et al., 2011; Sahoo et al., 2011b; Stefansson et al., 2014; Walsh et al., 2008; Wilson et al., 2011; Yeo et al., 2013), it was hypothesised that individuals with schizophrenia who carry risk CNVs are likely to be enriched for phenotypic features suggesting pre-existing neurodevelopmental compromise, earlier onset of psychotic symptoms, more severe illness course, or a positive family history of neurodevelopmental disorder.

This analysis was undertaken in a discovery sample of 1,215 individuals with schizophrenia and replicated in a sample of 479 individuals with schizophrenia.

Chapter 2: Investigation of psychiatric phenotypes associated with neurodevelopmental copy number variants in a cohort of youths with ASD

2.1 Introduction

2.1.1 Background

There are limited data describing psychiatric risk profiles associated with neurodevelopmental CNVs (ND CNVs) in youths with autism spectrum disorder (ASD). Clinical genetic guidelines recommend chromosomal microarray (CMA) testing as a first tier diagnostic investigation for individuals with ASD; around 15% of those tested will be diagnosed with a pathogenic CNV (Carter et al., 2013; Schaefer et al., 2013). Identifying psychiatric phenotypes associated with ND CNV status is important for providing informed genetic counselling for individuals who present clinically with ASD and are found to be ND CNV carriers and may help with psychiatric risk monitoring, prevention strategies and early interventions.

A number of population and case-control studies have identified increased rates of various psychopathologies associated with ND CNVs but these analyses have either excluded individuals with ASD from samples or assessed ASD or ASD traits as outcomes in analyses. Two large adult population studies have identified significant associations between depression and ND CNVs (OR 1.34, $p = 1.38 \times 10^{-7}$ (Kendall et al., 2019), OR 2.86, $p = 0.0017$ (Stefansson et al., 2014)); both of these studies excluded individuals with ASD (along with other neurodevelopmental disorders) from the samples analysed. Interestingly, Kendall et al. identified evidence of a higher rate of depression among female ND CNV carriers compared with male carriers (Kendall et al., 2019).

Childhood population studies examining associations between ND CNVs and psychopathologies have not excluded individuals with ASD from samples, but rather have examined ASD or ASD traits as outcomes in analyses. In a childhood population sample of 12,982 children, Martin et al. demonstrated that large ND CNVs presented with a higher frequency of "neurodevelopmental problems (NPs)" (NPs including: ADHD, ASD, motor problems, learning difficulties, and tic problems). No association was identified between CNVs and anxiety and depression in the sample, however findings indicated that large, rare CNVs may show sex-specific phenotypic effects; CNVs were enriched in females diagnosed with depression or anxiety, as compared to diagnosed males. (Martin et al., 2019). In another study of a childhood population cohort, Guyatt et al. identified schizophrenia associated (SCZ-associated) deletions to be associated with reduced cognitive

attainments and with a continuous measure of ASD. Insufficient case numbers prevented the assessment of association between SCZ-associated CNVs and anxiety, depression, ASD and ADHD binary measures (Guyatt et al., 2018).

ASD has also been included as an outcome in case-control ND CNV phenotypic association studies. Chawner et al. identified ND CNV carriers to be 14-times more likely to present with one or more psychiatric disorder compared with control siblings. ND CNV carriers presented with significantly higher rates of ASD, ADHD, anxiety disorders and oppositional defiant disorder (Chawner et al., 2019). A case-control study of 16p11.2 deletion carriers compared with familial controls found that 93% of carriers presented with psychiatric and/or developmental disorders, compared with 21% of non-carrier controls. The most common conditions were motor coordination disorders, speech and language disorders, enuresis, ASD and ADHD. The increased risk of psychiatric symptomatology in CNV carriers persisted when non-verbal IQ (NVIQ) and ASD diagnosis were controlled for in the sample (Hanson et al., 2015). In a case-control study of individuals with 22q11.2 deletion syndrome (22q11.2DS), more than half (54%) of children with the CNV met diagnostic criteria for one or more psychiatric disorder, compared with 10% of control siblings, with the affected children presenting with higher rates of ASD, ADHD, ODD and anxiety disorders. Mediation analysis indicated no significant relationship between psychopathology and IQ in children with 22q11.2DS (Niarchou et al., 2014).

These studies contributed valuable information to the growing body of knowledge indicating that ND CNV carriers may be at increased risk of a spectrum of psychiatric risks. However, there is considerable variation between studies in terms of sample selection, developmental stage examined and methodologies, limiting generalisability of findings particularly to populations with unique psychiatric and developmental profiles such as children with ASD.

Comorbid psychopathology rates are high in youths with ASD (estimated at ~70%) (Gjevik et al., 2011; Simonoff et al., 2008) and individuals with ASD are more likely to have psychiatric symptomatology than neurotypical populations (Abdallah et al., 2011; Lugo-Marín et al., 2019; Skokauskas et al., 2012a). Psychiatric comorbidity is associated with significant clinical impairment and additional needs for youths with ASD and their families (Leyfer et al., 2006; Mattila et al., 2010). Many children and adolescents with ASD experience significant difficulties in being able to process and describe their experiences, emotions and mental states (Leyfer et al., 2006; Mazzone et al., 2012), resulting in substantial challenges to being able to identify psychiatric comorbidities in these young people. Core features of ASD can also be difficult to differentiate from psychiatric symptomatology (Cholemkery et al., 2014) presenting further diagnostic challenges. These diagnostic challenges can result in missed or delayed diagnosis of psychopathology with significant impacts in terms of outcomes.

In view of the evidence from population and case-control studies suggesting that ND CNV carriers are at increased risk of having a range of psychopathologies, ND CNV carriers with ASD may also be at increased risk of having psychopathologies compared with ASD probands who do not carry ND CNVs. Psychiatric risk profiles associated with ND CNVs in individuals with ASD have been understudied to date. Addressing this deficit in knowledge is essential as ND CNVs are regularly identified in individuals with ASD. An improved understanding of psychiatric phenotypes associated with ND CNVs would improve clinical awareness of these risks and could facilitate targeted screening and therapeutic input for ND CNV carriers with ASD.

2.1.2 Aims and hypotheses

The aim of this analysis was to assess psychiatric risk profiles associated with ND CNV carrier status in youths with ASD. It was hypothesised that youths with ND CNVs would have higher rates of psychopathologies compared with youths without ND CNVs. Based on recent evidence suggesting that females with depression and anxiety may be enriched for CNVs compared with males with these diagnoses (Martin et al., 2019) and that females with ND CNVs may be at increased risk of depression compared with males carriers (Kendall et al., 2019), it was hypothesised that female carriers of ND CNVs would have higher rates of depression and anxiety compared with males.

A follow up analysis was conducted in a sample of ASD-unaffected siblings, to explore whether the identified associations between ND CNV carriers and psychopathology in individuals with ASD also presented in the absence of ASD. It was hypothesised that increased risk of psychopathology associated with ND CNV status in the ASD sample would also present in the ASD-unaffected siblings.

2.2 Materials and Methods

2.2.1 Sample

The data for this analysis were obtained from the Simons Simplex Collection (SSC, version 15, data access approval available in Appendix 3: Data Access Approval, Figure 7-5, Figure 7-7, Figure 7-8) (Fischbach et al., 2010). The SSC is a large, multi-site study that was collected for the purposes of an autism genomics study. The study aimed to recruit “simplex families” defined as a child with ASD with no other first to third-degree family member with a known or suspected diagnosis of ASD. Each family recruited consisted of a child with autism (aged 4-18 years), an unaffected sibling and both parents. Individuals were ascertained through clinics located at Michigan, Yale, Emory, Columbia, Vanderbilt, McGill Washington, and Harvard Universities (Children's Hospital of Boston), and at the Universities of Washington, Illinois (Chicago), Missouri, UCLA, and the Baylor

College of Medicine in the United States. A comprehensive phenotyping assessment battery was conducted with both children in each family. The dataset downloaded for this analysis contains phenotype data collected from 2,644 autism simplex families.

Autism affected youths were evaluated with gold standard autism research diagnostic assessments, Autism Diagnostic Interview-Revised (ADI-R) (Rutter et al.) and Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 2012) as well as other instruments providing additional information on the core features of autism, measures of intellectual and adaptive functioning, psychiatric and behavioural problems and motor functioning and language (Fischbach et al., 2010). Autism affected youths were excluded from the dataset if they were younger than 4 years or older than 18. Youths were also excluded if they presented with a condition that could compromise the ASD diagnosis using the above outlined instruments, for example non-verbal age estimated below 18 months, severe sensory or motor difficulties, genetic evidence of fragile X or Down syndrome or history of severe psychological deprivation (Fischbach et al., 2010).

Unaffected siblings closest in age to the child with ASD were selected for the collection. They were at least 4 years of age. A diagnosis of ASD or suspicion of the disorder was an exclusion criterion for siblings in the SSC, with siblings screened for ASD symptomatology and additional psychiatric conditions by trained clinicians (Fischbach et al., 2010; Schanding et al., 2012). Additionally, SSC protocol excluded siblings who were identified as having mental retardation or an adaptive behaviour standard score on the Vineland-II of less than 70, if they were diagnosed with schizophrenia or if they were diagnosed with a psychiatric disorder requiring treatment with more than one psychotropic medication (Fischbach et al., 2010).

2.2.2 Phenotype measures used in analysis

2.2.2.1 Measure of intellectual functioning

In the SSC phenotype data, measures of cognitive functioning were available for the majority of ASD affected individuals in the dataset but not for the siblings. Data from the Vineland Adaptive Behavior Scale, Second Edition (VABS-II) was available for both ASD affected individuals and siblings. Adaptive behaviour profiles are correlated with cognitive ability in typically developing youths (Pathak et al., 2019) and the VABS composite score has been shown to correlate strongly with IQ in youths with ASD ($r = 0.58$, $p < 0.0001$) (Pathak et al., 2019). VABS composite score (described further in section 2.2.2.1.1) was used in this analysis as an alternative to a formalised measure of cognitive ability.

2.2.2.1.1 Vineland Adaptive Behaviour Scale (VABS-II)

The Vineland Adaptive Behaviour Scale (VABS-II) is an informant based standardised interview that was administered by clinical or research staff with parents via phone interview. The interview measures adaptive behaviour in individuals from birth to 90 years. The VABS-II Parent/Caregiver rating form has been standardised based on a sample of 3,695 individuals nationally representative of the USA. Proportional random sampling was implemented for the normed sample according to demographic variables including sex, race/ethnicity, socioeconomic status, and geographic region. Data for norming and standardisation were collected from a range of neurodevelopmentally diverse groups: attention deficit/hyperactivity disorder, autism-nonverbal, autism-verbal, emotional or behavioural disturbance, deafness/hard of hearing, learning disability, cognitively delayed-mild (child and adult samples), cognitively delayed-moderate (child and adult samples), cognitively delayed severe/profound (adult sample) and visual impairment (Sparrow et al., 2005). The adaptive behaviour composite score describes the individual's overall level of adaptive functioning. The normative mean is 100 with a normative standard deviation of 15 (Sparrow et al., 2005).

2.2.2.2 Psychiatric phenotype outcome measures

2.2.2.2.1 Child Behaviour Checklist

The Child Behaviour Checklist (CBCL/6-18) is a standardised instrument frequently used in clinical and research settings which assesses a broad range of psychopathology in youths ranging in ages from 6-18. The CBCL is completed by parents and/or caregivers and describes the child's emotional and behavioural functioning during the previous six months. Each of 120 problem items are measured on a three-point Likert scale (0= "Not True," 1= "Somewhat or Sometimes True," or 2= "Very True or Often True"). The CBCL contains two empirically derived broadband scales: eight syndrome scales and six DSM-oriented scales. Raw scores for each scale are converted to norm-referenced T-scores (M = 50, SD = 10) (Achenbach, 2001). Cut points are provided for normal, borderline, clinical range scores on each scale. For the syndrome and DSM-oriented scales: T-scores < 65 (<93rd percentile) are classified as "Normal", T-scores 65-69 (93-97th percentile) are classified as "Borderline" and T-scores >69 (>97th percentile) are classified as "Clinical range" (Achenbach, 2001). There is evidence to support the use of the CBCL as a valid measure for screening and as part of diagnostic assessment of psychiatric symptomatology in youths with ASD and neurotypical youths (Bérubé, 2014; Pandolfi et al., 2012).

In this analysis, the aim was to assess clinically relevant psychopathology outcomes. Several studies have suggested that the CBCL syndrome scales yield only modest associations with DSM-IV disorders, have limited positive predictive values, and do not map well onto specific diagnoses (Ferdinand, 2008; Nakamura et al., 2009). Achenbach et al. developed the CBCL DSM-oriented

scales to align more closely to DSM classification. DSM-oriented scales were derived through agreement in experts' ratings of pre-existing items' consistency with DSM-IV diagnostic criteria (Achenbach et al., 2001). There are six separate DSM-oriented scales- affective problems scale (items corresponding to symptoms of depression, dysthymia), anxiety problems scale (items corresponding to symptoms of generalised anxiety, separation anxiety and specific phobia), somatic problems scale (items corresponding to somatization symptoms), attention-deficit/hyperactivity problems scale (items corresponding to primarily hyperactive, primarily inattentive and combined subtypes), oppositional defiant problems (items corresponding to oppositional defiant disorder), and conduct problems (items corresponding to symptoms of conduct disorder) (Ebesutani et al., 2010). Nakamura et al. evaluated psychometric properties of the CBCL DSM-oriented scales in a clinic-referred sample of youths and adolescents (n=673) and found good reliability (reliability coefficients ranging from 0.71 to 0.89), favourable internal consistency and excellent convergent and divergent validity of the scales (Nakamura et al., 2009).

The CBCL DSM-oriented scales were used in this analysis as the outcomes of interest reflecting psychiatric phenotypes. All of the outcome measures in the CBCL were highly positively skewed, prohibiting simple transformation. Due to the highly skewed nature of the data and the objective of using clinically significant cutpoints, the variables were dichotomised for this analysis. A cutoff point was set at a T-score of 65 to identify a clinical risk group for each DSM-oriented scale. A score of above 65 indicated that the individual was in the clinical risk group. A score of below 65 indicated that the individual was not in the clinical risk group for the outcome.

Table 2-1 presents a correlation matrix with tetrachoric correlations of the phenotypic variables assessed in the SSC proband dataset. The tetrachoric correlation coefficient (r_t) measures the covariation between two dichotomous variables (El-Hashash et al., 2018). Several of the psychiatric phenotypic outcomes in the analysis were correlated, ranging from the lowest correlation between attention-deficit/hyperactivity problems and somatic problems ($r_t = 0.23$) to the highest between oppositional defiant problems and conduct problems ($r_t = 0.76$). Co-occurrence of psychiatric symptoms and optimal classification of psychiatric disorder is a well-known and widely-discussed issue in psychiatric clinical practice (Pincus et al., 2004) and research (Batstra et al., 2002) and is discussed further in the limitations section 2.4.1.3. As the six psychiatric outcomes measured by the DSM-oriented scales in this analysis are recognised as separate diagnostic entities in current classification systems (American Psychiatric Association, 2013; *Diagnostic and statistical manual of mental disorders : DSM-IV*, 1994; World Health, 2004), these outcomes were assessed as independent outcomes in this analysis.

Table 2-1. Tetrachoric correlations between the psychiatric phenotype variables in the SSC proband sample.

	Affective Problems	Anxiety Problems	Somatic Problems	ADHD Problems	ODD Problems	Conduct Problems
Affective Problems	-					
Anxiety Problems	0.56	-				
Somatic Problems	0.47	0.38	-			
ADHD Problems	0.40	0.36	0.23	-		
ODD Problems	0.52	0.45	0.32	0.54	-	
Conduct Problems	0.50	0.33	0.34	0.48	0.76	-

Note: Correlation coefficients measured by tetrachoric correlation coefficient t_r (El-Hashash et al., 2018), calculated using "tetrachoric" function in the "psych" package in R (W, 2019). ADHD, attention deficit hyperactivity disorder; ODD, oppositional defiant disorder.

2.2.3 CNV Selection

A list of neurodevelopmentally-associated CNVs was selected based on a list of ND CNVs previously described by Martin et al. (Martin et al., 2019), including CNVs previously implicated in autism spectrum disorder (Sanders et al., 2015), intellectual disability (Coe et al., 2014) and schizophrenia (Marshall et al., 2017). The CNVs associated with ASD were identified from *de novo* CNV findings by Sanders et al. from the Autism Genome Project (Pinto et al., 2014) in addition to the SSC findings (Sanders et al., 2015), yielding 12 *de novo* CNVs implicated in ASD with a false discovery rate (FDR) ≤ 0.1 (CNVs listed in table 2 of Sanders et al. 2015 (Sanders et al., 2015)). The ID associated CNVs were identified from Coe et al. and were derived from comparison of 29,085 youths with developmental delay and 19,584 healthy controls (Coe et al., 2014). The list included 63 autosomal CNVs, which are listed in the supplementary data of another publication by Rees et al. (Rees et al., 2016b). The schizophrenia associated CNVs were identified from the Psychiatric Genomics Consortium (PGC) centralised analysis of 21,094 cases and 22,227 controls, identifying eight CNV loci with genome-wide significance (included in Table 1 of that paper (Marshall et al., 2017)). This yielded a list of 29 duplications and 40 deletions associated with at least one of these neurodevelopmental disorders. Full details of the ND CNV loci are presented in Appendix 1: Supplementary Tables, Table 7-1.

2.2.4 CNV data

Rare *de novo* and inherited CNVs identified in SSC patients were downloaded from Sanders et al. 2015 (Sanders et al., 2015). Rare CNVs were defined occurring at a population frequency ≤ 0.1% in either the Database of Genomic Variation (MacDonald et al., 2014) or among all 5,382 SSC parents (Sanders et al., 2015). CNVs were identified using three calling algorithms- PennCNV (Wang et al., 2007), QuantiSNP (Colella et al., 2007) and GNOSIS (Sanders et al., 2011).

The proband and sibling CNV data that was included in the Sanders et al. publication was cross-referenced with the SSC phenotype data which resulted in a total of n=3600 individuals for whom both CNV and phenotype data was available.

Individuals were called as carriers of ND CNVs if there was greater than 50% overlap with one of the genomic loci of interest. All ND CNVs identified in individuals in the analysis passed a pCNV[2] cut off point set by Sanders et al. The pCNV cut off point refers to a metric developed by Sanders et al. to improve the specificity of CNV predictions, the fact that all ND CNVs in the analysis passed the pCNV cut off point indicated that they were unlikely to be false positive or false negative calls (Sanders et al., 2015).

2.2.5 Data analysis

In the sample of youths with ASD, multiple logistic regression models were used to test associations between ND CNVs and six psychiatric outcomes analysed using the DSM-oriented scales of the CBCL: affective problems, anxiety problems, somatic problems, attention-deficit/hyperactivity problems, oppositional defiant problems, conduct problems. The dependent variables assessed all had dichotomous outcomes. VABS composite score, sex and age at assessment were included as covariates in the models. Predictor coefficients were tested using Wald tests and confidence intervals were obtained using the Wald method. Model fit was assessed using Nagelkerke pseudo R square index.

Based on previous evidence suggesting sex differences in depression and anxiety phenotypic associations with ND CNVs [13, 21], an interaction variable (sex*NDD CNV status) was introduced

[2] There are many methods for predicting CNVs in SNP genotyping data. However, many of these methods can be prone to large numbers of false positives and false negatives, particularly when trying to identify extremely rare variants such as *de novo* CNVs. Sanders et al. developed a statistical approach to assess the accuracy of each predicted CNV. This method estimates a p-value for the null-hypothesis that there is no deviation from the expected distribution of data in the SNPs within a predicted CNV. This p-value is estimated by assessing the ratio of the likelihood of the observed deviation in log R ratio or B allele frequency for each SNP in a CNV in a region with two copies versus a region with one or three copies, depending the on type of CNV. The metric was optimized to identify rare and potentially *de novo* CNVs; an overview of the methodology is discussed by Sanders et al. in (Sanders et al., 2015).

to examine for differences in rates of these disorders in male and female carriers of ND CNVs in the ASD sample.

Based on positive findings of significant associations between ND CNVs and specific psychiatric outcomes in the ASD sample, these associations were also tested for significance in the sample of ASD-unaffected siblings. A multiple logistic regression was used in this analysis, including VABS composite score, gender and age as relevant covariates. Model fit was assessed using Nagelkerke pseudo R square index.

Statistical significance was set at $p < 0.05$. The results presented are not corrected for multiple comparisons as the analysis was considered exploratory. All analyses were completed in R version 3.2.3 ("R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.," 2013).

2.3 Results

2.3.1 ASD Proband Analysis

2.3.1.1 Characteristics of Sample

The sample consisted of a total of 1966 youths with ASD. There was a preponderance of males (86.8%) in the sample. Data collection for the SSC took place in stages, therefore the ages at which various instruments were completed on youths varied. Youths in the sample ranged in age from approximately 4-18 years. The median age of youths in the probands sample was 8.8 years (IQR 6.7-11.3). Mean VABS composite score for the sample was 72.1 (SD = 11.8).

2.3.1.2 ND CNV carrier status

Neurodevelopmental copy number variants presented in n=63 (3.2%) of individuals in the sample. The ND CNVs identified in individuals in the SSC are presented in Appendix 1: Supplementary Tables,

Table 7-2. The characteristics of the proband ND CNV carriers compared with non-carriers are presented in Table 2-2. ND CNV carriers and non-carriers did not differ significantly by sex, age or VABS composite score.

Table 2-2. Comparison of characteristics of proband ND CNV carriers and non-carriers in the SSC.

	ND CNV status		**df**	**Test statistic**	**p**
	ND CNV carriers (n =63)	**Non CNV carriers (n = 1903)**			
Sex, M (%)	51 (81.0)	1656 (87.0)	1	1.47 [a]	0.23
Age, median (IQR)	8.34 (6.42-11.67)	8.76 (6.67-11.34)	-	62,272 [b]	0.60
VABS C, mean (SD)	71.52 (11.38)	72.10 (11.79)	66.5	0.39[c]	0.70

Note: ND CNV, neurodevelopmental copy number variant; VABS C, Vineland Adaptive Behaviour Scale composite score; IQR, interquartile range; SD, standard deviation; df, degrees of freedom. [a]*Chi-square.* [b]*Mann-Whitney U test.* [c]*Welch two sample t-test*

2.3.1.3 Psychiatric phenotypes

Table 2-3 summarises the descriptive data for the psychiatric phenotypes analysed in the SSC proband sample, including sample numbers and the total counts of individuals presenting with clinical level symptoms on CBCL DSM-oriented scales (as discussed in section 2.2.2.2.1, psychiatric clinical risk groups were defined on CBCL DSM-oriented scales by a T-score above 65).

Table 2-3. Summary of psychiatric phenotypes analysed.

CBCL DSM-oriented scale	Total N (max = 1,966)	In clinical risk group N (%)
Affective Problems	1,962	756 (38.5)
Anxiety Problems	1,963	888 (45.2)
Somatic Problems	1,962	293 (14.9)
ADHD Problems	1,961	765 (39.0)
ODD Problems	1,963	509 (25.9)
Conduct Problems	1,963	380 (19.4)

Note: ADHD, attention deficit hyperactivity disorder; ODD, oppositional defiant disorder.

Table 2-4 presents the proportions of individuals in psychiatric clinical risk groups in the SSC proband sample with classification according to ND CNV status and sex.

Table 2-4. Proportions of individuals in psychiatric clinical risk groups in SSC proband sample.

	ND CNV Carriers N in clinical risk group/ N total[1] (%)		Non-Carriers N in clinical risk group/ N total[1] (%)	
CBCL DSM-oriented scale	**Male**	**Female**	**Male**	**Female**
Affective Problems	12/50 (24.0%)	7/11 (63.6%)	640/1654 (38.7%)	97/247 (39.3%)
Anxiety Problems	21/50 (42.0%)	3/11 (27.3%)	763/1655 (46.1%)	101/247 (40.9%)
Somatic Problems	9/50 (18.0%)	0/11 (0%)	248/1654 (15.0%)	36/247 (14.6%)
ADHD Problems	23/50 (46.0%)	6/11 (54.5%)	634/1653 (38.4%)	102/247 (41.3%)
ODD Problems	14/50 (28.0%)	5/11 (45.6%)	427/1655 (25.8%)	63/247 (25.5%)
Conduct Problems	12/50 (24.0%)	5/11 (45.6%)	308/1655 (18.6%)	55/247 (22.3%)

Note: [1] N total refers to total in category with available phenotypic data. ADHD; attention deficit hyperactivity disorder; ODD, oppositional defiant disorder.

2.3.1.4 Phenotype Analysis Results

2.3.1.4.1 Affective Problems

ND CNV was not a significant predictor of affective problems as a main effect (Table 2-5).

Table 2-5. Multiple logistic regression model assessing ND CNV association with affective problems.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	0.408	0.382	1.010	0.285	1.50 (0.71-3.18)
ND CNV Status	-0.340	0.281	-1.211	0.226	0.71 (0.40-1.22)
Sex (Male)	-0.066	0.137	-0.478	0.633	0.94 (0.72-1.23)
VABS C	**-0.011**	**0.004**	**-2.706**	**0.007**	**0.99 (0.98-1.00)**
Age	0.535	0.015	0.004	0.997	1.00 (0.97-1.03)

Note: Significant variables highlighted in bold. VABS C- Vineland adaptive behaviour scale composite score. Nagelkerke pseudo r squared = 0.01.

There was a significant effect of sex as an interaction variable in the model (p=0.02) suggesting sex differences in affective problems associated with ND CNV carrier status (Table 2-6). Females with an ND CNV were more likely to be in the clinical risk group for affective problems compared to males with (OR 5.39, 95% CI 1.02 - 32.79)[3] or without (OR 2.70, 95% CI 1.04 – 8.00) an ND CNV. Females with ND CNVs did not have statistically significant higher odds of being in the clinical risk group for affective problems compared with females without ND CNVs (OR 2.70, 95% CI 0.79 – 10.57). The proportions of individuals in the affective problems clinical risk group are presented in Table 2-4, with classification according to ND CNV status and sex. Of eleven female carriers of ND CNVs in the sample, 7 (63.6%) were in the clinical risk group for affective problems. There were 50 male carriers of ND CNVs with affective phenotype data in the sample, 12 (24%) of whom were in the affective problems clinical risk group. The proportions of males and females in the affective

3 Chen emphasised the importance of careful interpretation of statistical interaction effects for accurate reporting of complex statistical models. Methods for correct interpretation of interaction effects from multiple logistic regression models are provided in (Chen, 2003). An example of the method used for calculation of the odds ratios for affective problems in males and females depending on ND CNV status is presented in Appendix 2: Supplementary Figures, Figure 7-1.

problems clinical risk group were similar (males - 38.7%, females - 39.3%) within the non-ND CNV carrier group (Figure 2-1).

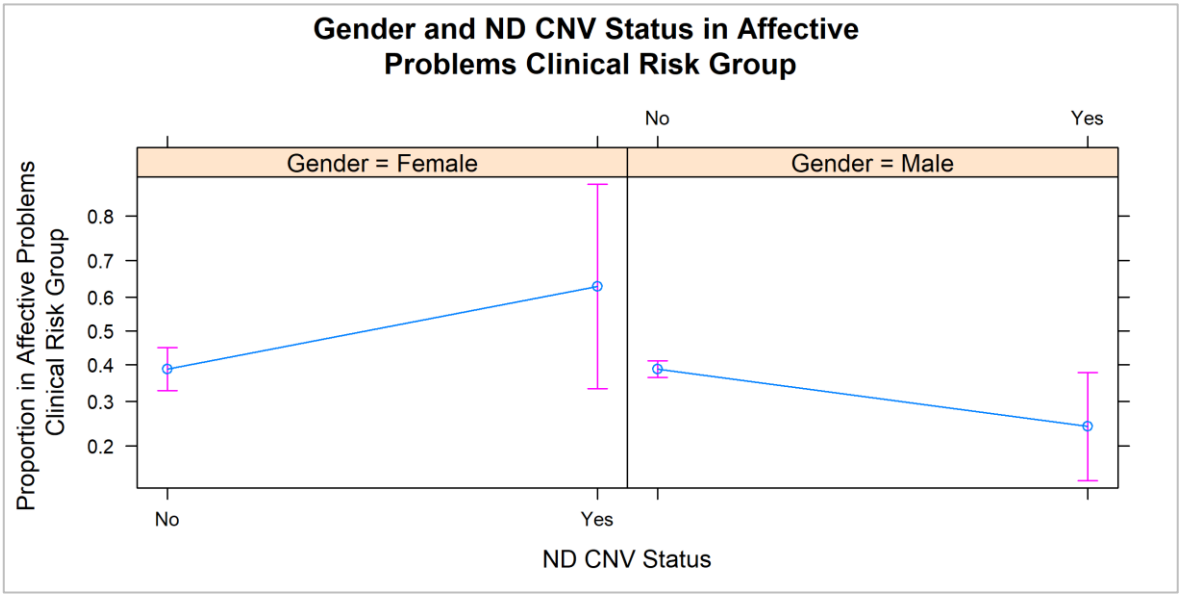
Odds of being in the clinical risk group for affective problems decreased with increase in level of adaptive behaviour (per point increase in VABS Composite Score, odds of affective problems decreased by OR 0.99 (95% CI 0.98-1.00), p = 0.007).

Table 2-6. Multiple logistic regression model assessing ND CNV association with affective problems including ND CNV status and sex interaction variable.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	0.342	0.383	0.893	0.372	1.41 (0.66-2.99)
ND CNV status	0.995	0.641	1.552	0.121	2.71 (0.79-10.57)
Sex (Male)	0.001	0.140	0.005	0.996	1.00 (0.76-1.32)
VABS C	**-0.011**	**0.004**	**-2.691**	**0.007**	**0.99 (0.98-1.00)**
Age	0.001	0.015	0.042	0.967	1.00 (0.97-1.03)
ND CNV status*Sex (Male)	**-1.687**	**0.724**	**-2.332**	**0.020**	**0.19 (0.04-0.74)**

Note: Significant variables highlighted in bold. VABS C- Vineland adaptive behaviour scale composite score. Nagelkerke pseudo r squared = 0.01.

Figure 2-1. Proportion of individuals in affective problems clinical risk group by ND CNV status and sex.



2.3.1.4.2 Anxiety Problems

ND CNV status was not significantly associated with anxiety problems in the main effects model (Table 2-7) and there was no significant interaction between ND CNV status and sex in association with anxiety problems (Table 2-8). Age predicted anxiety problems significantly; the odds of being in the clinical risk group for anxiety problems increased with each year (OR 1.03, 95% CI 1.00-1.06, p = 0.036 in the main effects model) (Table 2-7).

Table 2-7. Multiple logistic regression model assessing ND CNV association with anxiety problems.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-0.902	0.375	-2.405	0.016	0.41 (0.19-0.85)
ND CNV status	-0.227	0.267	-0.851	0.395	0.80 (0.47-1.34)
Sex (Male)	0.232	0.136	1.700	0.089	1.26 (0.97-1.65)
VABS C	0.003	0.004	0.785	0.433	1.00 (1.00-1.01)
Age	**0.031**	**0.015**	**2.102**	**0.036**	**1.03 (1.00-1.06)**

Note: Significant variables highlighted in bold. VABS C- Vineland adaptive behaviour scale composite score. Nagelkerke pseudo r squared = 0.01.

Table 2-8. Multiple logistic regression model assessing ND CNV association with anxiety problems including ND CNV status and sex interaction variable.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-0.886	0.376	-2.357	0.018	0.41 (0.20-0.86)
ND CNV status	-0.582	0.690	-0.844	0.399	0.56 (0.12-1.99)
Sex (Male)	0.217	0.139	1.559	0.119	1.24 (0.95-1.63)
VABS C	0.003	0.004	0.781	0.435	1.00 (1.00-1.01)
Age	**0.031**	**0.015**	**2.093**	**0.036**	**1.03 (1.00-1.06)**
ND CNV status*Sex (Male)	0.423	0.749	0.565	0.572	1.53 (0.38-7.76)

Note: Significant variables highlighted in bold. VABS C- Vineland adaptive behaviour scale composite score. Nagelkerke pseudo r squared = 0.01.

2.3.1.4.3 Somatic Problems

ND CNV status was not significantly associated with somatic problems in the sample. Sex, age and adaptive behaviour also were not significant predictors of somatic problems in the sample.

Table 2-9. Multiple logistic regression model assessing ND CNV association with somatic problems.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-1.332	0.520	-2.563	0.010	0.26 (0.09-0.73)
ND CNV status	-0.018	0.367	-0.050	0.960	0.98 (0.45-1.92)
Sex (Male)	0.092	0.193	0.477	0.633	1.10 (0.76-1.62)
VABS C	-0.004	0.006	-0.683	0.494	1.00 (0.99-1.01)
Age	-0.023	0.021	-1.104	0.270	0.98 (0.94-1.02)

Note: Significant variables highlighted in bold. VABS C- Vineland adaptive behaviour scale composite score. Nagelkerke pseudo r squared < 0.01.

2.3.1.4.4 Attention-deficit/Hyperactivity Problems

ND CNV status was not significantly associated with attention-deficit/hyperactivity problems in this sample. Adaptive behaviour and age were significantly associated with attention-deficit/hyperactivity problems. Odds of being in the clinical risk group for attention-deficit/hyperactivity problems decreased as VABS composite score increased (per point increase in VABS composite Score, odds decreased by OR 0.98 (95% CI 0.97-0.99, p =$1.38e^{-05}$). Odds of being in the clinical risk group for attention-deficit/hyperactivity problems also decreased with increasing age (per year increased age, odds decreased by OR 0.95 (95% CI 0.92-0.98, p = $6.93e^{-04}$) (Table 2-10).

Table 2-10. Multiple logistic regression model assessing ND CNV association with attention-deficit/hyperactivity problems.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	1.423	0.385	3.697	$2.18e^{-04}$	4.15 (1.96-8.84)
ND CNV Status	0.340	0.262	1.295	0.195	1.40 (0.84-2.35)
Sex (Male)	-0.110	0.137	-0.806	0.420	0.90 (0.69-1.17)
VABS C	**-0.018**	**0.004**	**-4.347**	**$1.38e^{-05}$**	**0.98 (0.97-0.99)**
Age	**-0.052**	**0.015**	**-3.393**	**$6.93e^{-04}$**	**0.95 (0.92-0.98)**

Note: Significant variables highlighted in bold. VABS C- Vineland adaptive behaviour scale composite score. Nagelkerke pseudo r squared = 0.02.

2.3.1.4.5 Oppositional Defiant Problems

ND CNV status was not significantly associated with odds of being in the clinical risk group for oppositional defiant problems. None of the other variables in the model predicted odds of having oppositional defiant problems (Table 2-11).

Table 2-11. Multiple logistic regression model assessing ND CNV association with oppositional defiant problems.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-0.697	0.423	-1.648	0.099	0.50 (0.22-1.14)
ND CNV status	0.253	0.281	0.897	0.370	1.29 (0.73-2.20)
Sex (Male)	-0.030	0.152	-0.195	0.846	0.97 (0.72-1.32)
VABS C	-0.001	0.005	-0.173	0.863	1.00 (0.99-1.01)
Age	-0.030	0.017	-1.789	0.074	0.97 (0.94-1.00)

Note: Significant variables highlighted in bold. VABS C- Vineland adaptive behaviour scale composite score. Nagelkerke pseudo r squared < 0.01.

2.3.1.4.6 Conduct problems

ND CNV status was not a significant predictor of odds of having conduct problems in the sample. Adaptive behaviour and age were both significant predictors of conduct disorder problems in the sample. Odds of being in the clinical risk group for conduct problems decreased as VABS composite score increased (per point increase in VABS composite Score, odds decreased by OR 0.98 (95% CI 0.97-0.99, $p = 4.96e^{-05}$). Odds of being in the clinical risk group for conduct problems also decreased with increasing age (per year increased age, odds decreased by OR 0.92 (95% CI 0.88-0.95, $p = 1.09e^{-05}$) (Table 2-12).

Table 2-12. Multiple logistic regression model assessing ND CNV association with conduct problems.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	1.036	0.467	2.218	0.027	2.82 (1.13-7.04)
ND CNV status	0.458	0.295	1.554	0.120	1.58 (0.86-2.77)
Sex (Male)	-0.252	0.162	-1.559	0.119	0.78 (0.57-1.07)
VABS C	**-0.021**	**0.005**	**-4.057**	**$4.96e^{-05}$**	**0.98 (0.97-0.99)**
Age	**-0.085**	**0.019**	**-4.398**	**$1.09e^{-05}$**	**0.92 (0.88-0.95)**

Note: Significant variables highlighted in bold. VABS C- Vineland adaptive behaviour scale composite score. Nagelkerke pseudo r squared = 0.03.

2.3.2 Sibling Analysis

Increased likelihood of clinical risk for affective problems was identified for females carrying ND CNVs in youths with ASD in the SSC sample. The association of ND CNV with affective problems was explored in the SSC sibling cohort. I hypothesised that there would also be an interaction effect of sex and ND CNV status with affective problems.

2.3.2.1 Characteristics of sample

The SSC sibling sample consisted of a total of 1634 siblings of youths with ASD. There was a lower proportion of males (n=750, 45.9%) than females in the sibling sample. As noted for the proband data, data collection for the SSC took place in stages, therefore the ages at which various instruments

were completed on youths varied. Youths in the sample ranged in age from approximately 4-18 years. The median age of youths in the sibling sample was 9.2 (IQR 7.1-12.0). Siblings were excluded from the sample if they presented with intellectual disability or a VABS composite score of less than 70. Mean VABS composite score for the sample was 103.3 (SD=11.2).

2.3.2.2 ND CNV carrier status

Neurodevelopmental copy number variants presented in n=18 (1.1%) of individuals in the sample. The ND CNVs identified in individuals in the SSC are presented in Appendix 1: Supplementary Tables,

Table 7-2. The characteristics of the sibling ND CNV carriers compared with non-carriers are presented in Table 2-13. ND CNV carriers and non-carriers did not differ significantly by sex, age or VABS composite score.

Table 2-13. Comparison of characteristics of sibling ND CNV carriers and non-carriers in the SSC.

	ND CNV status		**df**	**Test statistic**	**p**
	ND CNV carriers (n =18)	**Non CNV carriers (n = 1616)**			
Sex, M (%)	7 (38.9)	743 (46.0)	1	0.13[a]	0.72
Age, median (IQR)	7.71 (6.71-10.36)	9.18 (7.09-12.01)	-	17116[b]	0.20
VABS C, median (IQR)	103.00 (92.25-111.75)	102.00 (96.00-110.00)	-	14810[b]	0.87

Note: ND CNV, neurodevelopmental copy number variant; VABS C, Vineland adaptive behaviour scale composite score; IQR, interquartile range; df, degrees of freedom. [a] Chi square. [b] Mann Whitney U test

2.3.2.3 Psychiatric phenotypes

Affective problems presented in 5.5% of the SSC sibling sample, a much lower prevalence than in the proband sample where 38.5% of individuals presented in the affective problems clinical risk

group. Table 2-14 presents the proportions of individuals in the affective clinical risk group in the SSC sibling sample with classification according to ND CNV status and sex.

Table 2-14. Proportions of individuals in affective clinical risk group in SSC sibling sample.

	ND CNV Carriers N in clinical risk group/ N total[1] (%)		Non-Carriers N in clinical risk group/ N total[1] (%)	
CBCL DSM-oriented scale	**Male**	**Female**	**Male**	**Female**
Affective Problems	0/7 (0%)	3/11 (27.3%)	45/741 (6.1%)	42/872 (4.8%)

Note:[1] N total refers to total in category with available phenotypic data.

2.3.2.4 Phenotype Analysis Results

2.3.2.4.1 Affective problems

None of the n = 7 males with ND CNVs in the sibling sample were in the clinical risk group for affective problems prohibiting estimation of meaningful parameters for the ND CNV and sex interaction variable for this outcome. However, omitting the ND CNV and sex interaction variable identified a main effect of ND CNV status for odds of affective problems in the sample (Table 2-15). ND CNV carriers were more likely to be in the clinical risk group for affective problems (OR 4.14, 95% CI 1.15-14.83, p = 0.029) than non-carriers. It is notable that this result is based on a small sample size of ND CNV carriers (ND CNV carriers=18). The effect was largely attributable to three of 11 (27.3%) female sibling ND CNV carriers who were in the clinical risk group for affective problems (Table 2-14).

Adaptive behaviour and age were also significant predictors of odds of having affective problems in the sibling sample. Odds of being in the clinical risk group for affective problems decreased with increasing level of adaptive behaviour (per point increase in VABS composite Score, odds of affective problems decreased by OR 0.98 (95% CI 0.96-0.99, p = 0.012). Odds of being in the clinical risk group for affective problems increased in the sibling sample with age (per year increased age, odds increased by OR 1.12, (95% CI 1.05-1.19, p = $4.23e^{-04}$) (Table 2-15).

Table 2-15. Multiple logistic regression model assessing ND CNV main effect association with affective problems in the SSC sibling sample.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-1.521	1.091	-1.394	0.163	0.22 (0.03-1.86)
ND CNV status	**1.420**	**0.651**	**2.181**	**0.029**	**4.14 (1.15-14.83)**
Sex (Male)	0.115	0.222	0.517	0.605	1.12 (0.73-1.73)
VABS C	**-0.025**	**0.010**	**-2.509**	**0.012**	**0.98 (0.96-0.99)**
Age	**0.114**	**0.032**	**3.526**	**$4.23e^{-04}$**	**1.12 (1.05-1.19)**

Note: Significant variables highlighted in bold. VABS- Vineland adaptive behaviour scale. Nagelkerke pseudo r squared = 0.05.

2.4 Discussion

A substantial proportion of youths with ASD (~15%) are diagnosed with pathogenic CNVs on routine CMA testing. Although a body of knowledge is developing regarding psychiatric risk associated with ND CNVs, there is a significant deficit in the understanding of psychiatric risk associated with ND CNVs in youths with ASD and in particular whether risk arises in addition to ASD as an outcome. An improved understanding of psychiatric risk profiles associated with ND CNVs in youths with ASD can inform genetic counselling and optimal surveillance of psychiatric risk in these youths.

In this analysis of youths with ASD, sex differences were identified in the effect of ND CNV status on odds of having affective problems. Females with an ND CNV were more likely to present in a clinical risk group for affective problems compared with males with or without ND CNVs. Previous studies have also reported evidence of sex-specific associations between CNVs and depression outcomes. Martin et al. reported that females with anxiety or depression were more likely to carry large, rare CNVs than males in a population sample of youths (Martin et al., 2019). Kendall et al. observed higher rates of depression among adult female carriers of ND CNVs than among adult male carriers (Kendall et al., 2019). In the study described by Kendall et al., individuals with ASD and other NDDs were removed from the sample, indicating that the association between ND CNVs and depression may be independent of neurodevelopmental co-morbidity. In a sample of ASD-unaffected siblings assessed here, ND CNV status was significantly associated with increased likelihood of having affective problems (OR 4.14, $p = 0.029$); this effect was driven by female carriers of ND CNVs only. Large confidence intervals were associated with the effects of ND CNVs in this analysis, due to small sample sizes of carriers of ND CNVs in both cohorts.

No evidence was identified of association between ND CNVs and other psychopathological outcomes, including anxiety, somatic, attention-deficit/hyperactivity, oppositional defiant or conduct problems in youths with ASD.

The high rates of psychopathology reported in the youths with ASD in this sample are striking (Table 2-3). Rates of anxiety problems, affective problems and attention-deficit/hyperactivity problems were particularly high at 45.2%, 38.5% and 39.0% respectively. These high rates of comorbidities are consistent with previous studies (Bitsika et al., 2015; Gjevik et al., 2011; Simonoff et al., 2008; Stevens et al., 2016; Strang et al., 2012; van Steensel et al., 2011) and emphasise the burden of mental health problems experienced by this group of young people.

2.4.1 Limitations

2.4.1.1 Sample size and representativeness

A limitation that frequently presents in phenotype analyses of ND CNVs is the low prevalence of these genotypes in populations, even in neurodevelopmental cohorts. In the samples analysed in this study 63 ND CNV carriers (3.2%) were identified in the sample of ASD probands and 18 ND CNV carriers (1.1%) were identified in the sample of ASD-unaffected siblings. The small sample sizes of ND CNV carriers limit the power to detect associations with increased risk of psychopathologies, if such associations exist.

A lack of association was identified in the analysis between ND CNVs and anxiety, somatic, attention-deficit/hyperactive, oppositional defiant and conduct problems. These findings may reflect a true lack of association between these variables. If associations do exist however, larger sample size would be required to facilitate identification of associations.

The SSC was designed for simplex pedigrees specifically and therefore may not be generalizable to the ASD population as a whole. Siblings of children with ASD are known to be at increased risk of having psychiatric disorders (Jokiranta-Olkoniemi et al., 2016) and are therefore not representative of a general population of ASD-unaffected youths.

2.4.1.2 Secondary analysis of a retrospective cohort

The data used in this analysis was obtained from a retrospective cohort collection. A wealth of data was collected for the Simons Simplex Collection, which has yielded opportunities for many research groups to conduct a wide range of interesting studies on the data. However, one of the limitations of secondary analysis of existing datasets is that the data cannot be tailored to the aims of every research study.

In this analysis, psychiatric phenotypes were the primary outcome for analysis. The measure used to identify psychopathology in this analysis was the Child Behaviour Checklist (CBCL), a screening questionnaire which is used to identify those who are likely to be at clinical risk for psychiatric disorders but does not provide a reliable diagnosis of disorder. The CBCL has limitations in terms of accuracy for identifying psychiatric disorders in children with ASD (Gjevik et al., 2015). Identification of affective disorders has been reported to be good, but limitations have been described in specificity for identifying ADHD, ODD and anxiety disorders. Additionally, sensitivity for anxiety disorders is reported to be poor (Gjevik et al., 2015). These limitations may have contributed to heterogeneity in the outcomes and may have compromised the power of the analysis to detect associations.

The use of more comprehensive structured or semi-structured diagnostic interviews adapted for youths with ASD such as the Autism Comorbidity Interview (ACI) (Leyfer et al., 2006) may facilitate reliability in determining clinical cases improving homogeneity in samples, which may facilitate identification of any existing associations with ND CNVs.

2.4.1.3 Psychiatric phenotypic outcomes

Several of the psychiatric phenotypic outcomes in the analysis were correlated, ranging from the lowest correlation between attention-deficit/hyperactivity problems and somatic problems ($r_t = 0.23$) to the highest between Oppositional defiant disorder problems and conduct problems ($r_t = 0.76$). Co-occurrence of psychiatric symptoms and optimal classification of psychiatric disorder is a well-known and widely-discussed issue in psychiatric clinical practice (Pincus et al., 2004) and research (Batstra et al., 2002; Newman et al., 1998). The current approach of the most widely used classification systems in psychiatric clinical practice (ICD (World Health, 2004) and DSM (American Psychiatric Association, 2013; *Diagnostic and statistical manual of mental disorders : DSM-IV*, 1994)) is to classify psychiatric disorders as discrete disorders characterised by particular symptom sets. As the six DSM-oriented scale outcomes in this analysis are recognised as separate diagnostic entities in current classification systems, these outcomes were assessed as independent outcomes in this analysis.

The outcomes assessed in this analysis were categorical dichotomous outcomes. Although categorization of variables is common in clinical research, dichotomization of variables can impact negatively on the power of the analysis (Royston et al., 2006). Dimensional approaches to research, focusing on measures that may present within and across disorders, such as the Research Domain Criteria (RDoC) project under development with the National Institute of Mental Health (NIMH) (Cuthbert, 2014) are an alternative approach to consider in examining phenotypic associations with ND CNVs in the future.

2.4.1.4 Power of analyses

Post hoc power analyses confirmed limitations in the power of this study to detect associations between ND CNV status and the psychiatric outcomes assessed in the SSC ASD proband sample. With the effect sizes identified in the analyses, the power for detecting associations varied from 5-37% in the outcomes assessed, indicating that much larger samples will be required to detect associations in future studies (Table 7-3). The analysis in the SSC sibling sample had power of 56% to detect associations, also indicating a need for larger samples to reliably identify associations (Table 7-3).

2.4.1.5 Exclusion criteria for sibling recruitment

The SSC protocol excluded siblings who had been diagnosed with any psychiatric disorder that required treatment with two or more psychotropic medications (Fischbach et al., 2010). Therefore, the siblings most affected by psychiatric disorders were not part of the SSC sample and psychiatric disorder was likely under-represented in the sibling sample. The SSC sibling sample likely does not generalise to all siblings of children with ASD, or to more general neurotypical populations of children.

Siblings were also excluded if they were identified as having mental retardation or an adaptive behaviour standard score on the Vineland-II of less than 70 (Fischbach et al., 2010). A well-established phenotype associated with neurodevelopmental CNVs is a negative impact on cognitive functioning (Kendall et al., 2019; Sanders et al., 2015; Stefansson et al., 2014). In excluding siblings with ID or an adaptive behaviour score of less than 70, individuals more likely to have neurodevelopmental CNVs may have been excluded from the sample. ND CNVs may have been under-represented in the sibling group.

2.4.1.6 Multiple testing

Corrections were not made for multiple testing in this analysis. The analysis was considered exploratory in nature and replication of the outcomes will be pursued in future work.

2.4.2 Impact

The findings from this study add further support to possible sex differences in the phenotypic effects of ND CNVs, indicating an increased likelihood of depression in females with ND CNVs in a sample of youths with ASD. An increased likelihood of having depressive symptomatology associated with ND CNVs was also identified in a sample of ASD unaffected siblings; driven by female ND CNV carriers. The small sample sizes of ND CNV carriers in this analysis preclude confident interpretation of the findings, however in the context of previous studies also reporting sex-specific phenotypic effects of ND CNVs (Kendall et al., 2019; Martin et al., 2019), the finding of this analysis warrant further investigation in larger cohorts of youths with ASD.

The potential implications of a sex-specific association between ND CNVs and affective problems in youths with ASD are important to consider. Depression is a disabling comorbidity for youths with ASD that can result in significant functional impairment and distress; depression is also a major risk factor for suicidal ideation in youths with ASD (Mayes et al., 2013). Diagnosis of depression in youths with ASD can be complicated by phenotypic overlap between the two conditions. Core features of ASD may mask depressive symptomatology and depression presentations may be atypical

in youths with ASD (Magnuson et al., 2011). These clinical challenges in identifying depression in youths with ASD can delay diagnosis and therapeutic input. In the sample of youths with ASD analysed here, 64% (n = 7/11) of the females with ND CNVs were in the clinical risk group for affective problems. Females with an ND CNV were significantly more likely to present with depressive symptoms than males with (OR 5.39, 95% CI 1.02 - 32.79) or without (OR 2.70, 95% CI 1.04 – 8.00) an ND CNV. An increased risk of depressive symptoms in females with ASD who carry ND CNVs could be important information for the individuals themselves and their families to have, as well as their clinicians. Early identification of depressive symptomatology in these young people could facilitate early diagnosis and access to appropriate services and interventions, potentially mitigating prolonged periods of functional impairment and distress.

If the finding of an association between ND CNVs and risk of depressive symptomatology in the sibling sample in this analysis is replicated in larger cohorts of youths, this could also have important implications. Youths who carry ND CNVs who do not present with ASD may be at increased risk of developing depression as well. Siblings of youths with ASD with ND CNVs may also undergo genetic testing for ND CNV carrier status in some contexts and those diagnosed with ND CNVs may be at increased risk of developing depression, even in the absence of an ASD diagnosis. This could facilitate early diagnosis and intervention for these young people as well. Establishing sex-specific effects will be important in ND CNV carriers who are unaffected by ASD, as females may be at increased risk in this group based on our preliminary findings and previous findings in child (Martin et al., 2019) and adult population studies (Kendall et al., 2019).

2.5 Conclusion

The data from this analysis indicate that for youths with ASD, females with ND CNVs may be at higher risk of presenting with depression symptomatology than males with or without ND CNVs. These findings require replication in larger datasets. If replicated, the findings would indicate a need for early targeted psychiatric screening of female ASD ND CNV carriers. Early detection of psychiatric disorders could facilitate earlier treatment and potentially limit associated morbidity. Analysis of a sample of ASD-unaffected siblings indicated an association between ND CNVs and depressive symptomatology; this association may also be caused a higher risk in females only, but this cannot be confirmed without exploration of the finding in larger cohorts of youths.

Longitudinal studies will be crucial to an improved understanding of psychiatric co-morbidity in youths with ASD and the relationship with ND CNV status. Longitudinal studies would allow the careful characterisation over time of symptoms of depression in youths with ASD who carry ND CNVs and who do not carry ND CNVs. This could yield insights into aetiological subtypes of depression in youths with ASD and could be helpful in disentangling the complex clinical overlap between ASD and depression.

Chapter 3: Investigation of psychiatric phenotypes associated with neurodevelopmental copy number variants in a large population-based clinical cohort of youths

3.1 Introduction

3.1.1 Background

Neurodevelopmental copy number variants (ND CNVs) are associated with risk of depression (Kendall et al., 2019; Stefansson et al., 2014) and suicidal ideation (SI) (Stefansson et al., 2014) in general adult populations. Significant associations between ND CNVs and important psychopathologies such as depression, anxiety and subclinical psychotic symptoms have not been seen in childhood population samples (Guyatt et al., 2018; Martin et al., 2019). Case control studies suggest that ND CNV carriers (adult and child) may be at substantially increased risk of a range of psychopathologies compared with controls (Chawner et al., 2019; Hanson et al., 2015; Niarchou et al., 2014). Furthermore it appears that reported increases in the risk of psychopathology are not solely mediated by neurodevelopmental comorbidities such as autism spectrum disorder (ASD) and intellectual disability (ID) (Hanson et al., 2015; Niarchou et al., 2014). Intriguingly, evidence suggests that sex may moderate depression and anxiety phenotypes associated with ND CNVs (Kendall et al., 2019; Martin et al., 2019).

There have been limited investigations of the association between ND CNVs and childhood onset psychiatric disorders to date. In chapter 2, I reported an association between ND CNV status and depression symptomatology in female autistic youth. Similarly, there was an association between ND CNV and depression symptomatology in siblings of autistic youth, indicating that ND CNVs may also be associated with depression risk in youths without ASD.

There is a significant need to understand psychiatric risk that may be associated with ND CNVs in more general populations of youths. ND CNVs have incomplete penetrance for neurodevelopmental disorders (NDDs) such as ASD and ID in childhood and many carriers may not have major NDDs. Identifying whether a wider spectrum of psychiatric risk is associated with ND CNV status will be relevant for these individuals and their families. Increasingly, cytogenetic microarray genetic testing is completed in a range of children with mild developmental delays presenting to developmental paediatrics, not only those with moderate to severe ID or ASD (Fan et al., 2018; Moeschler et al., 2014; Vissers et al., 2010). Positive test results may cause concern for parents, given the uncertainty of neurodevelopmental and neuropsychiatric outcomes. More concerning, in many contexts routine

antenatal testing includes screening for ND CNVs which has the potential to influence pregnancy decisions based on limited evidence (Govaerts et al., 2017). It is crucial therefore to provide accurate information regarding psychiatric risk and ND CNVs to inform antenatal (where relevant) and paediatric genetic counselling.

This analysis aimed to address deficits in our clinical understanding of the impact of ND CNVs on NDD and psychiatric outcomes by investigating these relationships in a large clinical population cohort (n = 8,205) of youths recruited through the Children's Hospital of Philadelphia (CHOP), the Philadelphia Neurodevelopmental Cohort (PNC).

3.1.2 Aims and Hypotheses

The first aim of this analysis was to assess the relationship between ND CNV carrier status and four major outcomes reflective of significant psychopathology (internalising disorders, externalising disorders, subclinical psychotic symptoms and suicidal ideation) in the PNC cohort. It was hypothesised that youths with ND CNVs would have higher rates of psychopathology compared with those without ND CNVs.

The second aim of the analysis was to assess sex differences in the effects of ND CNVs on internalising disorders. Based on prior observations of sex differences in the effects of ND CNVs on anxiety (Martin et al., 2019) and depression phenotypes (Kendall et al., 2019; Martin et al., 2019), it was hypothesised that there would be observable sex differences in the effects of ND CNV on internalising psychiatric disorders.

The third aim was to identify whether NDD comorbidity was a major determinant for increased risk of psychopathology associated with ND CNVs in a clinical cohort of youths. Increased risk of ASD and ID are well-established in ND CNV carriers (Coe et al., 2014; Kirov et al., 2014; Sanders et al., 2015) and these are in turn associated with increased risk of psychiatric disorders compared with neurotypical populations (Abdallah et al., 2011; Lugo-Marín et al., 2019; Skokauskas et al., 2012a) (Einfeld et al., 2011). Therefore, increased risk of psychopathology associated with ND CNVs may occur as a result of increased rates of NDDs in ND CNV carriers. However, a number of studies have indicated that increased risk of psychopathology associated with ND CNV status is not solely attributable to neurodevelopmental comorbidity (Kendall et al., 2019; Niarchou et al., 2014; Stefansson et al., 2014). In this analysis, it was hypothesised that increased risk of psychopathology associated with ND CNVs would be conferred separately to ASD or ID comorbidity.

These hypotheses were tested using genetic and phenotype data from a publicly available large cohort of youths (n=8,205), the Philadelphia Neurodevelopmental Cohort (PNC).

3.2 Materials and methods

3.2.1 Sample

3.2.1.1 The Philadelphia Neurodevelopmental Cohort

The PNC is a population-based prospectively ascertained clinical sample of 8-21 year old youths (Calkins et al., 2015). The data in the PNC was collected through a collaboration between the Children's Hospital of Philadelphia (CHOP) and the University of Pennsylvania. The data collected were made publicly available through the National Institute of Mental Health's Database of Genotypes and Phenotypes (dbGaP) and were downloaded for use in this study from the dbGaP website, under phs000607.v3.p2 (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v3.p2). Data access approval available in Appendix 3: Data Access Approval, Figure 7-6.

Participants for the PNC were recruited by the Center for Applied Genomics at CHOP between 2006 and 2012 through a paediatric healthcare network in community sites in Pennsylvania, New Jersey and Delaware. The team recruited 50,293 youths who were genotyped and who provided written informed consent (for participants aged ≥ 18 years) and written assent and permission from a parent or legal guardian (for those aged < 18 years) for re-contact for future studies and authorizing access to their Electronic Medical Records (EMR). The majority of individuals (70%) were ascertained through outpatient medicine and 30% through preoperative surgery clinics. Participants were not recruited through psychiatric clinics (Merikangas et al., 2015b). The EMRs of the participants were reviewed for preliminary eligibility for the PNC study. Potential participants were included if they were between 8-21 years, had provided written informed consent/assent to be re-contacted for future studies, were proficient in the English language and would be able to complete the study procedures including computerized neurocognitive testing. Youths with significant developmental delays or physical conditions that would impair their ability to complete study procedures (including significant hearing loss) were not invited to participate in the study (Calkins et al., 2015; Merikangas et al., 2015b). As a result of the screening procedure, 19,161 youths were eligible for further participation in the study. Participants were enrolled on the study between November 2009 and December 2011. Participants were contacted via a letter introducing the study and were subsequently contacted by phone. The University of Pennsylvania and CHOP Institutional Review Boards approved the procedures associated with the PNC (Calkins et al., 2015). Of the 19,161 individuals who were screened as eligible for further participation, 13,598 were invited to participate further, 9,498 were enrolled and 9,421 completed the assessment. The recruitment flow chart and demographic details of the recruitment pool (n=19,161) of the PNC sample are available from Calkins et al. (Calkins et al., 2015) and are presented in Appendix 2: Supplementary Figures, Figure 7-2 and Appendix 1: Supplementary Tables, Table 7-4 respectively.

3.2.1.2 Phenotype Measures

3.2.1.2.1 Psychopathology Assessment in the PNC

Clinical phenotypic data was collected in the PNC through a computerised, structured interview-named "GOASSESS", an assessment developed from a modified version of the epidemiologic version of the NIMH Genetic Epidemiology Research Branch Kiddie-Schedule for Affective Disorders and Schizophrenia (K-SADS) (Calkins et al., 2015; Merikangas et al., 2009b). The K-SADS is a semi-structured interview designed to assess DSM-IV disorders. It consists of a screening interview covering a range of disorders and further interviews with questions relating to specific disorders. The K-SADS is widely used in clinical practice and research of psychiatric disorders in children (Kaufman et al., 1997). The K-SADS is also used in studies to assess comorbid psychiatric disorders in children with ASD (Gjevik et al., 2015). The GOASSESS interview was modified from the K-SADS with the aims of allowing rapid training and standardization across many assessors and allowing brief administration in order to facilitate high throughput of assessments (100-165 participants per week). Modifications of the K-SADS for the GOASSESS included the addition of diagnostic screening questions from the adolescent version of the WHO Composite International Diagnostic Interview (CIDI) to the start of each psychopathological domain, the inclusion of dimensional ratings of distress and impairment associated with symptoms and some variations from the K-SADS in the response options for symptoms (Calkins et al., 2015). The GOASSESS interview was administered to the youth and/or caregivers to collect clinical information. Additional response variables in the GOASSESS interview included demographics, medical and medication history, the Children's Global Assessment Scale (Shaffer et al., 1983), and interviewer observations.

The psychopathology assessment addressed psychiatric and psychological treatment history and lifetime occurrence of psychopathological domains including mood (major depressive episode, manic episode), anxiety (generalised anxiety disorder, separation anxiety disorder, specific phobia, social phobia, panic disorder, agoraphobia, obsessive compulsive disorder, agoraphobia, post-traumatic stress disorder), attention deficit hyperactivity disorder (ADHD), behavioural disorders (oppositional defiant disorder, conduct disorder), eating disorders (anorexia nervosa, bulimia nervosa) and death wish and suicidal ideation. The dataset included screening questions for relevant disorders and additional data consistent with Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria, including further relevant symptoms, duration of symptoms, symptom frequency. The severity of the relevant disorder was rated in terms of associated distress and/or functional impairment (each was measured on an individual 11-point Likert scale ranging from 0 (no bother/problems) to 10 (extremely serious bother/problems)) (Calkins et al., 2015). A sample of the GOASSESS proband screener section is provided by Calkins et al. is presented in Appendix 2: Supplementary Figures, Figure 7-3.

The psychosis spectrum was assessed via three screening tools in the GOASSESS programme. Positive subpsychotic symptoms in the last year were assessed using the 12-item PRIME Screen-Revised (Kobayashi et al., 2008; Miller, 2004), positive threshold psychotic symptoms were assessed with the K-SADS psychosis screen and negative/disorganised symptoms were assessed using six assessor-rated subscales of the Scale of Prodromal Symptoms (SOPS) from the Structured Interview for Prodromal Symptoms (SIPS) (McGlashan, 2003).

Further details on the training and interrater reliability of the PNC interviewers are provided by Calkins et al. in (Calkins et al., 2015).

3.2.1.2.2 Additional Measures in the PNC

The Wide Range Achievement Test (WRAT-4) measures the basic academic skills of reading, arithmetic and spelling (Wilkinson et al., 2006). Participants in the PNC completed the WRAT-4 Reading Subscale test in a Computerized Neurocognitive Battery (CNB) (Gur et al., 2012). The WRAT-4 Reading Subscale has a phenotypic correlation of ~0.4 with full-scale intelligent quotient (Shafee et al., 2018) and was used in the PNC to provide an estimate of IQ (Calkins et al., 2015). Standard scores for the WRAT-4 subtests range from 55-145 and have a mean of 100 and standard deviation of 15 (Wilkinson et al., 2006).

Autism spectrum disorder and pervasive developmental disorder diagnosis were assessed in the medical history section of the PNC with the question: “Autism or Pervasive Developmental Disorder - Do/did you have this problem?” (modified to be appropriate for proband or collateral questionnaire).

History of death wish and suicidal ideation were assessed with the questions: “Suicide: Have you ever thought a lot about death or dying?” and “Suicide: Have you ever thought about killing yourself?” (modified to be appropriate for proband or collateral questionnaire).

3.2.2 Clinical Phenotype Algorithms

The phenotypic data from the PNC provided through dbGaP was obtained in raw, uncoded format. A sample of the phenotype data available through dbGaP is available in - Appendix 1: Supplementary Tables, Table 7-5, showing the variables that were provided for ADHD. Clinical algorithms for the scoring of the GOASSESS data were not provided from the dbGaP dataset and are not included in PNC published literature. Contact was made with the research group who designed the GOASSESS and collected data for the PNC, however scoring algorithms could not be obtained from the group.

Scoring algorithms were created to identify participants with presentations consistent with clinical disorders as per DSM IV criteria or the PRIME Revised criteria for subclinical positive psychotic symptoms. Development of the clinical algorithms is described in sections 3.2.2.1 and 3.2.2.2.

3.2.2.1 DSM-IV Based Clinical Scoring Algorithms

Clinical scoring algorithms were developed for the mental health disorders presented in Table 3-1. Raw phenotype data were aligned with DSM IV criteria. A sample of the raw phenotype data for ADHD are presented in Appendix 1: Supplementary Tables, Table 7-5. The first step of each of the clinical algorithms integrated the screening questions for the disorder of the GOASSESS assessment. Further steps in the algorithms related to aspects of DSM-IV criteria including symptom frequency, duration of symptoms and episodes and level of associated distress and/or functional impairment. Computerised algorithms were applied to phenotypic data to generate diagnostic outcomes.

Diagnostic outcomes were dichotomous with "Yes" indicating that the individual positively endorsed all available DSM-IV criteria for the specific mental health disorder, strongly indicating that they met criteria for the disorder. "No" indicated that they negatively endorsed at least one DSM-IV criterion for the disorder, therefore the individual did not meet criteria for the diagnosis. "NA" indicated that at least one DSM-IV criterion was "unknown" or "NA" for the disorder, therefore the individual could not be conclusively identified as meeting or not meeting DSM-IV criteria for the disorder. The diagnostic algorithm for agoraphobia is presented in Appendix 2: Supplementary Figures, Figure 7-4 for illustration of the algorithmic processes. Proportions of outcomes for each disorder are presented in Appendix 1: Supplementary Tables, Table 7-6.

Table 3-1. Mental health disorders and diagnostic criteria used to formulate clinical scoring algorithms.

Mental Health Disorder	Diagnostic criteria
Anxiety Disorders	Diagnostic And Statistical Manual of Mental Disorders: DSM-IV (*Diagnostic and statistical manual of mental disorders : DSM-IV*, 1994)
Agoraphobia	
Obsessive Compulsive Disorder	
Generalised Anxiety Disorder	
Panic Disorder	
Post-Traumatic Stress Disorder	
Separation anxiety disorder	
Social phobia	
Affective disorders	
Major depression (Major depressive episode)	
Manic episode	
Behaviour Disorders	
Conduct Disorder	
Oppositional defiant disorder	
Attention Deficit Hyperactivity Disorder	
Predominantly inattentive type	
Predominantly hyperactive/impulsive type	
Psychotic Disorders	PRIME Screen Revised (Kobayashi et al., 2008; Miller, 2004)
Subclinical positive psychotic symptoms	

Separate scoring algorithms were created for adults (age ≥ 18 years) and children (age < 18 years) for a number of disorders (generalised anxiety disorder, social phobia and specific phobia) as the DSM-IV criteria differ for adults and children. The age-relevant algorithms for these disorders were applied to the appropriate groups in the cohort (children or adults) and observations were then classified into aggregate "Yes", "No" and "NA" groups for these disorders.

3.2.2.2 PRIME Screen-Revised Scoring Algorithm

The PRIME screen was developed by the Prevention through Risk Identification, Management, and Education (PRIME) group at Yale University as a screening instrument to identify individuals at risk of developing psychosis (Miller, 2004). The PRIME screen is a short self-administered questionnaire

based on the positive symptom portion of the Structured Interview for Psychosis-Risk Syndromes (SIPS), a more comprehensive psychosis risk screening tool (Miller et al., 2003).

There are twelve questions in the PRIME screen identifying presence or absence of: unusual thought content, delusional ideas, suspiciousness, persecutory ideas, grandiose ideas, perceptual abnormalities, hallucinations and loss of insight. Responses are measured on a Likert-scale of 0 (definitely disagree) to 6 (definitely agree) with a response of 'not sure' being 3. The PRIME screen has been shown to have high sensitivity, specificity and negative predictive value compared with other screening tools for identifying individuals at clinical high risk of psychosis (Addington et al., 2015).

Kobayashi et al. modified the PRIME screen by adding a "duration of symptoms" section to the PRIME Screen, developing the PRIME Screen-Revised (PS-R) to improve the specificity of the instrument in identifying prodromal or early onset psychosis in general populations (Kobayashi et al., 2008). The last item of the PRIME Screen was excluded in the PS-R to improve consistency as it did not refer to attenuated positive symptoms. The PS-R was validated in clinical and university populations of youths. Specificity and sensitivity of the PS-R, using the SIPS as a gold standard, were 0.74 and 1.00 respectively (Kobayashi et al., 2008).

Outcomes of the PS-R were classified into 11 levels by Kobayashi et al., integrating severity of symptoms, duration of symptoms and the total score of the PS-R (Table 3-2). Subjects with a rank of 4 or over were regarded as screening positive (Kobayashi et al., 2008). There were four permutations which resulted in receiving a rank of 4 or above:

- Selected one or more "definitely agree" response with a duration of more than one year ***or***
- Selected two or more "definitely agree" responses without regard to the duration ***or***
- Selected two or more "somewhat agree" responses with durations of more than one year ***or***
- Have a total PS-R score of 39 or over

Table 3-2. Severity ranking of the PRIME Screen-Revised scores by Kobayashi et al. (Kobayashi et al., 2008).

Rank	Definition
10	Selected three or more "definitely agree" responses with durations of more than one year
9	Selected two "definitely agree" responses with durations of more than one year
8	Selected two "definitely agree" responses with durations of more than one year or selected two or more "definitely agree" responses without regard to the duration and one or more "somewhat agree" response with a duration of more than one year
7	Selected one "definitely agree" response with a duration of more than one year *and* one or more "somewhat agree" response with a duration of more than one year
6	Selected two or more "definitely agree" responses without regard to the duration *or* selected three or more "somewhat agree" responses with durations of more than one year
5	selected one "definitely agree" response with a duration of more than one year *or* selected two "somewhat agree" response with durations of more than one year
4	Have a total PS-R score of 39 or over
3	Selected one "definitely agree" response without regard to the duration *or* selected one "somewhat agree" response with a duration of more than one year
2	Selected one or more "somewhat agree" response without regard to the duration *or* selected one or more "slightly agree" response with a duration of more than one year
1	Selected one or more "slightly agree" response without regard to the duration
0	Not selected any kind of "agree" response

The PS-R measure was used in this analysis to identify youths with subclinical psychotic symptoms. A scoring algorithm was formulated based on the criteria developed by Kobayashi et al. (Kobayashi et al., 2008) as described above. Outcomes were dichotomised, either meeting criteria by scoring above a rank of 4 on the PS-R (outcome "Yes") or not (outcome "No"). If relevant questions were unanswered, outcomes were classified as "NA". Proportions of outcomes are presented in Table 3-4.

3.2.3 Clinical Phenotype Categories

The statistical power of retrospective cohort studies of phenotypic associations with ND CNVs are often limited due to the low proportions of ND CNV carriers in samples. This was an important consideration in the planning of this study and in selecting the clinical phenotype categories for analysis. It was a priority to select clinical outcomes that were relevant to youths and to maximise statistical power for the analysis while minimising multiple testing of outcomes.

The majority of psychiatric diagnoses have been shown to converge onto internalising, externalising and psychotic dimensions in community and clinical samples (Kessler et al., 2011; Kotov et al., 2010; Kotov et al., 2017). The American Psychiatric Association has endorsed internalising and externalising groupings in the DSM-V (American Psychiatric Association, 2013). The internalising group are described as representing disorders with prominent anxiety, depressive and somatic symptoms and the externalising group as representing disorders with prominent impulsive, disruptive conduct and substance use symptoms. One goal of grouping the disorders in this way is to facilitate research on the genetic and neurobiological aetiologies of these dimensions as covariation/clustering of the disorders may occur as a result of shared underlying disease processes (Achenbach et al., 2016; American Psychiatric Association, 2013; Kotov et al., 2010). Within each group, the sharing of genetic and environmental risk factors have been cited as likely to explain comorbidities within the groups, at both a clinical and population level (American Psychiatric Association, 2013).

The majority of factor analytic studies examining internalising and externalising categories of mental health conditions have focussed on relatively common disorders in general populations and/or on participants who are relatively well (Kotov et al., 2010). Few studies have examined psychosis-spectrum disorders (Kotov et al., 2010) in this context. Kotov et al. assessed two samples, an inpatient sample with psychosis (Kotov et al., 2010) and a general outpatient sample (Kotov et al., 2011) and replicated the finding of the two dimensions of internalising and externalising groups in these clinical populations and identified a third dimension consistent with psychosis-spectrum disorders (Kotov et al., 2010; Kotov et al., 2011).

Analysis of internalising disorder, externalising disorder and subclinical psychotic symptom (dichotomous) outcomes were selected in this analysis, maximising statistical power and reducing the number of outcomes assessed overall (compared with assessing each diagnosis individually). Suicidal ideation was also selected as an outcome for analysis due to the clinical importance of this outcome (Rufino et al., 2019). An association has previously been identified between ND CNVs and SI in an adult population sample (Stefansson et al., 2014).

3.2.3.1 Internalising disorders

The internalising disorders variable was comprised of anxiety disorders and depression (symptoms consistent with a major depressive episode). The internalising variable was a categorical variable with a dichotomous outcome. Individuals who met criteria for at least one internalising disorder were classified into a "Yes" category. Those who obtained a "No" outcome for all internalising disorders were classified into a "No" category. Participants who could not be definitively classified into a "Yes" or "No" category were categorised as "NA" and were excluded from the analysis for the outcome. To illustrate this scoring system, the possible outcomes for the internalising disorders variable are presented in Appendix 1: Supplementary Tables, Table 7-7.

3.2.3.2 Externalising disorders

The externalising disorders variable was comprised of attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD) and conduct disorder and was a categorical variable with a dichotomous outcome. Individuals presenting with reported symptoms consistent with a diagnosis of an externalising disorder based on the diagnostic algorithms were classified into a "Yes" category. Those who obtained a "No" outcome for all externalising disorders were classified into a "No" category. Participants who could not be definitively classified into a "Yes" or "No" category were categorised as "NA" and were excluded from analysis for the outcome.

3.2.3.3 Subclinical psychotic symptoms

The subclinical psychotic symptom variable was formulated based on the PS-R as described in section 3.2.2.2. The variable was dichotomous, with individuals classified into "Yes" or "No" outcomes on the basis of the PS-R rank score. Participants who could not be definitively classified into a "Yes" or "No" category were categorised as "NA" and were excluded from analysis for the outcome.

3.2.3.4 Suicidal ideation

History of death wish and suicidal ideation were assessed with forced choice yes or no responses to two probe questions in the PNC (modified to be appropriate for proband or collateral questionnaire):

- "Suicide: Have you ever thought a lot about death or dying?"
- "Suicide: Have you ever thought about killing yourself?"

In this analysis, the probe “Have you ever thought about killing yourself?” was selected for analysis of suicidal ideation in the sample as this probe was more specific to suicidal ideation than the first probe. Individuals who did not provide an answer were excluded from the analysis for the outcome.

3.2.4 CNV selection

The same list of neurodevelopmentally-associated CNVs discussed in Chapter 2, section 2.2.3 was used for this analysis. In brief, this consisted of a list of 69 “neurodevelopmental” copy number variants that were previously implicated in autism spectrum disorder (Sanders et al., 2015), intellectual disability (Coe et al., 2014) and schizophrenia (Marshall et al., 2017). Full details of the ND CNV are described in Chapter 2, section 2.2.3 and all ND CNV loci are presented in Appendix 1: Supplementary Tables, Table 7-1**.**

3.2.5 CNV data

Subjects were genotyped using Illumina SNP-array platform at the Center for Applied Genomics at The Children' s Hospital of Philadelphia. Details of genotyping methods are described in (Glessner et al., 2010). Individuals were called as carriers of ND CNVs if there was greater than 50% overlap with one of the genomic loci of interest and there was agreement across at least two CNV calling platforms including PennCNV, QuantiSNP and dnaCopy (Xu et al., 2013). No CNVs at the 2p16.3 region met the criteria of at least 50% overlap due to the relatively small size of this CNV region. This is a known ND CNV that typically affects the neurexin1 gene (Nrxn1 deletions). Smaller CNVs in the Nrxn 1 region have previously been reported to be associated with neurodevelopmental disorders (Priebe et al., 2013; Rujescu et al., 2009a) and clinical microarrays have included additional probes for detection across this region (Kirov et al., 2009b). Hence the criteria for 2p16.3 calling in this analysis were relaxed. Criteria used in a previous analysis for detection of Nrxn1 CNVs required deletions to intersect at least one exon and duplications to cover the whole gene (Kendall et al., 2016). In this analysis, at least 10% overlap was required in the Nrxn1 region. Again, agreement on two of the CNV calling platforms was required.

Subjects were cross-referenced to identify those with genotype and phenotype data available which resulted in a total of n=8,205 individuals.

3.2.6 Data analysis

Multiple logistic regression models were used to test associations between ND CNVs and the four selected phenotypic outcomes: internalising disorders, externalising disorders, subclinical psychotic symptoms and suicidal ideation. Sex, age, ASD status and WRAT standardised score were included

as covariates in the models. For each of the regression models, the variance inflation factor (VIF) and tolerance statistics were examined to check for significant multicollinearity in the data (Cohen et al., 2013). The VIF values were below 10 and tolerance statistics were above 0.1, indicating that multicollinearity was not a significant issue in the models (Table 7-9, Table 7-10, Table 7-11, Table 7-12). Predictor coefficients were tested using Wald tests and confidence intervals obtained using the Wald method. Model fit was assessed using Nagelkerke pseudo R square index.

Based on previous evidence suggesting sex differences in depression and anxiety phenotypic associations with ND CNVs (Kendall et al., 2019; Martin et al., 2019), an interaction variable (sex*NDD CNV status) was added to the model assessing ND CNV association with internalising disorders, to examine for differential effects of ND CNVs on phenotypic outcomes between sexes.

Statistical significance was set at $p < 0.05$. The results presented are not corrected for multiple comparisons. All analyses were completed in R version 3.2.3 ("R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.," 2013).

3.3 Results

3.3.1 Characteristics of sample

The sample consisted of a total of n = 8,205 individuals. There was a slight preponderance of females (50.9%) in the PNC total sample. The median age at clinical assessment in the sample was 14.0 years (IQR 10.0-17.0). WRAT standardised scores ranged from 55.0 to 145. The median WRAT standardised score in the sample was 101.0 (IQR 92.0-112.0). A total of 283 individuals (3.4%) were reported as having autism spectrum disorder or pervasive developmental disorder in the sample.

3.3.1.1 Frequency of ND CNVs and characteristics of ND CNV carriers

A total of 158 (1.9%) carriers of ND CNV were identified in the sample. The ND CNVs identified in individuals are presented in Appendix 1: Supplementary Tables, Table 7-8.

The characteristics of the ND CNV carriers compared with non-carriers are presented in Table 3-3. There were no significant differences in sex distribution or mean age at clinical assessment between the ND CNV carriers and non-carriers. ND CNV carriers and non-carriers differed significantly in terms of estimated cognitive level, with carriers presenting with a median WRAT standard score 6 points lower than the non-carriers ($p=1.17e^{-05}$) which was expected given the well-established impact of ND CNVs on general cognitive functioning (Kendall et al., 2019; Sanders et al., 2015; Stefansson

et al., 2014). ASD presented at a higher frequency in ND CNV carriers (6.3%) compared to non-carriers (3.4%), but the difference was not statistically significant (p = 0.08) in the sample.

Table 3-3. Comparison of characteristics of ND CNV carriers and non-carriers in the PNC.

	ND CNV status		Test		
			df	statistic	p
	ND CNV carriers (n = 158)	Non CNV carriers (n = 8047)			
Sex, M (%)	84 (53.2)	3935 (48.9)	1	1.23[a]	0.27
Age, median (IQR)	14.0 (10.0-17.0)	14.0 (10.0-17.0)	-	635,000[b]	0.98
WRAT std, median (IQR)	95.0 (87.0-105.0)	101.0 (92.0-112.0)	-	759,000[b]	$1.17e^{-05}$
ASD present (%)	10 (6.3)	273 (3.4)	1	3.09[a]	0.08

Note: ND CNV, neurodevelopmental copy number variant; WRAT std, Wide ranging achievement test standardized score; ASD, autism spectrum disorder. Sd, standard deviation; df, degrees of freedom. [a]Chi-square, [b]Mann-Whitney U test

3.3.2 Psychopathology in PNC sample

The prevalence rates of the psychopathological categories analysed are presented in Table 3-4. The internalising and externalising variables were composed from specified DSM-IV disorders (as discussed in section 3.2.3.1, 3.2.3.2). To be classified as a "Yes" outcome for internalising or externalising disorders, participants met criteria for at least one relevant DSM-IV disorder. To be classified as a "No" outcome, participants definitively did not meet criteria for the relevant DSM-IV disorders. Participants who could not be classified into "Yes" or "No" outcomes were classified as "NA" and were excluded from analysis for the outcome. Prevalence of internalising disorders and externalising disorders was similar- 28.6% and 28.4% respectively. The proportion of missing data for internalising disorders was sizeable at 28.8%.

Subclinical psychotic symptoms presented in 7.3% of the sample. A positive history of suicidal ideation was reported by 7.9% of individuals in the sample.

Table 3-4. Prevalence of psychopathology outcomes in PNC.

Psychopathology outcome	No N (%)	Yes N (%)	NA N (%)
Internalising disorders	3,493 (42.6)	2,348 (28.6)	2,364 (28.8)
Externalising disorders	5,579 (68.0)	2,334 (28.4)	292 (3.6)
Subclinical psychotic symptoms	6,590 (80.3)	599 (7.3)	1,016 (12.4)
Suicidal ideation	7,458 (90.9)	649 (7.9)	98 (1.2)

Table 3-5 presents the proportions of individuals with positive presentations of the psychopathologies examined in the PNC with classification according to ND CNV status.

Table 3-5. Psychiatric phenotypic outcomes in ND CNV carriers and non-carriers in PNC (Total sample).

	N with positive history/ N total[1] (%)	
Psychiatric Outcome	**ND CNV Carriers**	**Non-Carriers**
Internalising disorders	48/112 (42.9%)	2300/5729 (40.1%)
Externalising disorders	49/154 (31.8%)	2285/7759 (29.4%)
Subclinical psychotic symptoms	21/135 (15.6%)	578/7054 (8.2%)
Suicidal ideation	13/155 (8.4%)	636/7952 (8.0%)

Note:[1] N total refers to total in category with available phenotypic data.

Correlations were observed between all of the variables of interest, however these outcomes are also individually associated with specific risk factors and psychiatric presentations (Hedley et al., 2018; Kelleher et al., 2011; Ormel et al., 2005; Segers et al., 2014; Williams et al., 2009; Woodman et al., 2016) and therefore were analysed as individual outcomes.

Table 3-6. Tetrachoric correlations between the psychopathology outcomes in PNC.

	Internalising disorders	Externalising disorders	Subclinical psychotic symptoms	Suicidal ideation
Internalising disorders	-			
Externalising disorders	0.58	-		
Subclinical psychotic symptoms	0.52	0.35	-	
Suicidal ideation	0.64	0.44	0.36	-

Note: Correlation coefficients measured by tetrachoric correlation coefficient t_r (El-Hashash et al., 2018), calculated using "tetrachoric" function in the "psych" package in R (W, 2019).

3.3.3 Phenotype analysis results

3.3.3.1 Internalising Disorders

ND CNV status was not a significant predictor of internalising disorders in the PNC sample (p = 0.683). Sex, age, estimated cognitive level and ASD status were significant predictors of internalising disorders in the sample. Males were less likely to have internalising disorders (OR 0.62, 95% CI 0.55-0.69, $p < 2e^{-16}$) compared with females. Likelihood of having an internalising disorder increased with age (per year of increased age, odds increased by OR 1.14 (95% CI 1.13-1.16, $p < 2e^{-16}$) and decreased with increasing estimated cognitive level (per point of increased WRAT standardised score, likelihood decreased by OR 0.99 (95% CI 0.99-0.99, $p = 3.91e^{-08}$). Youths with ASD were more likely to have internalising disorders (OR 4.28, 95% CI 3.10-5.94, $p < 2e^{-16}$) compared to youths without ASD (Table 3-7).

Table 3-7. Multiple logistic regression model assessing ND CNV association with internalising disorders.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-1.079	0.224	-4.823	$1.42e^{-06}$	0.34 (0.22-0.53)
ND CNV (Yes)	0.096	0.204	0.471	0.683	1.10 (0.74-1.64)
Sex (Male)	**-0.484**	**0.057**	**-8.511**	**$<2e^{-16}$**	**0.62 (0.55-0.69)**
Age	**0.135**	**0.008**	**17.178**	**$<2e^{-16}$**	**1.14 (1.13-1.16)**
WRAT Std Score	**-0.010**	**0.002**	**-5.495**	**$3.91e^{-08}$**	**0.99 (0.99-0.99)**
ASD Status	**1.453**	**0.166**	**8.772**	**$<2e^{-16}$**	**4.28 (3.10-5.94)**

Note: Significant variables highlighted in bold. WRAT std score- Wide range achievement test standardised score. Nagelkerke pseudo r squared = 0.13.

Sex differences in the effect of ND CNV status on risk of internalising disorders were examined by adding an interaction variable to the model. There was no evidence of an interaction effect between ND CNV status and sex on risk of internalising disorders (p = 0.807) (Table 3-8).

Table 3-8. Multiple logistic regression model assessing ND CNV association with internalising disorders including ND CNV status and sex interaction variable.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-1.077	0.224	-4.811	$1.50e^{-06}$	0.34 (0.22-0.53)
ND CNV (Yes)	0.041	0.305	0.133	0.894	1.04 (0.57-1.90)
Sex (Male)	**-0.486**	**0.057**	**-8.463**	**$<2e^{-16}$**	**0.62 (0.55-0.69)**
Age	**0.135**	**0.008**	**17.173**	**$<2e^{-16}$**	**1.14 (1.13-1.16)**
WRAT Std Score	**-0.010**	**0.002**	**-5.498**	**$3.84e^{-08}$**	**0.99 (0.99-0.99)**
ASD Status	**1.453**	**0.166**	**8.771**	**$<2e^{-16}$**	**4.27 (3.10-5.94)**
ND CNV (Yes)*Sex (Male)	0.100	0.408	0.244	0.807	1.10 (0.49-2.46)

Note: Significant variables highlighted in bold. WRAT std score- Wide range achievement test standardized score. Nagelkerke pseudo r squared = 0.13.

3.3.3.2 Externalising Disorders

ND CNV status was not a significant predictor of externalising disorders in the PNC sample (p = 0.732). Sex, age, estimated cognitive level and ASD status were significant predictors of externalising disorders in the sample. Males were more likely to have externalising disorders (OR 1.35, 95% CI 1.22-1.50, p = $5.16e^{-09}$) compared with females. Likelihood of having an externalising disorder increased with age (per year of increased age, risk increased by OR 1.03 (95% CI 1.01-1.04, p = $8.32e^{-05}$) and decreased with increasing estimated cognitive level (per point of increased WRAT standardised score, risk decreased by OR 0.98 (95% CI 0.97-0.98, p < $2e^{-16}$). Youths with ASD were more likely to have externalising disorders (OR 3.69, 95% CI 2.83-4.82, p <2e-16) compared to youths without ASD (Table 3-9).

Table 3-9. Multiple logistic regression model assessing ND CNV association with externalising disorders.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	1.065	0.209	5.085	$3.68e^{-07}$	2.90 (1.93-4.38)
ND CNV (Yes)	-0.062	0.183	-0.342	0.732	0.94 (0.65-1.34)
Sex (Male)	**0.301**	**0.052**	**5.842**	**$5.16e^{-09}$**	**1.35 (1.22-1.50)**
Age	**0.028**	**0.007**	**3.935**	**$8.32e^{-05}$**	**1.03 (1.01-1.04)**
WRAT Std Score	**-0.025**	**0.002**	**-14.563**	**<$2e^{-16}$**	**0.98 (0.97-0.98)**
ASD Status	**1.307**	**0.135**	**9.690**	**<$2e^{-16}$**	**3.69 (2.83-4.82)**

Note: Significant variables highlighted in bold. WRAT std score- Wide range achievement test standardized score. Nagelkerke pseudo r squared = 0.09.

3.3.3.3 Subclinical Psychotic Symptoms

ND CNV status was significantly associated with risk of subclinical positive psychotic symptoms in the PNC sample. Carriers of ND CNVs were more likely to have subclinical psychotic symptoms than non-carriers (OR 1.87, 95% CI 1.13-2.95, p = 0.01) (Table 3-10). A total of 15.6% (21/135) of youths with ND CNVs had subclinical positive psychotic symptoms, compared with 8.2% (578/7,054) (Appendix 1: Supplementary Tables, Table 3-5).

Estimated cognitive level and ASD status were also associated with risk of subclinical psychotic symptoms. As estimated cognitive level decreased, likelihood of having subclinical psychotic

symptoms increased (per point of increased WRAT standardised score, likelihood of having subclinical psychotic symptoms decreased by OR 0.98 (95% CI 0.98-0.99, p = $4.87e^{-10}$). Youths with ASD were more likely to have subclinical psychotic symptoms (OR 2.18, 95% CI 1.49-3.10, p = $3.10e^{-05}$) than youths without ASD.

Table 3-10. Multiple logistic regression model assessing ND CNV association with subclinical psychotic symptoms.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-0.990	0.353	-2.807	$5.0e^{-03}$	0.37 (0.19-0.74)
ND CNV (Yes)	**0.626**	**0.245**	**2.559**	**0.01**	**1.87 (1.13-2.95)**
Sex (Male)	0.175	0.088	1.989	0.05	1.19 (1.00-1.41)
Age	0.016	0.012	1.359	0.174	1.02 (0.99-1.04)
WRAT Std Score	**-0.018**	**0.003**	**-6.223**	**$4.87e^{-10}$**	**0.98 (0.98-0.99)**
ASD Status	**0.778**	**0.187**	**4.166**	**$3.10e^{-05}$**	**2.18 (1.49-3.10)**

Note: Significant variables highlighted in bold. WRAT std score- Wide range achievement test standardized score. Nagelkerke pseudo r squared = 0.04.

To identify whether ASD or likely ID comorbidity were the main determinants of the increased risk of subclinical psychotic symptoms in individuals with ND CNVs, we re-examined the relationship in a sample excluding those with reported ASD and with likely ID. Individuals with likely ID were defined as those with a WRAT standardised score of at least two standard deviations below the mean WRAT standardised score of 100 (WRAT std score < 70). This resulted in a sample of n = 7,736 individuals. A total of 142 (1.8%) individuals in this sample carried an ND CNV. Sample characteristics and comparisons of ND CNV carriers and non-carriers in this sample are presented in Appendix 1: Supplementary Tables, Table 7-13, Table 7-14).

ND CNV carrier status was associated with subclinical psychotic symptoms in the sample excluding individuals with ASD and likely ID. Youths with ND CNVs were more likely to have subclinical psychotic symptoms (OR 1.90, 95% CI 1.10-3.09, p = 0.01) than youths without ND CNVs in this sample (Table 3-11). Of the ND CNV carriers, 14.9% (18/121) had subclinical psychotic symptoms, compared to 7.8% (523/6683) of the non-carriers (Appendix 1: Supplementary Tables, Table 7-15). In this sample, estimated cognitive level remained significantly associated with risk of subclinical psychotic symptoms, with odds of psychotic symptoms decreased by OR 0.98 (95% CI 0.98-0.99, p = $8.49e^{-09}$) per point of increased WRAT standardised score.

Table 3-11. Multiple logistic regression model assessing ND CNV association with subclinical psychotic symptoms in sample with individuals with ASD and likely ID excluded.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-0.774	0.380	-2.038	0.04	0.46 (0.22-0.97)
ND CNV (Yes)	**0.641**	**0.262**	**2.451**	**0.01**	**1.90 (1.10-3.09)**
Sex (Male)	-0.178	0.091	-1.968	0.05	0.84 (0.70-1.00)
Age	0.016	0.012	1.273	0.20	1.02 (0.99-1.04)
WRAT Std Score	**-0.018**	**0.003**	**-5.758**	**$8.49e^{-09}$**	**0.98 (0.98-0.99)**

Note: Significant variables highlighted in bold. WRAT std score, Wide ranging achievement test standardized score. Nagelkerke pseudo r squared = 0.02.

3.3.3.4 Suicidal Ideation

ND CNV status was not a significant predictor of risk of suicidal ideation in the PNC sample. Risk of having suicidal ideation was associated with age and with ASD status. Youths with ASD were 3.6 times more likely to have suicidal ideation (95% CI 2.47-5.04, $p = 2.24e^{-12}$) than youths without ASD. Risk of suicidal ideation increased with increasing age (increase of OR 1.18 (95% CI 1.15-1.21, $p < 2e^{-16}$ per year of increased age).

Table 3-12. Multiple logistic regression model assessing ND CNV association with suicidal ideation.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-4.928	0.353	-13.958	$<2e^{-16}$	0.01 (0.00-0.01)
ND CNV (Yes)	0.037	0.298	0.126	0.90	1.04 (0.55-1.79)
Sex (Male)	-0.155	0.086	-1.808	0.07	0.86 (0.72-1.01)
Age	**0.166**	**0.012**	**13.343**	**$<2e^{-16}$**	**1.18 (1.15-1.21)**
WRAT Std Score	0.001	0.002	0.223	0.82	1.00 (1.00-1.01)
ASD Status	**1.272**	**0.181**	**7.019**	**$2.24e^{-12}$**	**3.57 (2.47-5.04)**

Note: Significant variables highlighted in bold. WRAT std score- Wide range achievement test standardized score. Nagelkerke pseudo r squared = 0.08.

3.4 Discussion

It is important to identify if ND CNVs increase psychiatric risk in clinical youth cohorts to provide greater certainty regarding outcomes to patients and families. ND CNVs have incomplete penetrance for NDDs such as ASD and ID in childhood and many carriers may not have major NDDs, however independently they may be more at risk of other disabling psychiatric disorders. ND CNVs are known to be associated with risk of schizophrenia (Marshall et al., 2017), depression (Kendall et al., 2019; Stefansson et al., 2014) and SI (Stefansson et al., 2014) in adult populations. There is also evidence from case control studies that ND CNV carriers may be at substantially increased risk of a range of psychopathologies including anxiety, affective and behavioural disorders compared with controls (Chawner et al., 2019; Hanson et al., 2015; Niarchou et al., 2014). Studies of childhood population samples have not to date identified significant associations between ND CNVs and psychopathologies such as depression, anxiety and subclinical psychotic symptoms (Guyatt et al., 2018; Martin et al., 2019); a common limitation in these studies have been small samples of ND CNV carriers. Greater certainties regarding these risks has implications for screening, diagnosis and management of psychiatric presentations in ND CNV carriers and for more accurate genetic counselling (Govaerts et al., 2017) given reduced penetrance of these outcomes.

In this analysis, the relationships between ND CNVs and four important child and adolescent mental health outcomes were investigated in a large clinical cohort of youths. ND CNV status was significantly associated with risk of subclinical psychotic symptoms in the sample. Youths with ND

CNVs were 1.9 times (p = 0.01) more likely to have subclinical psychotic symptoms than those without ND CNVs. The model testing this association indicated that ASD status and an estimate of cognitive ability (WRAT std score) were also associated with risk of subclinical psychotic symptoms. The effect of ND CNV status was also assessed in a sample that excluded individuals with ASD and likely ID (individuals with WRAT std score < 70). The association between ND CNV status and subclinical psychotic symptoms remained significant when individuals with ASD and with likely ID were excluded from the sample (OR 1.9, p = 0.01), suggesting that carriers of ND CNVs may still be at increased risk of having subclinical psychotic symptoms even if they do not have ASD or suspected intellectual disability. Notably, cognitive ability continued to predict subclinical psychotic symptoms significantly in the model excluding individuals with ASD and suspected ID. It is possible that, even for ND CNV carriers whose cognitive ability is within normal range, the association with subclinical psychotic symptoms is mediated by level of cognitive ability. This analysis was underpowered to be able to examine this possibility effectively.

Although associations between ND CNVs and schizophrenia are well established (Kirov, 2015), population and case-control studies of youths have not to date identified significant associations between ND CNVs and subclinical psychotic symptoms. Guyatt et al. found no evidence of association between schizophrenia-associated CNVs and psychotic experiences in a childhood population sample (n = 6,807), postulating that this outcome may have been attributable to a lack of power in the study or may have presented as a result of psychotic experiences in childhood and adolescence being more strongly attributable to environmental factors (e.g. childhood trauma, substance misuse) than to genetic predisposition to psychosis (Guyatt et al., 2018). In a case-control study of ND CNV carriers (n = 258) compared with sibling controls (n = 106), Chawner et al. found that subclinical psychotic experiences were present in ND CNV carriers, but the prevalence was not significantly increased compared with controls (Chawner et al., 2019). The rate of psychotic-like experiences was found to be similar between children with 22q11.2DS compared with control siblings in a study by Niarchou et al. (Niarchou et al., 2014).

This study provides novel evidence of an association between ND CNVs and subclinical psychotic symptoms in a large clinical cohort of youths. There are several possible explanations for the association identified. Prior studies showed that young people with psychotic symptoms were at increased risk for psychotic disorder in adulthood (Poulton et al., 2000; Welham et al., 2009b). Individuals with psychotic symptoms were also more likely to present with non-psychotic psychopathology, particularly depression in adults (Nishida et al., 2008; Scott et al., 2009; Wigman et al., 2011) and youths (Kelleher et al., 2012). Subclinical psychotic symptoms in this cohort may be related to severe and comorbid non-psychotic psychopathology. It is also possible that subclinical psychotic symptoms in ND CNV carriers may be a precursor to later development of psychotic disorder. The finding of ND CNV association with increased risk of subclinical psychotic symptoms requires further exploration in longitudinal cohorts to elucidate clinical and prognostic implications.

ND CNV status was not significantly associated with risk of internalising disorders in this analysis, a variable including anxiety and depression outcomes. Moreover, there was no significant interaction effect between ND CNV status and sex for risk of internalising disorders. The reported association between ND CNVs and internalising disorders more generally is variable, depending on age group studied (adult, child samples), study type (population studies, case control studies) and specific disorder (depression, anxiety). In the analysis described in chapter 2, ND CNV status was a significant predictor for affective problems in autistic females in a youth cohort and for affective problems in unaffected siblings. Depression is associated with ND CNV carrier status in adults but anxiety is not (Kendall et al., 2019; Stefansson et al., 2014). Neither anxiety nor depression were associated with ND CNV carrier status in children (Martin et al., 2019). Case control studies have however indicated increased risk in ND CNV carriers of internalising disorders (Reddy et al., 2007), anxiety (Chawner et al., 2019; Niarchou et al., 2014; Royston et al., 2017) and affective problems (Hanson et al., 2015) in samples of youths.

The lack of association between ND CNVs and internalising disorders in this study may be attributable to several possible factors. Phenotype measurement methods based on the algorithms I developed may have excluded "false negatives" disproportionately from the clinically affected groups impacting negatively on the power to detect association (discussed further in Limitations section 3.4.1.1). The analysis was also likely affected by the substantial proportion of missing data present within the internalising disorder phenotype; 28.8% of individuals were classified as "NA" due to missing phenotype data resulting in failure to definitively classify them (discussed further in Limitations section 3.4.1.2). The analysis of broad internalising and externalising psychopathological categories also presents with limitations. The rationale behind examining internalising and externalising disorders as hierarchical categories here was to try to limit phenotypic overlap between categories examined, while also reducing multiple testing in the analysis. The disadvantage of examining hierarchical categories is that if there are more subtle, specific phenotypic features associated with ND CNVs (that may be better represented by subcategories such as anxiety and depression diagnoses), analysis of hierarchical categories will not facilitate identification of such associations. Finally, limitations in statistical power due to insufficient sample size of ND CNV carriers may have contributed to the findings in this analysis.

ND CNV status was not significantly associated with risk of externalising disorders in this analysis. Externalising disorders, ADHD, ODD and conduct disorder have been significantly associated with specific neurodevelopmental CNVs in case control studies (Chawner et al., 2019; Hanson et al., 2015; Niarchou et al., 2014; Skokauskas et al., 2012b). There are few studies of externalising disorder associations with ND CNVs in large cohorts of youths. Guyatt et al. were unable to examine for association between schizophrenia-associated CNVs and ADHD traits in the Avon Longitudinal Study of Parents and Children (ALSPAC) population due to insufficient case numbers (Guyatt et al., 2018). The lack of association between ND CNVs and externalising disorders in this analysis may

be reflective of a true lack of association in general clinical populations of youths. If a true association does exist and was not identified in this analysis, factors including phenotype measurement, missing data, the general nature of hierarchical categorical classification and insufficient sample size may have contributed to the finding.

ND CNV status was not significantly associated with risk of having suicidal ideation (SI) in this sample. This may represent a true lack of association or may present as a result of lack of statistical power to identify associations due to limited sample size of ND CNV carriers. Discussion of the other variables (age, sex, estimate of cognitive level, ASD status) assessed in this study in relation to SI (a known predictor of suicide attempts (Wichstrøm, 2000)) are important in the context of increasing rates of suicide attempts and deaths in youths and an urgent need to better understand risk factors that contribute to suicide risk in this group (Rufino et al., 2019). A notable finding was the association between ASD and SI in the sample. Risk of SI was substantially increased in those who reported having an ASD (OR 3.57, $p = 2.24e^{-12}$). Systematic review has previously identified high rates of suicidality ~11–66% in ASD (adult and child) samples (Hedley et al., 2018). Rates of suicidal ideation in adults with ASD have been reported to be significantly higher than general adult populations (Cassidy et al., 2014; Cassidy et al., 2017) and there is an increased risk of premature death by suicide in people with ASD (Hedley et al., 2018; Hirvikoski et al., 2016; Kirby et al., 2019). Evidence is accumulating of increased prevalence of suicidal ideation and behaviour in children and adolescents with ASD compared with typically developing populations (Chen et al., 2017) (Aitken et al., 2016; Hunsche et al., 2020), however overall there is a dearth of studies on suicide risk in autistic youth. It is also notable in this analysis that autistic youths presented with increased risk of internalising disorders, externalising disorders and subclinical psychotic symptoms compared with non-autistic youth. Previous studies identified psychiatric co-morbidities as potential risk factors for suicidality in autistic youths (Horowitz et al., 2018; Oliphant et al., 2020), however further research is needed to better understand the relationship between ASD, psychiatric comorbidity and suicide risk (Oliphant et al., 2020). Some or all of the increased likelihood of having suicidal ideation in youths with ASD in this sample may have been mediated by psychiatric co-morbidities; analysis of this was beyond the scope of this study but future investigation of this would be important. The current findings contribute further evidence of higher risk of SI in youths with ASD and highlight the high rates of mental health problems experienced by these individuals.

Risk of SI increased with increasing age in this sample (OR 1.18, $p < 2e^{-16}$ per year of increased age). Previous studies have also identified increases in SI associated with increasing age (Peter et al., 2008; Rueter et al., 2005). Sex and estimate of cognitive level were not significant predictors of risk of having SI in this sample. Previous findings indicated that female youths presented with higher rates of SI compared with males (Beautrais, 2002; Souza et al., 2010); in this analysis, males were less likely to have SI than females, but this did not reach statistical significance ($p = 0.07$). A previous

study exploring the relationship between IQ and SI in adolescents did not find evidence of significant association between SI and the WRAT3 (Alati et al., 2009).

3.4.1 Limitations

3.4.1.1 Psychiatric assessment algorithms

Clinical phenotypic data was collected in the PNC through a computerised, structured interview (GOASSESS) which was based on a modified version of the K-SADS, a semi-structured interview designed to assess DSM-IV disorders (Calkins et al., 2015; Merikangas et al., 2009b). The GOASSESS interview was modified from the K-SADS with the aims of allowing rapid training and standardization across many assessors and allowing brief administration in order to facilitate high throughput of assessments. Modifications included the addition of diagnostic screening questions from another structured psychiatric interview, the inclusion of dimensional ratings of distress and impairment associated with symptoms and some variations from the K-SADS in the response options for symptoms (Calkins et al., 2015). The structural validity of the GOASSESS and psychosis spectrum tools were explored with factor analyses by the group that created it and were reported to be good. However, they noted that the structured and abbreviated format of the assessment tool may reduce sensitivity to clinically significant symptoms (Calkins et al., 2015).

I constructed scoring algorithms for this study to identify participants with presentations consistent with clinical disorders. The clinical algorithms were constructed based on DSM-IV criteria for internalising and externalising disorders and PRIME screen revised for subclinical psychotic symptoms. The aim of the approach used in constructing the clinical algorithms was to be consistent with psychiatric diagnostic classification systems used in clinical practice, where dichotomous diagnostic outcomes are used.

Current psychiatric nosological systems although clinically practical, present limitations in many areas of psychiatric research and clinical practice. For clinical use, diagnostic thresholds have been defined to differentiate between “normality” and “disorder” in systems such as ICD and DSM, however this can result in diagnostic errors that are relevant in clinical and research settings. As stated by Vella et al: *“The reality is that wherever the diagnostic threshold is set, diagnostic errors will still occur, with both false positive and negative diagnoses”* (Vella et al., 2013). It has been argued that false positives may be more harmful than false negatives in clinical practice (Vella et al., 2013) and extensive efforts have been made in classification systems to deal with “the false positives problem” (Spitzer et al., 1999). False negative diagnoses have been identified as a significant issue arising from strict utilisation of DSM-IV criteria (Spitzer et al., 1999). The scoring algorithms for this analysis were developed based on DSM-IV criteria. Individuals who were categorised as having a “Yes” outcome on the scoring algorithms in this analysis were likely to be true positives as they

clearly endorse *all* included criteria for the disorder. Negative endorsement of *any* criterion meant that an individual was classified in the "No" outcome; this may have resulted in some false negatives in this category (Robert L. Spitzer et al., 1999). Access to medical records to verify clinical diagnoses was not possible.

In genetic research precision and accuracy in the definition and measurement of phenotype groups (minimising false positives and false negatives in the "case" group) directly impact the ability to detect genetic associations (Gage et al., 2018). False negative cases may have impacted on the power to detect significant associations with ND CNVs in this analysis.

3.4.1.2 Missing data

There was substantial variability in the proportions of missing data for different DSM-IV disorders in the PNC (Table 7-6). Although contact was made with researchers involved in the collection of the data for the PNC, the reasons for missingness and variability between DSM-IV disorders in the PNC have not been reported.

The scoring criteria for the algorithms may also have contributed to the high proportions of "NA" outcomes for some of the psychopathology outcome variables. If individuals did not answer or answered "unknown" for any component of the DSM-IV diagnostic criteria for a disorder, it was deemed that they could not be reliably classified as meeting criteria or not and they were therefore categorised as "NA" for the outcome. The highest proportion (28.8%) of observations classified as "NA" was in the internalising disorders variable (Table 3-4). Outcomes from the analyses of this variable must be interpreted with caution.

Missing values for the outcome variables may have reduced statistical power in this analysis and, depending on the reasons for missing data, may have biased the estimates of the analysis (Kang, 2013). Consideration was given to utilising multiple imputation techniques for dealing with the missing data but since it has not been possible to date to confirm the cause of the missing data, proceeding with multiple imputation was not appropriate for the analysis.

In order to resolve this issue for future researchers who wish to use the PNC sample, I am now involved with a working group headed by Dr Kathleen Merikangas, Senior Investigator and Chief of the Genetic Epidemiology Research Branch in the Intramural Research Program at the National Institute of Mental Health (NIMH) to derive a final standardised set of clinical algorithms for the PNC cohort to enhance the usability of the data by researchers accessing these data. We plan to make the code for the algorithms freely available to other research groups to support further analysis of the dataset.

3.4.1.3 Dichotomous outcomes

In this sample, I constructed computerised clinical scoring algorithms, processing psychiatric phenotypic data to identify those in clinical risk groups. These clinical outcome variables were dichotomous. The data used in this analysis were collected based on a computerised version of the K-SADS interview (Orvaschel et al., 1987). The instrument provides dichotomous outcome variables and does not provide a quantitative measure of symptoms or severity. Although dichotomisation is intuitively favourable for medical researchers coming from a background where clinical diagnosis is generally based on dichotomous classification, dichotomisation results in a loss of power to detect relationships (Royston et al., 2006). Dichotomisation of phenotypic outcomes further limited power in this analysis. In the working group established for the development of standardised clinical scoring algorithms for the PNC data (discussed in section 3.4.1.2 above), development of a categorical scoring approach with ordinal outcomes for the psychopathology measures is under consideration with the aim of improving power for further studies examining associations with psychopathology in the PNC.

3.4.1.4 Sample size

The sample size of ND CNV carriers in this analysis may have limited the detection of effects of ND CNV on psychopathological outcomes. ND CNVs are rare genetic variants, individually presenting at frequencies of <0.3% in general populations (Kirov et al., 2014). The youth clinical cohort sample that was assessed here was relatively large (n=8,205), but the proportion of ND CNV carriers in the sample was 1.9% (n=158). Childhood cohort studies to date have been limited by relatively small samples (Guyatt et al., 2018; Martin et al., 2019) in their capacity to clearly establish relationships between ND CNVs and psychopathologies, as was likely the case for this analysis. Adult population samples that have previously identified significant associations between ND CNVs and psychopathologies (Kendall et al., 2019; Stefansson et al., 2014) have analysed samples of more than 100,000 individuals (Kendall et al., 2019), (Stefansson et al., 2014).

Much larger cohorts of youths are likely to be required for studies such as this one to reliably identify associations between ND CNVs and psychopathologies.

3.4.1.5 Power of analyses

Post hoc power analyses were carried out to evaluate the power of the phenotype analyses in the PNC sample. The power of the analyses ranged from 6% to 74% (Table 7-16), based on the effect sizes identified in the analyses. Much larger samples are required to effectively assess associations between ND CNV status and internalising symptoms, externalising symptoms and suicidal ideation in future studies of youth populations.

3.4.1.6 Clinical population

Although efforts were made through the recruitment approach used in collecting the PNC (Calkins et al., 2015), the sample was recruited through paediatric clinics. Psychiatric disorders are prevalent in youths with physical health problems and associations have been identified between severity of physical health conditions and mental health disorders (Merikangas et al., 2015b). Psychiatric disorders may be over-represented in this sample compared with general populations of youths. This limits the generalisability of the findings in terms of general youth populations.

3.4.1.7 Multiple testing

Corrections were not made for multiple testing in this analysis. This analysis was considered preliminary in nature and replication of the outcomes in further datasets is required.

3.4.2 Impact

Identifying whether youths with ND CNVs are at increased risk of developing psychopathologies is important as early screening and identification in those at risk may facilitate early diagnosis and intervention resulting in improved outcomes (Dadds et al., 1997; Snell et al., 2013) (Aos et al., 2004; Beitchman et al., 1992a; Beitchman et al., 1992b). Understanding the impact of ASD or ID co-morbidity and the role of sex in the relationship between childhood onset psychiatric disorder and ND CNVs can contribute to informed psychiatric screening and intervention strategies for carriers.

In this study, we identified children with ND CNVs to be at significantly increased risk of presenting with subclinical psychotic symptomatology. The effect of ND CNVs on risk of subclinical psychotic symptoms was present both in a sample including individuals with ASD and likely ID and in a sub-sample that excluded them. These findings require replication in further datasets for reliable interpretation, however there are important factors to consider based on these results. Screening for subclinical psychotic symptoms may be important in child and adolescent carriers of ND CNVs (including carriers with NDDs such as ASD/ID and carriers without these NDDs). Young people who present with subclinical psychotic symptoms are at higher risk of developing psychosis later in life (Poulton et al., 2000; Welham et al., 2009b) but psychotic like experiences may also be indicators of current severe non-psychotic psychopathology and co-morbidity (Nishida et al., 2008; Scott et al., 2009; Wigman et al., 2011). For ND CNV carriers in whom subclinical psychotic symptoms are identified, referral for further psychiatric assessment and follow up may be helpful for further assessment and monitoring. Further clinical characterisation of subclinical psychotic symptoms in ND CNV carriers in longitudinal cohorts will be helpful in exploring the clinical and prognostic significance of this association. ND CNV status should be considered as a candidate predictor variable for studies of transition to psychosis in youths at clinical high risk.

Although a lack of association was identified between ND CNVs and risk of internalising and externalising disorders and suicidal ideation in this study, previous case-control studies and adult population studies have identified significant associations between ND CNVs and various internalising and externalising disorders as well as suicidal ideation. Further exploration of psychiatric risk profiles in the context of ND CNVs will need to be explored in larger, well phenotyped cohorts of youths to provide more conclusive data on these relationships.

3.5 Conclusion

The data from this analysis indicate that youths with ND CNVs are at higher risk of having subclinical psychotic symptoms than youths without ND CNVs. The data suggest that this association is present even in carriers without an ASD diagnosis or significant cognitive impairment. The analysis contributes information to the growing base of knowledge on psychiatric risk profiles associated with ND CNVs in youths. A better understanding of the association between ND CNVs and subclinical psychotic symptoms will have considerable implications for the screening, diagnosis and clinical support of young people with ND CNVs. The findings of this analysis require replication, preferably in large, population-based samples of youths. Clinical characterisation of subclinical psychotic symptoms in ND CNV carriers in longitudinal cohorts will be helpful in exploring the clinical and prognostic significance of the association.

The lack of association between ND CNVs and risk of having internalising, externalising disorders and suicidal ideation identified in this study also require exploration in larger, well phenotyped population cohorts of youths to get more conclusive insights into these relationships.

Chapter 4: Identifying schizophrenia patients who carry pathogenic genetic copy number variants using standard clinical assessment: a retrospective cohort study

Parts of this chapter have been adapted from the following article:

Foley C, Heron EA, Harold D, Walters J, Owen M, O'Donovan M, Gallagher L, Corvin A et al. Identifying schizophrenia patients who carry pathogenic genetic copy number variants using standard clinical assessment: retrospective cohort study. The British Journal of Psychiatry. 2020:1-5.

I worked with E.H., M.G., L.G., A.C. on the study conception and design. I undertook and supervised the compilation, collection and coding of the phenotype data from existing research cohorts for the discovery dataset (supervised A.D.), analysed the phenotype data and wrote and revised the manuscript for publication under the supervision of A.C., L.G. and E.H., C.M., E.K., D.H., D.M., C.P., P.C., G.D. were responsible for the collection and processing of genetic and phenotype data from the pre-existing discovery dataset. J.W., M.O. and M.O'D contributed the data for the replication dataset and provided feedback with regard to the data analysis and interpretation. J.S. contributed to the core analysis of the genetic data used in the study. C.F., E.H., L.G., A.C. contributed to the data analysis and interpretation. All authors reviewed and approved the final manuscript.

4.1 Introduction

4.1.1 Background

Schizophrenia is a heterogeneous clinical syndrome characterised by positive symptoms (delusions, hallucinations, formal thought disorder), negative symptoms (anhedonia, flat affect, alogia) and cognitive impairments. There is marked variability between individuals with the disorder in terms of clinical course and symptomatology. The disorder affects ~0.5-1% of populations worldwide (Fischer et al., 2017; McGrath et al., 2008) and is a leading cause of disability globally (Murray et al., 1996). Schizophrenia is a devastating disorder for those it affects and their families and presents considerable social and economic demands for society (Knapp et al., 2004; Thornicroft et al., 2004).

The aetiology of schizophrenia is complex, both common and rare genetic risk factors are likely to augment risk, as well as environmental factors (Fusar-Poli et al., 2017). In 2014, a large genome wide association study conducted by the Psychiatric Genomics Consortium (PGC) identified 108 genomic risk loci for schizophrenia (Ripke et al., 2014). Implicated genes are enriched in pathways of glutamatergic neurotransmission, neuronal calcium signalling, broad synaptic function and immune function (Network et al., 2015; Ripke et al., 2014). For most affected individuals, schizophrenia has a polygenic architecture in which hundreds, or even thousands of variants collectively contribute to risk (Gratten et al., 2014; Ripke et al., 2014). Common genetic risk factors (polygenic risk) have been estimated to collectively account from some 30% of heritability (Loh et al., 2015); the great proportion of schizophrenia heritability is yet to be explained.

Copy number variants (CNVs) have a larger role to play in a small subset of cases. There is evidence genome-wide significant association with schizophrenia at eight CNVs (1q21.1, 2p16.3 (*NRXN1*), 3q29, 7q11.2, 15q13.3, distal 16p11.2, proximal 16p11.2 and 22q11.2) and suggestive support for eight more (2017). Presence of a schizophrenia-associated CNV in an individual increases the risk of developing schizophrenia by OR 2-30 (Rees et al., 2014b). These rare but recurrent copy number variants may be inherited or present *de novo*. High selection pressure has been identified against schizophrenia associated CNVs in the population, suggesting that these recurrent CNVs are likely caused by relatively recent *de novo* mutations (they appear to persist in populations for only a few generations) and recur in populations due to relatively high mutation rates (Marshall et al., 2017; Rees et al., 2011). Collectively, known schizophrenia associated CNVs are carried by <2.5% of patients (Kirov, 2015).

Family and comorbidity studies have long suggested shared genetic risk factors between neurodevelopmental disorders (NDDs). There is an increased risk of autism spectrum disorder (ASD) in children of parents with schizophrenia spectrum disorders (Larsson et al., 2005; Sullivan et al., 2012b) and more significant ASD traits have been identified in the siblings of children with childhood onset schizophrenia (Sporn et al., 2004). There is also an increased risk of intellectual

disability (ID) in children of mothers with schizophrenia, bipolar disorder or unipolar major depression (Morgan et al., 2012). Evidence is accumulating of both common and rare variants shared between NDDs (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Kirov et al., 2014). All known schizophrenia-associated CNVs are pleiotropic, carrying increased risk not only for schizophrenia but also for other NDDs such as ASD, ID or other developmental delays (Rees et al., 2014b). Penetrance of the CNVs is variable, with schizophrenia penetrance ranging from 2-18% and penetrance of developmental delay/ASD/congenital malformation ranging from 8-88% depending on the specific CNV (Kirov et al., 2014). Even in the absence of a psychiatric diagnosis, carriers of CNVs associated with schizophrenia have significant, but variable cognitive deficits (Kendall et al., 2016; Stefansson et al., 2014). These CNVs are therefore potentially pathogenic and clinically significant, but outcomes range from subtle cognitive effects to severe NDDs. Despite significant progress in our understanding of the genetics of schizophrenia, the progress in schizophrenia genetic discovery has yet to be leveraged in the context of clinical psychiatry.

Chromosomal microarray analysis (CMA) is a first line routine test in ASD and ID; 10-15% of individuals with these disorders carry identifiable genetic aetiologies that may have implications for clinical management and genetic counselling (Bass et al., 2018; Carter et al., 2013; Schaefer et al., 2013). For schizophrenia, the diagnostic yield is unlikely to be as high. CNVs with known schizophrenia risk are found in ~2.5% of individuals with the disorder (Kirov, 2015); clinically significant CNVs may present in up to 8% of individuals with schizophrenia (Costain et al., 2013).

Although CMA testing costs are modest compared with other clinical investigations such as neuroimaging (Baker et al., 2014) and costs of testing continue to decline, the relatively low prevalence of known schizophrenia-risk CNVs would result in low pick up rates in general schizophrenia populations, limiting cost-effectiveness of genetic testing. In ASD and ID, identification of ‘syndromal’ cases who have additional phenotypes and dysmorphic features can increase diagnostic yield of genetic testing (Miles et al., 2008a). The identification of clinical symptoms or demographic features that differentiate schizophrenia patients that carry risk CNVs may be helpful in clarifying who might benefit most from testing. Recent studies have suggested that identifying schizophrenia patients with co-morbid ID, multiple congenital malformations or dysmorphic features is likely to be helpful in identifying subsets of individuals with genomic disorders (Christian G. Bouwkamp et al., 2017; Lowther et al., 2017; Thygesen et al., 2018). The utility of other developmental indices in identifying such subsets of patients is less well explored.

4.1.2 Aims and Hypotheses

The objective of this work was to determine whether clinically identifiable phenotypic features were predictive of SCZ-associated CNV carrier status in a large schizophrenia cohort. Based on the known overlap with other NDDs and previously reported phenotype studies (Ahn et al., 2014; Derks et al., 2013; Kirov et al., 2014; Philip et al., 2011; Rujescu et al., 2009b; Sahoo et al., 2011b; Stefansson et

al., 2014; Walsh et al., 2008; Yeo et al., 2013), it was hypothesised that individuals with schizophrenia who carry risk CNVs are likely to be enriched for phenotypic features suggesting pre-existing neurodevelopmental compromise, earlier onset of psychotic symptoms, more severe illness course, or a positive family history of neurodevelopmental disorder.

4.2 Methods

4.2.1 Selection of phenotypic variables

A literature review was conducted in PubMed to identify clinical and phenotypic features reported to be associated with copy number variation in schizophrenia using the search terms "schizophrenia", "copy number variant" and "phenotype" from January 2008 to February 2016 when the study commenced. Publications that specifically described CNV associated clinical and phenotypic features in schizophrenia were selected to identify neurodevelopmental phenotypic categories. Identified phenotypic domains included early onset of psychosis; premorbid cognitive difficulties; delays in developmental milestones; and syndromal characteristics (dysmorphic features, congenital malformations). Having a first degree relative with schizophrenia is one of the greatest risk factors for the disorder, increasing risk by ~10 times (Rutkowski et al., 2017). Since copy number variants may be inherited (Rutkowski et al., 2017) and schizophrenia-risk CNVs are also associated with increased risk of early neurodevelopmental disorders such as ASD and ID (Kirov et al., 2014), it was hypothesised that individuals with a family history of neurodevelopmental disorder would be at increased risk of carrying an inherited schizophrenia-risk CNV and family history of neurodevelopmental disorder was therefore included as a variable in the analysis. Eight phenotypic variables were identified through expert clinical consensus (research psychiatrist, principal investigators) that are readily identifiable in a standard clinical evaluation and therefore ultimately of clinical utility and acceptability. Consequently 'dysmorphic features' and 'congenital anomalies' were excluded as reliable identification of these features requires additional training or clinical tools (Miles et al., 2008b). The phenotypic variables selected for analysis and associated references are presented in Table 4-1.

Table 4-1. Phenotypic variables selected for analysis.

Phenotypic Variable	Definition	References
Early onset of symptoms	Onset of symptoms < 18 years.	(Ahn et al., 2014; Walsh et al., 2008)
History of learning difficulties	Any difficulties with learning reported in school (excluding behavioural difficulties).	(Kendall et al., 2016; Stefansson et al., 2014)
Specific learning disorder	Identified as report of diagnosed dyslexia, dyscalculia or dysgraphia.	(Carrion-Castillo et al., 2013; Stefansson et al., 2014)
Remedial school support	Reported learning support in school, within class, separate classes, special school.	(Kendall et al., 2016; Stefansson et al., 2014)
Low educational attainment	Attained primary school education only.	(Kendall et al., 2016; Stefansson et al., 2014)
History of developmental delay	Delayed milestones- motor, speech, toilet training.	(Kirov et al., 2014; Sahoo et al., 2011b)
Comorbid neurodevelopmental diagnosis	Diagnosis of ASD, ID or epilepsy.	(Kirov et al., 2014; Sahoo et al., 2011b)
Family history of Neurodevelopmental disorder	Reported diagnosis of schizophrenia, ASD, ID or epilepsy in a first/second degree family member.	(Costain et al., 2014; Wilson et al., 2011)

4.2.2 Clinical subjects

4.2.2.1 Discovery dataset

4.2.2.1.1 Sample

The discovery dataset consisted of 1,215 individuals of Irish ancestry where both clinical phenotype and genome-wide SNP array data was available. Cases were all over 18 years of age and had a diagnosis of Schizophrenia or Schizoaffective Disorder after a structured clinical assessment (as described in (First, 2002)). Diagnosis was made based on the consensus lifetime best estimate method

using all available information (interview, family or staff report, chart review) with DSM-IV criteria as per the Structured Clinical Interview for DSM-IV-TR Axis I Disorders, research edition (SCID-P) (First, 2002). The SCID-P is the research version of the SCID, a semi-structured diagnostic interview, to ascertain DSM diagnostic criteria. Diagnostic assessments involved structured clinical interview and preparation of a clinical vignette based on all available information including chart review and collateral history. These were rated blind by the assessor. For all cases the vignettes were also rated blind by Professor Corvin. Where there were discrepancies across diagnoses, there was a consensus panel (principal investigator, a research psychiatrist, research nurses) who reached a consensus diagnosis based on all available information (there were few cases where this was required).

4.2.2.1.2 Ethical approval

All subjects provided written informed consent. Each referral center obtained local Research Ethics Committee (REC) approval.

4.2.2.1.3 Development of phenotype dataset

Phenotypic data were collected retrospectively from previous research (Irish Schizophrenia Genomics Consortium, 2012). The phenotypic data were collected from diagnostic assessments including information from structured clinical interview, clinical vignette, chart review and collateral history. The phenotype variable details that were collected and included in the analyses are presented in Table 4-2, including probe questions in structured assessments and details on coding criteria. Phenotypic data were coded as categorical variables. Missing information is described in Table 4-3.

Table 4-2. Phenotype definitions, probe questions, specific details and outcome codes used in discovery dataset.

Phenotype Variable	Source	Probe Questions in Clinical Interview	Specifying details	Coding
Early onset of symptoms	Structured clinical interview, clinical vignette, chart review and collateral history.	"When was the first time you saw someone for psychiatric problems (YEAR)? How old were you?"	Age specified in years.	0 = Age ≥ 18 years 1 = Age < 18 years "Don't know" or not answered- coded as missing.
History of learning difficulties	Structured clinical interview, clinical vignette, chart review and collateral history.	"Any problems with schoolwork?"	Details specified. If behavioural problems reported without learning problems, coded as "No".	0 = No, no difficulties with schoolwork indicated. 1 = Yes, difficulties with schoolwork indicated. "Don't know" or not answered- coded as missing.
Specific learning disorder	Structured clinical interview, clinical vignette, chart review and collateral history.	"Any problems with schoolwork?"	Details specified- dyslexia, dyscalculia or dysgraphia.	0= No, no specific learning difficulty indicated. 1 = Yes, specific learning difficulty including dyslexia, dyscalculia or dysgraphia specified. "Don't know" or not answered- coded as missing.
Remedial school support	Structured clinical interview, clinical vignette, chart review and collateral history.	"Any problems with schoolwork?"	Details specified- Reported learning support in school, within class, separate classes, special school.	0 = No, no specific learning support indicated. 1 = Yes, specified learning support, within class,

Phenotype Variable	Source	Probe Questions in Clinical Interview	Specifying details	Coding
				separate classes or special school. "Don't know" or not answered- coded as missing.
Low educational attainment	Structured clinical interview, clinical vignette, chart review and collateral history.	"At what stage did you leave school?"	Details specified. Primary, Secondary, Tertiary.	0 = Completed education past primary school level. 1 = Completed school to primary school level only. "Don't know" or not answered- coded as missing.
History of developmental delay	Structured clinical interview, clinical vignette, chart review and collateral history.	"Did you have any early learning difficulties?"	If answered affirmatively, details were specified: "Late walking?", "Late talking?", "Problems with toilet training?"	0 = No early learning difficulties indicated. 1 = Early difficulties indicated and specified as late walking, talking or toilet training. "Don't know" or not answered- coded as missing.
Comorbid neurodevelopmental diagnosis	Structured clinical interview, clinical vignette, chart review and collateral history.	"Did you ever attend your GP or another doctor for assessment of any developmental problems?"- Details specified	Details specified. Specific report of history of autism spectrum disorder, autism, Asperger's disorder, pervasive developmental disorder,	0 = No evidence comorbid neurodevelopmental diagnosis. 1 = Specified an autism spectrum disorder, intellectual disability or

Phenotype Variable	Source	Probe Questions in Clinical Interview	Specifying details	Coding
		"Have you ever been in hospital/attended your GP for a medical problem?" Specified: "Epilepsy?"	mild/moderate/severe intellectual disability. Specified a diagnosis of epilepsy or reference to specific seizures including grand mal (tonic-clonic) seizures, temporal lobe seizures, absence seizures, myoclonic seizures. Febrile seizures or psychogenic seizures not coded as positive for variable.	epilepsy diagnosis or epileptiform seizure. "Don't know" or not answered- coded as missing.
Family history of neurodevelopmental disorder	Structured clinical interview, clinical vignette, chart review and collateral history.	"Any history of mental health problems in your family?" "Any history of medical problems in your family?" "Family history of epilepsy?" "Family history of learning disability?" "Family history of pervasive developmental disorder?"	Details specified. Specific report of family member with history of schizophrenia, autism spectrum disorder, intellectual disability or epilepsy. First and second-degree relatives included.	0 = No reported family history of neurodevelopmental disorder. 1 = Specific report of family history of schizophrenia, ASD, ID or epilepsy in first or second degree relative. "Don't know" or not answered- coded as missing.

Table 4-3. Proportions of missing data in the Irish cohort for the phenotypes included in the analysis (discovery dataset).

Variable	Total N	Data available N	% Complete
Early age of onset of symptoms	1,215	1,199	98.7
History of learning difficulties	1,215	1,195	98.4
History of specific learning disorder	1,215	1,192	98.1
History of school supports	1,215	1,195	98.4
Primary school attainment only	1,215	1,204	99.1
History of developmental delay	1,215	1,162	95.6
Comorbid neurodevelopmental diagnosis	1,215	1,195	98.4
Family history of neurodevelopmental disorder	1,215	1,087	89.5

4.2.2.2 Replication sample

The identified replication dataset was based on 19,879 schizophrenia cases published by the Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC) cohorts (representing 40 cohorts excluding the Irish data) (Marshall et al., 2017). Contributors of the constituent datasets were approached to request access to additional phenotypic data to replicate the discovery findings. Only one cohort (Cardiff dataset) was identified with the requisite phenotype data and adequate sample size for replication (many of the well phenotyped cohorts were small and consequently had no CNV carriers).

4.2.2.2.1 Sample

The Cardiff dataset (n=479) consisted of participants from the previously reported CardiffCOGS study (Hamshere et al., 2013; Rees et al., 2014b) and was provided for this analysis by Dr James Walters and colleagues. Patients had a clinical diagnosis of schizophrenia; they were interviewed with the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) and case-notes were reviewed to derive a best-estimate lifetime diagnosis according to DSM-IV criteria.

4.2.2.2.2 Ethical approval

The sample was recruited with REC approval from community, in-patient and voluntary sector mental health services in the UK. Written informed consent was obtained from all subjects/patients.

4.2.2.2.3 Phenotype data provided

The comparable phenotype variables investigated in the Cardiff dataset were: 1/ "History of developmental delay", which was directly comparable to the Irish cohort defined as "clinically relevant delays in speech, walking, co-ordination or diagnosed developmental problem" and 2/ A positive history of epilepsy, intellectual disability and/or ASD was included as "Co-morbid neurodevelopmental disorder". Intellectual disability referred to an IQ < 70 and clinical specialist service involvement (diagnosis was based on evidence from interview, clinical notes review and patient questionnaires and corroborative history from a clinical team). The ASD and epilepsy variables were interview self-report of a clinical diagnosis- with data obtained through the question "Has a doctor or other healthcare professional ever diagnosed you with the following conditions?"- "Autism", "Epilepsy". Missing information is described in Table 4-4. The other phenotypic variables selected in the initial analysis were not collected in this dataset.

Table 4-4 Proportions of missing data in the Cardiff cohort for the phenotypes included in the analysis (replication dataset).

	Total	Data available	
Variable	**N**	**N**	**% Complete**
History of developmental delay	479	337	70.4
Comorbid neurodevelopmental diagnosis	479	417	87.1

4.2.3 CNV list

The target CNVs used in the analysis were fifteen CNVs with the strongest evidence of association with schizophrenia analysed by Rees et al. in (Rees et al., 2014b) at the time of analysis (Table 4-5). Twelve of these were also identified in a large PGC-CNV meta-analysis (Malhotra et al., 2012) and an additional three loci were included by Rees et al.: exon disrupting deletions at the *NRXN1* gene, deletion at distal 16p11.2 and duplications at the *WBS* region based on expert consensus or evidence published after the meta-analysis (Rees et al., 2014b).

Table 4-5 Schizophrenia-associated risk CNVs, table was adapted from Rees et al. 2014 (Rees et al., 2014b).

Locus	Position in Mb
1q21.1 del	chr1:146,57-147,39
1q21.1 dup	chr1:146,57-147,39
NRXN1 del	chr2:50,15-51,26
3q29 del	chr3:195,73-197,34
WBS dup	chr7:72,74-74,14
VIPR2 dup	chr7:158,82-158,94
15q11.2 del	chr15:22,80-23,09
AS/PWS dup	chr15:24,82-28,43
15q13.3 del	chr15:31,13-32,48
16p13.11 dup	chr16:15,51-16,30
16p11.2 distal del	chr16:28,82-29,05
16p11.2 dup	chr16:29,64-30,20
17p12 del	chr17:14,16-15,43
17q12 del	chr17:34,81-36,20
22q11.2 del	chr22:19,02-20,26

Note: Copy number variant loci listed in Rees et al. (Rees et al., 2014b) with previous evidence for associations with schizophrenia. Copy number variation positions are in UCSC Build 37. Del, deletion; dup, duplication; WBS, Williams-Beuren Syndrome; AS/PWS, Angelman/Prader-Willi syndrome.

4.2.4 Genotyping and CNV calling

The Irish sample was genotyped on the Affymetrix 6.0 array (n=802) or the Illumina HumanCoreExome chip (n=413) (full details available in (Irish Schizophrenia Genomics Consortium, 2012)). The Cardiff samples were all genotyped using HumanOmniExpress-12v1-1_B arrays (Illumina) (Rees et al., 2016b). To control for platform effects, raw intensity data was provided to the PGC CNV Analysis Group. This provided a centralized pipeline for systematic CNV calling including multiple CNV callers run in parallel. The final CNV set was defined as those >20kb in length including at least 10 probes and <1% minor allele frequency (MAF) (Marshall et al., 2017).

4.2.5 Statistical analyses

4.2.5.1 Discovery dataset

Univariate analyses (Fisher's exact tests) were first performed to assess associations between phenotypic predictors and SCZ-associated CNV status in the Irish Cohort. Multiple logistic regression analysis was then carried out to examine the effects of significant phenotypic variables, identified on univariate analysis, in modelling SCZ-associated CNV status. The final independent variables included in the model were those with a significance level of 0.05 following backward elimination steps. Model fit was assessed using Nagelkerke pseudo R-squared index.

In view of the rarity of SCZ-associated CNVs in the samples, the risk of small-sample bias in the analysis was considered. To examine this, a penalised likelihood method (Firth method) of logistic regression was also applied to the discovery sample for comparison with the classical logistic regression model. Penalised likelihood is a general approach to reducing small-sample bias in maximum likelihood estimation (FIRTH, 1993). The "logistf" package in R was used for this analysis (Heinze et al., 2016).

Receiver operating characteristic (ROC) curve analysis was used to test validity, sensitivity and specificity of the logistic regression parameters for modelling SCZ-associated CNV carrier status in the discovery dataset. The "OptimalCutpoints" package in R was used for this analysis (López-Ratón et al., 2014).

4.2.5.2 Replication dataset

The Cardiff replication dataset collected data on two of the three phenotypic variables identified in the discovery dataset as significant in modelling SCZ-associated CNV status. A model was trained on the discovery dataset with the two variables that were common to both datasets. This two-variable model was then applied to the replication dataset. The missing predictor variable was omitted. ROC curve analysis was used to assess the accuracy of the predictor variables in modelling SCZ-associated CNV carrier status in the replication dataset.

The analyses were considered exploratory and corrections were not made for multiple testing.

All analyses were completed in R version 3.2.3 ("R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.," 2013).

4.3 Results

4.3.1 Discovery data set

Nineteen individuals (1.6%) carried one of the fifteen identified SCZ-associated CNVs (Rees et al., 2014b) from the total sample of 1,215 cases. No individuals carried more than one SCZ-associated CNV. The details of the copy number variants and positions are presented in Appendix 1: Supplementary Tables, Table 7-17. Proportions of individuals with positive history of phenotypic variables and SCZ-associated pathogenic CNV status are presented in Table 4-6.

Table 4-6 Proportion of individuals with positive history of phenotypic variables in discovery and replication samples.

	Discovery Sample			Replication Sample		
	CNV	**No CNV**	**Total**	**CNV**	**No CNV**	**Total**
Phenotypic Variable	**(n = 19)**	**(n = 1,196)**	**(n = 1,215)**	**(n = 8)**	**(n = 471)**	**(n = 479)**
Early onset of symptoms	6 (31.6%)	270 (22.6%)	276 (22.7%)	-	-	-
History of learning difficulties	8 (42.1%)	181 (15.1%)	189 (15.6%)	-	-	-
Specific learning disorder	2 (10.5%)	15 (1.3%)	17 (1.4%)	-	-	-
Remedial school support	1 (5.3%)	36 (3.0%)	37 (3.0%)	-	-	-
Low educational attainment	4 (21.1%)	150 (12.5%)	154 (12.7%)	-	-	-
History of developmental delay	5 (26.3%)	77 (6.4%)	82 (6.7%)	3 (37.5%)	92 (19.5%)	95 (19.8%)
Comorbid neurodevelopmental diagnosis	3 (15.8%)	43 (3.6%)	46 (3.8%)	4 (50.0%)	40 (8.5%)	44 (9.2%)
Family history of NDD	5 (26.3%)	343 (28.7%)	348 (28.6%)	-	-	-

Note: Proportions of individuals with positive history of phenotypic variables and SCZ-associated pathogenic CNV status. NDD, neurodevelopmental disorder.

Univariate analyses identified four phenotypic variables with significant associations with SCZ-associated CNV status: “History of developmental delay”, “Co-morbid neurodevelopmental disorder”, “History of learning difficulties” and “Specific learning difficulties” (Table 4-7).

Table 4-7. Univariate analyses assessing associations between selected phenotypic variables and SCZ-associated CNV status in the discovery cohort.

Phenotypic Variable	OR (95% CI)	P
Early onset of symptoms	1.55 (0.58 - 4.27)	0.41
History of learning difficulties	**3.99 (1.55 - 10.30)**	**0.005**
Specific learning disorder	**9.03 (1.38 - 39.98)**	**0.03**
Remedial school support	1.76 (0.08 - 10.66)	0.45
Low educational attainment	1.84 (0.55 - 5.74)	0.29
History of developmental delay	**5.76 (1.91 - 16.44)**	**0.005**
Comorbid neurodevelopmental diagnosis	**4.93 (1.18 - 16.73)**	**0.034**
Family history of Neurodevelopmental disorder	1.06 (0.35 - 3.28)	1

Note: Fisher's exact tests were used to assess associations between phenotypic variables and CNV status groups. Significant results are in bold.

Correlations between significant phenotypic variables are presented in Table 4-8. The variables “History of learning difficulties” and “Specific learning disorder” were correlated (tetrachoric correlation, $r_t = 0.68$) and were likely capturing similar phenotypic information.

Table 4-8. Tetrachoric correlations between four phenotypic variables significantly associated with schizophrenia-risk CNV status in the discovery cohort.

	HLD	SLD	HDD	CNDD
History of learning difficulties (HLD)	-			
Specific learning disorder (SLD)	0.68	-		
History of developmental delay (HDD)	0.20	0.22	-	
Comorbid NDD (CNDD)	0.25	0.08	0.07	-

Note: Correlation coefficients measured by tetrachoric correlation coefficient t_r (El-Hashash et al., 2018) calculated using "tetrachoric" function in the "psych" package in R (W, 2019). NDD, neurodevelopmental disorder

A multiple logistic regression model was fitted using the four identified variables with significant associations with schizophrenia-risk CNV status and backward elimination at this point removed with the variable "History of learning difficulties" from the model. The final independent variables in the model were "History of developmental delay", "Co-morbid neurodevelopmental disorder" and "Specific learning disorder". These variables had odds ratios of 5.19 (95% CI 1.58-14.76, p=0.003), 5.87 (95% CI 1.28 – 19.69, p=0.009) and 8.12 (95% CI 1.16 – 34.88, p=0.012) respectively when included in the logistic regression model (Table 4-9). Nagelkerke pseudo r square for the model was 0.196, indicating that the phenotypic variables accounted for 19.6% of the variance in SCZ-associated CNV status in this sample.

Table 4-9. Multiple logistic regression model to determine whether clinically identifiable phenotypic features could be used to model schizophrenia associated CNV carrier status

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-4.72	0.32	-14.607	< 2e-16	0.01 (0.01-0.02)
History of developmental delay	**1.65**	**0.56**	**2.95**	**0.003**	**5.19 (1.58-14.76)**
Comorbid neurodevelopmental disorder	**1.77**	**0.67**	**2.63**	**0.009**	**5.87 (1.28-19.69)**
Specific learning disorder	**2.09**	**0.83**	**2.53**	**0.012**	**8.12 (1.16-34.88)**

Note: Generalised linear model. Predictor coefficients were tested using Wald tests and confidence intervals obtained using the Wald method. Significant variables highlighted in bold. Nagelkerke pseudo R Square = 0.196.

The performance of the three significant independent variables in modelling SCZ-associated CNV carrier status were tested using Receiver Operating Characteristic (ROC) curve analysis. An area under the ROC (AUROC) curve of 74.2% (95% CI 61.9 - 86.4%) was achieved, accounting for 58.8% (95% CI 32.9 - 81.6%) sensitivity and 89.1% (95% CI 87.1 - 90.9%) specificity in modelling SCZ-associated CNV carrier status (Table 4-10). The positive predictive value (PPV) of the model was 7.5% (95% CI 6.3-20.1%) and negative predictive value (NPV) was 99.3% (95% CI 98.0-99.4%), influenced by the low prevalence (1.6%) of the SCZ-associated CNVs in the sample.

Table 4-10. ROC curve results for modelling schizophrenia associated CNV status in the Irish dataset.

Cutoff Value	Sensitivity % (95% CI)	Specificity % (95% CI)	AUC % (95% CI)	PPV % (95% CI)	NPV % (95% CI)
0.04	58.8 (32.9-81.6)	89.1 (87.1-90.9)	74.2 (61.9-86.4)	7.5 (6.3-20.1)	99.3 (98.0- 99.4)

Note: Optimal Cutoff Value, Sensitivity, Specificity, Area Under The Curve, and Predictive Values using three independent variables ("History of developmental delay", "Comorbid neurodevelopmental disorder", "Specific learning disorder") to model for SCZ-associated CNV status in Irish cohort (discovery dataset). CI, confidence interval; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value

The risk of small-sample bias in the analysis was considered in the analysis and a penalized likelihood method (Firth method) of logistic regression was applied to the discovery sample for comparison with the classical generalized linear models described above. The coefficients and significance values using the penalized likelihood model did not differ substantially from the generalized linear model indicating that small sample bias was not a notable issue in the results presented above. The results of the penalized likelihood method are presented in Appendix 1: Supplementary Tables, Table 7-18, but are not referred to further in discussion of the results of this chapter.

4.3.2 Cardiff data set

Eight individuals (1.7%) in the Cardiff dataset (n = 479) set carried one of the fifteen identified risk CNVs, including one 1q21.2 duplication, one *NRXN1* deletion, one Williams-Beuren region duplication, three 15q11.2 deletions and two 22q11.2 deletions. No individual carried more than one of these CNVs.

The Cardiff replication dataset included data on two of the phenotypic variables of interest: "History of developmental delay" and "Co-morbid neurodevelopmental disorder". Proportions of individuals with positive history of phenotypic variables and SCZ-associated pathogenic CNV status are presented in Table 4-6. The Irish discovery dataset was used to build a multiple logistic regression model using these two variables (Appendix 1: Supplementary Tables, Table 7-19, Table 7-20). Applying this model to the Cardiff study population gave an AUROC of 83% (95% CI 52.0-100.0%) in identifying SCZ-associated CNV status. The sensitivity and specificity were 75.0% (95% CI 19.4-99.4%) and 97.6% (95% CI 95.1-99.0%) respectively (Table 4-11). The PPV for the model in this sample was 30.0% (95% CI 17.0-95.7%).

Table 4-11. ROC curve results for the Cardiff dataset.

Cutoff Value	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	AUC (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)
0.241	75.0 (19.4-99.4)	97.6 (95.1-99.0)	83.0 (52.0-100)	30.0 (17.0-95.7)	99.6 (95.8-99.9)

Note: Optimal Cutoff Value, Sensitivity, Specificity, Area Under the Curve, and Predictive Values using two independent variables ("History of developmental delay", "Comorbid neurodevelopmental disorder") to model for SCZ-associated CNV status in Cardiff cohort (replication dataset). CI, confidence interval; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value

4.4 Discussion

In this analysis, phenotypic information that can be generated by a standard clinical assessment was investigated for utility in identifying schizophrenia patients at greater risk of carrying pathogenic CNVs. On initial investigation of univariate associations, four phenotypic variables were identified to have significant associations with schizophrenia-associated CNV status: history of learning difficulties (OR 3.99, p = 0.005), specific learning disorder (OR 9.03, p = 0.03), history of developmental delay (OR 5.76, p = 0.005) and comorbid neurodevelopmental diagnosis (OR 4.93, p = 0.034). Previous studies that have identified associations between schizophrenia-risk/neurodevelopmentally associated CNVs and cognitive impairments (Kendall et al., 2016; Stefansson et al., 2014), autism spectrum disorders (Kirov, 2015; Kirov et al., 2014), intellectual disabilities (Kirov et al., 2014) and epilepsy (Sahoo et al., 2011a; Stewart et al., 2011). Individual pathogenic CNVs, such as the 15q11.2(BP1-BP2) deletion have previously shown associations with specific learning disorders dyslexia and dyscalculia (Stefansson et al., 2014; Ulfarsson et al., 2017).

Other clinical features such as early onset of psychosis, low educational attainment or a family history of neurodevelopmental disorders were not associated with SCZ-associated CNV carrier status in this cohort. Early onset psychotic experiences have been associated with autistic symptomatology in longitudinal studies (Bevan Jones et al., 2012; De Crescenzo et al., 2019; Sullivan et al., 2013), with reports of 20-50% of individuals with childhood onset schizophrenia meeting criteria for premorbid autism spectrum disorder (De Crescenzo et al., 2019; Rapoport et al., 2009). On this basis, the hypothesis followed that individuals with early onset psychosis may be at higher risk of carrying neurodevelopmentally associated copy number variants. It is possible that the low number of CNV carriers in the sample contributed to the lack of significance in this association (OR 1.55, p = 0.41), however it is also possible that there is no association between ND CNVs and early onset psychosis. The finding may be better explored when larger well-phenotyped datasets are available. The lack of association with family history of neurodevelopmental disorders (OR 1.06, p = 1) was notable. High selection pressure has been identified against schizophrenia associated CNVs, suggesting that recurrent schizophrenia-associated CNVs are likely to be caused by relatively recent *de novo* mutations (Rees et al., 2011). Exome sequencing analysis comparing rare inherited and *de novo* mutations in schizophrenia cases has also indicated that *de novo* mutations are likely to account for at least 50% of sporadic cases of schizophrenia (Xu et al., 2011). This may account for the lack of association with family history of neurodevelopmental disorders in our sample. Again, large well-phenotyped datasets would provide further information in this context.

In the discovery cohort, having a specific learning disorder (OR 8.12, p = 0.012); developmental delay (OR 5.19, p = 0.003); or comorbid neurodevelopmental disorder (OR 5.87, p = 0.009) modelled positive carrier status for a SCZ-associated CNV with a relatively high specificity (89.1%, 95% CI 87.1-90.9%) but more modest sensitivity (58.8%, 95% CI 32.9-81.6%). The positive predictive value (PPV) of the model was 7.5% (95% CI 6.3-20.1%) and negative predictive value (NPV) was 99.3%

(95% CI 98.0-99.4). The PPV and NPV of a test are influenced by the prevalence of the disease/risk factor tested for in the sample examined (Maxim et al., 2014; Ranganathan et al., 2018). The low prevalence of the SCZ-associated CNVs in the discovery sample (1.6%) strongly influenced these outcomes. PPV is important because it indicates the probability that a positive test result represents the presence of the risk factor of interest. The relatively low PPV may limit the model's value in this regard, however given that identifying the relevant information for the model in the clinical setting would be simple for clinicians and inexpensive, the clinical utility of the model could still be relevant as a first step in increasing diagnostic yield for genetic testing in schizophrenia populations.

Information on "Specific learning disorders" was not available for the Cardiff sample. Based on the remaining two variables, "Comorbid neurodevelopmental disorder" and "History of developmental delay", a model trained from the discovery dataset was applied to the Cardiff sample. This too showed relatively high specificity (97.6%, 95% CI 95.1-99.0%) but more modest sensitivity (75.0%, 95% CI 19.4-99.4%) in modelling carrier status for a SCZ-associated CNV.

Recent studies have suggested that identifying schizophrenia patients with co-morbid intellectual disability is likely to be helpful in identifying subsets of individuals with genomic disorders. Thygesen and colleagues reported an approximately three-fold higher rate of pathogenic CNVs in patients with psychosis and intellectual disability compared to rates in the general schizophrenia population (Thygesen et al., 2018). Lowther et al. examined the genome-wide burden of pathogenic CNVs in a schizophrenia cohort (n=546) and demonstrated a significantly higher burden of pathogenic CNVs (OR 5.01, $p = 0.0001$) in patients with schizophrenia and low IQ ($IQ < 85$) compared with those with average IQ ($IQ \geq 85$). Based on their findings, the authors concluded that individuals with schizophrenia and low IQ should be prioritised for clinical microarray testing in clinical and research contexts (Lowther et al., 2017). The findings reported from this study provide further support to this recommendation, but indicate that other developmental indices, which could be captured by a clinical neurodevelopmental history, will likely be important in the development of any future guidelines for schizophrenia genetic testing.

4.4.1 Limitations

4.4.1.1 Limited sample size of rare CNV carriers

The strength of this study was the well-characterised phenotype data set that was compiled from extensive clinical and research data that was systematically collected from previous schizophrenia research studies. It was possible to test multiple phenotypic features for potential in identifying pathogenic CNV status and identify three variables that are easily clinically identified and that show considerable promise in identifying a high-risk group.

Recurrent schizophrenia-associated CNVs are rare events (~1:150 to 1:1,000 (Kirov et al., 2014)) and individual cohorts are likely to identify only a modest number of known CNVs, as demonstrated in our sample of 1,215 schizophrenia cases. The small sample size of schizophrenia-risk CNV carriers resulted in low precision for parameter estimates in the analysis, reflected in the wide confidence intervals. Identifying suitable replication datasets for the analysis was a significant challenge. Other than the dataset provided by the Cardiff team, datasets of sufficient size with the requisite phenotype data were not available for further replication.

4.4.1.2 Pathogenic copy number variant selection

The list of pathogenic CNVs associated with schizophrenia, autism and other neurodevelopmental disorders is continuously evolving in the context of ongoing research in the field and new investigative methods (e.g. whole genome sequencing) (Foley et al., 2017).

In schizophrenia populations, copy number variants associated with other developmental disorders may also be relevant, however to date it has not been possible to confidently identify such CNVs because they are too rare, have smaller effect sizes for the development of schizophrenia, or both (Rees et al., 2016b). The phenotypic variables selected in the study aimed to identify individuals with evidence of early neurodevelopmental compromise and may have been relevant in identifying individuals with other ND CNVs as well as SCZ-associated CNVs.

Clinical interpretation of CNVs remains challenging in view of their variable penetrance and heterogeneous phenotypic presentations. The purpose of this study was to identify whether there were specific phenotypic features easily identifiable within schizophrenia populations which are more commonly identified in the carriers of the currently best understood risk CNVs. In this analysis therefore, a list of 15 CNVs was selected which, on review of the literature at the time, showed the strongest evidence of association with schizophrenia.

4.4.1.3 Power of analyses

Post hoc power analyses were carried out to evaluate the power of the univariate analyses of SCZ-associated CNV status and the selected phenotypic variables. Based on the effect sizes identified in the analyses, the power analyses confirmed that larger samples would be required to effectively assess associations between SCZ-associated CNV status and some of the phenotypic variables including early onset of symptoms, remedial school support, low educational attainment and family history of neurodevelopmental disorder (Table 7-21).

4.4.1.4 Secondary analysis of existing data

The discovery and replication cohorts used retrospective phenotypic data from which variables were identified that provided estimates of sensitivity and specificity for modelling SCZ-associated CNV carrier status. Significantly larger, well-characterized phenotypic samples (e.g. prospective cohorts) will be required to provide more refined estimates of sensitivity and specificity to inform genetic screening guidelines.

4.4.1.5 Lack of patient perspective

Patients with serious mental illness and their families have expressed interest in attaining an aetiological genetic diagnosis (Costain et al., 2012a; Inglis et al., 2015). Genetic testing and counselling have been shown to decrease internalised stigma and self-blame experienced by some patients (Costain et al., 2012b) and genetic diagnoses may also improve therapeutic alliances between patients and health providers (Costain et al., 2012a). Although it was beyond the scope of this study, it will be vital to be informed by the patient and family perspective in the development of future guidelines for genetic testing in schizophrenia.

4.4.2 Implications

A small subset of schizophrenia patients (~2.5%) carry CNVs that substantially increase risk for schizophrenia but also other NDDs. The clinical benefits of identifying such patients have been demonstrated for other NDDs (Miller et al., 2010b; Schaefer et al., 2013). Similar benefits are likely to apply in schizophrenia, but as these events are rare, routine genetic testing for all individuals is probably not indicated. Previous studies suggest that targeting schizophrenia patients with co-morbid ID is likely to be more fruitful in identifying such cases (Lowther et al., 2017; Thygesen et al., 2018). The findings from this study suggest that careful clinical history taking to document developmental delay; reported learning disorders; or a co-morbid diagnosis of ASD or epilepsy may also be informative in screening for schizophrenia patients at higher risk of carrying known SCZ-associated CNV.

4.5 Conclusions

The results from this study indicate that a detailed neurodevelopmental history identifying specific neurodevelopmental features may be informative in screening for schizophrenia patients at higher risk of carrying known SCZ-associated CNVs. Identification of genomic disorders in these patients may have clinical benefits similar to those demonstrated for other neurodevelopmental disorders.

Chapter 5: Discussion

5.1 Introduction

Psychiatric and neurodevelopmental disorders have profound implications for the well-being of affected individuals, their families, and society (Arim et al., 2017; Boyle et al., 2011; Jensen et al., 2011; Ochoa et al., 2003; Tomlinson et al., 2009). The majority of these disorders have complex aetiologies, with a significant genetic component to their development. Rare, recurrent CNVs contribute substantially to the development of autism spectrum disorder (ASD), intellectual disability (ID) and schizophrenia in subsets of individuals with the disorders. The presence of a neurodevelopmental CNV (ND CNV) can increase an individual's risk for schizophrenia (penetrance 2-18%) and ASD or ID (penetrance 8-18%) (Kirov et al., 2014). Even in the absence of a major NDD, ND CNV carriers may manifest subtle cognitive deficits (Kendall et al., 2016; Stefansson et al., 2014), other psychiatric disorders (Chawner et al., 2019; Hanson et al., 2015; Kendall et al., 2019; Niarchou et al., 2019; Niarchou et al., 2014; Stefansson et al., 2014) or have no evidence of cognitive, neurodevelopmental or psychiatric symptoms.

Establishing a comprehensive and refined understanding of the phenotypes associated with ND CNVs in different populations can provide clinically important information for ND CNV carriers and their families and may guide neurodevelopmental and psychiatric risk monitoring and intervention. In this thesis, I used secondary data analysis techniques to leverage existing datasets for the identification of neurodevelopmental and psychiatric phenotypes associated with ND CNV status in specific populations that have been understudied to date.

5.2 Overview of thesis aims and results

5.2.1 Analysis 1: Investigation of psychiatric phenotypes associated with ND CNVs in a cohort of youths with ASD

A substantial proportion of youths with ASD (~15%) are diagnosed with pathogenic CNVs on routine CMA testing (Carter et al., 2013; Schaefer et al., 2013). Identifying whether ASD-affected youths with ND CNVs are at increased risk of presenting with specific psychiatric disorders has implications for screening, diagnosis and therapeutic inputs for these young people.

This analysis aimed to assess psychiatric risk profiles associated with ND CNV carrier status in youths with ASD and to identify whether there was a significant interaction effect between sex and

ND CNV status in terms of risk of depression and anxiety in the sample. A follow up analysis was conducted in a sample of ASD-unaffected siblings, to explore whether the identified associations in individuals with ASD also presented in the absence of ASD. Proband and sibling data from the Simons Simplex Collection (SSC) was used for this analysis.

Sex differences were identified in the effect of ND CNV status on risk of affective problems in ASD-affected youths. Females with an ND CNV were significantly more likely to present with depressive symptoms than males with (OR 5.39, 95% CI 1.02 - 32.79) or without (OR 2.70, 95% CI 1.04 – 8.00) an ND CNV. This finding was consistent with previous studies that have also reported evidence of sex-specific effects of CNVs on depression outcomes. Martin et al. reported that females with anxiety or depression were more likely to carry large, rare CNVs than males in a population sample of youths (Martin et al., 2019). In an adult population sample, Kendall et al. observed higher rates of depression among female carriers of ND CNVs than among male carriers (Kendall et al., 2019). In the study described by Kendall et al., individuals with ASD and other NDDs were excluded from the sample, indicating that the association between ND CNVs and depression may be independent of neurodevelopmental co-morbidity. In the sample of apparently typically developing siblings assessed here, ND CNV status was significantly associated with increased risk of affective problems; this effect was driven by female carriers of ND CNVs only, however an interaction effect between ND CNVs and sex could not be tested in this sample due to sample size limitations.

These findings are potentially important in highlighting increased risk of affective problems that is gender specific in youths with ASD who carry ND CNVs. This may enhance awareness among clinicians to screen for affective problems in female carriers of ND CNVs. Early diagnosis and treatment of this difficult to identify and disabling comorbidity could mitigate prolonged periods of functional impairment, distress and progression to urgent psychiatric presentations such as suicidal ideation.

The analysis in the sample of siblings indicated that typically developing siblings who are carriers of ND CNV may also be at increased risk for affective problems. Increased awareness of risk of affective problems in apparently typically developing youth carriers of ND CNV also has important implications for mental health screening. This provides some evidence of a relationship between ND CNV carriers and depression in typically developing youth, reflecting previous reports from adult studies (Kendall et al., 2019; Stefansson et al., 2014). However, given that these youths were ascertained on the basis of positive family history for ASD, the association and any potential interaction effects, such as sex, need to be further assessed in other samples.

The clinical interpretation of the findings from this analysis are limited at this point due to the small sample size on which the results are based; replication in larger datasets will be required to progress the translation of these findings to clinical practice. Studies involving longitudinal follow up of

youths with ASD with and without ND CNVs are required to fully investigate the relationships between ND CNV carrier status and psychopathology.

5.2.2 Analysis 2: Investigation of psychiatric phenotypes associated with ND CNVs in a large population-based clinical cohort of youths

Exploration of childhood onset psychiatric disorder association with ND CNVs in large cohorts of youths has been limited to date. There is a significant need to understand psychiatric risk that may be associated with ND CNVs in populations of youths, particularly for those individuals who present clinically with developmental delay and are found to be ND CNV carriers. However, ND CNVs have incomplete penetrance for ASD and ID in childhood and many carriers may not have major NDDs. In Analysis 1, evidence was identified of gender-specific increased risk of affective problems in youths with ASD carrying ND CNVs; exploratory analysis in a cohort of typically developing siblings indicated that individuals without major NDDs who carry ND CNV may also be at increased risk for affective problems. Further exploration of risk of childhood psychiatric disorders associated with ND CNVs in the presence or absence of comorbid major NDDs can provide important information in the context of genetic counselling for ND CNV carriers and may help with monitoring risks, prevention strategies and early interventions.

The first aim of this analysis was to assess the relationship between ND CNVs and four major outcomes reflective of significant psychopathology (internalising disorders, externalising disorders, subclinical psychotic symptoms and suicidal ideation) in a large clinical cohort of youths. The second aim of the analysis was to assess sex differences in the effects of ND CNVs on internalising disorders. The third aim was to identify whether NDD comorbidity was a major determinant for increased risk of psychopathology associated with ND CNVs. Genetic and phenotypic data from a publicly available large cohort of youths (n=8205), the Philadelphia Neurodevelopmental Cohort (PNC), was used.

The data from the analysis indicated that youths with ND CNVs were at higher risk of having subclinical psychotic symptoms than youths without ND CNVs. Youths with ND CNVs were nearly twice as likely (OR 1.9, 95% CI 1.1 – 3.0) to have subclinical psychotic symptoms compared with youths without ND CNVs in the sample; 15.6% of those with ND CNVs had subclinical psychotic symptoms, compared with 8.2% of non-carriers. The association between ND CNVs and subclinical psychotic symptoms was also present in a subset of individuals without an ASD diagnosis or significant general cognitive deficits. Although the association between ND CNVs and schizophrenia risk in adults is well-established, my study provides novel evidence of an association between ND CNVs and subclinical psychotic symptoms in a population of youths.

Screening for subclinical psychotic symptoms may be important in child and adolescent carriers of ND CNVs. Subclinical psychotic symptoms in youths may be a precursor to the development of psychotic illness later in life (Poulton et al., 2000; Welham et al., 2009b), but psychotic-like experiences may also be indicators of current severe non-psychotic psychopathology and co-morbidity (Kelleher et al., 2012; Nishida et al., 2008; Scott et al., 2009; Wigman et al., 2011). In this study, ND CNVs were not associated with other psychopathological outcomes including internalising disorders, externalising disorders and suicidal ideation. This lack of association with other psychiatric disorders may suggest that the association with subclinical psychotic symptoms is more reflective of a psychotic prodrome, however a number of limitations in the analysis, including a substantial proportion of missing data from the variable measuring internalising disorders preclude confident interpretation of this.

For the majority of patients with schizophrenic psychoses, frank psychotic symptoms do not present initially but are preceded by a slow emergence over an average period of 4-5 years with non-specific prodromal symptoms followed by attenuated psychotic symptoms (Riecher-Rössler et al., 2017). Early identification and intervention for first episode psychosis can have a profound impact on outcomes for individuals with psychosis (Anderson et al., 2018). Longitudinal studies monitoring youths with ND CNVs with subclinical psychotic symptoms will be crucial to estimate the risk of conversion to frank psychotic states over time; such estimates would inform needs with regard to routine screening, prevention strategies and early interventions for this group.

The lack of evidence of an interaction between ND CNVs and sex in the context of internalising disorders in this analysis was of particular interest given that a significant association was identified between depression (an internalising disorder) and ND CNVs in female carriers in the previous study. There are a number of possible reasons for this. Unfortunately, the internalising disorders variable in this analysis included a substantial proportion of missing data (29%); which reduced the statistical power of the analysis and may have biased estimates for the outcome (Kang, 2013). It is also possible that ND CNV associations may be specific to symptoms of depression and not anxiety. Stefansson et al. previously identified association between ND CNVs and depression in an adult population sample, but did not identify an association with anxiety in the population (Stefansson et al., 2014). The study may also have been underpowered to detect associations with the other psychopathological outcomes assessed (externalising disorders, suicidal ideation) due to small sample size of ND CNV carriers.

5.2.3 Analysis 3: Identifying schizophrenia patients who carry pathogenic genetic copy number variants using standard clinical assessment: a retrospective cohort study

Chromosomal microarray testing is a first line routine investigation in ASD and ID; 10-15% of individuals with these disorders carry identifiable genetic aetiologies that may have implications for clinical management and genetic counselling (Bass et al., 2018; Carter et al., 2013; Schaefer et al., 2013). Despite the association with ND CNV in 2.5% of individuals with schizophrenia (Kirov, 2015), routine clinical genetic testing with CMA is not available to this clinical group. This means that individuals with potentially identifiable genetic causes are not identified and miss the opportunity for associated interventions such as genetic counselling. Since ND CNV have been defined across disorders, with stronger penetrance for early onset NDDs, it is plausible that the presence of NDD-related traits in schizophrenia patients may indicate an increased risk of ND CNV carrier status. Few studies have investigated clinically identifiable indicators of SCZ-associated CNV status. Here I sought to identify specific clinical or demographic features that are predictive of ND CNV carrier status in the context of schizophrenia. Identification of such features would be useful in informing decisions regarding genetic testing.

Based on the known overlap with other neurodevelopmental disorders and previously reported phenotype studies (Ahn et al., 2014; Derks et al., 2013; Kirov et al., 2014; Philip et al., 2011; Rujescu et al., 2009b; Sahoo et al., 2011b; Stefansson et al., 2014; Walsh et al., 2008; Yeo et al., 2013), it was hypothesised that individuals with schizophrenia who carry SCZ-associated CNVs were likely to be enriched for phenotypic features suggesting pre-existing neurodevelopmental compromise, earlier onset of psychotic symptoms, more severe illness course, or a positive family history of NDD.

In a discovery cohort, specific learning disorder, developmental delay, and comorbid neurodevelopmental disorder were significant independent variables in modelling positive carrier status for a SCZ-associated CNV, with an AUROC of 74.2% (95% CI 61.9 - 86.4%). A model constructed from the discovery cohort including developmental delay and comorbid neurodevelopmental disorder variables resulted in an AUROC of 83% (95% CI 52.0-100.0%) in a replication cohort. The small sample size of schizophrenia-risk CNV carriers resulted in low precision (reflected by the wide confidence intervals) for the estimated parameters.

Recent studies have suggested that identifying schizophrenia patients with co-morbid intellectual disability is likely to be helpful in identifying subsets of individuals with genomic disorders. Thygesen and colleagues reported an approximately three-fold higher rate of pathogenic CNVs in patients with psychosis and intellectual disability compared to rates in the general schizophrenia population (Thygesen et al., 2018). Lowther et al. examined the genome-wide burden of pathogenic CNVs in a schizophrenia cohort and demonstrated a significantly higher burden of pathogenic CNVs in patients with schizophrenia and low IQ (IQ< 85) compared with those with average IQ (IQ ≥ 85). Based on their findings, the authors concluded that individuals with schizophrenia and low IQ should

be prioritised for clinical microarray testing in clinical and research contexts (Lowther et al., 2017). The findings reported from my study provide further support to this recommendation, but indicate that other developmental indices, which could be captured by a clinical neurodevelopmental history, will likely be important in the development of any future guidelines for schizophrenia genetic testing.

Significantly larger, well-characterized phenotypic samples (e.g. prospective cohorts) will be required to provide more refined estimates of sensitivity and specificity to inform genetic screening guidelines.

5.3 Implications

This thesis adds to the growing body of clinical knowledge on the phenotypic features of youth and adult carriers of ND CNVs and takes steps towards translating some important findings in psychiatric genetics into clinically impactful information.

The clinical knowledge of psychiatric risk profiles associated with the majority of ND CNVs in populations of youths is limited. As a result of this lack of knowledge, young people with these CNVs and their families cannot at this point be optimally counselled with regard to psychiatric risk; this is a particular issue for youths with ASD, up to 15% of whom may be diagnosed with a pathogenic CNV on CMA testing (Schaefer et al., 2013). The lack of knowledge around psychiatric risk associated with ND CNVs could also delay screening and early intervention for young people who may be at risk of developing serious psychiatric symptoms, the outcomes of which could be improved through early identification and management (Dadds et al., 1997; Snell et al., 2013) (Aos et al., 2004; Beitchman et al., 1992a; Beitchman et al., 1992b).

The first two studies in this thesis contribute information addressing the deficit in knowledge around psychiatric risk profiles associated with ND CNVs. Findings from the first study suggest that females with ASD who carry ND CNVs may be at increased risk of depressive psychopathology compared with males with or without ND CNVs. The findings of the study contribute support to previous findings of associations between depression and ND CNVs (Kendall et al., 2019; Stefansson et al., 2014). If replicated in future studies, the findings may be clinically useful in terms of developing screening and management strategies for a subgroup of youths with ASD at increased risk of depression.

Findings from the second analysis suggested that youths with ND CNVs are at increased risk of having subclinical psychotic symptoms and that this increased risk is present even in those without a major NDD. Screening for subclinical psychotic symptoms may be important in child and adolescent carriers of ND CNVs. Subclinical psychotic symptoms in youths may represent a precursor to later development of psychotic disorder or indicate severe non-psychotic symptomatology. The nature of subclinical psychotic symptoms for youths with ND CNVs cannot be inferred from this analysis. The next step in making this finding clinically useful will be to

replicate the finding. Further clinical characterisation of subclinical psychotic symptoms in ND CNV carriers and assessment in longitudinal cohorts will then be useful in further determining clinical and prognostic implications.

Although significant progress has been made identifying specific CNVs as aetiological risk factors for schizophrenia, there is inadequate knowledge to be able to inform the practice of clinical genetic testing for this disorder. The final study in this thesis aimed to identify clinical phenotypic indicators that will be relevant in informing genetic testing in schizophrenia. Results from the study indicate that evaluation of specific developmental indices, identifiable through clinical neurodevelopmental history taking will be informative in screening for adult schizophrenia patients at higher risk of carrying known SCZ-associated CNVs. Incorporation of this information will likely be important in the development of any future guidelines for schizophrenia genetic testing.

5.4 Limitations

Limitations of each study described in this thesis are discussed in detail in respective chapters. A significant challenge that affected all three studies was the attainment of samples of sufficient size, with sufficient depth of phenotyping to be able to explore the hypotheses of interest and replicate findings.

The magnitude of sample sizes required to identify significant genetic associations in complex disorders have been illustrated by genome wide association studies (GWAS). In schizophrenia GWA studies, significant genome-wide significant associations were elusive until sample sizes reached a tipping point of ~15,000 cases, beyond which the number of significant genome-wide significant associations identified increased linearly (Levinson et al., 2014; Smoller et al., 2019). Although rare variants can present with much larger effect sizes than common variants, their low prevalence and variable penetrance also result in the need for large sample sizes to amass adequate carrier counts. ND CNVs are rare, even in neurodevelopmental populations (individual frequency ~ 0.6/1,000-8/1,000 in people with NDDs (Kirov et al., 2014)). Individual cohorts are likely to identify only a modest number of known CNVs.

In the studies assessing psychiatric risk profiles associated with ND CNV status described in this thesis, the power limitations were reflected in large confidence intervals of the effects associated with CNVs. For replication of these analyses, much larger samples are likely to be required to robustly support the findings identified. Adult population studies that have successfully identified associations between ND CNVs and depression in the last number of years were composed of very large, national samples (n > 100,000 (Kendall et al., 2019; Stefansson et al., 2014)). The collection of similarly sized youth population samples will require considerable resources, planning and organisation but will likely be required to progress research of rare ND CNVs at a population level.

Although a large sample is critical for sufficient statistical power of phenotypic associations with ND CNVs, sufficient detail in the phenotypic data collected is also of key importance. Collection of detailed phenotypic data can facilitate homogeneity within samples, another important factor in reaching sufficient statistical power for genetic association studies (Traylor et al., 2015). Collecting detailed phenotypic data is time consuming and requires substantial research resources to collect, which in practical terms slows the rate of ascertainment and can limit the scope of sample sizes in practice. Each of the samples used in the studies described in this thesis had strengths and limitations in terms of available phenotypic data. The Simons Simplex Collection contained a wealth of phenotypic data on a diverse set of variables, however the psychiatric phenotypic data collected was based on a psychiatric screening questionnaire (Achenbach et al., 1983; Fischbach et al., 2010). The CBCL has shown good sensitivity but limited specificity for identifying psychiatric disorders in youths with ASD, which may have impacted on the homogeneity within samples. More structured, in-depth interviews specifically developed for identification of psychiatric disorders in youths with ASD may be favourable in terms of their sensitivity and specificity for identifying psychiatric disorders in youths with ASD and improving the psychiatric homogeneity within samples. However, of course collection of data through structured interviews require much more time and resources.

The principal investigators of Philadelphia Neurodevelopmental Cohort sought to facilitate rapid assessor training and assessment of high volumes of participants by creating a computerised, modified version of a structured psychiatric interview, creating the GOASSESS analysis and they succeeded in collecting a large dataset with an abundance of phenotype data. A detailed description of the modifications made for the GOASSESS system and the scoring algorithms were not made publicly available and therefore it was necessary to create a set of diagnostic algorithms based on DSM-IV criteria to analyse the data. Although the abbreviated format of the GOASSESS facilitated rapid data collection, it also may have limited sensitivity of the tool for identifying clinically significant symptoms (Calkins et al., 2015).

The well characterised phenotype discovery data set used in the schizophrenia study was built based on access to extensive clinical and research data compiled from previous schizophrenia research studies. There were challenges in attaining further samples for replication which were of adequate size and had the required neurodevelopmental phenotype data. Member groups of the Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC) (representing 40 cohorts excluding the Irish data) were approached to request access to additional phenotypic data to replicate the discovery findings (Marshall et al., 2017). Only one cohort (Cardiff dataset) was identified with the requisite phenotype data and adequate sample size for replication (many of the well phenotyped cohorts were small and consequently had no CNV carriers).

Collecting adequately detailed phenotype data in the context of the large samples required for ND CNV research is an ongoing challenge for researchers in this field.

A further challenge that was relevant to the studies analysing psychiatric risk profiles associated with ND CNVs in youths was the use of current psychiatric classification systems. The two major psychiatric classification systems most widely used in clinical and research practice, the Diagnostic and Statistical Manual of Mental Disorders (DSM) (American Psychiatric Association, 2013) and International Classification of Diseases (ICD) (World Health, 2004) classification systems use categorical approaches to the diagnosis of psychiatric disorders. These categorical approaches to diagnosis have been criticised for generating substantial phenotypic heterogeneity within diagnostic categories (Allsopp et al., 2019; Hengartner et al., 2017; Lilienfeld et al., 2016). Frequent co-occurrence of psychiatric disorders and symptomatic overlap between diagnoses is also frequently observed (Allsopp et al., 2019). The dichotomous outcomes associated with current categorical classification systems are also unfavourable in the context of genetic association studies due to the negative impact that they have on statistical power compared with quantitative outcomes (Royston et al., 2006). The added power limitations of imprecise phenotype definitions and sample heterogeneity can be detrimental in phenotypic analyses of ND CNVs where power is already challenged by low carrier frequency. Dichotomous outcome variables based on a standard psychiatric classification systems (DSM) were utilised in the analyses in this thesis, which may have posed issues to statistical power of the analyses. Alternative classification systems such as dimensional approaches may offer advantages in terms of increased statistical power from the quantification of traits (Helzer et al., 2006; Hengartner et al., 2017; Insel et al., 2010; Kelly et al., 2018; Simmons et al., 2014). However, the ultimate aim of the research described in this thesis was to improve outcomes for individuals who carry neurodevelopmental CNVs by identifying clinically applicable psychiatric insights; in that context the phenotypic definitions selected in this work were informed by clinical relevance and applicability.

It is notable that the samples used in this thesis were all derived from Western countries, limiting international generalisability of the findings. Most neurodevelopmental research at present comes from Western countries, resulting in an under-representation of lower- and middle-income countries in spite of the considerable burden of neurodevelopmental disorders in these countries (Bitta et al., 2017; Tromans et al., 2020). International collaborations will be important in progressing neurodevelopmental research at a global level. The University of Liverpool's INDIGO study, for example, is bringing together professionals from the UK, Canada, Pakistan, Malawi and Uganda in a network of clinicians and researchers working to advance neurodevelopmental research in lower- and middle-income countries. Such approaches will be essential to achieve the United Nations mandate of full child health and neurodevelopment as a basic human right for all (Boivin et al., 2015; Unicef, 2013; United Nations, 2002).

5.5 Future Directions

Understanding how ND CNVs contribute to the development of psychiatric and neurodevelopmental conditions may add key insights to the overall understanding of mental health, potentially leading to improved strategies for treatment and prevention of psychopathology. Illuminating the pathways from ND CNV to neurobiology to clinical outcomes will be highly challenging. There are many more challenges to be considered as well, such as where ND CNVs fit in with the broader genetic architecture of psychiatric and neurodevelopmental disorders and how environmental factors may influence their effects. Where translating research findings into clinical practice is the ultimate goal, consideration of pathways for dissemination of research evidence and clinical care requirements will also be essential to successful translation.

5.5.1 Collaboration

It is important to highlight collaborative work as an integral factor for progress in studying neurodevelopmental and psychiatric disorders. Ascertaining sufficient samples of ND CNV carriers for well-powered analyses is especially challenging for individual research groups and international collaboration is essential to move progression in ND CNV research. It will be vital to ensure that sampling methods, phenotyping batteries, genotyping and coding methods are systematised and standardised so that samples can be efficiently combined for analysis. Sharing of data and methodologies will also be central to ensure efficient progress in the field.

The work of the Psychiatric Genomics Consortium (PGC) exemplifies the progress that can be made with combined research efforts at an international level. In 2014, the PGC published a genome wide association study of schizophrenia, providing robust evidence of association at 108 risk loci (Ripke et al., 2014). The work of the PGC also enabled the identification of eight CNV loci as having genome-wide significant association with schizophrenia (Marshall et al., 2017). The PGC attributes much of the success of these important studies to the level of teamwork that is committed and to the rigorous, harmonised methodologies used across groups (Corvin et al., 2016). The PGC has an open-source approach, making data widely available, within the limits of national laws and ethical review restrictions.

Networks specifically focussing on NDDs in the context of CNV such as the Maximising Impact of research in Neurodevelopmental Disorders (MINDDS) group can have a huge impact in facilitating collaborative advances in understanding ND CNVs (https://mindds.eu/). MINDDS is a pan-European partnership between researchers, clinicians and patient organizations funded by the European Cooperation in Science and Technology (COST) network set up to facilitate research on NDDs associated with pathogenic CNV. The MINDDS initiative aims to focus on the development of

improved assessment, standardisation of research protocols and methodologies for data sharing and analysis in NDD research.

Involvement with international collaborative groups will likely accelerate the ascertainment of large samples, with standardised data and evaluation methods which will progress research in ND CNVs and in neurodevelopmental and psychiatric disorders in general. Specifically, for the work described in this thesis, a priority will be to identify suitable cohorts for replication and further refinement of findings of the studies. The TCD Neuropsychiatric Genetics group are actively involved with both of the above described collaborations which may facilitate access to suitable cohorts, or collaborative efforts to assemble such cohorts.

A significant limitation of one of the studies in this thesis, the study of the Philadelphia Neurodevelopmental Cohort, was the lack of access to standardised scoring methods for the dataset. I am now involved with a working group headed by Dr Kathleen Merikangas, Senior Investigator and Chief of the Genetic Epidemiology Research Branch in the Intramural Research Program at the National Institute of Mental Health (NIMH) to devise a final standardised set of clinical algorithms for the PNC cohort to enhance the usability of the data by researchers accessing these data through the National Center for Biotechnology Information. We plan to make the code for the algorithms freely available to other research groups to support further analysis of the dataset.

5.5.2 Further clinical characterisation of ND CNV phenotypes

In this thesis, the studies focussed on identifying neurodevelopmental and psychiatric phenotypic associations with ND CNVs using secondary data analysis of available data on large cohorts of individuals. It will be very important to carefully characterise the psychiatric and neurodevelopmental associations identified. For example, in the analysis of the Philadelphia Neurodevelopmental Cohort, individuals with ND CNVs were observed to present with increased risk of having subclinical psychotic symptoms. Careful clinical characterisation of these symptoms as well as other psychopathology could help to understand whether the subclinical psychotic symptoms identified in this group are prodromal psychotic symptoms or are indicative of non-psychotic psychopathology.

A number of national and international collaborative groups have been established to accelerate the collection of standardised, detailed phenotypic data for large neurodevelopmental cohorts. The Intellectual Disability and Mental Health: Assessing the Genomic Impact on Neurodevelopment (IMAGINE-ID) consortium is a multi-centre research team based in the UK studying associations between rare genetic variants and mental health outcomes in children with ID. The IMAGINE-ID team recruited over 3,400 families over a 5-year period. Using multi-informant deep phenotyping methods, the group identified a range of cognitive, behavioural and psychiatric domains affected by

ND CNVs. As a result of the large sample of individuals with ND CNVs and the depth of phenotypic data collected, they were also able to successfully identify evidence of qualitative and quantitative differences between phenotypes associated with different ND CNVs (Chawner et al., 2019). Other national collaborations focussed on deep phenotyping of ND CNVs include the Simons Searchlight initiative (formerly termed the Simons Variation in Individuals Project (Simons VIP)) and the Cardiff University Experiences of Children With Copy Number Variants (ECHO) Study. International collaboration between these groups have also yielded important insights into the details of psychiatric phenotypes associated with ND CNVs (Niarchou et al., 2019).

Very large cohorts of genotyped schizophrenia patients are becoming available and it is likely that whole genome sequence analysis of >30,000 patients will soon be completed. As this data is analysed, the subset of schizophrenia patients who carry rare mutations and CNVs of likely clinical significance will increase, as has been the case for other NDDs. There is a dearth of phenotypic information available for many schizophrenia cohorts. For future cohorts, having detailed phenotypic information with neurodevelopmental and medical history, will likely be helpful in refining variables that are predictive of ND CNV status, which ultimately may inform guidelines for genetic testing for schizophrenia patients.

Analysis of large cohorts of individuals with NDDs with rich phenotype data available will be essential to the improved understanding of the complex and subtle phenotypes associated with ND CNVs and will contribute to the deeper understanding of psychopathological and neurodevelopmental aetiologies.

5.5.3 Longitudinal studies

Longitudinal studies will allow us to closely follow the developmental trajectories of individuals with ND CNVs, detecting subtle changes in cognitive and behavioural presentations at different developmental stages. This will be essential to tracking the progressive appearance of clinical symptoms (or lack thereof) and the impact of environmental factors in ND CNV carriers. The European Autism Interventions - A Multicentre Study for Developing New Medications (EU-AIMS) Longitudinal European Autism Project (LEAP) is a large multi-centre, multi-disciplinary observational study operating worldwide that aims to identify and validate stratification biomarkers for ASD (Charman et al., 2017). There are huge opportunities from such a study to gain insights into the development of psychiatric comorbidities in children with ASD. It would be fascinating to be able to compare carriers of ND CNVs with non-carriers in this context if sample sizes permit sufficient statistical power.

Calkins et al. published a prospective evaluation of youths with early psychotic-like experiences from the PNC in 2017. They identified that symptom persistence was predicted by higher severity of

subclinical psychotic symptoms, lower global functioning, and prior psychiatric medication at baseline (Calkins et al., 2017). Examining the role of ND CNVs in the context of symptom persistence in this cohort would be fascinating; small sample sizes may limit the statistical power for such an analysis however the study by Calkins et al. demonstrates the important insights that can be gained from careful longitudinal follow up of young people with psychiatric or neurodevelopmental symptomatology.

5.5.4 Insights for Therapeutic Strategies

A further consideration for ND CNV research is to consider how phenotypic profiles of ND CNV carriers may impact on therapeutic strategies. Effects of therapeutic strategies for individuals with NDDs and psychiatric disorders are almost universally heterogeneous, likely as a result of the clinical and aetiological heterogeneity associated with these conditions. Recently, Tammimies et al. studied associations between copy number variation and response to social skills training in children with ASD. Carriers of clinically significant, large genic CNVs showed inferior outcomes post-intervention and at follow up (Tammimies et al., 2019).

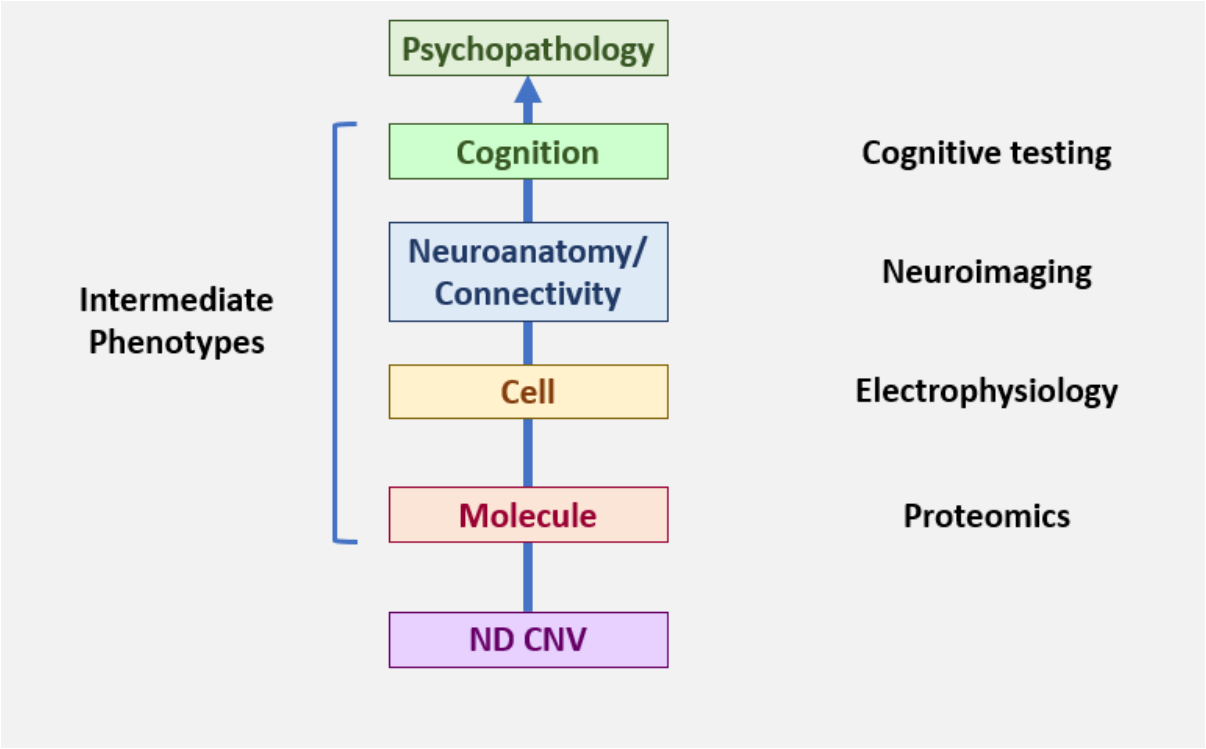
Copy number variation may also contribute to differences in how individuals with psychopathology respond to medications. Tansey et al. examined the role of CNVs in treatment response to antidepressants in depressed patients. No association was identified between antidepressant response for global number or burden of CNVs, but nominally significant associations with 15q13.3 duplications and exonic NRXN1 deletions were identified, with carriers of these CNVs having poorer response to treatments (Tansey et al., 2014).

ND CNV status could in the future be a factor in helping to guide treatment selection for patients.

5.5.5 Identifying intermediate phenotypes

Intermediate phenotypes are phenotypes that lie on the causal path between genetic variation and psychiatric or neurodevelopmental disorder (Figure 5-1), (Smoller et al., 2019). Investigation of intermediate phenotypes may help to illuminate the paths from ND CNVs to clinical outcomes.

Figure 5-1. Schematic diagram illustrating pathway from genotype to phenotype. Adapted from Sporns et al. (Sporns, 2013).



Mouse models can yield important insights into intermediate phenotypes associated with ND CNVs. A recent study of a mouse model of the 16p11.2 duplication identified evidence of relatively selective dysfunction of hippocampal-orbitofrontal-amygdaloid circuitry, modulated in part by GABAergic interneuron dysfunction, along with corresponding behavioural effects that resemble characteristics of patients with schizophrenia (Bristow et al., 2020). Mouse models of the 22q11.2 microdeletion have indicated deficits in the synchronization of the prefrontal cortex (PFC) and hippocampal networks linked with working memory demands, likely secondary to deficiencies in the growth of pyramidal cell axons in the PFC at the perinatal stage (del Pino et al., 2018).

Given that associations with ND CNVs transcend conventional diagnostic boundaries, it will be interesting to assess intermediate phenotypes of ND CNVs and how they fit in with multiple NDDs and psychiatric disorders. Voineskos and colleagues analysed structural brain differences associated with Nrxn1 variants in a sample of 53 healthy controls. Compared to non-risk homozygotes, healthy individuals who were homozygous for a specific risk allele (rs1045881C) exhibited reduced frontal lobe white matter volume and thalamic volume (Voineskos et al., 2011). The variant also influenced sensorimotor performance, a neurocognitive function that is impaired in both ASD (Sigman et al., 1981) and schizophrenia (Rajji et al., 2008; Voineskos et al., 2011; Welham et al., 2009a). Interestingly, there is also accumulating evidence indicating that sensorimotor dysfunction can have modulatory effects on mood and depressive symptoms as well (Canbeyli, 2010).

5.5.6 ND CNVs in the aetiological framework of neurodevelopmental disorders

As well as studying the phenotypic outcomes and neurobiological pathways specific to ND CNVs, it will be essential to understand where ND CNVs fit in to the aetiological context of neurodevelopmental and psychiatric disorders and how they may interact with environmental or other genetic risk factors to produce pathological phenotypes.

The developmental brain is characterised by substantial plasticity and is critically responsive to environmental change. Research over the last decade has advanced the understanding of the genetic basis of neurodevelopmental disorders, identifying risk loci and suggesting biological pathways through which genetic risk may be conferred, but much remains unknown. Twin studies of neurodevelopmental disorders indicate that genetic risk factors are likely to contribute a significant proportion, but not all, of the risk for neurodevelopmental disorders, suggesting that developmental and environmental factors also have a role to play (Moran et al., 2016). Epidemiological studies have suggested that many different factors such as prenatal infection/immune activation, advanced parental age, perinatal complications and social stress and/or trauma in youth are associated with increased risk of developing neurodevelopmental disorders (Boat et al., 2015; Modabbernia et al., 2017; Moran et al., 2016; Pesta et al., 2020; Zwicker et al., 2018).

The hypothesis of "behavioural sensitization" has been postulated as an aetiological model for schizophrenia (Lardinois et al., 2011; Van Winkel et al., 2008). The model asserts that repetitive stress increases the biological and behavioural response to subsequent stressors. Dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis, with associated increased plasma cortisol has been suggested to contribute to subsequent dopamine dysregulation and sensitisation of the mesolimbic region of the brain. Mutations and polymorphisms of genes impacting catecholamine neurotransmission, neuroplasticity and stress activation (for example Catechol-O-methyltransferase, dopamine transmitters, brain-derived neurotrophic factor, HPA axis) could then impact the extent of the sensitization process. However, findings from studies examining genetic-environmental interaction hypotheses formulated on this basis are not robust at this point (Giegling et al., 2017; Lardinois et al., 2011; Van Winkel et al., 2008).

The role of life events in psychosis was examined by Beards et al. in a review and meta-analysis, identifying some evidence that adult life events were associated with increased risk of psychotic disorder and subclinical psychotic experiences (OR 3.19, 95% CI 2.15–4.75), however the authors urged caution in the interpretation of the results owing to low methodological quality of many of the studies (Beards et al., 2013). Miller et al. completed a review of studies examining prenatal inflammation and risk of schizophrenia, concluding that inflammation may be associated with abnormal neurodevelopment (Miller et al., 2013).

The most well-established environmental risk factor for ASD is advanced parental age, with most evidence indicating that this risk factor contributes to genetic risk (Hultman et al., 2011; Muhle et

al., 2018; Sandin et al., 2012). For every 10-year increase in the age of the parent, risk of ASD increases by 18% for mothers and 21% for fathers (Modabbernia et al., 2017). Maternal and paternal age appear to be independent risk factors (Modabbernia et al., 2017; Muhle et al., 2018). Premature birth is a further well-supported risk factor for ASD (Lyall et al., 2017; Muhle et al., 2018). Major cognitive and motor impairments are more likely to be associated with ASD in those with very low birth weight and extreme preterm birth (Movsas et al., 2013). Many other environmental factors including birth complications with hypoxia or trauma, maternal obesity, medication exposure during pregnancy and prenatal use of anticonvulsants have been associated with ASD risk however further evidence is needed to confirm associations robustly (Muhle et al., 2018).

Epigenetic factors that result in changes in gene expression and function may be inherited or acquired during an individual's lifetime as a result of various environmental influences (for example toxins, stress, medications). Epigenetic regulation of the genome has been identified as a mechanism likely to mediate prenatal environmental factors (such as viruses, toxins, nutritional deficiencies) influencing risk of schizophrenia (Giegling et al., 2017; Maric et al., 2012).

Twin studies have identified the influence of epigenetics on the modulation of ASD phenotype; for example, Wong et al. studied 50 pairs of monozygotic twins discordant for ASD and demonstrated a number of ASD-associated differentially methylated regions (Wong et al., 2014).A recent review of 215 ASD candidate genes estimated that 19.5% were epigenetic regulators, suggesting a significant role for genes with epigenetic-modulating functions in ASD susceptibility (Duffney et al., 2018) (Rylaarsdam et al., 2019).

A "multihit" model has been proposed for schizophrenia and autism spectrum disorder, which is likely to involve a combination of SNPs, rare penetrant mutations such as CNVs, and other genetic and environmental factors (Grice et al., 2015; Leblond et al., 2012; Moran et al., 2016; Rudd et al., 2014).

Mazina et al. assessed for interaction effects between CNVs and an established environmental risk factor for ASD- maternal immune activation (MIA). They found significant interaction effects between CNVs and MIA for social communication deficits and repetitive and restrictive behaviours in children with ASD, but not for cognitive or adaptive outcomes (Mazina et al., 2015). Weiner et al. examined the influence of common variant risk on ASD in individuals with rare deleterious *de novo* variants, finding an additive effect between the common and rare variants for ASD (Weiner et al., 2017).

To gain a comprehensive understanding of the complexities of neurodevelopmental and psychiatric aetiologies, it is likely that interactive effects of common and rare variants as well as environmental influences will need to be examined in well powered and designed studies of gene x environment, gene x gene and epigenetic studies.

5.5.7 Pathobiology of ND CNVs

The clinical phenotypes associated with ND CNVs are highly variable. The molecular pathologies underlying neurodevelopmental and psychiatric disorders are not yet known. Identifying specific cellular phenotypes associated with ND CNVs may be helpful in progressing our understanding of neurodevelopmental and psychiatric phenotypes associated with these CNVs.

Recently Khan et al., induced pluripotent stem cells from 15 individuals with a 22q11.2 deletion and 15 individuals without the deletion and established three-dimensional cortical organoids and cortical neurons from the stem cells. They assessed gene expression in the developing organoids and found changes in the expression of genes associated with neuronal excitability in the 22q11.2 deletion organoids. Neuronal excitability was examined further by imaging cells and recording the electrical activity of neurons obtained from the individuals with the deletions. Abnormalities were identified in calcium transportation systems in the cells that were related to a defect in the resting electrical potential of the cell membrane (Khan et al., 2020). The DGCR8 gene, contained within the 22q11.2 region has previously been implicated in neuronal abnormalities in rodent models of the syndrome (Schofield et al., 2011). Khan et al. identified that heterozygous loss of the gene was enough to induce the changes in excitability observed in the 22q11.2 deletion neurons. They also identified that treating these neurons with specific antipsychotic medications could restore deficits in resting membrane potential (Khan et al., 2020). This study demonstrates the value of developing ND CNV models to help elucidate the neurobiology of complex neuropsychiatric disorders and the potential for identifying therapeutic targets.

5.5.8 Clinical considerations

5.5.8.1 Accumulating Clinical and Phenotypic Data for Variant Interpretation

For some ND CNVs (e.g. 22q11.2 deletion, Prader-Willi syndrome/Angelman syndrome), clinical phenotypes are relatively well established and clinical guidelines inform the optimal management and investigation of patients (Bassett et al., 2011; McCandless, 2010). The evidence bases for clinical and phenotypic presentations of other recurrent ND CNVs are not as well-established, limiting the clinical information that can be provided to patients and families. International databases such as DECIPHER (Wright et al., 2018) and ClinVar (Landrum et al., 2016) have been established to support interpretation of variants (Feenstra et al., 2006; Firth et al., 2009; Kaminsky et al., 2011).

The collection of detailed phenotypic data is expensive and time consuming and it will be essential to consider ways of collecting data in a cost and resource-efficient manner. Electronic health records (EHRs) could offer a rich source of phenotypic data that could empower variant interpretation (Son

et al., 2018). Although EHRs are not yet widely used in the public health system in Ireland, a national EHR has been identified by the national health service in Ireland as a key requirement for the future delivery of healthcare ("National Electronic Health Record," 2016). There are ethical, legal, data security, and intellectual property issues to consider as EHRs are integrated into the clinical system in Ireland. It will be important for researchers and clinician scientists to be centrally involved in the integration of EHRs into clinical practice in Ireland so that the data collected can be leveraged efficiently into research. Phenotypic data collected through EHRs could provide important data that contribute to the clinical interpretation of rare variants for patients and could ultimately help to establish insights into the aetiologies of psychopathologies and NDDs.

5.5.8.2 Clinical care pathways

Neurodevelopmental CNVs are associated with multiple neurodevelopmental (Chawner et al., 2019; Coe et al., 2014; Marshall et al., 2017; Sanders et al., 2015), psychiatric (Chawner et al., 2019; Guyatt et al., 2018; Kendall et al., 2019; Niarchou et al., 2019; Niarchou et al., 2014; Stefansson et al., 2014) and medical conditions (Bassett et al., 2011; Mefford et al., 2010). Interventions required over a lifetime from various medical, developmental and psychiatric services as well as academic and psychosocial support likely result in high financial costs for families and the health system, particularly where these services are delivered at different sites in an uncoordinated manner (Lawlor et al., 2017). The European Union Committee of Experts on Rare Diseases have made the recommendation that care of patients with rare diseases should be delivered through centres of expertise; recommended inputs include psychosocial care and the development of clinical care pathways (Aymé et al., 2014; Lawlor et al., 2017). Co-ordinated multi-disciplinary care delivery through dedicated clinical care pathways could be beneficial to individuals with ND CNVs and their families with development of standardised screening, psychoeducation and management protocols. This could help to develop gold-standard care for this group of individuals and could also reduce cost of care (Battersby, 2005; Peter et al., 2011).

5.5.8.3 Psychoeducation and individual perspectives

Psychiatrists across multiple studies have reported concern that they do not feel they have the expertise necessary to manage genomic testing and findings (Ward et al., 2019). It will be important to provide clinicians with standardised guidelines and practice recommendations in this context. I am a member of the COST funded Enhancing Psychiatric Genetic Counselling, Testing, and Training in Europe (EnGagE) initiative. This group aims to develop standardised guidelines, practice recommendations and research protocols for psychiatric genetic testing and counselling in Europe.

5.6 Conclusion

Neurodevelopmental CNVs present with clinically pleiotropic effects, transcending psychiatric diagnostic borders. This thesis identified associations between neurodevelopmental and psychiatric phenotypes and ND CNVs that could have clinical impacts for a number of previously understudied populations. In an ASD sample, females with an ND CNV were significantly more likely to present with depressive symptoms than males with or without an ND CNV. ND CNV status was also significantly associated with increased risk of affective problems in an associated sample of typically developing siblings; this effect was driven by female carriers of ND CNVs only. In a large clinical cohort of youths, an increased risk of subclinical psychotic symptoms in youths with ND CNVs was identified, which was independent of neurodevelopmental comorbidity in the sample. In adults with schizophrenia three phenotypic variables indicative of early neurodevelopmental compromise were significant in modelling positive carrier status for schizophrenia associated CNVs; replication analysis in a separate cohort confirmed neurodevelopmental phenotypic variables as significant predictors of positive carrier status for schizophrenia associated CNV status. Exploration of these results in longitudinal cohorts could facilitate successful clinical translation of the findings. Gaining clarity on the neurodevelopmental and psychiatric phenotypic effects of ND CNVs in different populations is essential to providing clinically relevant information and services to carriers and may contribute to our understanding of the marked heterogeneity observed within disorders and frequent comorbidity between disorders.

References

Abdallah, M. W., Greaves-Lord, K., Grove, J., Norgaard-Pedersen, B., Hougaard, D. M., & Mortensen, E. L. (2011). Psychiatric comorbidities in autism spectrum disorders: findings from a Danish Historic Birth Cohort. *Eur Child Adolesc Psychiatry, 20*(11-12), 599-601. doi:10.1007/s00787-011-0220-2

Achenbach, T. (2001). Manual for the Child Behavior Checklist/6-18 and 2001 profile. *Burlington: University of Vermont*.

Achenbach, T. M., & Dumenci, L. (2001). Advances in empirically based assessment: Revised cross-informant syndromes and new DSM-oriented scales for the CBCL, YSR, and TRF: Comment on Lengua, Sadowski, Friedrich, and Fisher (2001).

Achenbach, T. M., & Edelbrock, C. S. (1983). Manual for the child behavior checklist and revised child behavior profile.

Achenbach, T. M., Ivanova, M. Y., Rescorla, L. A., Turner, L. V., & Althoff, R. R. (2016). Internalizing/Externalizing Problems: Review and Recommendations for Clinical and Research Applications. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(8), 647-656. doi:https://doi.org/10.1016/j.jaac.2016.05.012

Addington, J., Stowkowy, J., & Weiser, M. (2015). Screening tools for clinical high risk for psychosis. *Early Intervention in Psychiatry*, *9*(5), 345-356. doi:10.1111/eip.12193

Ahn, K., Gotay, N., Andersen, T. M., Anvari, A. A., Gochman, P., Lee, Y., . . . Rapoport, J. L. (2014). High rate of disease-related copy number variations in childhood onset schizophrenia. *Mol Psychiatry, 19*(5), 568-572. doi:10.1038/mp.2013.59

Aitken, M., VanderLaan, D. P., Wasserman, L., Stojanovski, S., & Zucker, K. J. (2016). Self-harm and suicidality in children referred for gender dysphoria. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(6), 513-520.

Aitman, T. J., Dong, R., Vyse, T. J., Norsworthy, P. J., Johnson, M. D., Smith, J., . . . Petretto, E. (2006). Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature, 439*(7078), 851-855.

Al Shehhi, M., Forman, E. B., Fitzgerald, J. E., McInerney, V., Krawczyk, J., Shen, S., . . . Lynch, S. A. (2019). NRXN1 deletion syndrome; phenotypic and penetrance data from 34 families. *European journal of medical genetics, 62*(3), 204-209. doi:https://doi.org/10.1016/j.ejmg.2018.07.015

Alati, R., Gunnell, D., Najman, J., Williams, G., & Lawlor, D. (2009). Is IQ in Childhood Associated with Suicidal Thoughts and Attempts? Findings from The Mater University Study of Pregnancy and Its Outcomes. *Suicide and Life-Threatening Behavior, 39*(3), 282-293. doi:10.1521/suli.2009.39.3.282

Aleman, A., Kahn, R. S., & Selten, J.-P. (2003). Sex differences in the risk of schizophrenia: evidence from meta-analysis. *Archives of General psychiatry, 60*(6), 565-571.

Allsopp, K., Read, J., Corcoran, R., & Kinderman, P. (2019). Heterogeneity in psychiatric diagnostic classification. *Psychiatry research, 279*, 15-22. doi:https://doi.org/10.1016/j.psychres.2019.07.005

Aman, H., Naeem, F., Farooq, S., & Ayub, M. (2016). Prevalence of nonaffective psychosis in intellectually disabled clients: systematic review and meta-analysis. *Psychiatric genetics, 26*(4), 145-155.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

Anderson, K. K., Norman, R., MacDougall, A., Edwards, J., Palaniyappan, L., Lau, C., & Kurdyak, P. (2018). Effectiveness of Early Psychosis Intervention: Comparison of Service Users and Nonusers in Population-Based Health Administrative Data. *American Journal of Psychiatry, 175*(5), 443-452. doi:10.1176/appi.ajp.2017.17050480

Anderson, P., Jané-llopis, E., & Hosman, C. (2011). Reducing the silent burden of impaired mental health. *Health Promotion International, 26*(suppl_1), i4-i9. doi:10.1093/heapro/dar051

Anney, R. J. L., Ripke, S., Anttila, V., Grove, J., Holmans, P., Huang, H., . . . The Autism Spectrum Disorders Working Group of The Psychiatric Genomics, C. (2017). Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Molecular Autism, 8*(1), 21. doi:10.1186/s13229-017-0137-9

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., . . . Murray, R. (2018). Analysis of shared heritability in common disorders of the brain. *Science, 360*(6395). doi:10.1126/science.aap8757

Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). Benefits and costs of prevention and early intervention programs for youth. In: Olympia, WA: Washington State Institute for Public Policy.

Appelbaum, P. S., & Benston, S. (2017). Anticipating the Ethical Challenges of Psychiatric Genetic Testing. *Current psychiatry reports, 19*(7), 39. doi:10.1007/s11920-017-0790-x

Arim, R. G., Miller, A. R., Guèvremont, A., Lach, L. M., Brehaut, J. C., & Kohen, D. E. (2017). Children with neurodevelopmental disorders and disabilities: a population-based study of healthcare service utilization using administrative data. *Developmental Medicine & Child Neurology, 59*(12), 1284-1290. doi:10.1111/dmcn.13557

Aymé, S., & Rodwell, C. (2014). The European Union Committee of Experts on Rare Diseases: three productive years at the service of the rare disease community. *Orphanet Journal of Rare Diseases, 9*(1), 30. doi:10.1186/1750-1172-9-30

Baker, K., Costain, G., Fung, W. L. A., & Bassett, A. S. (2014). Chromosomal microarray analysis—a routine clinical genetic test for patients with schizophrenia. *The Lancet Psychiatry, 1*(5), 329-331.

Baranne, M. L., & Falissard, B. (2018). Global burden of mental disorders among children aged 5-14 years. *Child and adolescent psychiatry and mental health, 12*, 19-19. doi:10.1186/s13034-018-0225-4

Bass, N., & Skuse, D. (2018). Genetic testing in children and adolescents with intellectual disability. *Curr Opin Psychiatry, 31*(6), 490-495. doi:10.1097/yco.0000000000000456

Bassett, A. S., Chow, E. W. C., Husted, J., Weksberg, R., Caluseriu, O., Webb, G. D., & Gatzoulis, M. A. (2005). Clinical features of 78 adults with 22q11 deletion syndrome. *American Journal of Medical Genetics Part A, 138A*(4), 307-313. doi:10.1002/ajmg.a.30984

Bassett, A. S., Hodgkinson, K., Chow, E. W., Correia, S., Scutt, L. E., & Weksberg, R. (1998). 22q11 deletion syndrome in adults with schizophrenia. *American journal of medical genetics, 81*(4), 328-337.

Bassett, A. S., McDonald-McGinn, D. M., Devriendt, K., Digilio, M. C., Goldenberg, P., Habel, A., . . . International 22q11.2 Deletion Syndrome, C. (2011). Practical guidelines for managing patients with 22q11.2 deletion syndrome. *The Journal of pediatrics, 159*(2), 332-339.e331. doi:10.1016/j.jpeds.2011.02.039

Batstra, L., Bos, E. H., & Neeleman, J. (2002). Quantifying psychiatric comorbidity--lessions from chronic disease epidemiology. *Soc Psychiatry Psychiatr Epidemiol, 37*(3), 105-111. doi:10.1007/s001270200001

Battaglia, A., Doccini, V., Bernardini, L., Novelli, A., Loddo, S., Capalbo, A., . . . Carey, J. C. (2013). Confirmation of chromosomal microarray as a first-tier clinical diagnostic test for individuals with developmental delay, intellectual disability, autism spectrum disorders and dysmorphic features. *European Journal of Paediatric Neurology, 17*(6), 589-599.

Battersby, M. W. (2005). Health reform through coordinated care: SA HealthPlus. *Bmj, 330*(7492), 662-665. doi:10.1136/bmj.330.7492.662

Beards, S., Gayer-Anderson, C., Borges, S., Dewey, M. E., Fisher, H. L., & Morgan, C. (2013). Life events and psychosis: a review and meta-analysis. *Schizophrenia Bulletin, 39*(4), 740-747. doi:10.1093/schbul/sbt065

Beautrais, A. L. (2002). Gender issues in youth suicidal behaviour. *Emergency Medicine, 14*(1), 35-42. doi:10.1046/j.1442-2026.2002.00283.x

Beitchman, J., Inglis, A., & Schachter, D. (1992a). Child psychiatry and early intervention: IV. The externalizing disorders. In: SAGE Publications Sage CA: Los Angeles, CA.

Beitchman, J. H., Inglis, A., & Schachter, D. (1992b). Child Psychiatry and Early Intervention: II. The Internalizing Disorders*. *The Canadian Journal of Psychiatry, 37*(4), 234-239. doi:10.1177/070674379203700405

Bérubé, R. (2014). Bibliography of published studies using the Achenbach System of Empirically Based Assessment (ASEBA). In: Burlington, VT: University of Vermont, Research Center for Children, Youth ….

Bevan Jones, R., Thapar, A., Lewis, G., & Zammit, S. (2012). The association between early autistic traits and psychotic experiences in adolescence. *Schizophrenia research, 135*(1), 164-169. doi:https://doi.org/10.1016/j.schres.2011.11.037

Bishop, S. L., Farmer, C., Bal, V., Robinson, E. B., Willsey, A. J., Werling, D. M., . . . Thurm, A. (2017). Identification of Developmental and Behavioral Markers Associated With Genetic Abnormalities in Autism Spectrum Disorder. *Am J Psychiatry, 174*(6), 576-585. doi:10.1176/appi.ajp.2017.16101115

Bitsika, V., & Sharpley, C. F. (2015). Differences in the Prevalence, Severity and Symptom Profiles of Depression in Boys and Adolescents with an Autism Spectrum Disorder versus Normally Developing Controls. *International Journal of Disability, Development and Education, 62*(2), 158-167. doi:10.1080/1034912X.2014.998179

Bitta, M., Kariuki, S. M., Abubakar, A., & Newton, C. R. J. C. (2017). Burden of neurodevelopmental disorders in low and middle-income countries: A systematic review and meta-analysis. *Wellcome open research, 2*, 121-121. doi:10.12688/wellcomeopenres.13540.3

Blacker, D., Racette, S. R., Sklar, P., & Smoller, J. W. (2005). Psychiatric genetics: a survey of psychiatrists' knowledge, opinions, and practice patterns. *J Clin Psychiatry, 66*(7), 821-830.

Boat, T. F., & Wu, J. T. (2015). Clinical characteristics of intellectual disabilities. In *Mental disorders and disabilities among low-income children*: National Academies Press (US), Sciences, Social National Academies of Sciences, Engineering, Medicine

Boivin, M. J., Kakooza, A. M., Warf, B. C., Davidson, L. L., & Grigorenko, E. L. (2015). Reducing neurodevelopmental disorders and disability through research and interventions. *Nature, 527*(7578), S155-S160. doi:10.1038/nature16029

Bourgeron, T. (2015). From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nature Reviews Neuroscience, 16*(9), 551-563.

Bourgeron, T. (2016). Current knowledge on the genetics of autism and propositions for future research. *Comptes Rendus Biologies, 339*(7), 300-307. doi:https://doi.org/10.1016/j.crvi.2016.05.004

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., . . . Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics, 127*(6), 1034-1042. doi:10.1542/peds.2010-2989

Brabbing-Goldstein, D., Reches, A., Svirsky, R., Bar-Shira, A., & Yaron, Y. (2018). Dilemmas in genetic counseling for low-penetrance neuro-susceptibility loci detected on prenatal chromosomal microarray analysis. *American journal of obstetrics and gynecology, 218*(2), 247. e241-247. e212.

Bristow, G. C., Thomson, D. M., Openshaw, R. L., Mitchell, E. J., Pratt, J. A., Dawson, N., & Morris, B. J. (2020). 16p11 Duplication Disrupts Hippocampal-Orbitofrontal-Amygdala

Connectivity, Revealing a Neural Circuit Endophenotype for Schizophrenia. *Cell Reports, 31*(3), 107536.

Bryson, S. E., Bradley, E. A., Thompson, A., & Wainwright, A. (2008). Prevalence of autism among adolescents with intellectual disabilities. *The Canadian Journal of Psychiatry, 53*(7), 449-459.

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

Buckley, P. F., Miller, B. J., Lehrer, D. S., & Castle, D. J. (2009). Psychiatric comorbidities and schizophrenia. *Schizophrenia Bulletin, 35*(2), 383-402. doi:10.1093/schbul/sbn135

Calkins, M. E., Merikangas, K. R., Moore, T. M., Burstein, M., Behr, M. A., Satterthwaite, T. D., . . . Mentch, F. D. (2015). The Philadelphia Neurodevelopmental Cohort: constructing a deep phenotyping collaborative. *Journal of Child Psychology and Psychiatry, 56*(12), 1356-1369.

Calkins, M. E., Moore, T. M., Satterthwaite, T. D., Wolf, D. H., Turetsky, B. I., Roalf, D. R., . . . Gur, R. E. (2017). Persistence of psychosis spectrum symptoms in the Philadelphia Neurodevelopmental Cohort: a prospective two-year follow-up. *World Psychiatry, 16*(1), 62-76. doi:10.1002/wps.20386

Canbeyli, R. (2010). Sensorimotor modulation of mood and depression: An integrative review. *Behavioural Brain Research, 207*(2), 249-264. doi:https://doi.org/10.1016/j.bbr.2009.11.002

Cardno, A. G., & Gottesman, I. I. (2000). Twin studies of schizophrenia: From bow-and-arrow concordances to Star Wars Mx and functional genomics. *American journal of medical genetics, 97*(1), 12-17. doi:10.1002/(sici)1096-8628(200021)97:1<12::Aid-ajmg3>3.0.Co;2-u

Carrasco, X., Castillo, S., Aravena, T., Rothhammer, P., & Aboitiz, F. (2005). Williams syndrome: pediatric, neurologic, and cognitive development. *Pediatr Neurol, 32*(3), 166-172. doi:10.1016/j.pediatrneurol.2004.09.013

Carrion-Castillo, A., Franke, B., & Fisher, S. E. (2013). Molecular genetics of dyslexia: an overview. *Dyslexia, 19*(4), 214-240.

Carter, M. T., & Scherer, S. W. (2013). Autism spectrum disorder in the genetics clinic: a review. *Clin Genet, 83*(5), 399-407. doi:10.1111/cge.12101

Cassidy, S., Bradley, P., Robinson, J., Allison, C., McHugh, M., & Baron-Cohen, S. (2014). Suicidal ideation and suicide plans or attempts in adults with Asperger's syndrome attending a specialist diagnostic clinic: a clinical cohort study. *The Lancet Psychiatry, 1*(2), 142-147.

Cassidy, S., & Rodgers, J. (2017). Understanding and prevention of suicide in autism. *The Lancet Psychiatry, 4*(6), e11.

Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: confirmation of high prevalence. *American Journal of Psychiatry, 162*(6), 1133-1141.

Charman, T., Loth, E., Tillmann, J., Crawley, D., Wooldridge, C., Goyard, D., . . . Buitelaar, J. K. (2017). The EU-AIMS Longitudinal European Autism Project (LEAP): clinical characterisation. *Molecular Autism, 8*(1), 27. doi:10.1186/s13229-017-0145-9

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: data from the Special Needs and Autism Project (SNAP). *Psychol Med, 41*(3), 619-627. doi:10.1017/s0033291710000991

Chatterjee, N., Shi, J., & García-Closas, M. (2016). Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet, 17*(7), 392-406. doi:10.1038/nrg.2016.27

Chawner, S. J. R. A., Owen, M. J., Holmans, P., Raymond, F. L., Skuse, D., Hall, J., & van den Bree, M. B. M. (2019). Genotype–phenotype associations in children with copy number variants associated with high neuropsychiatric risk in the UK (IMAGINE-ID): a case-control cohort study. *The Lancet Psychiatry, 6*(6), 493-505. doi:https://doi.org/10.1016/S2215-0366(19)30123-3

Chen, J. J. (2003). Communicating complex information: the interpretation of statistical interaction in multiple logistic regression analysis. *American journal of public health, 93*(9), 1376-a-1377.

Chen, M.-H., Pan, T.-L., Lan, W.-H., Hsu, J.-W., Huang, K.-L., Su, T.-P., . . . Chen, T.-J. (2017). Risk of Suicide Attempts Among Adolescents and Young Adults With Autism Spectrum Disorder: A Nationwide Longitudinal Follow-Up Study. *The Journal of clinical psychiatry, 78*(9), e1174-e1179.

Cholemkery, H., Mojica, L., Rohrmann, S., Gensthaler, A., & Freitag, C. M. (2014). Can Autism Spectrum Disorders and Social Anxiety Disorders be Differentiated by the Social Responsiveness Scale in Children and Adolescents? *Journal of Autism and Developmental Disorders, 44*(5), 1168-1182. doi:10.1007/s10803-013-1979-4

Chow, L., Garcia-Barcelo, M., Wing, Y., & Waye, M. M. (1999). Schizophrenia and hypocalcaemia: variable phenotype of deletion at chromosome 22q11. *Australian & New Zealand Journal of Psychiatry, 33*(5), 760-762.

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, 309*(16), 1696-1703. doi:10.1001/jama.2013.2270

Christian G. Bouwkamp, Ph.D. ,, Anneke J.A. Kievit, M.D., Ph.D. ,, Sander Markx, M.D. ,, Joseph I. Friedman, M.D. ,, Laura van Zutven, Ph.D. ,, Rick van Minkelen, Ph.D. ,, . . . Steven A. Kushner, M.D., Ph.D. (2017). Copy Number Variation in Syndromic Forms of Psychiatric Illness: The Emerging Value of Clinical Genetic Testing in Psychiatry. *American Journal of Psychiatry, 174*(11), 1036-1050. doi:10.1176/appi.ajp.2017.16080946

Clarke, T. K., Lupton, M. K., Fernandez-Pujals, A. M., Starr, J., Davies, G., Cox, S., . . . McIntosh, A. M. (2016). Common polygenic risk for autism spectrum disorder (ASD) is associated

with cognitive ability in the general population. *Molecular Psychiatry, 21*(3), 419-425. doi:10.1038/mp.2015.12

Coe, B. P., Witherspoon, K., Rosenfeld, J. A., Van Bon, B. W., Vulto-van Silfhout, A. T., Bosco, P., . . . Vissers, L. E. (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nature Genetics, 46*(10), 1063.

Cohen, D. J., & Volkmar, F. R. (1997). *Handbook of autism and pervasive developmental disorders*: John Wiley & Sons Inc.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*: Routledge.

Colella, S., Yau, C., Taylor, J. M., Mirza, G., Butler, H., Clouston, P., . . . Ragoussis, J. (2007). QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res, 35*(6), 2013-2025. doi:10.1093/nar/gkm076

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, 72*(5), 415-423. doi:10.1001/jamapsychiatry.2014.3028

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

Cooper, S.-A., Smiley, E., Morrison, J., Williamson, A., & Allan, L. (2007). Mental ill-health in adults with intellectual disabilities: prevalence and associated factors. *British Journal of Psychiatry, 190*(1), 27-35. doi:10.1192/bjp.bp.106.022483

Cooper, S. A., & Bailey, N. M. (2001). Psychiatric disorders amongst adults with learning disabilities - prevalence and relationship to ability level. *Ir J Psychol Med, 18*(2), 45-53. doi:10.1017/s0790966700006315

Corvin, A., & Sullivan, P. F. (2016). What Next in Schizophrenia Genetics for the Psychiatric Genomics Consortium? *Schizophrenia Bulletin, 42*(3), 538-541. doi:10.1093/schbul/sbw014

Costain, G., Chow, E. W. C., Ray, P. N., & Bassett, A. S. (2012a). Caregiver and adult patient perspectives on the importance of a diagnosis of 22q11.2 deletion syndrome. *Journal of Intellectual Disability Research, 56*(6), 641-651. doi:10.1111/j.1365-2788.2011.01510.x

Costain, G., Esplen, M. J., Toner, B., Scherer, S. W., Meschino, W. S., Hodgkinson, K. A., & Bassett, A. S. (2012b). Evaluating Genetic Counseling for Individuals With Schizophrenia in the Molecular Age. *Schizophrenia Bulletin, 40*(1), 78-87. doi:10.1093/schbul/sbs138

Costain, G., Lionel, A. C., Fu, F., Stavropoulos, D. J., Gazzellone, M. J., Marshall, C. R., . . . Bassett, A. S. (2014). Adult neuropsychiatric expression and familial segregation of 2q13 duplications. *Am J Med Genet B Neuropsychiatr Genet, 165b*(4), 337-344. doi:10.1002/ajmg.b.32236

Costain, G., Lionel, A. C., Merico, D., Forsythe, P., Russell, K., Lowther, C., . . . Speevak, M. (2013). Pathogenic rare copy number variants in community-based schizophrenia suggest a potential role for clinical microarrays. *Hum Mol Genet, 22*(22), 4485-4501.

Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The lancet, 381*(9875), 1371-1379.

Cuthbert, B. N. (2014). The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry, 13*(1), 28-35.

Dadds, M. R., Spence, S. H., Holland, D. E., Barrett, P. M., & Laurens, K. R. (1997). Prevention and early intervention for anxiety disorders: a controlled trial. *Journal of Consulting and Clinical Psychology, 65*(4), 627.

Dazzan, P., & Murray, R. M. (2002). Neurological soft signs in first-episode psychosis: a systematic review. *The British Journal of Psychiatry, 181*(S43), s50-s57.

De Crescenzo, F., Postorino, V., Siracusano, M., Riccioni, A., Armando, M., Curatolo, P., & Mazzone, L. (2019). Autistic Symptoms in Schizophrenia Spectrum Disorders: A Systematic Review and Meta-Analysis. *Frontiers in Psychiatry, 10*(78). doi:10.3389/fpsyt.2019.00078

de Vries, B. B. A., Pfundt, R., Leisink, M., Koolen, D. A., Vissers, L. E. L. M., Janssen, I. M., . . . Veltman, J. A. (2005). Diagnostic Genome Profiling in Mental Retardation. *The American Journal of Human Genetics, 77*(4), 606-616. doi:https://doi.org/10.1086/491719

del Pino, I., Rico, B., & Marín, O. (2018). Neural circuit dysfunction in mouse models of neurodevelopmental disorders. *Current Opinion in Neurobiology, 48*, 174-182. doi:https://doi.org/10.1016/j.conb.2017.12.013

Derks, E. M., Ayub, M., Chambert, K., Del Favero, J., Johnstone, M., MacGregor, S., . . . Blackwood, D. H. (2013). A genome wide survey supports the involvement of large copy number variants in schizophrenia with and without intellectual disability. *Am J Med Genet B Neuropsychiatr Genet, 162b*(8), 847-854. doi:10.1002/ajmg.b.32189

Diagnostic and statistical manual of mental disorders : DSM-IV. (1994). Fourth edition. Washington, DC : American Psychiatric Association, [1994] ©1994.

Dolcetti, A., Silversides, C. K., Marshall, C. R., Lionel, A. C., Stavropoulos, D. J., Scherer, S. W., & Bassett, A. S. (2013). 1q21. 1 Microduplication expression in adults. *Genetics in Medicine, 15*(4), 282-289.

Duffney, L. J., Valdez, P., Tremblay, M. W., Cao, X., Montgomery, S., McConkie-Rosell, A., & Jiang, Y.-h. (2018). Epigenetics and autism spectrum disorder: A report of an autism case with mutation in H1 linker histone HIST1H1E and literature review. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 177*(4), 426-433. doi:https://doi.org/10.1002/ajmg.b.32631

Dugoff, L., Norton, M. E., & Kuller, J. A. (2016). The use of chromosomal microarray for prenatal diagnosis. *American journal of obstetrics and gynecology, 215*(4), B2-B9. doi:https://doi.org/10.1016/j.ajog.2016.07.016

Ebesutani, C., Bernstein, A., Nakamura, B. J., Chorpita, B. F., Higa-McMillan, C. K., & Weisz, J. R. (2010). Concurrent Validity of the Child Behavior Checklist DSM-Oriented Scales: Correspondence with DSM Diagnoses and Comparison to Syndrome Scales. *J Psychopathol Behav Assess, 32*(3), 373-384. doi:10.1007/s10862-009-9174-9

Edwards, J. J., & Gelb, B. D. (2016). Genetics of congenital heart disease. *Current opinion in cardiology, 31*(3), 235-241. doi:10.1097/HCO.0000000000000274

Eichler, E. E. (2008). Copy Number Variation and Human Disease. *Nature Education, 1(3):1*.

Einfeld, S. L., Ellis, L. A., & Emerson, E. (2011). Comorbidity of intellectual disability and mental disorder in children and adolescents: A systematic review. *Journal of Intellectual and Developmental Disability, 36*(2), 137-143.

El-Hashash, E., & El-Absy, K. (2018). Methods for Determining the Tetrachoric Correlation Coefficient for Binary Variables. *Asian Journal of Probability and Statistics*, 1-12.

Emerson, E., & Hatton, C. (2007). Mental health of children and adolescents with intellectual disabilities in Britain. *British Journal of Psychiatry, 191*(6), 493-499. doi:10.1192/bjp.bp.107.038729

Erskine, H. E., Moffitt, T. E., Copeland, W. E., Costello, E. J., Ferrari, A. J., Patton, G., . . . Scott, J. G. (2015). A heavy burden on young minds: the global burden of mental and substance use disorders in children and youth. *Psychological medicine, 45*(7), 1551-1563. doi:10.1017/S0033291714002888

Falkenburg, J., & Tracy, D. K. (2014). Sex and schizophrenia: a review of gender differences. *Psychosis, 6*(1), 61-69. doi:10.1080/17522439.2012.733405

Fan, Y., Wu, Y., Wang, L., Wang, Y., Gong, Z., Qiu, W., . . . Yu, Y. (2018). Chromosomal microarray analysis in developmental delay and intellectual disability with comorbid conditions. *BMC medical genomics, 11*(1), 49-49. doi:10.1186/s12920-018-0368-4

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods, 41*(4), 1149-1160.

Feenstra, I., Fang, J., Koolen, D. A., Siezen, A., Evans, C., Winter, R. M., . . . Van Ravenswaaij, C. (2006). European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations (ECARUCA); an online database for rare chromosome abnormalities. *European journal of medical genetics, 49*(4), 279-291.

Ferdinand, R. F. (2008). Validity of the CBCL/YSR DSM-IV scales anxiety problems and affective problems. *Journal of anxiety disorders, 22*(1), 126-134.

Fine, S. E., Weissman, A., Gerdes, M., Pinto-Martin, J., Zackai, E. H., McDonald-McGinn, D. M., & Emanuel, B. S. (2005). Autism Spectrum Disorders and Symptoms in Children with

Molecularly Confirmed 22q11.2 Deletion Syndrome. *Journal of Autism and Developmental Disorders, 35*(4), 461-470. doi:10.1007/s10803-005-5036-9

First, M. B., Spitzer, Robert L, Gibbon Miriam, and Williams, Janet B.W. (2002). Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). *New York: Biometrics Research, New York State Psychiatric Institute*.

FIRTH, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika, 80*(1), 27-38. doi:10.1093/biomet/80.1.27

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., . . . Carter, N. P. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. *The American Journal of Human Genetics, 84*(4), 524-533.

Fischbach, G. D., & Lord, C. (2010). The Simons Simplex Collection: a resource for identification of autism genetic risk factors. *Neuron, 68*(2), 192-195. doi:10.1016/j.neuron.2010.10.006

Fischer, B. A., & Buchanan, R. W. (2017). Schizophrenia in adults: Epidemiology and pathogenesis. *UpToDate, Rose, BD (Ed), UpToDate, Waltham, MA*.

Foley, C., Corvin, A., & Nakagome, S. (2017). Genetics of schizophrenia: ready to translate? *Current psychiatry reports, 19*(9), 61.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric research, 65*(6), 591-598.

Freeman, J. L., Perry, G. H., Feuk, L., Redon, R., McCarroll, S. A., Altshuler, D. M., . . . Lee, C. (2006). Copy number variation: new insights in genome diversity. *Genome Res, 16*(8), 949-961. doi:10.1101/gr.3677206

Fusar-Poli, P., Tantardini, M., De Simone, S., Ramella-Cravaro, V., Oliver, D., Kingdon, J., . . . McGuire, P. (2017). Deconstructing vulnerability for psychosis: Meta-analysis of environmental risk factors for psychosis in subjects at ultra high-risk. *European Psychiatry, 40*, 65-75. doi:https://doi.org/10.1016/j.eurpsy.2016.09.003

Gage, J. L., de Leon, N., & Clayton, M. K. (2018). Comparing Genome-Wide Association Study Results from Different Measurements of an Underlying Phenotype. *G3 (Bethesda, Md.), 8*(11), 3715-3722. doi:10.1534/g3.118.200700

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), 881-885. doi:10.1038/ng.3039

Genovese, G., Fromer, M., Stahl, E. A., Ruderfer, D. M., Chambert, K., Landén, M., . . . McCarroll, S. A. (2016). Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. *Nat Neurosci, 19*(11), 1433-1441. doi:10.1038/nn.4402

Geschwind, D. H. (2011). Genetics of autism spectrum disorders. *Trends in cognitive sciences, 15*(9), 409-416.

Giegling, I., Hosak, L., Mössner, R., Serretti, A., Bellivier, F., Claes, S., . . . Gallo, C. (2017). Genetics of schizophrenia: A consensus paper of the WFSBP Task Force on Genetics. *The world journal of biological psychiatry, 18*(7), 492-505.

Gillberg, C., Persson, E., Grufman, M., & Themnér, U. (1986). Psychiatric disorders in mildly and severely mentally retarded urban children and adolescents: epidemiological aspects. *The British Journal of Psychiatry, 149*(1), 68-74.

Gilman, S. R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M., & Vitkup, D. (2011). Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron, 70*(5), 898-907.

Gjevik, E., Eldevik, S., Fjæran-Granum, T., & Sponheim, E. (2011). Kiddie-SADS Reveals High Rates of DSM-IV Disorders in Children and Adolescents with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 41*(6), 761-769. doi:10.1007/s10803-010-1095-7

Gjevik, E., Sandstad, B., Andreassen, O. A., Myhre, A. M., & Sponheim, E. (2015). Exploring the agreement between questionnaire information and DSM-IV diagnoses of comorbid psychopathology in children with autism spectrum disorders. *Autism, 19*(4), 433-442. doi:10.1177/1362361314526003

Glessner, J. T., Reilly, M. P., Kim, C. E., Takahashi, N., Albano, A., Hou, C., . . . Hakonarson, H. (2010). Strong synaptic transmission impact by copy number variations in schizophrenia. *Proceedings of the National Academy of Sciences, 107*(23), 10584-10589. doi:10.1073/pnas.1000274107

Glessner, J. T., Wang, K., Cai, G., Korvatska, O., Kim, C. E., Wood, S., . . . Bradfield, J. P. (2009). Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature, 459*(7246), 569-573.

Goldberg, R., Motzkin, B., Marion, R., Scambler, P. J., & Shprintzen, R. J. (1993). Velo-cardio-facial syndrome: A review of 120 patients. *American journal of medical genetics, 45*(3), 313-319.

Gonzalez, E., Kulkarni, H., Bolivar, H., Mangano, A., Sanchez, R., Catano, G., . . . Bamshad, M. J. (2005). The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science, 307*(5714), 1434-1440.

Govaerts, L., Srebniak, M., Diderich, K., Joosten, M., Riedijk, S., Knapen, M., . . . Toolenaar, T. (2017). Prenatal diagnosis of susceptibility loci for neurodevelopmental disorders–genetic counseling and pregnancy outcome in 57 cases. *Prenatal diagnosis, 37*(1), 73-80.

Gratten, J., Wray, N. R., Keller, M. C., & Visscher, P. M. (2014). Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat Neurosci, 17*(6), 782-790. doi:10.1038/nn.3708

Grayton, H. M., Fernandes, C., Rujescu, D., & Collier, D. A. (2012). Copy number variations in neurodevelopmental disorders. *Progress in Neurobiology, 99*(1), 81-91. doi:https://doi.org/10.1016/j.pneurobio.2012.07.005

Green, R. C., Berg, J. S., Grody, W. W., Kalia, S. S., Korf, B. R., Martin, C. L., . . . Ormond, K. E. (2013). ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genetics in Medicine, 15*(7), 565-574.

Grice, S. J., Liu, J.-L., & Webber, C. (2015). Synergistic interactions between Drosophila orthologues of genes spanned by de novo human CNVs support multiple-hit models of autism. *PLoS Genet, 11*(3), e1004998.

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., . . . Anney, R. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics, 51*(3), 431-444.

Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., . . . Gur, R. E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8-21. *Neuropsychology, 26*(2), 251-265. doi:10.1037/a0026712

Guyatt, A. L., Stergiakouli, E., Martin, J., Walters, J., O'Donovan, M., Owen, M., . . . Gaunt, T. R. (2018). Association of copy number variation across the genome with neuropsychiatric traits in the general population. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 177*(5), 489-502. doi:10.1002/ajmg.b.32637

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., . . . Risch, N. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry, 68*(11), 1095-1102. doi:10.1001/archgenpsychiatry.2011.76

Hamshere, M. L., Walters, J. T., Smith, R., Richards, A. L., Green, E., Grozeva, D., . . . O'Donovan, M. C. (2013). Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Mol Psychiatry, 18*(6), 708-712. doi:10.1038/mp.2012.67

Han, J., Walters, J. T. R., Kirov, G., Pocklington, A., Escott-Price, V., Owen, M. J., . . . Rees, E. (2016). Gender differences in CNV burden do not confound schizophrenia CNV associations. *Scientific Reports, 6*(1), 25986. doi:10.1038/srep25986

Hanson, E., Bernier, R., Porche, K., Jackson, F. I., Goin-Kochel, R. P., Snyder, L. G., . . . Simons Variation in Individuals Project, C. (2015). The cognitive and behavioral phenotype of the 16p11.2 deletion in a clinically ascertained population. *Biological psychiatry, 77*(9), 785-793. doi:10.1016/j.biopsych.2014.04.021

Harris, J. C. (2006). *Intellectual disability: Understanding its development, causes, classification, evaluation, and treatment*: Oxford University Press.

Hedley, D., & Uljarević, M. (2018). Systematic Review of Suicide in Autism Spectrum Disorder: Current Trends and Implications. *Current Developmental Disorders Reports, 5*(1), 65-76. doi:10.1007/s40474-018-0133-6

Hehir-Kwa, J. Y., Rodríguez-Santiago, B., Vissers, L. E., de Leeuw, N., Pfundt, R., Buitelaar, J. K., . . . Veltman, J. A. (2011). De novo copy number variants associated with intellectual disability have a paternal origin and age bias. *Journal of medical genetics, 48*(11), 776-778.

Heil, K. M., & Schaaf, C. P. (2012). The Genetics of Autism Spectrum Disorders – A Guide for Clinicians. *Current psychiatry reports, 15*(1), 334. doi:10.1007/s11920-012-0334-3

Heinze, G., Ploner, M., Dunkler, D., Southworth, H., & Heinze, M. G. (2016). Package 'logistf'. In.

Helzer, J. E., Kraemer, H. C., & Krueger, R. F. (2006). The feasibility and need for dimensional psychiatric diagnoses. *Psychological medicine, 36*(12), 1671-1680.

Hemmings, C. P. (2006). Schizophrenia spectrum disorders in people with intellectual disabilities. *Curr Opin Psychiatry, 19*(5), 470-474. doi:10.1097/01.yco.0000238472.22325.17

Hengartner, M. P., & Lehmann, S. N. (2017). Why Psychiatric Research Must Abandon Traditional Diagnostic Classification and Adopt a Fully Dimensional Scope: Two Solutions to a Persistent Problem. *Frontiers in Psychiatry, 8*(101). doi:10.3389/fpsyt.2017.00101

Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge, T. M., . . . Glenthøj, B. (2018). Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biological psychiatry, 83*(6), 492-498. doi:https://doi.org/10.1016/j.biopsych.2017.08.017

Hirvikoski, T., Mittendorfer-Rutz, E., Boman, M., Larsson, H., Lichtenstein, P., & Bölte, S. (2016). Premature mortality in autism spectrum disorder. *The British Journal of Psychiatry, 208*(3), 232-238.

Hoop, J. G., Roberts, L. W., Hammond, K. A. G., & Cox, N. J. (2008). Psychiatrists' attitudes, knowledge, and experience regarding genetics: a preliminary study. *Genetics in Medicine, 10*(6), 439-449. doi:10.1097/GIM.0b013e318177014b

Horowitz, L. M., Thurm, A., Farmer, C., Mazefsky, C., Lanzillo, E., Bridge, J. A., . . . Developmental Disorders Inpatient Research, C. (2018). Talking About Death or Suicide: Prevalence and Clinical Correlates in Youth with Autism Spectrum Disorder in the Psychiatric Inpatient Setting. *Journal of Autism and Developmental Disorders, 48*(11), 3702-3710. doi:10.1007/s10803-017-3180-7

Hultman, C. M., Sandin, S., Levine, S. Z., Lichtenstein, P., & Reichenberg, A. (2011). Advancing paternal age and risk of autism: new evidence from a population-based study and a meta-analysis of epidemiological studies. *Molecular Psychiatry, 16*(12), 1203-1212. doi:10.1038/mp.2010.121

Hunsche, M. C., Saqui, S., Mirenda, P., Zaidman-Zait, A., Bennett, T., Duku, E., . . . Kerns, C. M. (2020). Parent-Reported Rates and Clinical Correlates of Suicidality in Children with Autism Spectrum Disorder: A Longitudinal Study. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-020-04373-y

Inglis, A., Koehn, D., McGillivray, B., Stewart, S. E., & Austin, J. (2015). Evaluating a unique, specialist psychiatric genetic counseling clinic: uptake and impact. *Clin Genet, 87*(3), 218-224. doi:10.1111/cge.12415

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. In: Am Psychiatric Assoc.

International Schizophrenia Consortium. (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature, 455*(7210), 237.

Irish Schizophrenia Genomics Consortium. (2012). Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biol Psychiatry, 72*(8), 620-628. doi:10.1016/j.biopsych.2012.05.035

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

Jensen, V. K., & Spannagel, S. C. (2011). The Spectrum of Autism Spectrum Disorder: A Spectrum of Needs, Services, and Challenges. *Journal of Contemporary Psychotherapy, 41*(1), 1-9. doi:10.1007/s10879-010-9161-1

Jez, S., Martin, M., South, S., Vanzo, R., & Rothwell, E. (2015). Variants of unknown significance on chromosomal microarray analysis: parental perspectives. *Journal of community genetics, 6*(4), 343-349. doi:10.1007/s12687-015-0218-4

Jokiranta-Olkoniemi, E., Cheslack-Postava, K., Sucksdorff, D., Suominen, A., Gyllenberg, D., Chudal, R., . . . Sourander, A. (2016). Risk of Psychiatric and Neurodevelopmental Disorders Among Siblings of Probands With Autism Spectrum Disorders. *JAMA psychiatry, 73*(6), 622-629. doi:10.1001/jamapsychiatry.2016.0495

Joshi, G., Wozniak, J., Petty, C., Martelon, M. K., Fried, R., Bolfek, A., . . . Bourgeois, M. (2013). Psychiatric comorbidity and functioning in a clinically referred population of adults with autism spectrum disorders: a comparative study. *Journal of Autism and Developmental Disorders, 43*(6), 1314-1325.

Kaminsky, E. B., Kaul, V., Paschall, J., Church, D. M., Bunke, B., Kunig, D., . . . Warren, S. T. (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.

Kang, H. (2013). The prevention and handling of the missing data. *Korean journal of anesthesiology, 64*(5), 402-406. doi:10.4097/kjae.2013.64.5.402

Karayiorgou, M., Morris, M. A., Morrow, B., Shprintzen, R. J., Goldberg, R., Borrow, J., . . . Lasseter, V. K. (1995). Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. *Proceedings of the National Academy of Sciences, 92*(17), 7612-7616.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., . . . Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry, 36*(7), 980-988. doi:10.1097/00004583-199707000-00021

Kelleher, I., & Cannon, M. (2011). Psychotic-like experiences in the general population: characterizing a high-risk group for psychosis. *Psychological medicine, 41*(1), 1-6.

Kelleher, I., Keeley, H., Corcoran, P., Lynch, F., Fitzpatrick, C., Devlin, N., . . . Harley, M. (2012). Clinicopathological significance of psychotic experiences in non-psychotic young people: evidence from four population-based studies. *The British Journal of Psychiatry, 201*(1), 26-32.

Kelly, J. R., Clarke, G., Cryan, J. F., & Dinan, T. G. (2018). Dimensional thinking in psychiatry in the era of the Research Domain Criteria (RDoC). *Irish Journal of Psychological Medicine, 35*(2), 89-94. doi:10.1017/ipm.2017.7

Kendall, K., Hall, J., & Owen, M. (2020). Neurodevelopmental Disorders and Schizophrenia: Recent Developments. *Oxford Textbook of the Psychiatry of Intellectual Disability*, 213.

Kendall, K. M., Rees, E., Bracher-Smith, M., Legge, S., Riglin, L., Zammit, S., . . . Walters, J. T. R. (2019). Association of Rare Copy Number Variants With Risk of Depression. *JAMA psychiatry, 76*(8), 818-825. doi:10.1001/jamapsychiatry.2019.0566

Kendall, K. M., Rees, E., Escott-Price, V., Einon, M., Thomas, R., Hewitt, J., . . . Kirov, G. (2016). Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. *Biol Psychiatry*. doi:10.1016/j.biopsych.2016.08.014

Kessler, R. C., Ormel, J., Petukhova, M., McLaughlin, K. A., Green, J. G., Russo, L. J., . . . Üstün, T. B. (2011). Development of Lifetime Comorbidity in the World Health Organization World Mental Health Surveys. *Archives of General psychiatry, 68*(1), 90-100. doi:10.1001/archgenpsychiatry.2010.180

Khan, T. A., Revah, O., Gordon, A., Yoon, S.-J., Krawisz, A. K., Goold, C., . . . Paşca, S. P. (2020). Neuronal defects in a human cellular model of 22q11.2 deletion syndrome. *Nature Medicine*. doi:10.1038/s41591-020-1043-9

Kincaid, D. L., Doris, M., Shannon, C., & Mulholland, C. (2017). What is the prevalence of autism spectrum disorder and ASD traits in psychosis? A systematic review. *Psychiatry research, 250*, 99-105.

Kirby, A. V., Bakian, A. V., Zhang, Y., Bilder, D. A., Keeshin, B. R., & Coon, H. (2019). A 20-year study of suicide death in a statewide autism population. *Autism Research, 12*(4), 658-666. doi:10.1002/aur.2076

Kirov, G. (2015). CNVs in neuropsychiatric disorders. *Hum Mol Genet, 24*(R1), R45-49. doi:10.1093/hmg/ddv253

Kirov, G., Grozeva, D., Norton, N., Ivanov, D., Mantripragada, K. K., Holmans, P., . . . Owen, M. J. (2009a). Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum Mol Genet, 18*(8), 1497-1503.

Kirov, G., Pocklington, A. J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., . . . Owen, M. J. (2012). De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry, 17*(2), 142-153. doi:10.1038/mp.2011.154

Kirov, G., Rees, E., Walters, J. T., Escott-Price, V., Georgieva, L., Richards, A. L., . . . Owen, M. J. (2014). The penetrance of copy number variations for schizophrenia and developmental delay. *Biol Psychiatry, 75*(5), 378-385. doi:10.1016/j.biopsych.2013.07.022

Kirov, G., Rujescu, D., Ingason, A., Collier, D. A., O'Donovan, M. C., & Owen, M. J. (2009b). Neurexin 1 (NRXN1) deletions in schizophrenia. *Schizophrenia Bulletin, 35*(5), 851-854. doi:10.1093/schbul/sbp079

Knapp, M., Mangalore, R., & Simon, J. (2004). The global costs of schizophrenia. *Schizophrenia Bulletin, 30*(2), 279-293.

Kobayashi, H., Nemoto, T., Koshikawa, H., Osono, Y., Yamazawa, R., Murakami, M., . . . Mizuno, M. (2008). A self-reported instrument for prodromal symptoms of psychosis: testing the clinical validity of the PRIME Screen-Revised (PS-R) in a Japanese population. *Schizophr Res, 106*(2-3), 356-362. doi:10.1016/j.schres.2008.08.018

Kotov, R., Chang, S.-W., Fochtmann, L. J., Mojtabai, R., Carlson, G. A., Sedler, M. J., & Bromet, E. J. (2010). Schizophrenia in the Internalizing-Externalizing Framework: A Third Dimension? *Schizophrenia Bulletin, 37*(6), 1168-1178. doi:10.1093/schbul/sbq024

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., . . . Clark, L. A. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *Journal of abnormal psychology, 126*(4), 454.

Kotov, R., Ruggero, C. J., Krueger, R. F., Watson, D., Yuan, Q., & Zimmerman, M. (2011). New Dimensions in the Quantitative Classification of Mental Illness. *Archives of General psychiatry, 68*(10), 1003-1011. doi:10.1001/archgenpsychiatry.2011.107

Krumm, N., Turner, T. N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., . . . Eichler, E. E. (2015). Excess of rare, inherited truncating mutations in autism. *Nature Genetics, 47*(6), 582-588. doi:10.1038/ng.3303

Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., . . . Hoover, J. (2016). ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Research, 44*(D1), D862-D868.

Lardinois, M., Lataster, T., Mengelers, R., Van Os, J., & Myin-Germeys, I. (2011). Childhood trauma and increased stress sensitivity in psychosis. *Acta Psychiatrica Scandinavica, 123*(1), 28-35.

Larsson, H. J., Eaton, W. W., Madsen, K. M., Vestergaard, M., Olesen, A. V., Agerbo, E., . . . Mortensen, P. B. (2005). Risk Factors for Autism: Perinatal Factors, Parental Psychiatric

History, and Socioeconomic Status. *American Journal of Epidemiology, 161*(10), 916-925. doi:10.1093/aje/kwi123

Lawlor, A., Kerin, L., Orr, D., Leahy, R., Crotty, F., Kelleher, S., . . . Nicholas, F. (2017). Developing integrated care in the context of rare chromosomal conditions: 22q11 Deletion Syndrome; A parent/clinician collaboration. *International Journal of Integrated Care, 17*, 215. doi:10.5334/ijic.3525

Leblond, C. S., Cliquet, F., Carton, C., Huguet, G., Mathieu, A., Kergrohen, T., . . . Bourgeron, T. (2019). Both rare and common genetic variants contribute to autism in the Faroe Islands. *npj Genomic Medicine, 4*(1), 1. doi:10.1038/s41525-018-0075-2

Leblond, C. S., Heinrich, J., Delorme, R., Proepper, C., Betancur, C., Huguet, G., . . . Rastam, M. (2012). Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genet, 8*(2), e1002521.

Lee, P. H., Anttila, V., Won, H., Feng, Y.-C. A., Rosenthal, J., Zhu, Z., . . . Posthuma, D. (2019). Genome wide meta-analysis identifies genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *BioRxiv*(528117).

Leppa, V. M., Kravitz, S. N., Martin, C. L., Andrieux, J., Le Caignec, C., Martin-Coignard, D., . . . Geschwind, D. 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. doi:10.1016/j.ajhg.2016.06.036

Levinson, D. F., Mostafavi, S., Milaneschi, Y., Rivera, M., Ripke, S., Wray, N. R., & Sullivan, P. F. (2014). Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it? *Biol Psychiatry, 76*(7), 510-512. doi:10.1016/j.biopsych.2014.07.029

Levy, D., Ronemus, M., Yamrom, B., Lee, Y.-h., Leotta, A., Kendall, J., . . . Ye, K. (2011). Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron, 70*(5), 886-897.

Levy, S. E., Giarelli, E., Lee, L.-C., Schieve, L. A., Kirby, R. S., Cunniff, C., . . . Rice, C. E. (2010). Autism spectrum disorder and co-occurring developmental, psychiatric, and medical conditions among children in multiple populations of the United States. *Journal of Developmental & Behavioral Pediatrics, 31*(4), 267-275.

Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., . . . Lainhart, J. E. (2006). Comorbid Psychiatric Disorders in Children with Autism: Interview Development and Rates of Disorders. *Journal of Autism and Developmental Disorders, 36*(7), 849-861. doi:10.1007/s10803-006-0123-0

Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing Diagnostic Approaches: DSM-ICD Versus RDoC. *Annual Review of Clinical Psychology, 12*, 435-463. doi:10.1146/annurev-clinpsy-021815-093122

Linehan, C. (2019). *Epidemiology Disability of Intellectual.* Paper presented at the Seminars in the Psychiatry of Intellectual Disability.

Loh, P.-R., Bhatia, G., Gusev, A., Finucane, H. K., Bulik-Sullivan, B. K., Pollack, S. J., . . . Price, A. L. (2015). Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nature Genetics, 47*(12), 1385-1392. doi:10.1038/ng.3431

López-Ratón, M., Rodríguez-Álvarez, M. X., Cadarso-Suárez, C., & Gude-Sampedro, F. (2014). OptimalCutpoints: an R package for selecting optimal cutpoints in diagnostic tests. *J Stat Softw, 61*(8), 1-36.

Lord, C., DiLavore, P. C., & Gotham, K. (2012). *Autism diagnostic observation schedule*: Western Psychological Services Torrance, CA.

Lowther, C., Merico, D., Costain, G., Waserman, J., Boyd, K., Noor, A., . . . Bassett, A. S. (2017). Impact of IQ on the diagnostic yield of chromosomal microarray in a community sample of adults with schizophrenia. *Genome Med, 9*(1), 105. doi:10.1186/s13073-017-0488-z

Lugo-Marín, J., Magán-Maganto, M., Rivero-Santana, A., Cuellar-Pompa, L., Alviani, M., Jenaro-Rio, C., . . . Canal-Bedia, R. (2019). Prevalence of psychiatric disorders in adults with autism spectrum disorder: A systematic review and meta-analysis. *Research in Autism Spectrum Disorders, 59*, 22-33. doi:https://doi.org/10.1016/j.rasd.2018.12.004

Luo, R., Sanders, Stephan J., Tian, Y., Voineagu, I., Huang, N., Chu, Su H., . . . Geschwind, Daniel H. (2012). Genome-wide Transcriptome Profiling Reveals the Functional Impact of Rare De Novo and Recurrent CNVs in Autism Spectrum Disorders. *The American Journal of Human Genetics, 91*(1), 38-55. doi:https://doi.org/10.1016/j.ajhg.2012.05.011

Lupski, J. R. (1998). Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet, 14*(10), 417-422. doi:10.1016/s0168-9525(98)01555-8

Lupski, J. R., Wise, C. A., Kuwano, A., Pentao, L., Parke, J. T., Glaze, D. G., . . . Patel, P. I. (1992). Gene dosage is a mechanism for Charcot-Marie-Tooth disease type 1A. *Nat Genet, 1*(1), 29-33. doi:10.1038/ng0492-29

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. doi:10.1146/annurev-publhealth-031816-044318

MacCabe, J. H. (2008). Population-based Cohort Studies on Premorbid Cognitive Function in Schizophrenia. *Epidemiologic Reviews, 30*(1), 77-83. doi:10.1093/epirev/mxn007

MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L., & Scherer, S. W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res, 42*(Database issue), D986-992. doi:10.1093/nar/gkt958

Magnuson, K. M., & Constantino, J. N. (2011). Characterization of depression in children with autism spectrum disorders. *Journal of developmental and behavioral pediatrics : JDBP, 32*(4), 332-340. doi:10.1097/DBP.0b013e318213f56c

Malaspina, D., Harlap, S., Fennig, S., Heiman, D., Nahon, D., Feldman, D., & Susser, E. S. (2001). Advancing paternal age and the risk of schizophrenia. *Arch Gen Psychiatry, 58*(4), 361-367. doi:10.1001/archpsyc.58.4.361

Malhotra, D., & Sebat, J. (2012). CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell, 148*(6), 1223-1241. doi:10.1016/j.cell.2012.02.039

Männik, K., Mägi, R., Macé, A., Cole, B., Guyatt, A. L., Shihab, H. A., . . . Reymond, A. (2015). Copy number variations and cognitive phenotypes in unselected populations. *JAMA, 313*(20), 2044-2054. doi:10.1001/jama.2015.4845

Maric, N. P., & Svrakic, D. M. (2012). Why schizophrenia genetics needs epigenetics: a review. *Psychiatr Danub, 24*(1), 2-18.

Marshall, C. R., Howrigan, D. P., Merico, D., Thiruvahindrapuram, B., Wu, W., Greer, D. S., . . . Sebat, J. (2017). Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet, 49*(1), 27-35. doi:10.1038/ng.3725

Marshall, C. R., Noor, A., Vincent, J. B., Lionel, A. C., Feuk, L., Skaug, J., . . . Scherer, S. W. (2008). Structural Variation of Chromosomes in Autism Spectrum Disorder. *The American Journal of Human Genetics, 82*(2), 477-488. doi:https://doi.org/10.1016/j.ajhg.2007.12.009

Martin, J., Tammimies, K., Karlsson, R., Lu, Y., Larsson, H., Lichtenstein, P., & Magnusson, P. K. E. (2019). Copy number variation and neuropsychiatric problems in females and males in the general population. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 180*(6), 341-350. doi:10.1002/ajmg.b.32685

Mattila, M.-L., Hurtig, T., Haapsamo, H., Jussila, K., Kuusikko-Gauffin, S., Kielinen, M., . . . Moilanen, I. (2010). Comorbid Psychiatric Disorders Associated with Asperger Syndrome/High-functioning Autism: A Community- and Clinic-based Study. *Journal of Autism and Developmental Disorders, 40*(9), 1080-1093. doi:10.1007/s10803-010-0958-2

Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., & Saxena, S. (2011). Prevalence of intellectual disability: a meta-analysis of population-based studies. *Research in developmental disabilities, 32*(2), 419-436.

Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., & Saxena, S. (2013). "Prevalence of intellectual disability: A meta-analysis of population-based studies": Corrigendum. *Research in developmental disabilities, 34*(2), 729-729. doi:10.1016/j.ridd.2012.10.013

Maxim, L. D., Niebo, R., & Utell, M. J. (2014). Screening tests: a review with examples. *Inhalation toxicology, 26*(13), 811-828. doi:10.3109/08958378.2014.955932

Mayes, S. D., Gorman, A. A., Hillwig-Garcia, J., & Syed, E. (2013). Suicide ideation and attempts in children with autism. *Research in Autism Spectrum Disorders, 7*(1), 109-119. doi:https://doi.org/10.1016/j.rasd.2012.07.009

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. *Journal of Developmental & Behavioral Pediatrics, 36*(2), 61-67.

Mazzone, L., Ruta, L., & Reale, L. (2012). Psychiatric comorbidities in asperger syndrome and high functioning autism: diagnostic challenges. *Annals of General Psychiatry, 11*(1), 16. doi:10.1186/1744-859X-11-16

McCandless, S. E. (2010). Clinical report—health supervision for children with Prader-Willi syndrome. *Pediatrics*.

McCarroll, S. A., Hadnott, T. N., Perry, G. H., Sabeti, P. C., Zody, M. C., Barrett, J. C., . . . Altshuler, D. M. (2006). Common deletion polymorphisms in the human genome. *Nat Genet, 38*(1), 86-92. doi:10.1038/ng1696

McCarroll, S. A., Kuruvilla, F. G., Korn, J. M., Cawley, S., Nemesh, J., Wysoker, A., . . . Kirby, A. (2008). Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics, 40*(10), 1166.

McCarthy, S. E., Gillis, J., Kramer, M., Lihm, J., Yoon, S., Berstein, Y., . . . Corvin, A. (2014). De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Molecular Psychiatry, 19*(6), 652-658. doi:10.1038/mp.2014.29

McGlashan, T. H., Miller, T.J., Woods, S.W., Rosen, J.L., Hoffman, R.E., Davidson, L. (2003). Structured interview for prodromal syndromes, Version 4.0. . *New Haven, CT: Prime Clinic Yale School of Medicine*.

McGrath, J., Saha, S., Chant, D., & Welham, J. (2008). Schizophrenia: A Concise Overview of Incidence, Prevalence, and Mortality. *Epidemiologic Reviews, 30*(1), 67-76. doi:10.1093/epirev/mxn001

McGue, M., Bouchard Jr, T., & Iacono, W. (1993). Behavioral genetics of cognitive ability: A life-span perspective. *American Psychological Association*, 59-76.

Mefford, H. C., & Mulley, J. C. (2010). Genetically complex epilepsies, copy number variants and syndrome constellations. *Genome Med, 2*(10), 71.

Merikangas, A., Segurado, R., Heron, E., Anney, R., Paterson, A., Cook, E., . . . Gill, M. (2015a). The phenotypic manifestations of rare genic CNVs in autism spectrum disorder. *Molecular Psychiatry, 20*(11), 1366-1372.

Merikangas, A. K., Corvin, A. P., & Gallagher, L. (2009a). Copy-number variants in neurodevelopmental disorders: promises and challenges. *Trends in genetics, 25*(12), 536-544.

Merikangas, K. R., Avenevoli, S., Costello, E. J., Koretz, D., & Kessler, R. C. (2009b). National comorbidity survey replication adolescent supplement (NCS-A): I. Background and measures. *Journal of the American Academy of Child & Adolescent Psychiatry, 48*(4), 367-379.

Merikangas, K. R., Calkins, M. E., Burstein, M., He, J. P., Chiavacci, R., Lateef, T., . . . Gur, R. E. (2015b). Comorbidity of physical and mental disorders in the neurodevelopmental genomics cohort study. *Pediatrics, 135*(4), e927-938. doi:10.1542/peds.2014-1444

Miles, J. H., Takahashi, T. N., Hong, J., Munden, N., Flournoy, N., Braddock, S. R., . . . Farmer, J. E. (2008a). Development and validation of a measure of dysmorphology: useful for autism subgroup classification. *Am J Med Genet A, 146a*(9), 1101-1116. doi:10.1002/ajmg.a.32244

Miles, J. H., Takahashi, T. N., Hong, J., Munden, N., Flournoy, N., Braddock, S. R., . . . Farmer, J. E. (2008b). Development and validation of a measure of dysmorphology: Useful for autism subgroup classification. *American Journal of Medical Genetics Part A, 146A*(9), 1101-1116. doi:10.1002/ajmg.a.32244

Miller, B. J., Culpepper, N., Rapaport, M. H., & Buckley, P. (2013). Prenatal inflammation and neurodevelopment in schizophrenia: a review of human studies. *Prog Neuropsychopharmacol Biol Psychiatry, 42*, 92-100. doi:10.1016/j.pnpbp.2012.03.010

Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P., . . . Epstein, C. J. (2010a). Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *The American Journal of Human Genetics, 86*(5), 749-764.

Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P., . . . Ledbetter, D. H. (2010b). Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. *The American Journal of Human Genetics, 86*(5), 749-764. doi:https://doi.org/10.1016/j.ajhg.2010.04.006

Miller, T. (2004). The SIPS-Screen: a brief self-report screen to detect the schizophrenia prodrome. *Schizophr Res*, 78.

Miller, T. J., McGlashan, T. H., Rosen, J. L., Cadenhead, K., Cannon, T., Ventura, J., . . . Woods, S. W. (2003). Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. *Schizophr Bull, 29*(4), 703-715. doi:10.1093/oxfordjournals.schbul.a007040

Miny, P., Wenzel, F., Tercanli, S., & Filges, I. (2013). Chromosomal Microarrays in Prenatal Diagnosis: Time for a Change of Policy? *Microarrays (Basel, Switzerland), 2*(4), 304-317. doi:10.3390/microarrays2040304

Modabbernia, A., Velthorst, E., & Reichenberg, A. (2017). Environmental risk factors for autism: an evidence-based review of systematic reviews and meta-analyses. *Molecular Autism, 8*(1), 13. doi:10.1186/s13229-017-0121-4

Moeschler, J. B., & Shevell, M. (2014). Comprehensive evaluation of the child with intellectual disability or global developmental delays. *Pediatrics, 134*(3), e903-e918.

Moran, P., Stokes, J., Marr, J., Bock, G., Desbonnet, L., Waddington, J., & O'Tuathaigh, C. (2016). Gene× environment interactions in schizophrenia: evidence from genetic mouse models. *Neural Plasticity, 2016*.

Morgan, V. A., Croft, M. L., Valuri, G. M., Zubrick, S. R., Bower, C., McNeil, T. F., & Jablensky, A. V. (2012). Intellectual disability and other neuropsychiatric outcomes in high-risk children of mothers with schizophrenia, bipolar disorder and unipolar major depression. *The British Journal of Psychiatry, 200*(4), 282-289.

Morgan, V. A., Leonard, H., Bourke, J., & Jablensky, A. (2008). Intellectual disability co-occurring with schizophrenia and other psychiatric illness: population-based study. *The British Journal of Psychiatry, 193*(5), 364-372.

Movsas, T. Z., Pinto-Martin, J. A., Whitaker, A. H., Feldman, J. F., Lorenz, J. M., Korzeniewski, S. J., . . . Paneth, N. (2013). Autism Spectrum Disorder Is Associated with Ventricular Enlargement in a Low Birth Weight Population. *The Journal of pediatrics, 163*(1), 73-78. doi:https://doi.org/10.1016/j.jpeds.2012.12.084

Muhle, R. A., Reed, H. E., Stratigos, K. A., & Veenstra-VanderWeele, J. (2018). The Emerging Clinical Neuroscience of Autism Spectrum Disorder: A Review. *JAMA psychiatry, 75*(5), 514-523. doi:10.1001/jamapsychiatry.2017.4685

Mukaddes, N. M., & Fateh, R. (2010). High rates of psychiatric co-morbidity in individuals with Asperger's disorder. *World J Biol Psychiatry, 11*(2 Pt 2), 486-492. doi:10.1080/15622970902789130

Murray, C. J., Lopez, A. D., & Organization, W. H. (1996). *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary*: World Health Organization.

Myers, S. M., Challman, T. D., Bernier, R., Bourgeron, T., Chung, W. K., Constantino, J. N., . . . Mitchell, K. J. (2020). Insufficient Evidence for "Autism-Specific" Genes. *The American Journal of Human Genetics*.

Nakamura, B. J., Ebesutani, C., Bernstein, A., & Chorpita, B. F. (2009). A psychometric analysis of the child behavior checklist DSM-oriented scales. *Journal of Psychopathology and Behavioral Assessment, 31*(3), 178-189.

National Electronic Health Record. (2016). *https://www.ehealthireland.ie/Strategic-Programmes/Electronic-Health-Record-EHR-/Progress/National-Business-Case.pdf*.

Network, T., O'Dushlaine, C., Rossin, L., Lee, P. H., Duncan, L., Parikshak, N. N., . . . Purcell, S. M. (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci, 18*(2), 199.

Newman, D. L., Moffitt, T. E., Caspi, A., & Silva, P. A. (1998). Comorbid mental disorders: implications for treatment and sample selection. *Journal of abnormal psychology, 107*(2), 305.

Niarchou, M., Chawner, S. J. R. A., Doherty, J. L., Maillard, A. M., Jacquemont, S., Chung, W. K., . . . Bree, M. B. M. v. d. (2019). Psychiatric disorders in children with 16p11.2 deletion and duplication. *Translational psychiatry, 9*(1), 8. doi:10.1038/s41398-018-0339-8

Niarchou, M., Zammit, S., van Goozen, S. H. M., Thapar, A., Tierling, H. M., Owen, M. J., & van den Bree, M. B. M. (2014). Psychopathology and cognition in children with 22q11.2 deletion syndrome. *Br J Psychiatry, 204*(1), 46-54. doi:10.1192/bjp.bp.113.132324

Niklasson, L., Rasmussen, P., Óskarsdóttir, S., & Gillberg, C. (2001). Neuropsychiatric disorders in the 22q11 deletion syndrome. *Genetics in Medicine, 3*(1), 79-84. doi:10.1097/00125817-200101000-00017

Nishida, A., Tanii, H., Nishimura, Y., Kajiki, N., Inoue, K., Okada, M., . . . Okazaki, Y. (2008). Erratum to" Associations between psychotic-like experiences and mental health status and other psychopathologies among Japanese early teens.".

Nordenbæk, C., Jørgensen, M., Kyvik, K. O., & Bilenberg, N. (2014). A Danish population-based twin study on autism spectrum disorders. *European Child & Adolescent Psychiatry, 23*(1), 35-43. doi:10.1007/s00787-013-0419-5

Nordin, V., & Gillberg, C. (1996). Autism spectrum disorders in children with physical or mental disability or both. I: Clinical and epidemiological aspects. *Dev Med Child Neurol, 38*(4), 297-313. doi:10.1111/j.1469-8749.1996.tb12096.x

Nowakowska, B. (2017). Clinical interpretation of copy number variants in the human genome. *Journal of applied genetics, 58*(4), 449-457. doi:10.1007/s13353-017-0407-4

Obstetricians, A. C. o., & Gynecologists. (2013). Committee opinion no. 581: the use of chromosomal microarray analysis in prenatal diagnosis. *Obstet Gynecol, 122*(6), 1374-1377.

Ochoa, S., Haro, J. M., Autonell, J., Pendàs, A., Teba, F., Màrquez, M., & Group, t. N. (2003). Met and Unmet Needs of Schizophrenia Patients in a Spanish Sample. *Schizophrenia Bulletin, 29*(2), 201-210. doi:10.1093/oxfordjournals.schbul.a006998

Oliphant, R. Y. K., Smith, E. M., & Grahame, V. (2020). What is the Prevalence of Self-harming and Suicidal Behaviour in Under 18s with ASD, With or Without an Intellectual Disability? *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-020-04422-6

Organization, W. H. (1988). *Tenth revision of the International Classification of Diseases chapter V (F: mental, behavioural and developmental disorders, clinical descriptions and diagnostic guidelines*. Retrieved from

Organization, W. H. (2008). *The global burden of disease: 2004 update*: World Health Organization.

Ormel, J., Oldehinkel, A., Ferdinand, R., Hartman, C., De Winter, A., Veenstra, R., . . . Verhulst, F. (2005). Internalizing and externalizing problems in adolescence: general and dimension-specific effects of familial loadings and preadolescent temperament traits. *Psychological medicine, 35*(12), 1825-1835.

Orvaschel, H., & Puig-Antich, J. (1987). *Schedule for affective disorder and schizophrenia for school-age children: Epidemiologic version: Kiddie-SADS-E (K-SADS-E)*.

Owen, M. J., & O'Donovan, M. C. (2017). Schizophrenia and the neurodevelopmental continuum:evidence from genomics. *World Psychiatry, 16*(3), 227-235. doi:10.1002/wps.20440

Pandolfi, V., Magyar, C. I., & Dill, C. A. (2012). An initial psychometric evaluation of the CBCL 6–18 in a sample of youth with autism spectrum disorders. *Research in Autism Spectrum Disorders, 6*(1), 96-108. doi:https://doi.org/10.1016/j.rasd.2011.03.009

Pathak, M., Bennett, A., & Shui, A. M. (2019). Correlates of adaptive behavior profiles in a large cohort of children with autism: The autism speaks Autism Treatment Network registry data. *Autism, 23*(1), 87-99. doi:10.1177/1362361317733113

Pesta, B. J., Kirkegaard, E. O. W., te Nijenhuis, J., Lasker, J., & Fuerst, J. G. R. (2020). Racial and ethnic group differences in the heritability of intelligence: A systematic review and meta-analysis. *Intelligence, 78*, 101408. doi:https://doi.org/10.1016/j.intell.2019.101408

Peter, S., Chaney, G., Zappia, T., Van Veldhuisen, C., Pereira, S., & Santamaria, N. (2011). Care coordination for children with complex care needs significantly reduces hospital utilization. *Journal for Specialists in Pediatric Nursing, 16*(4), 305-312. doi:10.1111/j.1744-6155.2011.00303.x

Peter, T., Roberts, L. W., & Buzdugan, R. (2008). Suicidal Ideation among Canadian Youth: A Multivariate Analysis. *Archives of Suicide Research, 12*(3), 263-275. doi:10.1080/13811110802100882

Pharoah, P. D., Antoniou, A., Bobrow, M., Zimmern, R. L., Easton, D. F., & Ponder, B. A. (2002). Polygenic susceptibility to breast cancer and implications for prevention. *Nature Genetics, 31*(1), 33-36.

Philip, N., & Bassett, A. (2011). Cognitive, behavioural and psychiatric phenotype in 22q11. 2 deletion syndrome. *Behavior genetics, 41*(3), 403-412.

Pincus, H. A., Tew, J. D., & First, M. B. (2004). Psychiatric comorbidity: is more less? *World psychiatry : official journal of the World Psychiatric Association (WPA), 3*(1), 18-23. Retrieved from https://pubmed.ncbi.nlm.nih.gov/16633444

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1414654/

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

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

Poulton, R., Caspi, A., Moffitt, T. E., Cannon, M., Murray, R., & Harrington, H. (2000). Children's self-reported psychotic symptoms and adult schizophreniform disorder: a 15-year longitudinal study. *Archives of General psychiatry, 57*(11), 1053-1058.

Priebe, L., Degenhardt, F., Strohmaier, J., Breuer, R., Herms, S., Witt, S. H., . . . Moebus, S. (2013). Copy number variants in German patients with schizophrenia. *PloS one, 8*(7), e64035.

Qiao, Y., Mercier, E., Dastan, J., Hurlburt, J., McGillivray, B., Chudley, A. E., . . . Pavlidis, P. (2014). Copy number variants (CNVs) analysis in a deeply phenotyped cohort of individuals with intellectual disability (ID). *BMC medical genetics, 15*(1), 82.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2013). Retrieved from http://www.R-project.org/. http://www.R-project.org/.

Rajji, T. K., & Mulsant, B. H. (2008). Nature and course of cognitive function in late-life schizophrenia: a systematic review. *Schizophrenia research, 102*(1-3), 122-140.

Ramaswami, G., & Geschwind, D. H. (2018). Genetics of autism spectrum disorder. In *Handbook of clinical neurology* (Vol. 147, pp. 321-329): Elsevier.

Ranganathan, P., & Aggarwal, R. (2018). Common pitfalls in statistical analysis: Understanding the properties of diagnostic tests - Part 1. *Perspectives in clinical research, 9*(1), 40-43. doi:10.4103/picr.PICR_170_17

Rapoport, J., Chavez, A., Greenstein, D., Addington, A., & Gogtay, N. (2009). Autism Spectrum Disorders and Childhood-Onset Schizophrenia: Clinical and Biological Contributions to a Relation Revisited. *Journal of the American Academy of Child & Adolescent Psychiatry, 48*(1), 10-18. doi:https://doi.org/10.1097/CHI.0b013e31818b1c63

Reardon, T. C., Gray, K. M., & Melvin, G. A. (2015). Anxiety disorders in children and adolescents with intellectual disability: Prevalence and assessment. *Res Dev Disabil, 36c*, 175-190. doi:10.1016/j.ridd.2014.10.007

Reddy, L. A., & Pfeiffer, S. I. (2007). Behavioral and Emotional Symptoms of Children and Adolescents with Prader-Willi Syndrome. *Journal of Autism and Developmental Disorders, 37*(5), 830-839. doi:10.1007/s10803-006-0210-2

Rees, E., Kendall, K., Pardinas, A. F., Legge, S. E., Pocklington, A., Escott-Price, V., . . . Kirov, G. (2016a). Analysis of Intellectual Disability Copy Number Variants for Association With Schizophrenia. *JAMA psychiatry, 73*(9), 963-969. doi:10.1001/jamapsychiatry.2016.1831

Rees, E., Kendall, K., Pardiñas, A. F., Legge, S. E., Pocklington, A., Escott-Price, V., . . . O'Donovan, M. C. (2016b). Analysis of intellectual disability copy number variants for association with schizophrenia. *JAMA psychiatry, 73*(9), 963-969.

Rees, E., Moskvina, V., Owen, M. J., O'Donovan, M. C., & Kirov, G. (2011). De Novo Rates and Selection of Schizophrenia-Associated Copy Number Variants. *Biological psychiatry, 70*(12), 1109-1114. doi:https://doi.org/10.1016/j.biopsych.2011.07.011

Rees, E., Walters, J. T., Chambert, K. D., O'Dushlaine, C., Szatkiewicz, J., Richards, A. L., . . . Kirov, G. (2014a). CNV analysis in a large schizophrenia sample implicates deletions at 16p12.1 and SLC1A1 and duplications at 1p36.33 and CGNL1. *Hum Mol Genet, 23*(6), 1669-1676. doi:10.1093/hmg/ddt540

Rees, E., Walters, J. T. R., Georgieva, L., Isles, A. R., Chambert, K. D., Richards, A. L., . . . Kirov, G. (2014b). Analysis of copy number variations at 15 schizophrenia-associated loci. *The British Journal of Psychiatry, 204*(2), 108-114. doi:10.1192/bjp.bp.113.131052

Reichenberg, A., Cederlöf, M., McMillan, A., Trzaskowski, M., Kapra, O., Fruchter, E., . . . Lichtenstein, P. (2016). Discontinuity in the genetic and environmental causes of the intellectual disability spectrum. *Proceedings of the National Academy of Sciences, 113*(4), 1098-1103. doi:10.1073/pnas.1508093112

Riecher-Rössler, A., & Studerus, E. (2017). Prediction of conversion to psychosis in individuals with an at-risk mental state: a brief update on recent developments. *Curr Opin Psychiatry, 30*(3), 209-219. doi:10.1097/yco.0000000000000320

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K.-H., Holmans, P. A., . . . Huang, H. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature, 511*(7510), 421-427.

Robert L. Spitzer, M.D. , and, & Jerome C. Wakefield, D.S.W. (1999). DSM-IV Diagnostic Criterion for Clinical Significance: Does It Help Solve the False Positives Problem? *American Journal of Psychiatry, 156*(12), 1856-1864. doi:10.1176/ajp.156.12.1856

Roche, M. I., & Berg, J. S. (2015). Incidental Findings with Genomic Testing: Implications for Genetic Counseling Practice. *Curr Genet Med Rep, 3*(4), 166-176. doi:10.1007/s40142-015-0075-9

Royston, P., Altman, D. G., & Sauerbrei, W. (2006). Dichotomizing continuous predictors in multiple regression: a bad idea. *Statistics in medicine, 25*(1), 127-141. doi:10.1002/sim.2331

Royston, R., Howlin, P., Waite, J., & Oliver, C. (2017). Anxiety Disorders in Williams Syndrome Contrasted with Intellectual Disability and the General Population: A Systematic Review and Meta-Analysis. *Journal of Autism and Developmental Disorders, 47*(12), 3765-3777. doi:10.1007/s10803-016-2909-z

Rudd, D. S., Axelsen, M., Epping, E. A., Andreasen, N. C., & Wassink, T. H. (2014). A genome-wide CNV analysis of schizophrenia reveals a potential role for a multiple-hit model. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 165*(8), 619-626.

Rueter, M. A., & Kwon, H.-K. (2005). Developmental Trends in Adolescent Suicidal Ideation. *Journal of Research on Adolescence, 15*(2), 205-222. doi:10.1111/j.1532-7795.2005.00092.x

Rufino, K. A., & Patriquin, M. A. (2019). Child and adolescent suicide: contributing risk factors and new evidence-based interventions. *Children's Health Care, 48*(4), 345-350. doi:10.1080/02739615.2019.1666009

Rujescu, D., Ingason, A., Cichon, S., Pietiläinen, O. P., Barnes, M. R., Toulopoulou, T., . . . Bramon, E. (2009a). Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet, 18*(5), 988-996.

Rujescu, D., Ingason, A., Cichon, S., Pietilainen, O. P., Barnes, M. R., Toulopoulou, T., . . . Collier, D. A. (2009b). Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet, 18*(5), 988-996. doi:10.1093/hmg/ddn351

Rund, B. R. (2018). The research evidence for schizophrenia as a neurodevelopmental disorder. *Scandinavian journal of psychology, 59*(1), 49-58.

Rutkowski, T. P., Schroeder, J. P., Gafford, G. M., Warren, S. T., Weinshenker, D., Caspary, T., & Mulle, J. G. (2017). Unraveling the genetic architecture of copy number variants associated with schizophrenia and other neuropsychiatric disorders. *Journal of neuroscience research, 95*(5), 1144-1160. doi:10.1002/jnr.23970

Rutter, M., LeCouteur, A., & Lord, C. Autism Diagnostic Interview-Revised Manual. 2003. *Los Angeles: Western Psychological Services*.

Rylaarsdam, L., & Guemez-Gamboa, A. (2019). Genetic Causes and Modifiers of Autism Spectrum Disorder. *Frontiers in cellular neuroscience, 13*(385). doi:10.3389/fncel.2019.00385

Sahoo, T., Theisen, A., Rosenfeld, J. A., Lamb, A. N., Ravnan, J. B., Schultz, R. A., . . . Bejjani, B. A. (2011a). Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. *Genetics in Medicine, 13*(10), 868-880.

Sahoo, T., Theisen, A., Rosenfeld, J. A., Lamb, A. N., Ravnan, J. B., Schultz, R. A., . . . Shaffer, L. G. (2011b). Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. *Genet Med, 13*(10), 868-880. doi:10.1097/GIM.0b013e3182217a06

Sanders, S. J., Ercan-Sencicek, A. G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., . . . Thomson, S. A. (2011). Multiple recurrent de novo CNVs, including duplications of the 7q11. 23 Williams 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., . . . Dong, S. (2015). Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron, 87*(6), 1215-1233.

Sandin, S., Hultman, C. M., Kolevzon, A., Gross, R., MacCabe, J. H., & Reichenberg, A. (2012). Advancing Maternal Age Is Associated With Increasing Risk for Autism: A Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(5), 477-486.e471. doi:https://doi.org/10.1016/j.jaac.2012.02.018

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Hultman, C., Larsson, H., & Reichenberg, A. (2017). The Heritability of Autism Spectrum Disorder. *JAMA, 318*(12), 1182-1184. doi:10.1001/jama.2017.12141

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Larsson, H., Hultman, C. M., & Reichenberg, A. (2014). The Familial Risk of Autism. *JAMA, 311*(17), 1770-1777. doi:10.1001/jama.2014.4144

Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J.-Y., . . . Klei, L. (2020). Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell, 180*(3), 568-584. e523.

Schaefer, G. B., & Mendelsohn, N. J. (2013). Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. *Genet Med, 15*(5), 399-407.

Schanding, G. T., Nowell, K. P., & Goin-Kochel, R. P. (2012). Utility of the Social Communication Questionnaire-Current and Social Responsiveness Scale as Teacher-Report Screening Tools for Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 42*(8), 1705-1716. doi:10.1007/s10803-011-1412-9

Scherer, S. W., Lee, C., Birney, E., Altshuler, D. M., Eichler, E. E., Carter, N. P., . . . Feuk, L. (2007). Challenges and standards in integrating surveys of structural variation. *Nature Genetics, 39*(7), S7-S15. doi:10.1038/ng2093

Schofield, C. M., Hsu, R., Barker, A. J., Gertz, C. C., Blelloch, R., & Ullian, E. M. (2011). Monoallelic deletion of the microRNA biogenesis gene Dgcr8 produces deficits in the development of excitatory synaptic transmission in the prefrontal cortex. *Neural Dev, 6*, 11. doi:10.1186/1749-8104-6-11

Scott, J., Martin, G., Welham, J., Bor, W., Najman, J., O'Callaghan, M., . . . McGrath, J. (2009). Psychopathology during childhood and adolescence predicts delusional-like experiences in adults: a 21-year birth cohort study. *American Journal of Psychiatry, 166*(5), 567-574.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., . . . Kendall, J. (2007). Strong association of de novo copy number mutations with autism. *Science, 316*(5823), 445-449.

Segers, M., & Rawana, J. (2014). What Do We Know About Suicidality in Autism Spectrum Disorders? A Systematic Review. *Autism Research, 7*(4), 507-521. doi:10.1002/aur.1375

Shafee, R., Nanda, P., Padmanabhan, J. L., Tandon, N., Alliey-Rodriguez, N., Kalapurakkel, S., . . . Hill, S. K. (2018). Polygenic risk for schizophrenia and measured domains of cognition in individuals with psychosis and controls. *Translational psychiatry, 8*(1), 1-9.

Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A children's global assessment scale (CGAS). *Archives of General psychiatry, 40*(11), 1228-1231.

Shaffer, L. G., Dabell, M. P., Fisher, A. J., Coppinger, J., Bandholz, A. M., Ellison, J. W., . . . Rosenfeld, J. A. (2012). Experience with microarray-based comparative genomic hybridization for prenatal diagnosis in over 5000 pregnancies. *Prenatal diagnosis, 32*(10), 976-985. doi:10.1002/pd.3945

Shao, X., Lv, N., Liao, J., Long, J., Xue, R., Ai, N., . . . Fan, X. (2019). Copy number variation is highly correlated with differential gene expression: a pan-cancer study. *BMC medical genetics, 20*(1), 175. doi:10.1186/s12881-019-0909-5

Sharp, A. J., Locke, D. P., McGrath, S. D., Cheng, Z., Bailey, J. A., Vallente, R. U., . . . Eichler, E. E. (2005). Segmental duplications and copy-number variation in the human genome. *American journal of human genetics, 77*(1), 78-88. doi:10.1086/431652

Sheehan, R., Hassiotis, A., Walters, K., Osborn, D., Strydom, A., & Horsfall, L. (2015). Mental illness, challenging behaviour, and psychotropic drug prescribing in people with intellectual disability: UK population based cohort study. *BMJ : British Medical Journal, 351*, h4326. doi:10.1136/bmj.h4326

Shkedi-Rafid, S., Fenwick, A., Dheensa, S., Wellesley, D., & Lucassen, A. M. (2016). What results to disclose, when, and who decides? Healthcare professionals' views on prenatal chromosomal microarray analysis. *Prenatal diagnosis, 36*(3), 252-259. doi:10.1002/pd.4772

Shlien, A., & Malkin, D. (2009). Copy number variations and cancer. *Genome Med, 1*(6), 62.

Sigman, M., & Ungerer, J. (1981). Sensorimotor skills and language comprehension in autistic children. *Journal of abnormal child psychology, 9*(2), 149-165.

Simmons, J. M., & Quinn, K. J. (2014). The NIMH Research Domain Criteria (RDoC) Project: implications for genetics research. *Mammalian Genome, 25*(1), 23-31. doi:10.1007/s00335-013-9476-9

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(8), 921-929.

Singh, T., Kurki, M. I., Curtis, D., Purcell, S. M., Crooks, L., McRae, J., . . . Barrett, J. C. (2016). Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nat Neurosci, 19*(4), 571-577. doi:10.1038/nn.4267

Skokauskas, N., & Gallagher, L. (2010). Psychosis, affective disorders and anxiety in autistic spectrum disorder: prevalence and nosological considerations. *Psychopathology, 43*(1), 8-16.

Skokauskas, N., & Gallagher, L. (2012a). Mental health aspects of autistic spectrum disorders in children. *J Intellect Disabil Res, 56*(3), 248-257. doi:10.1111/j.1365-2788.2011.01423.x

Skokauskas, N., Sweeny, E., Meehan, J., & Gallagher, L. (2012b). Mental health problems in children with prader-willi syndrome. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent, 21*(3), 194-203. Retrieved from https://pubmed.ncbi.nlm.nih.gov/22876265

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3413469/

Smeland, O. B., Bahrami, S., Frei, O., Shadrin, A., O'Connell, K., Savage, J., . . . Andreassen, O. A. (2019). Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. *Molecular Psychiatry*. doi:10.1038/s41380-018-0332-x

Smoller, J. W., Andreassen, O. A., Edenberg, H. J., Faraone, S. V., Glatt, S. J., & Kendler, K. S. (2019). Psychiatric genetics and the structure of psychopathology. *Molecular Psychiatry, 24*(3), 409-420. doi:10.1038/s41380-017-0010-4

Snell, T., Knapp, M., Healey, A., Guglani, S., Evans-Lacko, S., Fernandez, J.-L., . . . Ford, T. (2013). Economic impact of childhood psychiatric disorder on public sector services in Britain: estimates from national survey data. *Journal of Child Psychology and Psychiatry, 54*(9), 977-985. doi:10.1111/jcpp.12055

Son, J. H., Xie, G., Yuan, C., Ena, L., Li, Z., Goldstein, A., . . . Wang, K. (2018). Deep Phenotyping on Electronic Health Records Facilitates Genetic Diagnosis by Clinical Exomes. *American journal of human genetics, 103*(1), 58-73. doi:10.1016/j.ajhg.2018.05.010

Souza, L. D. d. M., Silva, R. A. d., Jansen, K., Kuhn, R. P., Horta, B. L., & Pinheiro, R. T. (2010). Suicidal ideation in adolescents aged 11 to 15 years: prevalence and associated factors. *Brazilian Journal of Psychiatry, 32*, 37-41. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-44462010000100008&nrm=iso

Sparrow, S. S., Cicchetti, D. V., Balla, D. A., & Doll, E. A. (2005). *Vineland adaptive behavior scales: Survey forms manual*: American Guidance Service.

Spitzer, R. L., & Wakefield, J. C. (1999). DSM-IV diagnostic criterion for clinical significance: does it help solve the false positives problem? *Am J Psychiatry, 156*(12), 1856-1864. doi:10.1176/ajp.156.12.1856

Sporn, A. L., Addington, A. M., Gogtay, N., Ordoñez, A. E., Gornick, M., Clasen, L., . . . Lenane, M. (2004). Pervasive developmental disorder and childhood-onset schizophrenia: comorbid disorder or a phenotypic variant of a very early onset illness? *Biological psychiatry, 55*(10), 989-994.

Sporns, O. (2013). Structure and function of complex brain networks. *Dialogues in clinical neuroscience, 15*, 247-262.

Stefansson, H., Meyer-Lindenberg, A., Steinberg, S., Magnusdottir, B., Morgen, K., Arnarsdottir, S., . . . Stefansson, K. (2014). CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature, 505*(7483), 361-366. doi:10.1038/nature12818

Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O. P. H., Ingason, A., Steinberg, S., . . . Group. (2008). Large recurrent microdeletions associated with schizophrenia. *Nature, 455*(7210), 232-236. doi:10.1038/nature07229

Stessman, H. A., Xiong, B., Coe, B. P., Wang, T., Hoekzema, K., Fenckova, M., . . . Cosemans, N. (2017). Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nature Genetics, 49*(4), 515.

Stevens, T., Peng, L., & Barnard-Brak, L. (2016). The comorbidity of ADHD in children diagnosed with autism spectrum disorder. *Research in Autism Spectrum Disorders, 31*, 11-18. doi:https://doi.org/10.1016/j.rasd.2016.07.003

Stewart, L. R., Hall, A. L., Kang, S.-H. L., Shaw, C. A., & Beaudet, A. L. (2011). High frequency of known copy number abnormalities and maternal duplication 15q11-q13 in patients with combined schizophrenia and epilepsy. *BMC medical genetics, 12*(1), 154.

Strang, J. F., Kenworthy, L., Daniolos, P., Case, L., Wills, M. C., Martin, A., & Wallace, G. L. (2012). Depression and anxiety symptoms in children and adolescents with autism spectrum disorders without intellectual disability. *Research in Autism Spectrum Disorders, 6*(1), 406-412. doi:https://doi.org/10.1016/j.rasd.2011.06.015

Sullivan, P. F., Daly, M. J., & O'Donovan, M. (2012a). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature reviews. Genetics, 13*(8), 537-551. doi:10.1038/nrg3240

Sullivan, P. F., Kendler, K. S., & Neale, M. C. (2003). Schizophrenia as a Complex Trait: Evidence From a Meta-analysis of Twin Studies. *Archives of General psychiatry, 60*(12), 1187-1192. doi:10.1001/archpsyc.60.12.1187

Sullivan, P. F., Magnusson, C., Reichenberg, A., Boman, M., Dalman, C., Davidson, M., . . . Långström, N. (2012b). Family history of schizophrenia and bipolar disorder as risk factors for autism. *Archives of General psychiatry, 69*(11), 1099-1103.

Sullivan, S., Rai, D., Golding, J., Zammit, S., & Steer, C. (2013). The Association Between Autism Spectrum Disorder and Psychotic Experiences in the Avon Longitudinal Study of Parents and Children (ALSPAC) Birth Cohort. *Journal of the American Academy of Child & Adolescent Psychiatry, 52*(8), 806-814.e802. doi:https://doi.org/10.1016/j.jaac.2013.05.010

Tammimies, K., Li, D., Rabkina, I., Stamouli, S., Becker, M., Nicolaou, V., . . . Bölte, S. (2019). Association between Copy Number Variation and Response to Social Skills Training in Autism Spectrum Disorder. *Scientific Reports, 9*(1), 9810. doi:10.1038/s41598-019-46396-1

Tansey, K. E., Rees, E., Linden, D. E., Ripke, S., Chambert, K. D., Moran, J. L., . . . O'Donovan, M. C. (2016). Common alleles contribute to schizophrenia in CNV carriers. *Mol Psychiatry, 21*(8), 1085-1089. doi:10.1038/mp.2015.143

Tansey, K. E., Rucker, J. J. H., Kavanagh, D. H., Guipponi, M., Perroud, N., Bondolfi, G., . . . Uher, R. (2014). Copy number variants and therapeutic response to antidepressant medication in major depressive disorder. *The Pharmacogenomics Journal, 14*(4), 395-399. doi:10.1038/tpj.2013.51

Thapar, A., Cooper, M., & Rutter, M. (2017). Neurodevelopmental disorders. *Lancet Psychiatry, 4*(4), 339-346. doi:10.1016/s2215-0366(16)30376-5

Thornicroft, G., Tansella, M., Becker, T., Knapp, M., Leese, M., Schene, A., & Vazquez-Barquero, J. L. (2004). The personal impact of schizophrenia in Europe. *Schizophrenia research, 69*(2), 125-132. doi:https://doi.org/10.1016/S0920-9964(03)00191-9

Thygesen, J. H., Wolfe, K., McQuillin, A., Viñas-Jornet, M., Baena, N., Brison, N., . . . Ribas-Vidal, N. (2018). Neurodevelopmental risk copy number variants in adults with intellectual

disabilities and comorbid psychiatric disorders. *The British Journal of Psychiatry, 212*(5), 287-294.

Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijsdijk, F. (2016). Heritability of autism spectrum disorders: a meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry, 57*(5), 585-595. doi:10.1111/jcpp.12499

Tiihonen, J., Lonnqvist, J., Wahlbeck, K., Klaukka, T., Niskanen, L., Tanskanen, A., & Haukka, J. (2009). 11-year follow-up of mortality in patients with schizophrenia: a population-based cohort study (FIN11 study). *Lancet, 374*(9690), 620-627. doi:10.1016/s0140-6736(09)60742-x

Tomlinson, M., Rudan, I., Saxena, S., Swartz, L., Tsai, A. C., & Patel, V. (2009). Setting priorities for global mental health research. *Bulletin of the World Health Organization, 87*, 438-446.

Tonnsen, B. L., Boan, A. D., Bradley, C. C., Charles, J., Cohen, A., & Carpenter, L. A. (2016). Prevalence of autism spectrum disorders among children with intellectual disability. *American journal on intellectual and developmental disabilities, 121*(6), 487-500.

Torres, F., Barbosa, M., & Maciel, P. (2016). 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. doi:10.1136/jmedgenet-2015-103366

Traylor, M., Markus, H., & Lewis, C. M. (2015). Homogeneous case subgroups increase power in genetic association studies. *European journal of human genetics : EJHG, 23*(6), 863-869. doi:10.1038/ejhg.2014.194

Tromans, S., Chester, V., Gemegah, E., Roberts, K., Morgan, Z., Yao, G. L., & Brugha, T. (2020). Autism identification across ethnic groups: a narrative review. *Advances in Autism.*

Turner, T. H. (1989). Schizophrenia and mental handicap: an historical review, with implications for further research. *Psychol Med, 19*(2), 301-314. doi:10.1017/s0033291700012344

Turner, T. N., Coe, B. P., Dickel, D. E., Hoekzema, K., Nelson, B. J., Zody, M. C., . . . Eichler, E. E. (2017). Genomic Patterns of De Novo Mutation in Simplex Autism. *Cell, 171*(3), 710-722.e712. doi:https://doi.org/10.1016/j.cell.2017.08.047

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-e1109. doi:10.1038/tp.2017.77

Unicef. (2013). *The state of the world's children: children with disabilities*: UN.

United Nations. (2002). Road Map Towards the Implementation of the United Nations Millennium Declaration. *Report No. A56/326*.

van Ommen, G.-J. B. (2005). Frequency of new copy number variation in humans. *Nature Genetics, 37*(4), 333-334.

van Steensel, F. J. A., Bögels, S. M., & Perrin, S. (2011). Anxiety Disorders in Children and Adolescents with Autistic Spectrum Disorders: A Meta-Analysis. *Clinical Child and Family Psychology Review, 14*(3), 302. doi:10.1007/s10567-011-0097-0

Van Winkel, R., Stefanis, N. C., & Myin-Germeys, I. (2008). Psychosocial stress and psychosis. A review of the neurobiological mechanisms and the evidence for gene-stress interaction. *Schizophrenia Bulletin, 34*(6), 1095-1105.

Velinov, M. (2019). Genomic Copy Number Variations in the Autism Clinic-Work in Progress. *Frontiers in cellular neuroscience, 13*, 57-57. doi:10.3389/fncel.2019.00057

Vella, S., & Pai, N. (2013). False positives and false negatives: is the answer relatively simple? *Acta Psychiatr Scand, 127*(1), 83. doi:10.1111/acps.12036

Vicari, S., Napoli, E., Cordeddu, V., Menghini, D., Alesi, V., Loddo, S., . . . Tartaglia, M. (2019). Copy number variants in autism spectrum disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 92*, 421-427. doi:https://doi.org/10.1016/j.pnpbp.2019.02.012

Vissers, L. E., Gilissen, C., & Veltman, J. A. (2016). Genetic studies in intellectual disability and related disorders. *Nature Reviews Genetics, 17*(1), 9-18.

Vissers, L. E. L. M., de Vries, B. B. A., & Veltman, J. A. (2010). Genomic microarrays in mental retardation: from copy number variation to gene, from research to diagnosis. *Journal of medical genetics, 47*(5), 289-297. doi:10.1136/jmg.2009.072942

Voineskos, A. N., Lett, T. A. P., Lerch, J. P., Tiwari, A. K., Ameis, S. H., Rajji, T. K., . . . Kennedy, J. L. (2011). Neurexin-1 and frontal lobe white matter: an overlapping intermediate phenotype for schizophrenia and autism spectrum disorders. *PloS one, 6*(6), e20982-e20982. doi:10.1371/journal.pone.0020982

Volkmar, F. R., Lord, C., Bailey, A., Schultz, R. T., & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 45*(1), 135-170.

W, R. (2019). psych: Procedures for Psychological, Psychometric, and Personality Research. *Northwestern University, Evanston, Illinois., R package version 1.9.12*.

Walsh, T., McClellan, J. M., McCarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M., . . . Sebat, J. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science, 320*(5875), 539-543. doi:10.1126/science.1155174

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F., . . . Bucan, M. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res, 17*(11), 1665-1674. doi:10.1101/gr.6861907

Ward, E. T., Kostick, K. M., & Lázaro-Muñoz, G. (2019). Integrating Genomics into Psychiatric Practice: Ethical and Legal Challenges for Clinicians. *Harvard review of psychiatry, 27*(1), 53-64. doi:10.1097/HRP.0000000000000203

Weiner, D. J., Wigdor, E. M., Ripke, S., Walters, R. K., Kosmicki, J. A., Grove, J., . . . Bybjerg-Grauholm, J. (2017). Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nature Genetics, 49*(7), 978-985.

Welham, J., Isohanni, M., Jones, P., & McGrath, J. (2009a). The antecedents of schizophrenia: a review of birth cohort studies. *Schizophrenia Bulletin, 35*(3), 603-623.

Welham, J., Scott, J., Williams, G., Najman, J., Bor, W., O'callaghan, M., & McGrath, J. (2009b). Emotional and behavioural antecedents of young adults who screen positive for non-affective psychosis: a 21-year birth cohort study. *Psychological medicine, 39*(4), 625-634.

Wichstrøm, L. (2000). Predictors of adolescent suicide attempts: a nationally representative longitudinal study of Norwegian adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(5), 603-610.

Wigman, J. T., Vollebergh, W. A., Raaijmakers, Q. A., Iedema, J., Van Dorsselaer, S., Ormel, J., . . . van Os, J. (2011). The structure of the extended psychosis phenotype in early adolescence—a cross-sample replication. *Schizophrenia Bulletin, 37*(4), 850-860.

Wilkinson, G. S., & Robertson, G. J. (2006). Wide range achievement test (WRAT4). *Lutz, FL: Psychological Assessment Resources*.

Willcocks, L. C., Lyons, P. A., Clatworthy, M. R., Robinson, J. I., Yang, W., Newland, S. A., . . . Chilvers, E. R. (2008). Copy number of FCGR3B, which is associated with systemic lupus erythematosus, correlates with protein expression and immune complex uptake. *The Journal of experimental medicine, 205*(7), 1573-1582.

Williams, L. R., Degnan, K. A., Perez-Edgar, K. E., Henderson, H. A., Rubin, K. H., Pine, D. S., . . . Fox, N. A. (2009). Impact of behavioral inhibition and parenting style on internalizing and externalizing problems from early childhood through adolescence. *Journal of abnormal child psychology, 37*(8), 1063-1075.

Wilson, N. K., 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 Rep Genet, 2011*, 585893. doi:10.1155/2011/585893

Wolfe, K., Strydom, A., Morrogh, D., Carter, J., Cutajar, P., Eyeoyibo, M., . . . Bass, N. (2017). Chromosomal microarray testing in adults with intellectual disability presenting with comorbid psychiatric disorders. *European Journal of Human Genetics, 25*(1), 66-72. doi:10.1038/ejhg.2016.107

Wong, C. C. Y., Meaburn, E. L., Ronald, A., Price, T. S., Jeffries, A. R., Schalkwyk, L. C., . . . Mill, J. (2014). Methylomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural traits. *Molecular Psychiatry, 19*(4), 495-503. doi:10.1038/mp.2013.41

Woodberry, K. A., Giuliano, A. J., & Seidman, L. J. (2008). Premorbid IQ in schizophrenia: a meta-analytic review. *American Journal of Psychiatry, 165*(5), 579-587.

Woodbury-Smith, M., & Scherer, S. W. (2018). Progress in the genetics of autism spectrum disorder. *Developmental Medicine & Child Neurology, 60*(5), 445-451. doi:10.1111/dmcn.13717

Woodman, A. C., Mailick, M. R., & Greenberg, J. S. (2016). Trajectories of internalizing and externalizing symptoms among adults with autism spectrum disorders. *Development and psychopathology, 28*(2), 565-581. doi:10.1017/S095457941500108X

World Health, O. (2004). ICD-10 : international statistical classification of diseases and related health problems : tenth revision. In (2nd ed ed.). Geneva: World Health Organization.

Wright, C. F., FitzPatrick, D. R., & Firth, H. V. (2018). Paediatric genomics: diagnosing rare disease in children. *Nature Reviews Genetics, 19*(5), 253.

Xu, B., Roos, J. L., Dexheimer, P., Boone, B., Plummer, B., Levy, S., . . . Karayiorgou, M. (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nature Genetics, 43*(9), 864-868. doi:10.1038/ng.902

Xu, L., Hou, Y., Bickhart, D. M., Song, J., & Liu, G. E. (2013). Comparative Analysis of CNV Calling Algorithms: Literature Survey and a Case Study Using Bovine High-Density SNP Data. *Microarrays (Basel, Switzerland), 2*(3), 171-185. doi:10.3390/microarrays2030171

Yeo, R. A., Gangestad, S. W., Liu, J., Ehrlich, S., Thoma, R. J., Pommy, J., . . . Calhoun, V. D. (2013). The impact of copy number deletions on general cognitive ability and ventricle size in patients with schizophrenia and healthy control subjects. *Biol Psychiatry, 73*(6), 540-545. doi:10.1016/j.biopsych.2012.10.013

Yuen, R. K., Merico, D., Bookman, M., Howe, J. L., Thiruvahindrapuram, B., Patel, R. V., . . . Wang, Z. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci, 20*(4), 602.

Zarrei, M., MacDonald, J. R., Merico, D., & Scherer, S. W. (2015). A copy number variation map of the human genome. *Nat Rev Genet, 16*(3), 172-183. doi:10.1038/nrg3871

Zwicker, A., Denovan-Wright, E. M., & Uher, R. (2018). Gene-environment interplay in the etiology of psychosis. *Psychol Med, 48*(12), 1925-1936. doi:10.1017/s003329171700383x

Appendices

7.1 Appendix 1: Supplementary Tables

Supplementary Tables- Chapter 2

Table 7-1. Neurodevelopmental copy number variant list. Based on list previously compiled by Martin et al. (Martin et al., 2019) and contains ND CNVs with association with ASD (Sanders et al., 2015), ID (Coe et al., 2014; Rees et al., 2016b) and schizophrenia (Marshall et al., 2017).

Locus/Syndrome	Locus	Chr	Start (hg18)	End (hg 18)	Start (hg19)	End(hg19)	Reference
1p36 del	1p36	1	3588	2,474,467	0	2500000	(Coe et al., 2014; Rees et al., 2016b)
1p36 dup	1p36	1	3588	2474467	0	2500000	(Coe et al., 2014; Rees et al., 2016b)
TAR del	1q21.1	1	144087069	144461205	145394955	145807817	(Coe et al., 2014; Rees et al., 2016b)
TAR dup	1q21.1	1	144087069	144461205	145394955	145807817	(Coe et al., 2014; Rees et al., 2016b)
1q21.1 del	1q21.1	1	145045482	145863214	146527987	147394444	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b; Sanders et al., 2015)
1q21.1 dup	1q21.1	1	145045482	145863214	146527987	147394444	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b; Sanders et al., 2015)

Locus/Syndrome	Locus	Chr	Start (hg18)	End (hg 18)	Start (hg19)	End(hg19)	Reference
1q24 (FMO and DNM3) del	1q24	1	167946957	171569960	169680333	173303337	(Coe et al., 2014; Rees et al., 2016b)
NRXN1 del	2p16.3	2	50000992	51113178	50145643	51259674	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b; Sanders et al., 2015)
2p15-16.1 proximal (PEX13 to AHSA2) dup	2p15-16.1	2	61098792	61268076	61245288	61414572	(Coe et al., 2014; Rees et al., 2016b)
2q11.2 del	2q11.2	2	96107648	97075984	96742409	97677516	(Coe et al., 2014; Rees et al., 2016b)
2q13 del	2q13	2	111109777	111727681	111394040	112012649	(Coe et al., 2014; Rees et al., 2016b)
2q13 dup	2q13	2	111109777	111727681	111394040	112012649	(Coe et al., 2014; Rees et al., 2016b)
2q33.1 (SATB2) del	2q33.1	2	199842469	200033500	200134224	200325255	(Coe et al., 2014; Rees et al., 2016b)
2q37 (HDAC4) del	2q37	2	239370000	242120000	239716679	243199373	(Coe et al., 2014; Rees et al., 2016b)
3p25.3 (JAGN1 to TATDN2) dup	3p25.3	3	9907271	10297902	9932271	10322902	(Coe et al., 2014; Rees et al., 2016b)
3p11.2 (CHMP2B to POU1F1) del	3p11.2	3	87350302	87614321	87267612	87531631	(Coe et al., 2014; Rees et al., 2016b)

Locus/Syndrome	Locus	Chr	Start (hg18)	End (hg 18)	Start (hg19)	End(hg19)	Reference
3q13 (GAP43) del	3q13	3	116815024	116986728	115332334	115504038	(Coe et al., 2014; Rees et al., 2016b)
3q28-29 (FGF12) del	3q28-29	3	193342422	193608706	191859728	192126012	(Coe et al., 2014; Rees et al., 2016b)
3q29 del	3q29	3	197244323	198830949	195720167	197354826	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b; Sanders et al., 2015)
Wolf-Hirschhorn del	4p16.3	4	1840000	1980000	1552030	2091303	(Coe et al., 2014; Rees et al., 2016b)
Wolf-Hirschhorn dup	4p16.3	4	1840000	1980000	1552030	2091303	(Coe et al., 2014; Rees et al., 2016b)
4q21 (BMP3) del	4q21	4	82164501	82204351	81945477	81985327	(Coe et al., 2014; Rees et al., 2016b)
5q14 (MEF2C) del	5q14	5	88047410	88236459	88011654	88200703	(Coe et al., 2014; Rees et al., 2016b)
Sotos syndrome del	5q35.3	5	175664037	176981275	175720924	177052594	(Coe et al., 2014; Rees et al., 2016b)
Williams-Beuren syndrome (WBS) del	7q11.23	7	72383764	73783369	72744915	74142892	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b)

Locus/Syndrome	Locus	Chr	Start (hg18)	End (hg 18)	Start (hg19)	End(hg19)	Reference
WBS dup	7q11.23	7	72383764	73783369	72744915	74142892	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b)
7q11.23	7q11.23	7	72411506	72795997	72773570	73158061	(Sanders et al., 2015)
7q11.23	7q11.23	7	72411506	72795997	72773570	73158061	(Sanders et al., 2015)
7q11.23	7q11.23	7	73616737	73782113	73978801	74144177	(Sanders et al., 2015)
7q11.23	7q11.23	7	73616737	73782113	73978801	74144177	(Sanders et al., 2015)
8p23.1 del	8p23.1	8	8142910	11902028	8098990	11872558	(Coe et al., 2014; Rees et al., 2016b)
8p23.1 dup	8p23.1	8	8142910	11902028	8098990	11872558	(Coe et al., 2014; Rees et al., 2016b)
9p13 dup	9p13	9	32638800	38798255	32648800	38808255	(Coe et al., 2014; Rees et al., 2016b)
9q34 dup	9q34	9	137600518	140156247	138460697	141036426	(Coe et al., 2014; Rees et al., 2016b)
10q11.21q11.23 dup	10q11.21-q11.23	10	49063225	50726582	49390199	51058796	(Coe et al., 2014; Rees et al., 2016b)
10q23 del	10q23	10	82020665	88871284	82045472	88931651	(Coe et al., 2014; Rees et al., 2016b)
Potocki-Shaffer syndrome del	11p11.2	11	43940000	46020000	43940000	46020000	(Coe et al., 2014; Rees et al., 2016b)

Locus/Syndrome	Locus	Chr	Start (hg18)	End (hg 18)	Start (hg19)	End(hg19)	Reference
12p13 (SCNN1A to PIANP) dup	12p13	12	6342220	6696216	6471959	6825955	(Coe et al., 2014; Rees et al., 2016b)
Prader-Willi syndrome/Angelman syndrome (PWS/AS) del	15q11.2-13.1	15	20306099	26244113	22805313	28390339	(Coe et al., 2014; Rees et al., 2016b)
PWS/AS dup	15q11.2-13.1	15	20306099	26244113	22805313	28390339	(Coe et al., 2014; Rees et al., 2016b)
15q11.2 BP1-BP2 del	15q11.2	15	20353880	20643728	22805313	23094530	(Coe et al., 2014; Rees et al., 2016b; Sanders et al., 2015)
15q12	15q12	15	24522927	25131566	26971834	27548820	(Sanders et al., 2015)
15q13.2-13.3	15q13.2-13.3	15	28730804	30303141	30943512	32515849	(Sanders et al., 2015)
15q13.3 del	15q13.3	15	28863083	30234145	31080645	32462776	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b; Sanders et al., 2015)
15q24 del	15q24	15	70761609	75943228	72900171	78151253	(Coe et al., 2014; Rees et al., 2016b)
15q24 dup	15q24	15	70761609	75943228	72900171	78151253	(Coe et al., 2014; Rees et al., 2016b)
15q25 del	15q25	15	82940819	83517628	85139815	85716624	(Coe et al., 2014; Rees et al., 2016b)

Locus/Syndrome	Locus	Chr	Start (hg18)	End (hg 18)	Start (hg19)	End(hg19)	Reference
16p13.11 del	16p13.11	16	15420294	16201433	15511655	16293689	(Coe et al., 2014; Rees et al., 2016b)
16p13.11 dup	16p13.11	16	15420294	16201433	15511655	16293689	(Coe et al., 2014; Rees et al., 2016b)
16p12.1 del	16p12.1	16	21857379	22338610	21950135	22431889	(Coe et al., 2014; Rees et al., 2016b)
16p11.2 distal del	16p11.2	16	28730590	28954235	28823196	29046783	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b)
16p11.2 distal dup	16p11.2	16	28730590	28954235	28823196	29046783	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b)
16p11.2 del	16p11.2	16	29558245	30108138	29650840	30200773	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b; Sanders et al., 2015)
16p11.2 dup	16p11.2	16	29558245	30108138	29650840	30200773	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b; Sanders et al., 2015)
17p13.3 (YWHAE and PAFAH1B1) del	17p13.3	17	1194584	2535659	1247834	2588909	(Coe et al., 2014; Rees et al., 2016b)

Locus/Syndrome	Locus	Chr	Start (hg18)	End (hg 18)	Start (hg19)	End(hg19)	Reference
17p13.3 (YWHAE and PAFAH1B1) dup	17p13.3	17	1194584	2535659	1247834	2588909	(Coe et al., 2014; Rees et al., 2016b)
Smith-Magenis syndrome del	17p11.2	17	16752468	20157424	16812771	20211017	(Coe et al., 2014; Rees et al., 2016b)
Potocki-Lupski syndrome dup	17p11.2	17	16752468	20157424	16812771	20211017	(Coe et al., 2014; Rees et al., 2016b)
17q11.2 del	17q11.2	17	26132227	27288567	29107491	30265075	(Coe et al., 2014; Rees et al., 2016b)
17q11.2 dup	17q11.2	17	26132227	27288567	29107491	30265075	(Coe et al., 2014; Rees et al., 2016b)
17q12 del	17q12	17	31890685	33289785	34815904	36217432	(Coe et al., 2014; Rees et al., 2016b)
17q12 dup	17q12	17	31890685	33289785	34815904	36217432	(Coe et al., 2014; Rees et al., 2016b)
17q21.31 (Koolen-de Vries syndrome) del	17q21.31	17	41061902	41520974	43705356	44164691	(Coe et al., 2014; Rees et al., 2016b)
22q11.2 (DiGeorge/VCFS syndrome) del	22q11.2	22	17404860	19799135	19037332	21466726	(Coe et al., 2014; Marshall et al., 2017; Rees et al., 2016b; Sanders et al., 2015)
22q11.2 dup	22q11.2	22	17404860	19799135	19037332	21466726	(Coe et al., 2014; Marshall et al., 2017;

Locus/Syndrome	Locus	Chr	Start (hg18)	End (hg 18)	Start (hg19)	End(hg19)	Reference
							Rees et al., 2016b; Sanders et al., 2015)
distal 22q11.2 del	22q11.2	22	20250486	21982202	21920127	23653646	(Coe et al., 2014; Rees et al., 2016b)
distal/22q11.2 dup	22q11.2	22	20250486	21982202	21920127	23653646	(Coe et al., 2014; Rees et al., 2016b)
Phelan-McDermid syndrome del	22q13.33	22	49459936	49518507	51113070	51171640	(Coe et al., 2014; Rees et al., 2016b; Sanders et al., 2015)
Phelan-McDermid syndrome dup	22q13.33	22	49459936	49518507	51113070	51171640	(Coe et al., 2014; Rees et al., 2016b; Sanders et al., 2015)

Note: hg18, human reference genome version 18; hg19, human reference genome version 19; del, deletion; dup, duplication.

Table 7-2. Copy number variants identified in individuals in SSC (probands and siblings).

Case	Band	Chr	Size	CNV	Position in Mb	
					Start	Stop
Case 1	1q21.1	1	1668719	Dup	146007521	147920973
Case 2	1q21.1	1	2743667	Dup	145900678	149089078
Case 3	1q21.1	1	1402895	Dup	146467203	147870096
Case 4	1q21.1	1	1327988	Dup	146494246	147822234
Case 5	1q21.1	1	1841899	Dup	145771576	147858208
Case 6	1q21.1	1	360062	Del	145388137	145748199
Case 7	1q21.1	1	2674563	Del	146126818	149246114
Case 8	1q21.1	1	1354367	Del	146503841	147858208
Case 9	1q21.1	1	3922572	Dup	143663775	147911246
Case 10	1q21.1	1	3922572	Dup	143663775	147911246
Case 11	1q21.1	1	376734	Dup	145386225	145762959
Case 12	1q21.1	1	58801	Del	143663775	143722576
Case 13	1q21.1	1	1577259	Dup	146089254	147911246
Case 14	2q11.2	2	1453239	Del	96658588	98109123
Case 15	2q13	2	1723146	Dup	111376870	113100014
Case 16	2q13	2	1713783	Dup	111386233	113100014

Case	Band	Chr	Size	CNV	Position in Mb	
					Start	Stop
Case 17	2q11.2	2	1459095	Del	96737083	98193473
Case 18	2q11.2	2	1429521	Del	96741757	98168573
Case 19	2q11.2	2	1525675	Del	96499474	98022737
Case 20	3q29	3	1631717	Del	195734915	197366632
Case 21	3q29	3	1599573	Del	195747398	197346971
Case 22	3q28	3	125557	Del	191921546	192047103
Case 23	3q29	3	1612056	Del	195734915	197346971
Case 24	7q11.23	7	1421196	Dup	72722981	74144177
Case 25	7q11.23	7	509974	Dup	73978801	74488775
Case 26	7q11.23	7	1421196	Dup	72722981	74144177
Case 27	7q11.23	7	1370607	Dup	72773570	74144177
Case 28	7q11.23	7	2019505	Dup	74455447	76274952
Case 29	7q11.23	7	2384025	Del	74455447	76639472
Case 30	10q11.21,10q11.22, 10q11.23	10	5557765	Dup	46518615	51826380
Case 31	15q25.2,15q25.3	15	792026	Del	84947899	85729925
Case 32	15q25.2,15q25.3	15	814178	Del	84952634	85756812
Case 33	15q13.2,15q13.3	15	1579688	Dup	30936285	32515973
Case 34	15q11.1,15q11.2,15 q12,15q13.1	15	8452637	Dup	20015396	28929005
Case 35	15q13.2,15q13.3	15	1796251	Del	30732837	32529088

Case	Band	Chr	Size	CNV	Position in Mb	
					Start	Stop
Case 36	15q13.2,15q13.3	15	1581629	Del	30934220	32515849
Case 37	15q11.2	15	1499046	Del	21021056	23228712
Case 38	15q11.2,15q12,15q 13.1	15	4909512	Dup	23683783	28471141
Case 39	15q11.2,15q12,15q 13.1	15	5964103	Dup	22652330	28494202
Case 40	15q11.2,15q12,15q 13.1	15	6547570	Dup	22750305	29050198
Case 41	15q11.1,15q11.2,15 q12,15q13.1,15q13. 2,15q13.3	15	12142124	Dup	20005287	32620127
Case 42	16p11.2	16	581055	Del	29624247	30205302
Case 43	16p11.2	16	568279	Del	29647342	30215621
Case 44	16p13.11	16	1333059	Del	14975292	16308351
Case 45	16p13.11	16	906289	Dup	15445228	16351517
Case 46	16p11.2	16	592186	Del	29647342	30239528
Case 47	16p13.11	16	1193494	Dup	15069867	16263361
Case 48	16p11.2	16	684861	Dup	29647342	30332203
Case 49	16p11.2	16	557960	Dup	29647342	30205302
Case 50	16p11.2	16	552463	Del	29647342	30199805
Case 51	16p11.2	16	584367	Del	29624247	30208614

Case	Band	Chr	Size	CNV	Position in Mb	
					Start	Stop
Case 52	16p12.3,16p13.11	16	2851345	Dup	15459174	18310519
Case 53	16p11.2	16	575558	Del	29624247	30199805
Case 54	16p11.2	16	568279	Del	29647342	30215621
Case 55	16p12.3,16p13.11	16	2804168	Del	15479879	18284047
Case 56	16p13.11	16	1239237	Dup	15052746	16291983
Case 57	16p11.2	16	596312	Dup	29603493	30199805
Case 58	16p11.2	16	575558	Del	29624247	30199805
Case 59	16p11.2	16	567497	Dup	29655864	30223361
Case 60	16p13.11	16	921498	Del	15383857	16305355
Case 61	16p13.11	16	1199205	Del	15092778	16291983
Case 62	16p13.11	16	1316691	Dup	14975292	16291983
Case 63	16p11.2	16	547706	Dup	29647342	30195048
Case 64	16p11.2	16	220677	Del	28822773	29043450
Case 65	16p11.2	16	604322	Dup	29595483	30199805
Case 66	16p11.2	16	631004	Dup	29592357	30223361
Case 67	16p11.2	16	552463	Dup	29647342	30199805
Case 68	16p11.2	16	599565	Dup	29595483	30195048
Case 69	16p13.11	16	1269997	Dup	15021986	16291983
Case 70	16p13.11	16	1239237	Dup	15052746	16291983
Case 71	16p12.1	16	586069	Del	21839340	22425409
Case 72	16p11.2	16	633857	Dup	29581764	30215621

Case	Band	Chr	Size	CNV	Position in Mb	
					Start	Stop
Case 73	16p13.11	16	1212577	Dup	15092778	16305355
Case 74	16p13.11	16	825476	Dup	15479879	16305355
Case 75	17q12	17	1448856	Del	34815551	36264149
Case 76	17q11.2	17	1320156	Dup	29060212	30380381
Case 77	17q12	17	1448856	Del	34815551	36264149
Case 78	22q11.21	22	2127049	Del	18874965	21052014
Case 79	22q11.21	22	2875076	Dup	18886915	21811991
Case 80	22q11.21	22	2539514	Dup	18874965	21464479
Case 81	22q11.21	22	2708385	Del	18874965	21633350

Note: Copy number variation positions are in UCSC build hg19. Del, deletion; dup, duplication.

Table 7-3. Power analyses for associations between psychiatric phenotypes and ND CNV status in SSC sample. Power analyses for associations between psychiatric phenotypes and ND CNV status in SSC sample.

	SSC ASD Proband Sample						SSC Sibling Sample
Input parameter	**Affective problems**	**Anxiety problems**	**Somatic problems**	**ADHD problems**	**Oppositional defiant problems**	**Conduct problems**	**Affective problems**
Tail(s)	Two	Two	Two	Two	Two	Two	Two
Odds ratio	0.71	0.80	0.98	1.4	1.29	1.58	4.14
Pr(Y=1|X=1) H0	0.39	0.45	0.15	0.39	0.26	0.19	0.05
α err prob	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Total sample size	1966	1966	1966	1966	1966	1966	1634
R^2 other X	0.001	0.001	0.001	0.001	0.001	0.001	0.001
X distribution	Binomial	Binomial	Binomial	Binomial	Binomial	Binomial	Binomial
X parm π	0.03	0.03	0.03	0.03	0.03	0.03	0.01
Output parameter							
Critical Z	-1.96	-1.96	-1.96	1.96	1.96	1.96	1.96
Power	0.22	0.13	0.05	0.25	0.16	0.35	0.56

*Note: Test family, z tests; Statistical test, Logistic regression; Type of power analysis, Post hoc. Power analyses completed using G*Power (Faul et al., 2009).*

Supplementary Tables- Chapter 3

Table 7-4. Demographic characteristics of recruitment pool (n=19,161) by enrolment status, from Calkins et al. (Calkins et al., 2015).

	Recruitment Status				Comparison			
	Enrolled	**Not Invited[a]**	**Declined**	**Excluded[b]**	**Statistical Test**	**df**	**Result**	**p**
Age								
Mean	14.27	15.04*	14.70*	14.90*	ANOVA	3,18497	53.67	0.001
s.d.	3.67	3.84	3.63	3.83				
Cohen's d	--	0.21	0.12	0.17				
Sex								
M:F %	48:52	49:51	52:48	50:50	Chi-Square	6	11.51	n.s.
Race								
EA:non-EA %	58:42	44:56*	65:36*	42:58*	Chi-square	3	468.43	0.001

*Note: Table used directly from supplementary information of Calkins et al.(Calkins et al., 2015). Following details provided from Calkins et al.(Calkins et al., 2015): "Note: Analyses excluded those in the pool who were deceased or outside the study age range of 8-21. *Post-hoc pairwise comparison significantly different from enrolled (all p's <0.05). Cohen's d = effect size of mean difference between enrolled and comparison group. EA=European-American. [a]As shown in Figure 1, the majority of individuals in the not invited category were unreachable, and this was disproportionately so in the non-EA group (57.1%) compared to the EA group (42.9%). [b]Among those excluded, a disproportionately high number of non-EA (65.1%) compared to EA (34.9%) were excluded due to multiple no-shows, cancellations or re-schedulings."*

Table 7-5. Phenotype Variables for ADHD in PNC provided through dbGaP.

Variable Code	Variable	Outcome	Code
ADD011	Attention Deficit Disorder: Did you often have trouble paying attention or keeping your mind on your school, work, chores, or other activities that you were doing?	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD012	Attention Deficit Disorder: Did you often have problems following instructions and often fail to finish school, work, or other things you meant to get done?	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD013	Attention Deficit Disorder: Did you often dislike, avoid, or put off school or homework (or any other activity requiring concentration)	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD014	Attention Deficit Disorder: Did you often lose things you needed for school or projects at home (assignments or books) or make careless mistakes in schoolwork or other activities?	No	0
		Yes	1

Variable Code	Variable	Outcome	Code
		Not Available/Pending Validation	9
		Unknown	NA
ADD015	Attention Deficit Disorder: Did you often have trouble making plans, doing things that had to be done in a certain kind of order, or that had a lot of different steps?	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD016	Attention Deficit Disorder: Did you often have people tell you that you did not seem to be listening when they spoke to you or that you were daydreaming?	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD017	Attention Deficit Disorder: Did you have these difficulties at home (with your parents)?	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD018	Attention Deficit Disorder: Did you have these difficulties at school (with your teachers)?	No	0
		Yes	1

Variable Code	Variable	Outcome	Code
		Not Available/Pending Validation	9
		Unknown	NA
		No	0
		Yes	1
ADD019	Attention Deficit Disorder: Did you have these difficulties anywhere else (with other adults)?	Not Available/Pending Validation	9
		Unknown	NA
		No	0
		Yes	1
ADD020	Attention Deficit Disorder: Did you often have difficulty sitting still for more than a few minutes at a time, even after being asked to stay seated, or did you often fidget with your hands or feet or wiggle in your seat or were you "always on the go"?	Not Available/Pending Validation	9
		Unknown	NA
		No	0
		Yes	1
ADD021	Attention Deficit Disorder: Did you often blurt out answers to other people's questions before they finished speaking or interrupt people abruptly?	Not Available/Pending Validation	9
		Unknown	NA
		No	0
ADD022	Attention Deficit Disorder: Did you often join other people's conversations or have trouble waiting your turn (e.g., waiting in line, waiting for a teacher to call on you in class)?	Yes	1

Variable Code	Variable	Outcome	Code
		Not Available/Pending Validation	9
		Unknown	NA
		No	0
		Yes	1
ADD023	Attention Deficit Disorder: Did you have these difficulties at home (with your parents)?	Not Available/Pending Validation	9
		Unknown	NA
		No	0
		Yes	1
ADD024	Attention Deficit Disorder: Did you have these difficulties at school (with your teachers)?	Not Available/Pending Validation	9
		Unknown	NA
		No	0
		Yes	1
ADD025	Attention Deficit Disorder: Did you have these difficulties anywhere else (with others adults)?	Not Available/Pending Validation	9
		Unknown	NA
ADD026	Attention Deficit Disorder: When did these difficulties with attention and/or restlessness begin? (age)	Numerical value	0-21

Variable Code	Variable	Outcome	Code
ADD027	Attention Deficit Disorder: You told me that you (list endorsed attentional problems). How much did having difficulties with concentration, attention, or being restless or impatient upset or bother you?	Value range (0-10, NA). 0=None, 5=Significant, 10=Extremely Serious	0-10, NA
ADD028	Attention Deficit Disorder: Did family members seem upset, angry, or annoyed with you because of your difficulties?	Value range (0-10, NA) 0=None, 5=Significant, 10=Extremely Serious	0-10, NA
ADD029	Attention Deficit Disorder: Did these behaviors/inattention bother your friends?	Value range (0-10, NA) 0=None, 5=Significant, 10=Extremely Serious	0-10, NA
ADD030	Attention Deficit Disorder: Did teachers or classmates (or co-workers/supervisors) complain about your inattention/behavior?	Value range (0-10, NA) 0=None, 5=Significant - 10=Extremely Serious	0-10, NA
ADD032	Attention Deficit Disorder: Did you stay home from school or work because of your difficulties with inattention or overactivity? (For example: Were you sent home?)	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD033	Attention Deficit Disorder: How many days of school or work did you miss because of your difficulties with attention or overactivity? (lifetime)	Numerical Value	0-780, NA

Variable Code	Variable	Outcome	Code
ADD034	Attention Deficit Disorder: How old were you the first time you did these (list behaviors)? (Age)	Numerical Value	0-18
ADD034A	Attention Deficit Disorder: How old were you the first time you did these (list behaviors)? (Always)	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD035	Attention Deficit Disorder: Do you still have any of these (list behaviors)?	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA
ADD036	Attention Deficit Disorder: How old were you the last time you did this behavior?	Numerical Value	0-25, NA
ADD050	Attention Deficit Disorder: Did a teacher ever talk to your parents about problems paying attention or sitting still?	No	0
		Yes	1
		Not Available/Pending Validation	9
		Unknown	NA

Table 7-6. Proportions of outcomes for DSM-IV disorders in PNC.

DSM-IV Disorder	*No* *N (%)*	*Yes* *N (%)*	*NA* *N (%)*
Major depression (Major depressive episode)	5421 (65.7)	738 (9.5)	2046 (24.8)
Anxiety disorder	4351 (53.0)	2061 (25.1)	1793 (21.9)
ADHD	7110 (86.7)	749 (9.1)	346 (4.2)
ODD	6298 (76.8)	1797 (21.9)	110 (1.3)
Conduct Disorder	7532 (91.8)	604 (7.4)	69 (0.8)

Note: Yes- indicates proportion of individual who met DSM-IV criteria for disorder based on clinical scoring algorithm; No- indicates proportion of individual who did not meet DSM-IV criteria for disorder based on clinical scoring algorithm; NA- indicates insufficient data to classify into Yes or No outcome. ADHD, attention deficit hyperactivity disorder; ODD, oppositional defiant disorder

Table 7-7. Possible classification outcomes for internalising disorders variable.

Met DSM-IV criteria for major depressive episode	Met DSM-IV criteria for an anxiety disorder		Internalising disorder classification	Explanation for classification
Yes	Yes	→	Yes	Meets criteria for at least one internalising disorder.
Yes	No	→	Yes	Meets criteria for at least one internalising disorder.
No	Yes	→	Yes	Meets criteria for at least one internalising disorder.
Yes	NA	→	Yes	Meets criteria for at least one internalising disorder.
NA	Yes	→	Yes	Meets criteria for at least one internalising disorder.
No	No	→	No	Does not meet criteria for either internalising disorder.
No	NA	→	NA	Does not meet criteria for one internalising disorders. Status regarding one internalising disorder not known. Cannot conclusively outrule internalising disorder.
NA	No	→	NA	Does not meet criteria for one internalising disorders. Status regarding one internalising disorder not known. Cannot conclusively outrule internalising disorder.
NA	NA	→	NA	Status regarding both internalising disorders not known. Cannot conclusively outrule internalising disorder.

Table 7-8. Neurodevelopmental copy number variants identified in individuals in the Philadelphia Neurodevelopmental Cohort.

Case	Locus/syndrome	Chr	Size	CNV	Position in Mb	
					Start(hg19)	End(hg19)
Case 1	1p36	dup	1	1314749	391411	1706160
Case 2	1p36	dup	1	1314749	391411	1706160
Case 3	1p36	dup	1	1314749	391411	1706160
Case 4	1q21.1	del	1	1333583	146494246	147827829
Case 5	1q21.1	del	1	1368748	146501348	147870096
Case 6	1q21.1	del	1	1350132	146476526	147826658
Case 7	1q21.1	del	1	1326481	146501348	147827829
Case 8	1q21.1	del	1	1330896	146499479	147830375
Case 9	1q21.1	dup	1	1351303	146476526	147827829
Case 10	1q21.1	dup	1	1706685	146089254	147795939
Case 11	1q21.1	dup	1	1350132	146476526	147826658
Case 12	TAR	del	1	391163	145397947	145789110
Case 13	TAR	del	1	391163	145397947	145789110
Case 14	TAR	del	1	368004	145394955	145762959
Case 15	TAR	del	1	391163	145397947	145789110
Case 16	TAR	dup	1	352508	145394955	145747463
Case 17	TAR	dup	1	352508	145394955	145747463
Case 18	TAR	dup	1	517637	145394955	145912592
Case 19	TAR	dup	1	352508	145394955	145747463

Case	Locus/syndrome	Chr	Size	CNV	Position in Mb	
					Start(hg19)	End(hg19)
Case 20	TAR	dup	1	394155	145394955	145789110
Case 21	TAR	dup	1	368004	145394955	145762959
Case 22	TAR	dup	1	368004	145394955	145762959
Case 23	TAR	dup	1	436205	145394955	145831160
Case 24	TAR	dup	1	346903	145416056	145762959
Case 25	TAR	dup	1	612566	145394955	146007521
Case 26	2p15-16.1	dup	2	95331	61273201	61368532
Case 27	2q11.2	del	2	1456454	96166556	97623010
Case 28	2q13	del	2	479605	111451816	111931421
Case 29	2q13	dup	2	426399	111392259	111818658
Case 30	NRXN1	del	2	564201	50741932	51306133
Case 31	NRXN1	del	2	291803	50932288	51224091
Case 32	NRXN1	del	2	157294	51136802	51294096
Case 33	NRXN1	del	2	356346	50184394	50540740
Case 34	NRXN1	del	2	139226	50943528	51082754
Case 35	3p25.3	dup	3	267953	9947997	10215950
Case 36	3q29	del	3	1609261	195734915	197344176
Case 37	8p23.1	dup	8	4744051	7718061	12462112
Case 38	10q11.21q11.23	dup	10	1491003	49636507	51127510
Case 39	10q11.21q11.23	dup	10	976933	50075461	51052394
Case 40	10q11.21q11.23	dup	10	2910510	49456053	52366563

Case	Locus/syndrome	Chr	Size	CNV	Position in Mb	
					Start(hg19)	End(hg19)
Case 41	10q11.21q11.23	dup	10	3011636	49402084	52413720
Case 42	15q11.2	del	15	464126	22763396	23227522
Case 43	15q11.2	del	15	346461	22769771	23116232
Case 44	15q11.2	del	15	208512	22853808	23062320
Case 45	15q11.2	del	15	472337	22755185	23227522
Case 46	15q11.2	del	15	208512	22853808	23062320
Case 47	15q11.2	del	15	522428	22750305	23272733
Case 48	15q11.2	del	15	472337	22755185	23227522
Case 49	15q11.2	del	15	348691	22787103	23135794
Case 50	15q11.2	del	15	472337	22755185	23227522
Case 51	15q11.2	del	15	372398	22763396	23135794
Case 52	15q11.2	del	15	327150	22769771	23096921
Case 53	15q11.2	del	15	352836	22763396	23116232
Case 54	15q11.2	del	15	522428	22750305	23272733
Case 55	15q11.2	del	15	522428	22750305	23272733
Case 56	15q11.2	del	15	475949	22750305	23226254
Case 57	15q11.2	del	15	208512	22853808	23062320
Case 58	15q11.2	del	15	346461	22769771	23116232
Case 59	15q11.2	del	15	366023	22769771	23135794
Case 60	15q11.2	del	15	441204	22750305	23191509
Case 61	15q11.2	del	15	352836	22763396	23116232

Case	Locus/syndrome	Chr	Size	CNV	Position in Mb	
					Start(hg19)	End(hg19)
Case 62	15q11.2	del	15	472337	22755185	23227522
Case 63	15q11.2	del	15	522428	22750305	23272733
Case 64	15q11.2	del	15	329258	22755185	23084443
Case 65	15q11.2	del	15	348691	22787103	23135794
Case 66	15q11.2	del	15	475949	22750305	23226254
Case 67	15q11.2	del	15	472337	22755185	23227522
Case 68	15q11.2	del	15	372398	22763396	23135794
Case 69	15q11.2	del	15	477217	22750305	23227522
Case 70	15q13.2-13.3	dup	15	1626431	30918248	32544679
Case 71	15q13.2-13.3	dup	15	1578641	30936285	32514926
Case 72	15q13.2-13.3	dup	15	1578641	30936285	32514926
Case 73	15q13.2-13.3	dup	15	1551408	30963518	32514926
Case 74	15q13.2-13.3	dup	15	1676615	30943512	32620127
Case 75	15q13.2-13.3	dup	15	1918508	30936285	32854793
Case 76	15q13.3	del	15	1587716	30927210	32514926
Case 77	15q25	del	15	653406	85075428	85728834
Case 78	Prader-Willi	del	15	4851483	23683783	28535266
Case 79	Prader-Willi	del	15	4914407	23620859	28535266
Case 80	PWS/AS	dup	15	5592496	22751742	28344238
Case 81	PWS/AS	dup	15	4685929	23683783	28369712
Case 82	16p11.2	del	16	547091	29652488	30199579

Case	Locus/syndrome	Chr	Size	CNV	Position in Mb	
					Start(hg19)	End(hg19)
Case 83	16p11.2	del	16	555405	29644174	30199579
Case 84	16p11.2	del	16	563133	29652488	30215621
Case 85	16p11.2	del	16	217401	28826049	29043450
Case 86	16p11.2	del	16	217845	28825605	29043450
Case 87	16p11.2	del	16	155044	28855727	29010771
Case 88	16p11.2	del	16	160482	28837515	28997997
Case 89	16p11.2	del	16	205935	28837515	29043450
Case 90	16p11.2	del	16	205935	28837515	29043450
Case 91	16p11.2	dup	16	217401	28826049	29043450
Case 92	16p11.2	dup	16	217401	28826049	29043450
Case 93	16p11.2	dup	16	185805	28857645	29043450
Case 94	16p11.2	dup	16	205935	28837515	29043450
Case 95	16p11.2	dup	16	205935	28837515	29043450
Case 96	16p11.2	dup	16	217401	28826049	29043450
Case 97	16p11.2	dup	16	521148	29653836	30174984
Case 98	16p11.2	dup	16	338061	29595483	29933544
Case 99	16p11.2	dup	16	485393	29692414	30177807
Case 100	16p11.2	dup	16	530465	29647342	30177807
Case 101	16p12.1	del	16	474576	21949122	22423698
Case 102	16p12.1	del	16	474576	21949122	22423698
Case 103	16p12.1	del	16	596471	21839340	22435811

Case	Locus/syndrome	Chr	Size	CNV	Position in Mb	
					Start(hg19)	End(hg19)
Case 104	16p12.1	del	16	474576	21949122	22423698
Case 105	16p12.1	del	16	486689	21949122	22435811
Case 106	16p12.1	del	16	474576	21949122	22423698
Case 107	16p13.11	del	16	1165658	15125441	16291099
Case 108	16p13.11	del	16	825476	15479879	16305355
Case 109	16p13.11	dup	16	661948	15593292	16255240
Case 110	16p13.11	dup	16	1280143	15383857	16664000
Case 111	16p13.11	dup	16	1210610	15092778	16303388
Case 112	16p13.11	dup	16	1164091	15125441	16289532
Case 113	16p13.11	dup	16	1164091	15125441	16289532
Case 114	16p13.11	dup	16	1164091	15125441	16289532
Case 115	16p13.11	dup	16	811220	15479879	16291099
Case 116	16p13.11	dup	16	1118688	15125441	16244129
Case 117	16p13.11	dup	16	1164091	15125441	16289532
Case 118	16p13.11	dup	16	809653	15479879	16289532
Case 119	16p13.11	dup	16	698836	15593292	16292128
Case 120	16p13.11	dup	16	1164091	15125441	16289532
Case 121	16p13.11	dup	16	809653	15479879	16289532
Case 122	17q11.2	del	17	1017662	29107708	30125370
Case 123	17q12	dup	17	1396108	34815551	36211659
Case 124	17q12	dup	17	1424585	34824845	36249430

Case	Locus/syndrome	Chr	Size	CNV	Position in Mb	
					Start(hg19)	End(hg19)
Case 125	17q12	dup	17	1430217	34815551	36245768
Case 126	17q12	dup	17	1500928	34815551	36316479
Case 127	17q12	dup	17	1433879	34815551	36249430
Case 128	17q12	dup	17	1379321	34815551	36194872
Case 129	22q11.2	del	22	2519388	18945091	21464479
Case 130	22q11.2	del	22	2574989	18889490	21464479
Case 131	22q11.2	del	22	2584566	18877787	21462353
Case 132	22q11.2	del	22	2523986	18938367	21462353
Case 133	22q11.2	del	22	2584566	18877787	21462353
Case 134	22q11.2	del	22	2742984	18718623	21461607
Case 135	22q11.2	del	22	2584566	18877787	21462353
Case 136	22q11.2	del	22	2584566	18877787	21462353
Case 137	22q11.2	del	22	1292734	20170996	21463730
Case 138	22q11.2	del	22	2584566	18877787	21462353
Case 139	22q11.2	del	22	2584566	18877787	21462353
Case 140	22q11.2	del	22	2584566	18877787	21462353
Case 141	22q11.2	del	22	1437703	18874965	20312668
Case 142	22q11.2	del	22	2584566	18877787	21462353
Case 143	22q11.2	del	22	2545597	18916756	21462353
Case 144	22q11.2	dup	22	2574685	18889794	21464479
Case 145	22q11.2	dup	22	2574685	18889794	21464479

Case	Locus/syndrome	Chr	Size	CNV	Position in Mb	
					Start(hg19)	End(hg19)
Case 146	22q11.2	dup	22	2586949	18877530	21464479
Case 147	22q11.2	dup	22	2974458	18738296	21712754
Case 148	22q11.2	dup	22	2576815	18886915	21463730
Case 149	22q11.2	dup	22	2647461	18877787	21525248
Case 150	22q11.2	dup	22	2587993	18877787	21465780
Case 151	22q11.2	dup	22	2584566	18877787	21462353
Case 152	22q11.2	dup	22	2582709	18881770	21464479
Case 153	distal/22q11.2	dup	22	1057116	21914652	22971768
Case 154	distal/22q11.2	dup	22	3065984	21964761	25030745
Case 155	22q11.2	del	22	1061218	21899201	22960419
Case 156	22q11.2	del	22	1032093	21939675	22971768
Case 157	Phelan-McDermid syndrome	del	22	73473	51078251	51151724
Case 158	Phelan-McDermid syndrome	del	22	90742	51087348	51178090

Note: Copy number variation positions are in UCSC build hg19. Del, deletion; dup, duplication. TAR, thrombocytopenia-absent radius; BP, breakpoint; VCFS, Velocardiofacial syndrome.

Table 7-9. Results of multicollinearity tests for internalising disorders model.

Variable	Variance Inflation Factor (VIF)	Tolerance
ND CNV status	1.00	1.00
Sex	1.02	0.98
Age	1.02	0.98
WRAT Std Score	1.01	0.99
ASD Status	1.03	0.97

Table 7-10. Results of multicollinearity tests for externalising disorders model.

Variable	Variance Inflation Factor (VIF)	Tolerance
ND CNV status	1.00	1.00
Sex	1.02	0.98
Age	1.03	0.97
WRAT Std Score	1.02	0.98
ASD Status	1.02	0.98

Table 7-11. Results of multicollinearity test for subclinical psychotic symptoms model.

Variable	Variance Inflation Factor (VIF)	Tolerance
ND CNV status	1.00	1.00
Sex	1.03	0.97
Age	1.04	0.96
WRAT Std Score	1.03	0.98
ASD Status	1.03	0.97

Table 7-12. Results of multicollinearity test for suicidal ideation model.

Variable	Variance Inflation Factor (VIF)	Tolerance
ND CNV status	1.00	1.00
Sex	1.03	0.97
Age	1.05	0.95
WRAT Std Score	1.03	0.97
ASD Status	1.05	0.96

Table 7-13. Characteristics and relevant variables in the PNC sub-sample excluding individuals with ASD and likely ID.

Variable	PNC sub-sample (n=7736)
Sex, M (%)	3702 (48.0)
Age in years, median (IQR)	14.0 (11.0-17.0)
WRAT std score, median (IQR)	102.0 (93.0-112.0)

Note: IQR, interquartile range; WRAT std score, wide ranging achievement test standardised score.

Table 7-14. Comparison of characteristics of ND CNV carriers and non-carriers in PNC sample with individuals ASD and likely ID excluded.

	ND CNV status		df	Test Statistic	p
	ND CNV carriers (n = 142)	**Non CNV carriers (n = 7594)**			
Sex, M (%)	75 (53.5%)	3627 (47.9%)	1	1.55[a]	0.21
Age, median (IQR)	14.0 (10.0-17.0)	14.0 (11.0-17.0)	-	533000[b]	0.82
WRAT std score, median (IQR)	96.0 (87.0-105.8)	102.0 (93.0-112.0)	-	649000[b]	$3.12e^{-05}$

Note: WRAT std score, Wide ranging achievement test standardized score; Sd, standard deviation; df, degrees of freedom. [a]Chi square. [b]Mann-Whitney U Test.

Table 7-15. Psychiatric phenotypic outcomes in ND CNV carriers and non-carriers in PNC (sample excluding individuals with likely ASD and ID).

	N with positive history/ N total[1] (%)	
Psychiatric Outcome	**ND CNV Carriers**	**Non-Carriers**
Subclinical psychotic symptoms	18/121 (14.9%)	523/6683 (7.8%)

Note: [1] N total refers to total in category with available phenotypic data.

Table 7-16. Power analyses for associations between psychopathological phenotypes and ND CNV status in Philadelphia Neurodevelopmental Cohort sample.

	PNC Sample			
Input parameter	**Internalising symptoms**	**Externalising symptoms**	**Subclinical psychotic symptoms**	**Suicidal ideation**
Tail(s)	Two	Two	Two	Two
Odds ratio	1.10	0.94	1.87	1.04
Pr(Y=1|X=1) H0	0.40	0.29	0.08	0.08
α err prob	0.05	0.05	0.05	0.05
Total sample size	8205	8205	8205	8205
R^2 other X	0.004	0.003	0.004	0.003
X distribution	Binomial	Binomial	Binomial	Binomial
X parm π	0.02	0.02	0.02	0.02
Output parameter				
Critical Z	1.96	-1.96	1.96	1.96
Power	0.09	0.06	0.74	0.06

*Note: Test family, z tests; Statistical test, Logistic regression; Type of power analysis, Post hoc. Power analyses completed using G*Power (Faul et al., 2009).*

Table 7-17 Copy number variants identified in individuals in discovery dataset.

Case	Locus	CNV	Chr	Position in Mb	
				Start	Stop
Case 1	1q21.1	Del	1	145932468	147831184
Case 2	1q21.1	Del	1	146330584	147825662
Case 3	1q21.1	Del	1	146717564	146970946
Case 4	NRXN1	Del	2	50870373	50971478
Case 5	15q11.2	Del	15	22751094	23487547
Case 6	15q11.2	Del	15	22770994	23236972
Case 7	15q11.2	Del	15	22770994	23236972
Case 8	15q13.3	Del	15	30755047	32489254
Case 9	17p12	Del	17	14098277	15475088
Case 10	WBS	Dup	7	72732833	74142105
Case 11	AS/PWS	Dup	15	24727231	25107593
Case 12	16p13.11	Dup	16	14929751	16396513
Case 13	16p13.11	Dup	16	15481934	18314234
Case 14	16p13.11	Dup	16	15481934	17296550
Case 15	16p13.11	Dup	16	15481934	16340441

Case	Locus	CNV	Chr	Position in Mb	
				Start	Stop
Case 16	16p11.2	Dup	16	29517711	30177808
Case 17	16p11.2	Dup	16	29517711	30306969
Case 18	16p11.2	Dup	16	29580034	30191908
Case 19	16p11.2	Dup	16	29595483	30017620

Note: Copy number variation positions are in UCSC Build 37. Del, deletion; dup, duplication; WBS, Williams-Beuren Syndrome; AS/PWS, Angelman/Prader-Willi syndrome.

Table 7-18. Penalised likelihood method of logistic regression applied for comparison with generalised linear model, to examine for effects of small sample bias.

Variable	β	95% CI	SE	Chi Square	P
Intercept	-4.66	-5.33,-4.10	0.31	Inf	$< 2e^{-16}$
History of developmental delay	**1.67**	**0.54,2.68**	**0.54**	**7.69**	**0.006**
Comorbid NDD	**1.85**	**0.45,3.00**	**0.64**	**6.18**	**0.013**
Specific learning disorder	**2.21**	**0.49,3.58**	**0.78**	**5.79**	**0.016**

Note: Significant variables highlighted in bold. NDD, neurodevelopmental disorder.

In view of the rarity of SCZ-associated CNVs in the samples, the risk of small-sample bias in the analysis was considered and a penalised likelihood method (Firth method) of logistic regression was applied to the discovery sample for comparison. The coefficients and significance values using the penalised likelihood model did not differ substantially from the generalised linear model Table 7-18.

Table 7-19. Multiple logistic regression model using two phenotypic variables to model SCZ-associated CNV carrier status in Irish data set.

Variable	β	SE	Wald statistic (Z)	P	OR (95% CI)
Intercept	-4.63	0.31	-14.89	$< 2e^{-16}$	0.01 (0.01-0.02)
Comorbid NDD	**1.76**	**0.67**	**2.63**	**0.009**	**5.81 (1.28-19.29)**
History of developmental delay	**1.72**	**0.55**	**3.13**	**0.002**	**5.60 (1.73-15.74)**

Note: Predictor coefficients were tested using Wald tests and confidence intervals obtained using the Wald method. Significant variables highlighted in bold. Nagelkerke pseudo R Square = 0.172). NDD, neurodevelopmental disorder.

Table 7-20. ROC curve results for modelling SCZ-associated CNV status using two phenotypic variables in the Irish dataset.

Cutoff Value	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	AUC (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)
0.05	47.1 (23.0-72.2)	90.1 (88.3-91.8)	68.6 (56.3-80.9)	6.7 (5.5-17.3)	99.1 (97.4-99.3)

Note: Optimal Cutoff Value, Sensitivity, Specificity, Area Under the Curve, and Predictive Values using two independent variables ("History of developmental delay", "Comorbid neurodevelopmental disorder") to model for SCZ-associated CNV status in Irish cohort (discovery dataset). CI, confidence interval; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value

Table 7-21. Power analyses for associations between SCZ-associated CNV status and phenotype variables in discovery cohort.

	Discovery sample							
Input parameter	**Early onset of symptoms**	**History of learning difficulties**	**Specific learning disorder**	**Remedial school support**	**Low educational attainment**	**History of developmental delay**	**Comorbid NDD**	**Family history of NDD**
Tail(s)	Two	Two	Two	Two	Two	Two	Two	Two
Proportion of positive phenotype in SCZ-associated CNV carriers (p2)	0.32	0.42	0.11	0.05	0.21	0.26	0.16	0.26
Proportion of positive phenotype in non CNV carriers (p1)	0.23	0.15	0.01	0.03	0.13	0.06	0.04	0.29
α error probability	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Sample size group 1	1196	1196	1196	1196	1196	1196	1196	1196
Sample size group 2	19	19	19	19	19	19	19	19
Output parameter								
Critical z	1.96	1.96	1.96	1.96	1.96	1.96	1.96	-1.96
Power	0.18	0.83	0.78	0.15	0.23	0.81	0.64	0.05

*Note: Test family, z tests; Statistical test, Proportions: Difference between two independent proportions; Type of power analysis, Post hoc. Power analyses completed using G*Power (Faul et al., 2009).*

7.2 Appendix 2: Supplementary Figures

Supplementary Figures- Chapter 2

Figure 7-1. Example of calculation of odds ratios comparing ND CNV effects in sexes

Compare: Female with ND CNV vs Male with ND CNV:

β_{CNV}: CNV Yes = 1, CNV No = 0

β_{sex}: Male = 1, Female = 0

Female with ND CNV:

Outcome = $\alpha + \beta_{CNV} + \beta_{sex} + \beta_{CNV*sex}$

= $\alpha + \beta_{CNV}(1) + \beta_{sex}(0) + \beta_{CNV*sex}(0)$

= $\alpha + \beta_{CNV}$

Male with ND CNV:

Outcome = $\alpha + \beta_{CNV} + \beta_{sex} + \beta_{CNV*sex}$

= $\alpha + \beta_{CNV}(1) + \beta_{sex}(1) + \beta_{CNV*sex}(1)$

= $\alpha + \beta_{CNV} + \beta_{sex} + \beta_{CNV*sex}$

Outcome: Female with ND CNV – Male with ND CNV

= $[\alpha + \beta_{CNV}] - [\alpha + \beta_{CNV} + \beta_{sex} + \beta_{CNV*sex}]$

= $-\beta_{sex} - \beta_{CNV*sex}$

Comparing Female with ND CNV vs Male with ND CNV

- Effect: $-\beta_{sex} - \beta_{CNV*sex}$ = 0.001 - (-1.687) = 1.686
  OR -> exp(1.686) = 5.39
- CI_upper: $-\beta_{sex} - \beta_{CNV*sex}$ = -(log(0.76)) – (log(0.04)) = -(-0.27) – (-3.22) = 3.49
  OR -> exp(3.49) = 32.79
- CI_lower: $-\beta_{sex} - \beta_{CNV*sex}$ = -(log(1.32))-(log(0.74)) = -(0.28) – (-0.30) = 0.02
  OR -> exp(0.02) = 1.02

Figure 7-2. Recruitment flow of Philadelphia Neurodevelopmental Cohort, from Calkins et al.(Calkins et al., 2015).

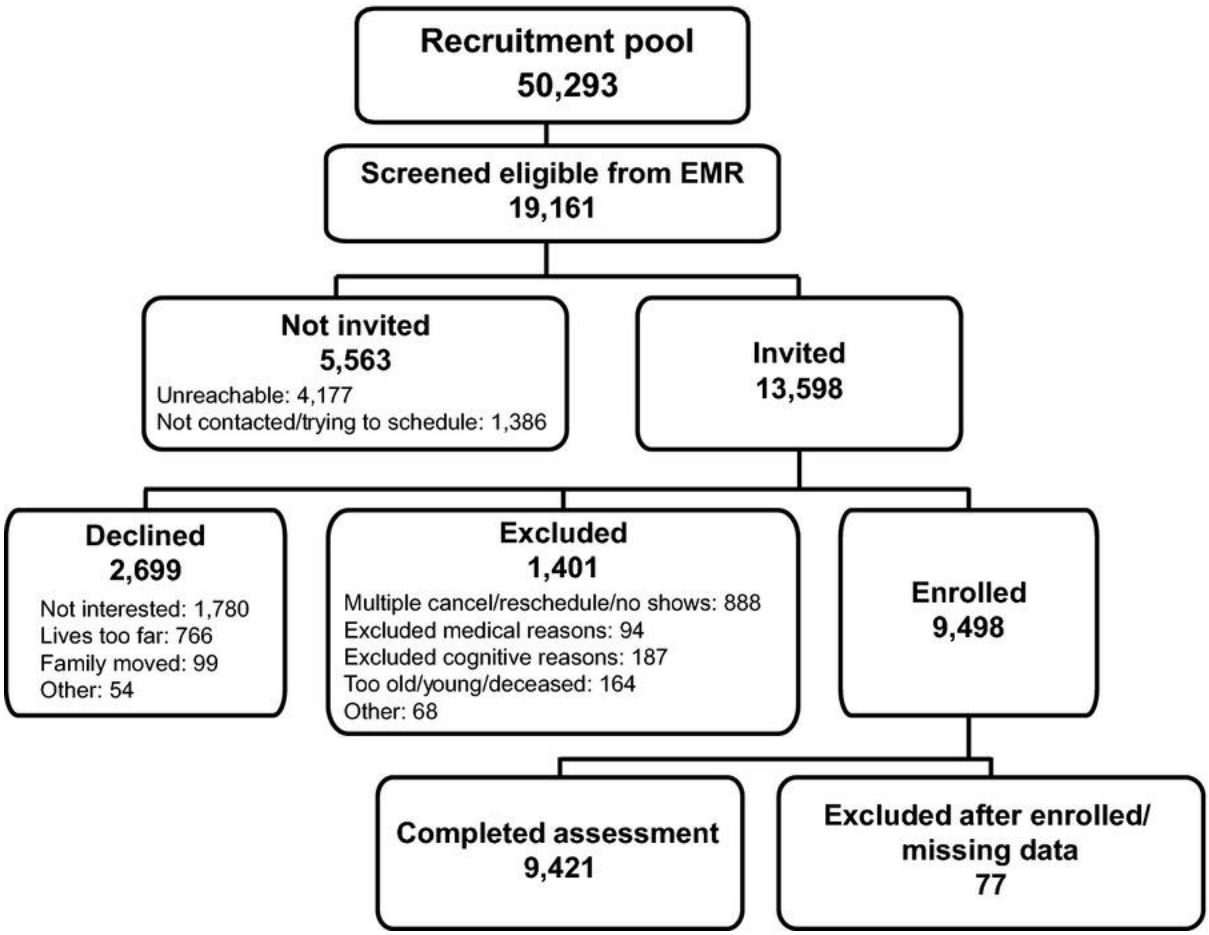


Figure 7-3. Sample GOASSESS proband screener section provided in supplementary material from Calkins et al. (Calkins et al., 2015).

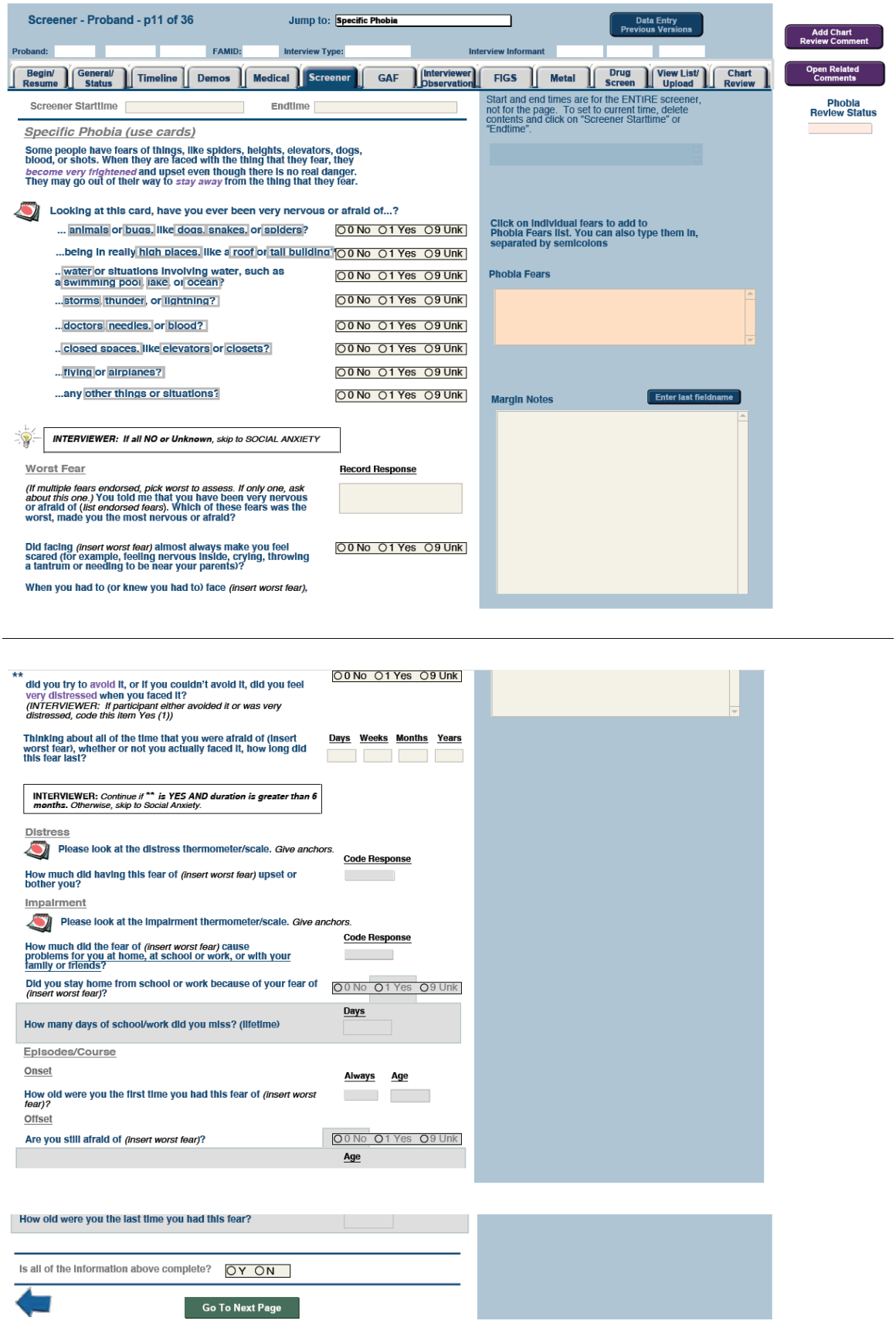
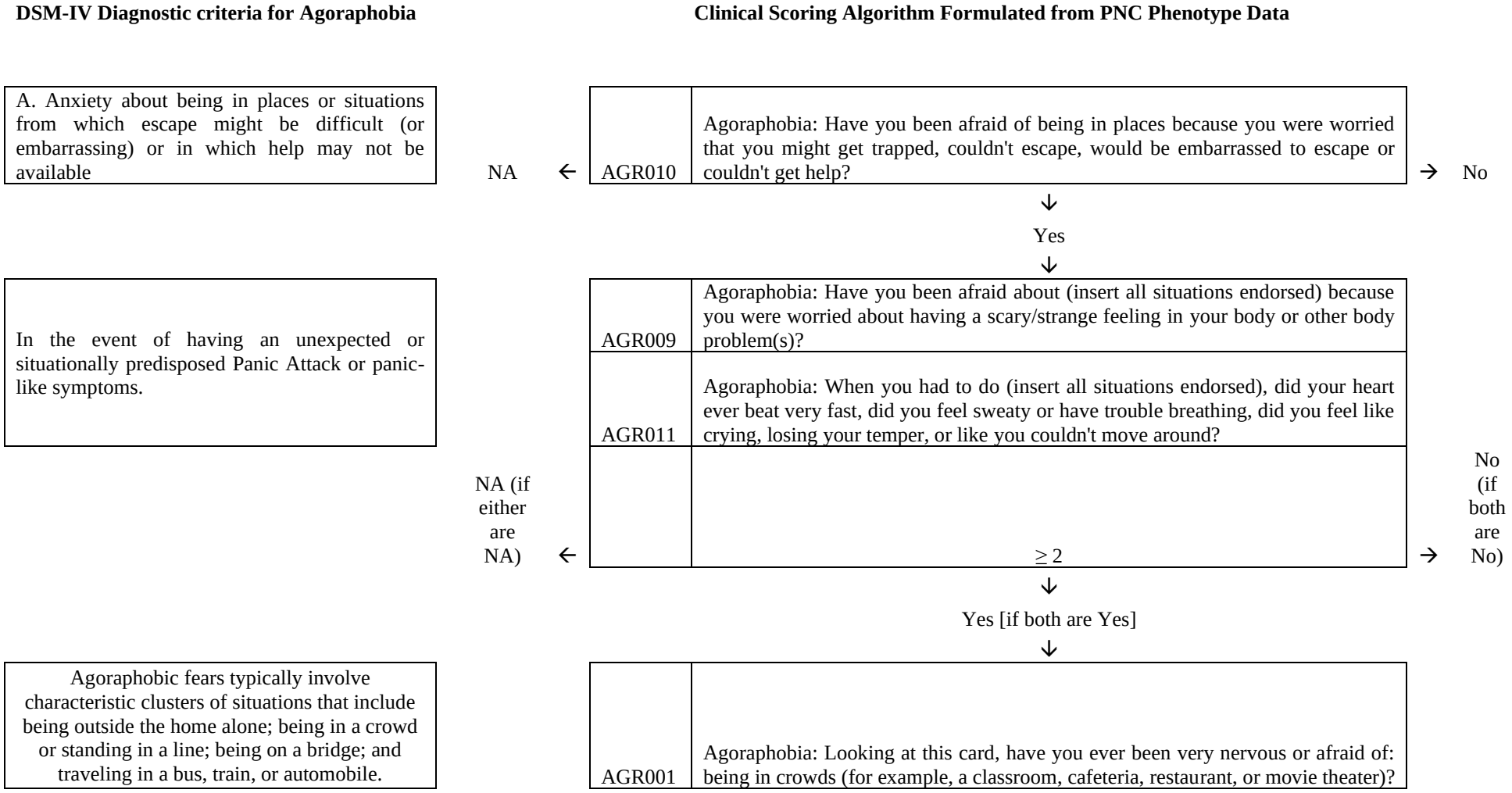


Figure 7-4. Clinical scoring algorithm formulated for scoring of PNC raw phenotype data for Agoraphobia based on DSM IV Diagnostic Criteria.



			AGR002	Agoraphobia: Looking at this card, have you ever been very nervous or afraid of: going to public places (such as a store or shopping mall)?		
			AGR003	Agoraphobia: Looking at this card, have you ever been very nervous or afraid of: being in an open field?		
			AGR004	Agoraphobia: Looking at this card, have you ever been very nervous or afraid of: going over bridges or through tunnels?		
			AGR005	Agoraphobia: Looking at this card, have you ever been very nervous or afraid of: traveling by yourself?		
			AGR006	Agoraphobia: Looking at this card, have you ever been very nervous or afraid of: traveling away from home?		
			AGR007	Agoraphobia: Looking at this card, have you ever been very nervous or afraid of: traveling in a car?		
			AGR008	Agoraphobia: Looking at this card, have you ever been very nervous or afraid of: using public transportation like a bus or SEPTA?		
Note: Consider the diagnosis of Specific Phobia if the avoidance is limited to one or only a few specific situations, or Social Phobia if the avoidance is limited to social situations.	NA	←		≥ 2	→	No
				↓		
				Yes		
				↓		
B. The situations are avoided (e.g., travel is restricted) or else are endured with marked distress or with anxiety about having a Panic Attack or panic-like symptoms or require the presence of a companion	NA	←	AGR013	Agoraphobia: When you had to face (feared situations), did you try to avoid it or if you couldn't avoid it, did you feel very distressed or need someone with you when you faced it?	→	No
				↓		
				Yes		
				↓		
			AGR014	Agoraphobia: How much did having this fear upset or bother you?		

		Or
	AGR015	Agoraphobia: How much did the fear you told me about cause problems for you at home, at school or work, or with your family or friends?
NA ←		≥ 5 → No

↓

Yes

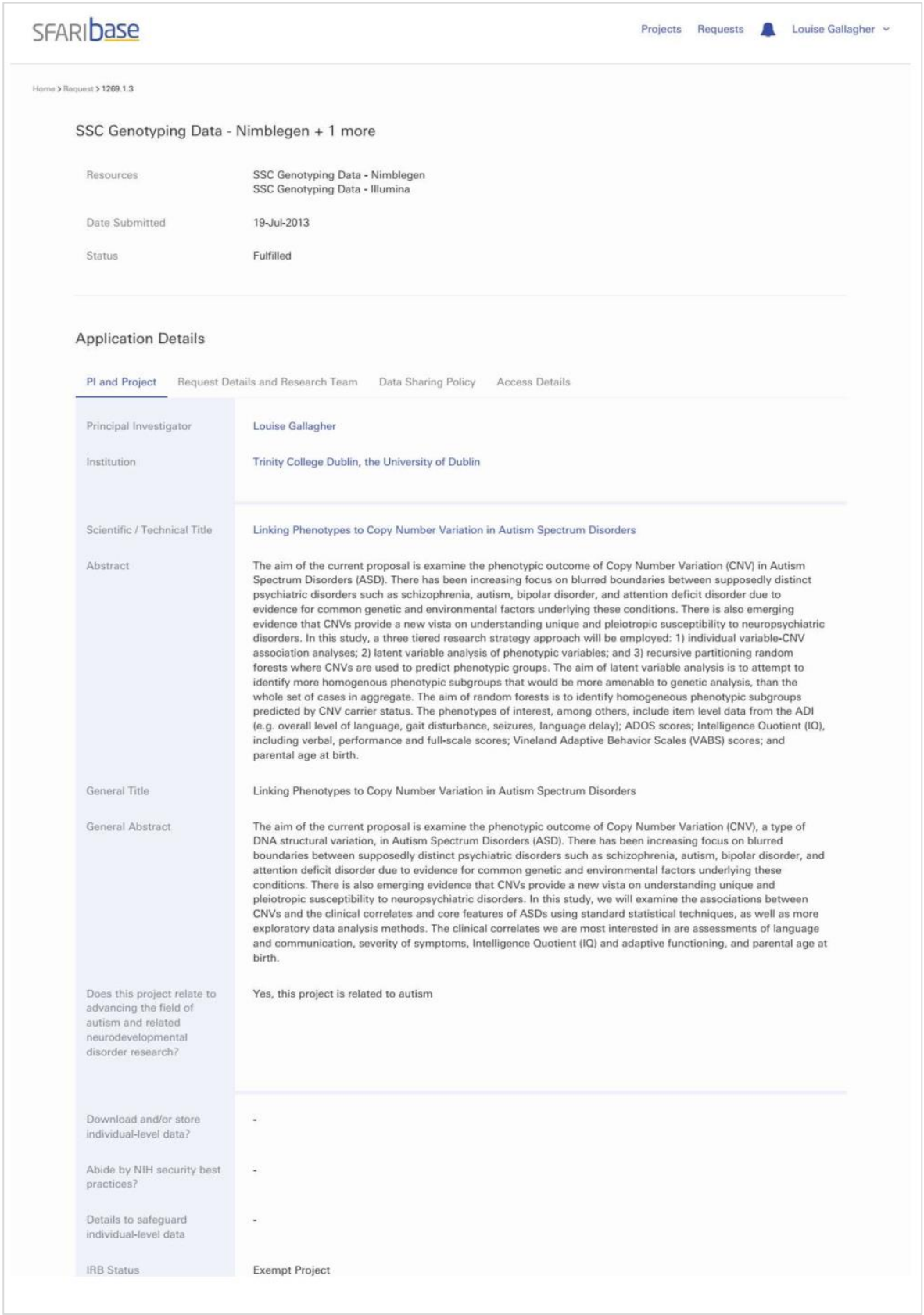
c. The anxiety or phobic avoidance is not better accounted for by another mental disorder, such as Social Phobia (e.g., avoidance limited to social situations because of fear of embarrassment), Specific Phobia (e.g., avoidance limited to a single situation like elevators), Obsessive Compulsive Disorder (e.g., avoidance of dirt in someone with an obsession about contamination), Posttraumatic Stress Disorder (e.g., avoidance of stimuli associated with a severe stressor), or Separation Anxiety Disorder (e.g., avoidance of leaving home or relatives).

↓

Suggestive of Agoraphobia Diagnosis based on DSM-IV Criteria

7.3 Appendix 3: Data Access Approval

Figure 7-5. Data Access Approval- Simons Simplex Collection



SFARIbase

Projects Requests Louise Gallagher

Home > Request > 1269.1.3

SSC Genotyping Data - Nimblegen + 1 more

Resources	SSC Genotyping Data - Nimblegen SSC Genotyping Data - Illumina
Date Submitted	19-Jul-2013
Status	Fulfilled

Application Details

PI and Project | Request Details and Research Team | Data Sharing Policy | Access Details

Principal Investigator	Louise Gallagher
Institution	Trinity College Dublin, the University of Dublin
Scientific / Technical Title	Linking Phenotypes to Copy Number Variation in Autism Spectrum Disorders
Abstract	The aim of the current proposal is examine the phenotypic outcome of Copy Number Variation (CNV) in Autism Spectrum Disorders (ASD). There has been increasing focus on blurred boundaries between supposedly distinct psychiatric disorders such as schizophrenia, autism, bipolar disorder, and attention deficit disorder due to evidence for common genetic and environmental factors underlying these conditions. There is also emerging evidence that CNVs provide a new vista on understanding unique and pleiotropic susceptibility to neuropsychiatric disorders. In this study, a three tiered research strategy approach will be employed: 1) individual variable-CNV association analyses; 2) latent variable analysis of phenotypic variables; and 3) recursive partitioning random forests where CNVs are used to predict phenotypic groups. The aim of latent variable analysis is to attempt to identify more homogenous phenotypic subgroups that would be more amenable to genetic analysis, than the whole set of cases in aggregate. The aim of random forests is to identify homogeneous phenotypic subgroups predicted by CNV carrier status. The phenotypes of interest, among others, include item level data from the ADI (e.g. overall level of language, gait disturbance, seizures, language delay); ADOS scores; Intelligence Quotient (IQ), including verbal, performance and full-scale scores; Vineland Adaptive Behavior Scales (VABS) scores; and parental age at birth.
General Title	Linking Phenotypes to Copy Number Variation in Autism Spectrum Disorders
General Abstract	The aim of the current proposal is examine the phenotypic outcome of Copy Number Variation (CNV), a type of DNA structural variation, in Autism Spectrum Disorders (ASD). There has been increasing focus on blurred boundaries between supposedly distinct psychiatric disorders such as schizophrenia, autism, bipolar disorder, and attention deficit disorder due to evidence for common genetic and environmental factors underlying these conditions. There is also emerging evidence that CNVs provide a new vista on understanding unique and pleiotropic susceptibility to neuropsychiatric disorders. In this study, we will examine the associations between CNVs and the clinical correlates and core features of ASDs using standard statistical techniques, as well as more exploratory data analysis methods. The clinical correlates we are most interested in are assessments of language and communication, severity of symptoms, Intelligence Quotient (IQ) and adaptive functioning, and parental age at birth.
Does this project relate to advancing the field of autism and related neurodevelopmental disorder research?	Yes, this project is related to autism
Download and/or store individual-level data?	-
Abide by NIH security best practices?	-
Details to safeguard individual-level data	-
IRB Status	Exempt Project

SFARIbase

Projects Requests Louise Gallagher

Home > Request > 1269.4.4

Simons Simplex Collection Dataset

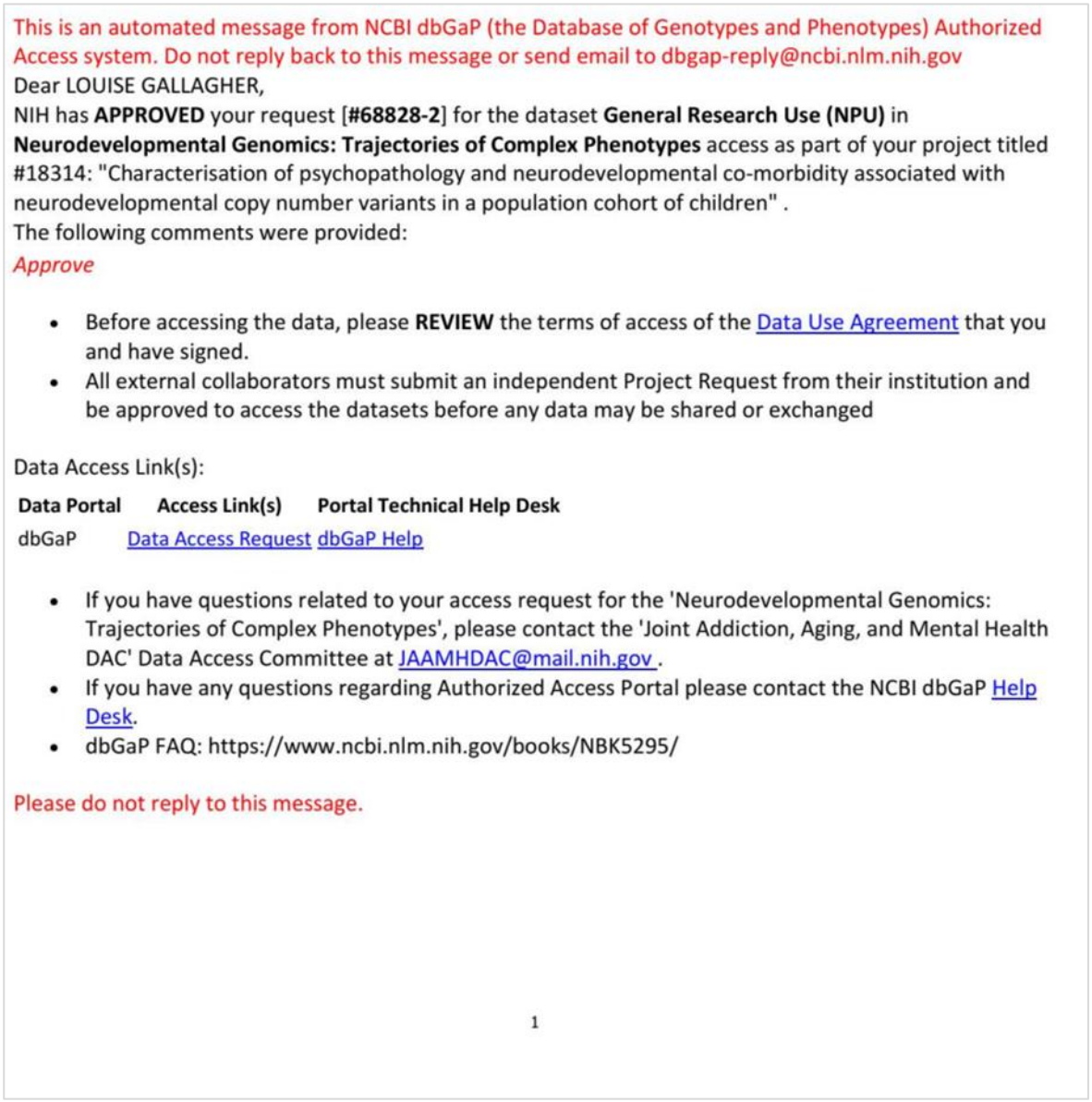
Resources	Simons Simplex Collection Dataset
Date Submitted	30-Apr-2020
Status	Approved

Application Details

PI and Project | Request Details and Research Team | Data Sharing Policy | Access Details

Principal Investigator	Louise Gallagher
Institution	Trinity College Dublin, the University of Dublin
Scientific / Technical Title	Characterisation of the psychopathology of neurodevelopmental copy number variant carriers in the SSC
Abstract	Neurodevelopmentally associated copy number variants are an important pathological contributor to neurodevelopmental and neuropsychiatric conditions such as autism spectrum disorder, intellectual disability and schizophrenia. Evidence from population studies of adults indicate that neurodevelopmental CNVs may also contribute to risk of developing depression and suicidal ideation. Childhood population studies to date have failed to identify significant associations between neurodevelopmental CNVs and psychopathologies such as depression, anxiety and subclinical psychotic symptoms. There is evidence from case control studies that neurodevelopmental CNV carriers may be at substantially higher risk of a range of psychopathologies than controls and that increased risk is not solely mediated by neurodevelopmental comorbidities autism spectrum disorder and intellectual disability. There is also evidence of the role of gender as a moderator of phenotypic effects of neurodevelopmental copy number variants. The first aim of this study is to examine the relationship between ND CNVs and risk of a range of psychopathologies presenting at clinical levels of impairment. We hypothesise that ND CNV carriers will have higher rates of clinical psychiatric symptomatology compared with non-carriers. The second aim is to examine the roles of ASD and intellectual/adaptive functioning in this relationship. We hypothesise that ND CNVs confer a distinct risk for psychopathology that is independent of intellectual disability or autism co-morbidity. The third aim of this study is to assess whether sex moderates the effect of ND CNVs on risk of psychopathology. We hypothesise that there will be sex differences in the risk of clinical psychiatric symptomatology in ND CNV carriers. The phenotypes of interest in both siblings and probands, include Child Behaviour Checklist (CBCL) Data, Vineland Adaptive Behaviour Scales (VABS), gender, age at clinical assessment.
General Title	Characterisation of the psychopathology of neurodevelopmental copy number variant carriers
General Abstract	Neurodevelopmentally associated copy number variants are an important pathological contributor to neurodevelopmental and neuropsychiatric conditions such as autism spectrum disorder, intellectual disability and schizophrenia. Evidence from population studies of adults indicate that neurodevelopmental CNVs may also contribute to risk of developing depression and suicidal ideation. Childhood population studies to date have failed to identify significant associations between neurodevelopmental CNVs and psychopathologies such as depression, anxiety and subclinical psychotic symptoms. There is evidence from case control studies that neurodevelopmental CNV carriers may be at substantially higher risk of a range of psychopathologies than controls and that increased risk is not solely mediated by neurodevelopmental comorbidities autism spectrum disorder and intellectual disability. There is also evidence of the role of gender as a moderator of phenotypic effects of neurodevelopmental copy number variants. The first aim of this study is to examine the relationship between ND CNVs and risk of a range of psychopathologies presenting at clinical levels of impairment. We hypothesise that ND CNV carriers will have higher rates of clinical psychiatric symptomatology compared with non-carriers. The second aim is to examine the roles of ASD and intellectual/adaptive functioning in this relationship. We hypothesise that ND CNVs confer a distinct risk for psychopathology that is independent of intellectual disability or autism co-morbidity. The third aim of this study is to assess whether sex moderates the effect of ND CNVs on risk of psychopathology. We hypothesise that there will be sex differences in the risk of clinical psychiatric symptomatology in ND CNV carriers. The phenotypes of interest in both siblings and probands, include Child Behaviour Checklist (CBCL) Data, Vineland Adaptive Behaviour Scales (VABS), gender, age at clinical assessment.
Does this project relate to advancing the field of autism and related neurodevelopmental disorder research?	Yes, this project is related to autism

Figure 7-6. Data Access Approval- Philadelphia Neurodevelopmental Cohort at dbGaP (weblink below).



This is an automated message from NCBI dbGaP (the Database of Genotypes and Phenotypes) Authorized Access system. Do not reply back to this message or send email to dbgap-reply@ncbi.nlm.nih.gov

Dear LOUISE GALLAGHER,

NIH has **APPROVED** your request [**#68828-2**] for the dataset **General Research Use (NPU)** in **Neurodevelopmental Genomics: Trajectories of Complex Phenotypes** access as part of your project titled #18314: "Characterisation of psychopathology and neurodevelopmental co-morbidity associated with neurodevelopmental copy number variants in a population cohort of children" .

The following comments were provided:

Approve

- Before accessing the data, please **REVIEW** the terms of access of the Data Use Agreement that you and have signed.
- All external collaborators must submit an independent Project Request from their institution and be approved to access the datasets before any data may be shared or exchanged

Data Access Link(s):

Data Portal	Access Link(s)	Portal Technical Help Desk
dbGaP	Data Access Request	dbGaP Help

- If you have questions related to your access request for the 'Neurodevelopmental Genomics: Trajectories of Complex Phenotypes', please contact the 'Joint Addiction, Aging, and Mental Health DAC' Data Access Committee at JAAMHDAC@mail.nih.gov .
- If you have any questions regarding Authorized Access Portal please contact the NCBI dbGaP Help Desk.
- dbGaP FAQ: https://www.ncbi.nlm.nih.gov/books/NBK5295/

Please do not reply to this message.

1

Weblink:

https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v3.p2#authorized-requests-section

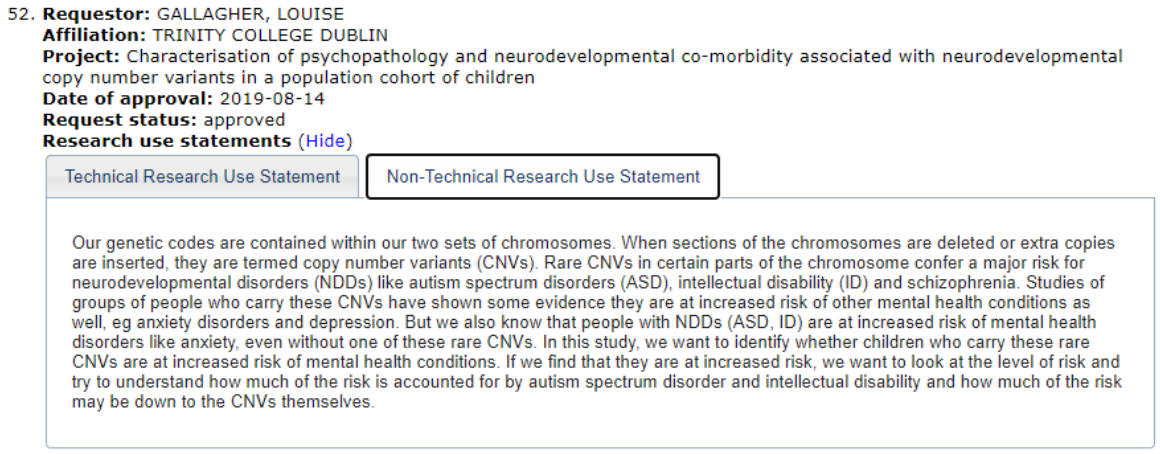
Figure 7-7. Philadelphia Neurodevelopmental Cohort at dbGaP, Data Access Approval Technical Research Use Statement.

52. **Requestor:** GALLAGHER, LOUISE
Affiliation: TRINITY COLLEGE DUBLIN
Project: Characterisation of psychopathology and neurodevelopmental co-morbidity associated with neurodevelopmental copy number variants in a population cohort of children
Date of approval: 2019-08-14
Request status: approved
Research use statements (Hide)

Technical Research Use Statement | Non-Technical Research Use Statement

A number of recurrent copy number variants (CNVs) increase the risk for multiple neurodevelopmental disorders (NDDs): autism spectrum disorder (ASD), intellectual disability (ID), schizophrenia, and for more subtle neurocognitive deficits. Specific neurodevelopmental CNVs (ND CNVs) have also demonstrated increased risk for other psychiatric and behavioural disorders in carriers. Population studies examining risk of psychopathology in groups of ND CNV carriers have demonstrated conflicting results, with evidence of increased risk of depressive symptoms/suicidal ideation in adult control carriers of ND CNVs, but no association with anxiety/depression diagnoses in child carriers of ND CNVs demonstrated to date. An increased risk of psychopathology in ND CNV carriers may be expected, given that they are at higher risk of NDDs and that NDDs themselves confer an increased risk of psychiatric and behavioural disorders occurring co-morbidly. However, a number of studies have presented results indicating that increased risk of psychopathology in those with ND CNVs is not solely mediated by IQ level or ASD symptomatology. In this study, we aim to quantify and characterise psychiatric and behavioural disorder association with ND CNVs in a population cohort of children. We aim to assess the role of ID and ASD comorbidity within this context. We hypothesise that ND CNV carriers will have higher rates of psychopathologies compared with non-carriers. We hypothesise that ND CNVs confer a distinct risk for psychopathology that is independent of ID or ASD co-morbidity. Using phenotypic data from GOASSESS (psychopathology measures, demographic, medical, GAF data), we will estimate the risk of psychopathologies in carriers of a set of specific ND CNVs. We have established a collaboration with Professor Hakonarson's group in the Children's Hospital of Philadelphia to identify the ND CNV carriers within the dataset. Professor Hakonarson's group have been added as external collaborators on this project. Professor Hakonarson's group submitted the original genetic data for this dbGaP dataset. The group are contributing processed genetic data to this research project directly, therefore we will not be using the raw genetic data in the dbGaP dataset and will use the phenotype data from the dbGaP dataset only. We will use multivariate regression models to examine ND CNV status, ASD-status and ID-status as independent variables contributing to risk of psychopathology. We continue to use de-identified data for the research purposes outline in the initial Research Use Statement in accordance with data use limitations outlined by dbGAP. Data is stored on secure servers, with only authorized personnel granted access.

Figure 7-8. Philadelphia Neurodevelopmental Cohort at dbGaP, Data Access Approval, Non-Technical Research Use Statement.



7.4 Appendix 4: Papers published during this work

BJPsych | The British Journal of Psychiatry (2020)
216, 275–279. doi: 10.1192/bjp.2019.262

Identifying schizophrenia patients who carry pathogenic genetic copy number variants using standard clinical assessment: retrospective cohort study

Claire Foley, Elizabeth A. Heron, Denise Harold, James Walters, Michael Owen, Michael O'Donovan, Jonathan Sebat, Eric Kelleher, Christina Mooney, Amy Durand, Carlos Pinto, Paul Cormican, Derek Morris, Gary Donohoe, Michael Gill, Louise Gallagher and Aiden Corvin

Background

Copy number variants (CNVs) play a significant role in disease pathogenesis in a small subset of individuals with schizophrenia (~2.5%). Chromosomal microarray testing is a first-tier genetic test for many neurodevelopmental disorders. Similar testing could be useful in schizophrenia.

Aims

To determine whether clinically identifiable phenotypic features could be used to successfully model schizophrenia-associated (SCZ-associated) CNV carrier status in a large schizophrenia cohort.

Method

Logistic regression and receiver operating characteristic (ROC) curves tested the accuracy of readily identifiable phenotypic features in modelling SCZ-associated CNV status in a discovery data-set of 1215 individuals with psychosis. A replication analysis was undertaken in a second psychosis data-set ($n = 479$).

Results

In the discovery cohort, specific learning disorder (OR = 8.12; 95% CI 1.16–34.88, $P = 0.012$), developmental delay (OR = 5.19; 95% CI 1.58–14.76, $P = 0.003$) and comorbid neurodevelopmental disorder (OR = 5.87; 95% CI 1.28–19.69, $P = 0.009$) were significant independent variables in modelling positive carrier status for a SCZ-associated CNV, with an area under the ROC (AUROC) of 74.2% (95% CI 61.9–86.4%). A model constructed from the discovery cohort including developmental delay and comorbid neurodevelopmental disorder variables resulted in an AUROC of 83% (95% CI 52.0–100.0%) for the replication cohort.

Conclusions

These findings suggest that careful clinical history taking to document specific neurodevelopmental features may be informative in screening for individuals with schizophrenia who are at higher risk of carrying known SCZ-associated CNVs. Identification of genomic disorders in these individuals is likely to have clinical benefits similar to those demonstrated for other neurodevelopmental disorders.

Declaration of interest

None.

Keywords

Genetics; schizophrenia; developmental disorders; autistic spectrum disorders; intellectual disability.

Schizophrenia is a clinically heterogeneous syndrome with substantial heritability. Common small genetic risk factors (polygenic risk) collectively account for some 30% of heritability.[1] Copy number variants (CNVs) (DNA segments of >1 kilobase, present at higher (duplication) or lower (deletion) number than in a reference genome) have a larger role to play in a small subset of cases. There is evidence of genome-wide significant association for at least eight CNV loci in schizophrenia.[2] Individually these events are of moderate penetrance for schizophrenia, with reported odds ratios (ORs) of 2–30, which indicates that they are likely to have a substantial role in disease aetiology, at least for a small group of patients.[3] Almost all of these CNVs are pleiotropic; for carriers of any one of the schizophrenia-associated (SCZ-associated) CNVs, the risk of developing *any* early developmental disorder (e.g. intellectual disability, autism spectrum disorder, developmental delay) is significantly higher than the risk of developing schizophrenia itself.[3] Even in the absence of a psychiatric diagnosis, carriers of CNVs associated with schizophrenia have significant but variable cognitive deficits.[4] These CNVs are therefore potentially pathogenic and clinically significant, but outcomes range from subtle cognitive effects to severe neurodevelopmental disorders. Despite significant progress in our understanding of the genetics of schizophrenia, the process of translating SCZ genetic discovery into clinical impact is in its infancy.

Chromosomal microarray (CMA) testing is recommended as a first-tier genetic test in autism, developmental delay and intellectual disability.[5] Because a smaller proportion of people with schizophrenia (~2.5%) carry known pathogenic CNVs, routine testing is not currently recommended. However, a genetic diagnosis may be empowering for patients and their families, it can inform screening for relevant medical comorbidities and help in reproductive planning.[6] The identification of clinical symptoms or demographic features that differentiate people with schizophrenia who carry SCZ-associated CNVs may be helpful in clarifying who might benefit most from testing. On the basis of the known overlap with other neurodevelopmental disorders and previously reported phenotype studies[3,7–15] we hypothesised that individuals with schizophrenia who carry SCZ-associated CNVs are more likely to have phenotypic features suggestive of pre-existing neurodevelopmental compromise, earlier onset of psychotic symptoms or a positive family history of neurodevelopmental disorder. The objective of this work was to determine whether clinically identifiable phenotypic features could be used to model SCZ-associated CNV carrier status in a large schizophrenia cohort.

Method

Selection of phenotypic variables

A literature review was conducted in PubMed from January 2008 to February 2016 to identify clinical and phenotypic features reported

to be associated with copy number variation in schizophrenia using the search terms 'schizophrenia', 'copy number variant' and 'phenotype'. Publications that specifically described CNV-associated clinical and phenotypic features in schizophrenia were selected to identify neurodevelopmental phenotypic categories. Identified phenotypic domains included early onset of psychosis; premorbid cognitive difficulties; delays in developmental milestones; family history of neurodevelopmental disorder; and syndromal characteristics (dysmorphic features, congenital malformations). Eight specific features falling within these domains were identified through expert clinical consensus that are readily identifiable in a standard clinical evaluation and therefore ultimately of clinical utility and acceptability. Subsequently, 'dysmorphic features' and 'congenital anomalies' were excluded because reliable identification of these features requires additional training or clinical tools.[16] The phenotypic variables selected for analysis are outlined in Table 1.

Clinical sample

The discovery data-set

The discovery data-set consisted of 1215 individuals of Irish ancestry for whom both clinical phenotype and genome-wide SNP array data were available.[17] The individuals were all over 18 years of age and had a diagnosis of schizophrenia or schizoaffective disorder after a structured clinical assessment (as described by First *et al*[18]). Written informed consent was obtained from all participants. Diagnosis was made on the basis of the consensus lifetime best estimate method using all available information (interview, family or staff report, chart review) with DSM-IV criteria as per the Structured Clinical Interview for DSM-IV-TR Axis I Disorders, research version, patient edition (SCID-I/P). Each referral centre obtained local research ethics committee (REC) approval. There was a preponderance of males in this sample (64%).

Phenotypic data were collected retrospectively from an existing research cohort.[17] The phenotypic data were collected from the SCID-I/P and consisted of interview self-reports. The definitions applied to identify a positive history of the phenotypic variables are outlined in Table 1. Phenotypic data were coded as categorical variables (missing information is described in supplementary Table 1, available at https://doi.org/10.1192/bjp.2019.262).

The replication data-set

The replication data-set was obtained from 19 879 schizophrenia cases published by the Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC) cohorts (representing 40 cohorts excluding data on Irish individuals).[2] Contributors of the constituent data-sets were approached to request access to additional phenotypic data to replicate the discovery findings. Only one cohort (the Cardiff data-set) was identified with the requisite phenotype data and adequate sample size for replication (many of the well-phenotyped cohorts were small and consequently had no CNV carriers).

The Cardiff data-set ($n = 479$) consisted of participants from the previously reported Cardiff Cognition in Schizophrenia (CardiffCOGS) study.[19,20] In brief, the sample was recruited with REC approval from community, in-patient and voluntary-sector mental health services in the UK. Written informed consent was obtained from all participants. Participants had a clinical diagnosis of schizophrenia and were interviewed using the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) and case-note review to derive a best-estimate lifetime diagnosis according to DSM-IV criteria. Similar to the discovery set, there was a preponderance of males in the sample (61.2%). The comparable phenotype variables investigated in the Cardiff data-set were: (a) 'history of developmental delay', which was directly comparable to the Irish data-set variable and was defined as 'clinically relevant delays in speech, walking, coordination or diagnosed developmental problem' and (b) a positive history of epilepsy, intellectual disability and/or autism spectrum disorder, which was included as 'comorbid neurodevelopmental diagnosis'. Intellectual disability referred to an IQ <70 and clinical specialist service involvement. The autism spectrum disorder and epilepsy variables were interview self-report of a clinical diagnosis. Missing information is described in supplementary Table 2. The other phenotypic variables selected for in the initial analysis were not collected in this data-set.

Ethical approval

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involving human participants were approved by the relevant local research ethics committees in Ireland and the UK, as outlined above.

CNV list

The target CNVs used in the analysis were fifteen CNVs with the strongest evidence of association with schizophrenia (supplementary Table 3) analysed by Rees *et al*.[20] Twelve of these were also identified in the large PGC CNV meta-analysis[21] and the other three were exon-disrupting deletions at the *NRXN1* gene, deletion at distal 16p11.2 and duplications at the Williams–Beuren region identified on the basis of expert consensus or evidence published after the meta-analysis.[20]

Genotyping and CNV calling

The Irish sample was genotyped on the Affymetrix 6.0 array ($n = 802$) or the Illumina HumanCoreExome chip ($n = 413$) (full details are available in the literature[17]). The Cardiff sample was

Table 1 Univariate analyses assessing associations[a] between selected phenotypic variables and schizophrenia-associated copy number variant status

Phenotypic variable	Definition	OR (95% CI)	*P*
Early onset of symptoms[8,9]	Onset of symptoms at <18 years of age	1.55 (0.58–4.27)	0.41
History of learning difficulties[4,11]	Any difficulties with learning reported in school (excluding behavioural difficulties)	3.99 (1.55–10.30)	**0.005**
Specific learning disorder[4,11]	Identified as report of diagnosed dyslexia, dyscalculia or dysgraphia	9.03 (1.38–39.98)	**0.03**
Remedial school support[4,11]	Reported learning support at school, within class, separate classes or special school	1.76 (0.08–10.66)	0.45
Low educational attainment[4,11]	Attained primary school education only	1.84 (0.55–5.74)	0.29
History of developmental delay[3,13]	Delayed milestones: motor, speech, toilet training	5.76 (1.91–16.44)	**0.005**
Comorbid neurodevelopmental diagnosis[3,13]	Diagnosis of ASD, intellectual disability or epilepsy	4.93 (1.18–16.73)	**0.034**
Family history of neurodevelopmental disorder[14,15]	Reported diagnosis of schizophrenia, ASD, intellectual disability or epilepsy in a first- or second-degree family member	1.06 (0.35–3.28)	1

ASD, autism spectrum disorder.
a. Fisher's exact tests were used to assess associations. Significant results are in bold.

Table 2 Multiple logistic regression model to determine whether clinically identifiable phenotypic features could be used to model schizophrenia-associated copy number variant carrier status[a]

Variable	β	s.e.	Wald statistic, *Z*	*P*	OR (95% CI)
History of developmental delay	1.65	0.56	2.95	0.003	5.19 (1.58–14.76)
Comorbid neurodevelopmental disorder	1.77	0.67	2.63	0.009	5.87 (1.28–19.69)
Specific learning disorder	2.09	0.83	2.53	0.012	8.12 (1.16–34.88)

a. Predictor coefficients were tested using Wald tests and confidence intervals were obtained using the Wald method. Nagelkerke pseudo $R^2 = 0.196$.

Table 3 Receiver operating characteristic (ROC) curve results[a] for modelling schizophrenia-associated copy number variant status in the discovery (Irish) data-set

Cut-off value	Sensitivity, % (95% CI)	Specificity, % (95% CI)	AUC, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
0.04	58.8 (32.9–81.6)	89.1 (87.1–90.9)	74.2 (61.9–86.4)	7.5 (6.3–20.1)	99.3 (98.0–99.4)

AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.
a. Optimal cut-off value, sensitivity, specificity, AUC and predictive values were calculated using three independent variables ('history of developmental delay', 'comorbid neurodevelopmental disorder', 'specific learning disorder').

genotyped using HumanOmniExpress-12v1-1_B arrays (Illumina).[22] To control for platform effects, raw intensity data were provided to the PGC CNV analysis group. This provided a centralised pipeline for systematic CNV calling including multiple CNV callers run in parallel. The final CNV set was defined as those >20 kb in length and including at least 10 probes and <1% minor allele frequency (MAF).[2]

Statistical analyses

Univariate analyses (Fisher's exact tests) were performed first, to assess associations between phenotypic predictors and SCZ-associated CNV status in the Irish cohort. Multiple logistic regression analysis was then carried out to examine the effects of significant phenotypic variables, identified on univariate analysis, in modelling SCZ-associated CNV status. The final independent variables included in the model were those with a significance level of 0.05 following backward elimination steps. Model fit was assessed using Nagelkerke pseudo R^2 index.

Receiver operating characteristic (ROC) curve analysis was used to test the validity, sensitivity and specificity of the logistic regression parameters for modelling SCZ-associated CNV carrier status in the Irish discovery data-set.

The Cardiff replication data-set included data on two of the phenotypic variables of interest. A multiple logistic regression model including these two variables was trained from the Irish discovery data-set and then applied to the Cardiff data-set. ROC curve analysis was used to assess the accuracy of the neurodevelopmental variables in modelling SCZ-associated CNV carrier status in the replication data-set.

The results presented are not corrected for multiple comparisons and all analyses were completed in R version 3.2.3 for Windows.[23]

Results

Discovery data-set

From the total sample of 1215 individuals in the Irish discovery data-set, 19 (1.6%) carried one of the 15 identified SCZ-associated CNVs.[20] No individuals carried more than one SCZ-associated CNV. The details of the CNVs and positions are listed in supplementary Table 4. The proportions of individuals with a positive history of phenotypic variables and SCZ-associated pathogenic CNV status are available in supplementary Table 5.

Univariate analyses identified four phenotypic variables with significant associations with SCZ-associated CNV status: 'history of developmental delay', 'comorbid neurodevelopmental disorder', 'history of learning difficulties' and 'specific learning disorder' (Table 1). A multiple logistic regression model was fitted using these four variables. The variables 'history of learning difficulties' and 'specific learning disorder' were correlated (phi coefficient $\varphi = 0.22$) and were likely capturing similar phenotypic information. Backward elimination at this point removed the variable 'history of learning difficulties' from the model. The final independent variables in the model were 'history of developmental delay', 'comorbid neurodevelopmental disorder' and 'specific learning disorder'. These variables had odds ratios of 5.19 (95% CI 1.58–14.76, $P = 0.003$), 5.87 (95% CI 1.28–19.69, $P = 0.009$) and 8.12 (95% CI 1.16–34.88, $P = 0.012$) respectively when included in the logistic regression model (Table 2). Nagelkerke pseudo R^2 for the model was 0.196, indicating that the phenotypic variables accounted for 19.6% of the variance in SCZ-associated CNV status in this sample.

The performance of the three significant independent variables in modelling SCZ-associated CNV carrier status was tested using ROC curve analysis. An area under the ROC (AUROC) curve of 74.2% (95% CI 61.9–86.4%) was achieved, accounting for 58.8% (95% CI 32.9–81.6%) sensitivity and 89.1% (95% CI 87.1–90.9%) specificity in modelling SCZ-associated CNV carrier status (Table 3).

Replication data-set

Eight individuals (1.7%) in the Cardiff replication data-set ($n = 479$) set carried one of the 15 identified risk CNVs, including one 1q21.2 duplication, one *NRXN1* deletion, one Williams–Beuren region duplication, three 15q11.2 deletions and two 22q11.2 deletions. No individual carried more than one of these CNVs.

The Cardiff replication data-set included data on two of the phenotypic variables of interest: 'history of developmental delay' and 'comorbid neurodevelopmental disorder'. The Irish discovery data-set was used to build a multiple logistic regression model using these two variables (supplementary Tables 6 and 7). Applying this model to the Cardiff study population gave an AUROC of 83% (95% CI 52.0–100.0%) in identifying SCZ-associated CNV status. The sensitivity and specificity were 75.0% (95% CI 19.4–99.4%) and 97.6% (95% CI 95.1–99.0%) respectively (Table 4).

Discussion

We investigated whether phenotype information generated by a standard clinical assessment could identify people with schizophrenia at greater risk of carrying pathogenic CNVs. In a discovery cohort of 1215 people with schizophrenia, having a specific learning disorder (OR = 8.12, $P = 0.012$), developmental delay (OR = 5.19, $P = 0.003$) or a comorbid neurodevelopmental disorder (OR =

Table 4 Receiver operating characteristic (ROC) curve results[a] for modelling schizophrenia-associated copy number variant status in the replication (Cardiff) data-set

Cut-off value	Sensitivity, % (95% CI)	Specificity, % (95% CI)	AUC, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
0.241	75.0 (19.4–99.4)	97.6 (95.1–99.0)	83.0 (52.0–100)	30.0 (17.0–95.7)	99.6 (95.8–99.9)

AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.
a. Optimal cut-off value, sensitivity, specificity, AUC and predictive values were calculated using two independent variables ('history of developmental delay', 'comorbid neurodevelopmental disorder').

5.87, $P = 0.009$) successfully modelled positive carrier status for a SCZ-associated CNV. Other clinical features, such as early onset of psychosis, low educational attainment and a family history of neurodevelopmental disorders, were not associated with SCZ-associated CNV carrier status in this cohort. The three 'neurodevelopmental' variables showed a relatively high specificity (89.1% (95% CI 87.1–90.9%)) but more modest sensitivity (58.8% (95% CI 32.9–81.6%)) in modelling carrier status for a SCZ-associated CNV in the Irish discovery sample. Information on 'specific learning disorders' was not available for the Cardiff replication sample. On the basis of the remaining two variables, 'comorbid neurodevelopmental disorder' and 'history of developmental delay', we applied a model from the original data-set to the Cardiff sample. This too showed relatively high specificity (97.6% (95% CI 95.1–99.0%)) but more modest sensitivity (75.0% (95% CI 19.4–99.4%)) in modelling carrier status for a SCZ-associated CNV.

Recent studies have suggested that identifying people with schizophrenia who have comorbid intellectual disability is likely to be helpful in identifying subsets of individuals with genomic disorders. Thygesen and colleagues reported an approximately threefold higher rate of pathogenic CNVs in people with psychosis and intellectual disability compared with rates in the general schizophrenia population.[24] Lowther *et al* examined the genome-wide burden of pathogenic CNVs in a schizophrenia cohort ($n = 546$) and demonstrated a significantly higher burden of pathogenic CNVs (OR = 5.01, $P = 0.0001$) in people with schizophrenia and low IQ (IQ <85) compared with those with average IQ (IQ ≥85). On the basis of their findings, the authors concluded that individuals with schizophrenia and low IQ should be prioritised for clinical microarray testing in clinical and research contexts.[25] We believe that our study provides further support for this recommendation, but that other developmental indices, which could be captured by a clinical neurodevelopmental history, should also be considered in the development of any future guidelines.

Implications

A small subset of people with schizophrenia (~2.5%) carry CNVs that substantially increase the risk for schizophrenia but also for other neurodevelopmental disorders. The clinical benefits of identifying such people been demonstrated for other neurodevelopmental disorders.[5,6] Similar benefits are likely to apply in schizophrenia, but as these events are rare, routine genetic testing for all individuals is probably not indicated. Previous studies suggest that targeting people with schizophrenia and comorbid intellectual disability is likely to be more fruitful in identifying such cases.[24,25] Our findings suggest that careful clinical history taking to document developmental delay, reported learning disorders or a comorbid diagnosis of autism spectrum disorder or epilepsy may also be informative in screening for people with schizophrenia at higher risk of carrying known SCZ-associated CNVs.

These are rare events, but very large cohorts of genotyped people with schizophrenia are available and it is likely that whole genome sequence analysis of >30 000 such individuals will soon be completed. As these data are analysed the subset of people with schizophrenia who carry rare mutations and CNVs of likely clinical significance will increase, as has been the case for other neurodevelopmental disorders. Regrettably, there is a dearth of phenotype information available from many of the contributory cohorts. We strongly support efforts by the PGC to collect and standardise such phenotype information where it is available. For future cohorts, having detailed phenotype information together with neurodevelopmental and medical history will likely be helpful in refining predictor variables that ultimately may inform guidelines for genetic testing for people with schizophrenia.

Strengths and limitations

The strength of our study lies in the fact that we were able to build a well-characterised phenotype data-set, based on extensive clinical and research data compiled from previous schizophrenia research studies. We were able to test multiple phenotypic features for potential in identifying pathogenic CNV status and identify three variables that are easily clinically identified and that show considerable promise in identifying a high-risk group.

Recurrent SCZ-associated CNVs are rare events (~1:150–1:1000)[3] and individual cohorts are likely to identify only a modest number of known CNVs, as demonstrated in our sample of 1215 people with schizophrenia. The study highlighted the relative limitations of phenotypic information across schizophrenia cohorts and suggested phenotypes derived from a standard clinical interview that could inform future studies. Further analysis of a wider psychosis population and other cross-disorder analyses are also likely to be valuable.

Our discovery and replication cohorts used retrospective phenotypic data from which we identified variables that provided estimates of sensitivity and specificity for modelling SCZ-associated CNV carrier status. Significantly larger, well-characterised phenotypic samples (e.g. prospective cohorts) will be required to provide more refined estimates of sensitivity and specificity to inform genetic screening guidelines. It will be important to consider the patient and family perspective to inform any future guidelines for genetic testing, but that was beyond the scope of the current investigation.

Claire Foley, MB, MRCPsych, Clinical Research Fellow, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Elizabeth A. Heron**, PhD, Assistant Professor, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Denise Harold**, PhD, Assistant Professor, School of Biotechnology, Dublin City University, Ireland; **James Walters**, MRCPsych, PhD, Professor, MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, UK; **Michael Owen**, FRCPsych, PhD, Director, MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, UK; **Michael O'Donovan**, FRCPsych, PhD, Professor, MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, UK; **Jonathan Sebat**, PhD, Chief, Beyster Center for Genomics of Psychiatric Diseases, Departments of Psychiatry, Cellular and Molecular Medicine and Pediatrics, University of California, San Diego; and Professor, Institute for Genomic Medicine, University of California, San Diego; and Department of Pediatrics, University of California, San Diego, USA; **Eric Kelleher**, MRCPsych, PhD, Honorary Clinical Senior Lecturer, Department of Psychiatry and Neurobehavioural Science, University College Cork; and Visiting Research Fellow, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Christina Mooney**, HDip in Mental Health Nursing, Clinical Research Nurse, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Amy Durand**, Medical Student, University of Texas Health Science Center at Houston, McGovern Medical School, Texas, USA; and

Research Assistant, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Carlos Pinto**, PhD, Research Fellow, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Paul Cormican**, PhD, Lecturer, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Derek Morris**, PhD, Lecturer, Cognitive Genetics and Cognitive Therapy Group, Neuroimaging, Cognition and Genomics (NICOG) Centre, School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Ireland; **Gary Donohoe**, PhD, Professor, Cognitive Genetics and Cognitive Therapy Group, Neuroimaging, Cognition and Genomics (NICOG) Centre, School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Ireland; **Michael Gill**, MD, MRCPsych, FTCD, Professor, Head of School of Medicine, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Louise Gallagher**, MB, MRCPsych, PhD, Director of Research, School of Medicine, Trinity College Dublin; and Professor, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland; **Aiden Corvin**, MB, MRCPsych, PhD, Professor, Head of Discipline, Neuropsychiatric Genetics Research Group, Department of Psychiatry, School of Medicine, Trinity College Dublin, Ireland.

Correspondence: Professor Aiden Corvin. Email: acorvin@tcd.ie

First received 27 Jan 2019, final revision 18 Oct 2019, accepted 22 Nov 2019

Funding

This work was supported by: the National Institutes of Health (J.S., MH119746 and MH109501; A.C., NIH grant MH109501); the National Institute of Mental Health (A.C., NIMH grants MH 41953 and MH083094); Science Foundation Ireland (A.C., 12.IP.1359 and 08/IN.1/B1916); and the Wellcome Trust Case Control Consortium 2 (A.C., 085475/B/08/Z and 085475/Z/08/Z).

Acknowledgements

We thank all the study participants, participating professionals, investigators and recruitment sites. We gratefully acknowledge the work of the Schizophrenia Working Group of the Psychiatric Genomics Consortium, whose ongoing collaborative efforts are essential to continuing progress in understanding the genetic architecture of schizophrenia.

Data availability

Data are available from the authors on reasonable request.

Author contributions

C.F., E.A.H., M.G., L.G. and A.C. were responsible for the study conception and design. C.M., E.K., C.F. and A.D. were responsible for collection and coding of primary phenotypic data from the discovery data-set. C.M., E.K., D.H., D.M., C.P., P.C. and G.D. were responsible for the collection and processing of the genetic data from the discovery data-set. J.W., M.O. and M.O'D. contributed the data for the replication data-set and provided feedback with regard to the data analysis and interpretation. J.S. contributed to the core analysis of the genetic data used in the study. C.F., E.A.H., L.G. and A.C. contributed to the data analysis and interpretation and drafted the manuscript. All authors reviewed and approved the final manuscript.

Supplementary material

Supplementary material is available online at https://doi.org/10.1192/bjp.2019.262.

References

1 International Schizophrenia Consortium, Purcell SM, Wray NR, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–52.

2 Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet* 2017; **49**: 27–35.

3 Kirov G, Rees E, Walters JT, Escott-Price V, Georgieva L, Richards AL, et al. The penetrance of copy number variations for schizophrenia and developmental delay. *Biol Psychiatry* 2014; **75**: 378–85.

4 Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, et al. Cognitive performance among carriers of pathogenic copy number variants: analysis of 152,000 UK biobank subjects. *Biol Psychiatry* 2017; **82**: 103–10.

5 Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010; **86**: 749–64.

6 Schaefer GB, Mendelsohn NJ. Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. *Genet Med* 2013; **15**: 399–407.

7 Philip N, Bassett A. Cognitive, behavioural and psychiatric phenotype in 22q11.2 deletion syndrome. *Behav Genet* 2011; **41**: 403–12.

8 Ahn K, Gotay N, Andersen TM, Anvari AA, Gochman P, Lee Y, et al. High rate of disease-related copy number variations in childhood onset schizophrenia. *Mol Psychiatry* 2014; **19**: 568–72.

9 Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008; **320**: 539–43.

10 Yeo RA, Gangestad SW, Liu J, Ehrlich S, Thoma RJ, Pommy J, et al. The impact of copy number deletions on general cognitive ability and ventricle size in patients with schizophrenia and healthy control subjects. *Biol Psychiatry* 2013; **73**: 540–5.

11 Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* 2014; **505**: 361–6.

12 Derks EM, Ayub M, Chambert K, Del Favero J, Johnstone M, MacGregor S, et al. A genome wide survey supports the involvement of large copy number variants in schizophrenia with and without intellectual disability. *Am J Med Genet B Neuropsychiatr Genet* 2013; **162B**: 847–54.

13 Sahoo T, Theisen A, Rosenfeld JA, Lamb AN, Ravnan JB, Schultz RA, et al. Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. *Genet Med* 2011; **13**: 868–80.

14 Wilson NK, Lee Y, Long R, Hermetz K, Rudd MK, Miller R, et al. A novel micro-duplication in the neurodevelopmental gene SRGAP3 that segregates with psychotic illness in the family of a COS proband. *Case Rep Genet* 2011; **2011**: 585893.

15 Costain G, Lionel AC, Fu F, Stavropoulos DJ, Gazzellone MJ, Marshall CR, et al. Adult neuropsychiatric expression and familial segregation of 2q13 duplications. *Am J Med Genet B Neuropsychiatr Genet* 2014; **165B**: 337–44.

16 Miles JH, Takahashi TN, Hong J, Munden N, Flournoy N, Braddock SR, et al. Development and validation of a measure of dysmorphology: useful for autism subgroup classification. *Am J Med Genet A* 2008; **146A**: 1101–16.

17 Irish Schizophrenia Genomics Consortium, Wellcome Trust Case Control Consortium 2. Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biol Psychiatry* 2012; **72**: 620–8.

18 First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P)*. Biometrics Research, New York State Psychiatric Institute, 2002.

19 Hamshere ML, Walters JT, Smith R, Richards AL, Green E, Grozeva D, et al. Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Mol Psychiatry* 2013; **18**: 708–12.

20 Rees E, Walters JTR, Georgieva L, Isles AR, Chambert KD, Richards AL, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. *Br J Psychiatry* 2014; **204**: 108–14.

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

22 Rees E, Kendall K, Pardiñas AF, Legge SE, Pocklington A, Escott-Price V, et al. Analysis of intellectual disability copy number variants for association with schizophrenia. *JAMA Psychiatry* 2016; **73**: 963–9.

23 R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2013 (http://www.R-project.org/).

24 Thygesen JH, Wolfe K, McQuillin A, Viñas-Jornet M, Baena N, Brison N, et al. Neurodevelopmental risk copy number variants in adults with intellectual disabilities and comorbid psychiatric disorders. *Br J Psychiatry* 2018; **212**: 287–94.

25 Lowther C, Merico D, Costain G, Waserman J, Boyd K, Noor A, et al. Impact of IQ on the diagnostic yield of chromosomal microarray in a community sample of adults with schizophrenia. *Genome Med* 2017; **9**(1): 105.

EXTRA CONTENT ONLINE

Curr Psychiatry Rep (2017) 19: 61
DOI 10.1007/s11920-017-0807-5

GENETIC DISORDERS (F GOES, SECTION EDITOR)

Genetics of Schizophrenia: Ready to Translate?

Claire Foley[1] · **Aiden Corvin**[1] · **Shigeki Nakagome**[1]

Published online: 24 July 2017

Abstract
Purpose of Review This is an era where we have significantly advanced the understanding of the genetic architecture of schizophrenia. In this review, we consider how this knowledge may translate into advances that will improve patient care.
Recent Findings Large-scale genome-wide association studies (GWAS) have identified more than a hundred loci each making a small contribution to illness risk. Meta-analysis of copy number variants (CNVs) in the Psychiatric Genomics Consortium (PGC) dataset has confirmed that some variants have a moderate or large impact on risk, although these are rare in the population. Genome sequencing advances allow a much more comprehensive evaluation of genomic variation. We describe the key findings from whole exome studies to date. These studies are happening against a backdrop of growing understanding of the regulation and expression of genes and better functional tools to investigate molecular mechanisms in model systems.
Summary We provide an overview of how recent approaches in schizophrenia genetics are converging and consider how they could impact on diagnostics, the development of personalized medicine, and drug discovery.

Keywords Schizophrenia · Genomics · Translational medicine · Polygene score · Mutations

This article is part of the Topical Collection on *Genetic Disorders*

✉ Aiden Corvin
acorvin@tcd.ie

[1] Department of Psychiatry and Neuropsychiatric Genetics Research Group, Trinity College Dublin, Dublin, Ireland

Introduction

Schizophrenia is almost uniquely challenging in forcing patients to confront their perception and understanding of the world around them. The core symptoms affecting perception of reality, thinking, behavior, and motivation typically emerge in early adulthood and presage a long-term course characterized by episodic or continuous illness, with significant functional impairment in many cases [1]. The disorder is largely heritable, but social factors including childhood adversity and migration also contribute to risk [2, 3]. Many (if not most cases) are likely to be of neurodevelopmental etiology, and the syndrome is associated with structural, functional, and neurochemical brain changes, particularly involving the dopaminergic system [4]. Antipsychotic medications, targeting the dopamine D2 receptor, are effective in treating psychotic symptoms and reducing the risk of relapse [5], that is the good news.. However, treatment has little impact on debilitating behavioral and cognitive deficits and overall illness outcomes have not improved substantially since their introduction [6].

Schizophrenia is a public health problem, as it affects almost 1% of adults and is associated with significant morbidity and premature mortality (by 10–20 years) [7]. Despite more than a century of accrued wisdom, we lack a sufficiently cohesive understanding of pathophysiology to improve patient care. The diagnosis, although reliable, remains a clinical one. This has hampered international efforts to develop early intervention strategies as it is difficult to predict which individuals will go on to develop the full syndrome among those at risk. It is also likely that the syndromal diagnosis captures a heterogeneous population of patients with a range of clinical symptom profiles and outcomes. Precious little progress has been made in developing new therapies and at a fundamental level it remains to be seen if the syndrome captures one or many underlying disease mechanisms.

By providing better understanding of disease etiology, genomics is a powerful enabler of drug development [8]. Witness the identification of targets for the development of drugs to treat dyslipidemia based on the discovery of rare genetic sequence variants that increase or decrease non-HDL cholesterol levels [9•]. The last decade has seen significant progress in our understanding of the genetic architecture of schizophrenia. The aim of this review is to consider how this knowledge may translate into advances that will improve patient care. We provide a brief overview of the most recent discoveries and consider how they could impact on diagnostics, the development of personalized medicine and drug discovery (Fig. 1).

Recent Genetic Discovery

Genome-Wide Association Studies (GWAS)

Genome-wide association studies (GWAS) represent a powerful, hypothesis-free approach to dissecting the genetic architecture of complex traits. GWAS studies employ array-based methods to comprehensively assay common single nucleotide polymorphisms (SNPs) that occur predictably across the human genome. Significant progress has been made in understanding the common genetic effects that contribute to many complex traits, by testing for differences in SNP frequency between cases and controls to identify risk or protective alleles [10]. The Psychiatric Genomics Consortium (PGC) published the most extensive GWAS report on schizophrenia in 2014, including data on up to 36,989 cases and 113,075 controls. This identified 108 independent genomic risk loci, localizing the search to genes that are current or promising targets for treatment (*DRD2* or *GRM3*), genes more widely involved in glutamatergic neurotransmission (*GRIN2A*, *SRR*, *CLCN3*, and *GRIA1*), and unexpected candidate mechanisms involving neuronal calcium signaling (e.g., *CACNA1C*, *CACNA1I*, *CACNB2*, *RIMS1*) and broader synaptic function (*KCTD13*, *CNTN4*, *PAK6*) [11•]. Eighty-three of these loci were newly implicated in schizophrenia, highlighting the gene discovery power of GWAS once a sufficiently large sample size became available.

Perhaps, one of the more surprising discoveries from the GWAS was identification of enrichment of genetic variants in immunological pathways [12]. The top signals of the association with schizophrenia come from the major histocompatibility complex (MHC) locus spanning four megabases of chromosome 6 [11•, 13–16]. The MHC is one of the most genetically diverse regions of the genome, and in its extended form encodes more than 400 genes critical to immune function but also involved in many other functions. The challenges of mapping this locus have been discussed elsewhere, but promising findings, detailed below, could impact on our understanding of schizophrenia etiology [17, 18•]. The evidence for association with immunological pathways extends beyond the MHC making immune mechanisms a new focus for schizophrenia research.

As the number of common risk loci identified approaches 150 (per comm), they collectively explain <5% of the disease variance: the great proportion of schizophrenia heritability is yet to be explained. SNP heritability across the genome is largely uniform, and it is likely that more than 71% of 1 Mb genomic regions contain at least one risk variant [19]. For most affected individuals, schizophrenia has a polygenic architecture in which hundreds or even thousands of variants collectively contribute to risk [20]. Many more of these are likely to be identified by ongoing GWAS efforts, but as effect sizes and allele frequencies become smaller, study power will diminish. Other analytical approaches can extract useful information and provide an estimate of the likely total contribution of common variants to susceptibility. The polygenic risk score (PRS) method is given as the weighted sum of risk alleles with the weights specified by association coefficients [21, 22]. Since large GWAS datasets are publicly available for

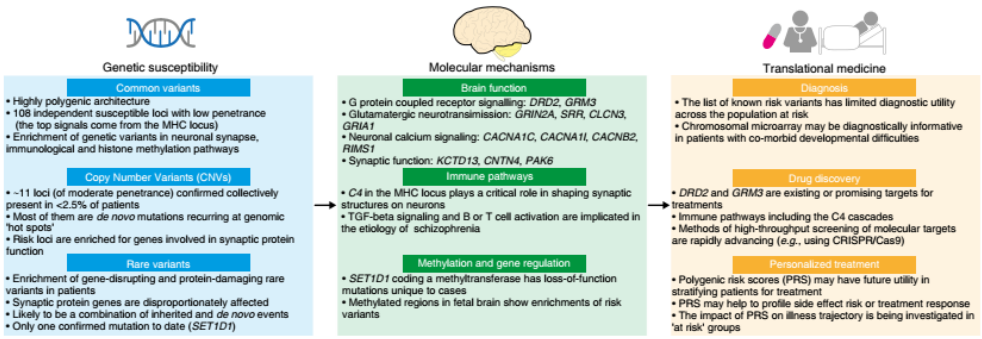


Fig. 1 A schematic flow of translational research. Diagrams show a pipeline from genetic discoveries via biological insights to potential applications for translational medicine

schizophrenia, it has been possible to calculate PRS from SNPs chosen for different thresholds of significance [11•, 16]. This allows estimation of the aggregate contribution of SNPs to schizophrenia and may be a useful method of stratifying and subgrouping based on disease status or other phenotypes like illness severity or treatment response (considered more fully below). Rather than testing the significance of individual SNPs, variance-component methods fit all SNPs available in samples to explain variance of a particular phenotype [23]. The method has been applied to a dataset of 22,177 schizophrenia cases and 27,629 controls and estimated 27.4% of the heritability to be explained by SNPs with minor allele frequency >2% [19]. This suggests that common variants explain a significant fraction, but not all of the heritability, begging the question where does the remainder come from?

Rare Structural Risk Variants

Just over a decade ago, it became evident that rare chromosomal rearrangements involving deletion, duplication, inversion, or translocation of DNA (termed copy number variants (CNVs)) make a sizable contribution to human genetic variation [24, 25]. Typically defined as being greater than 1 kb in size, in truth they follow a continuous distribution from very large or complex events to simple insertions/deletions [26]. With the transition from conventional karyotyping to array-based methods, it became clear that these sub-microscopic events make a sizable and clinically significant contribution to neurodevelopmental disorders. Predating this era, it was already known that deletions at 22q11.2 substantially increased schizophrenia risk [27], but by using array technology, Walsh and colleagues identified a global increase in deletions and duplications (>100 kb) in schizophrenia patients [28].

Larger consortia studies and the recent PGC-centralized analysis of 21,094 cases and 20,227 controls confirmed this finding and have shown that this effect is driven by CNVs that overlap genes [29–31]. The implicated loci overlap CNV regions known to be involved in intellectual disability (ID) [32], and genes involved in *N*-methyl-*D*-aspartate receptor (NMDAR) and neuronal activity-regulated cytoskeleton-associated (ARC) postsynaptic signaling complexes [33]. These studies have also identified specific loci that were significantly associated with schizophrenia. Confirming the known association with chr22q11.2 deletions, the early consortia studies also reported novel associations with deletions at 1q21.1, 15q11.2, and 15q13.3. Subsequent studies implicated at least ten other regions with either risk or protective loci (reviewed in [34]). The PGC analysis provided genome-wide significant evidence for eight of these (1q21.1, 2p16.3 (*NRXN1*), 3q29, 7q11.2, 15q13.3, distal 16p11.2, proximal 16p11.2, and 22q11.2) and suggestive support for eight more [30•].

Collectively, this list of known/probable risk loci is carried by <2.5% of patients [34] and is of moderate penetrance for schizophrenia (OR = 2–30) [26, 35–39]. Most represent new events probably explained by the higher rate of de novo CNVs reported in schizophrenia compared to controls [33]. Structurally, most of the implicated CNVs recur in the population at the same genomic locations mediated by non-allelic homologous recombination (NAHR) where repeat sequences flanking the critical region make it prone to rearrangement in the population. But this is only one of a number of proposed mechanisms for genome rearrangements (reviewed in [40]) and deletions at the genes *NRXN1* and *VIPR2* were identified because of different, disruptive CNV events with different breakpoints in individual carriers rather than a more universal genomic mechanism.

Genome Sequencing Studies

Next generation sequencing (NGS) opens up the huge reservoir of rare coding point mutations and small insertions or deletions (indels) in the human genome to investigators [41]. Gene-disrupting and protein-damaging rare variants (rare variants—alternative forms of a gene that are present with a minor allele frequency (MAF) of less than 1% [42]), or ultra-rare variants (dURVs—damaging variants unique to a single individual) are known to play a substantial role in other neurodevelopmental disorders (e.g., intellectual disability [43, 44], autism [45, 46], and epilepsy [47]). A large combined analysis of case-control (n = 6669 cases) and trio (n = 1077) exome data identified association between rare variants in *SETD1A* and risk of schizophrenia [48]. In more than 20,000 exomes, the investigators observed seven loss-of-function (LoF) variants and three de novo mutations in cases and none in controls. *SETD1A* encodes a methyltransferase that catalyzes the methylation of lysine residues in histone H3 and is a regulator of gene transcription. Other studies have yet to generate statistically robust findings and this may be because purifying selection prevents high-risk variants from reaching even modest allele frequencies in the population. If so, it is likely that reported case-control studies have been underpowered [49, 50•, 51] and scant sequencing data is published on extended pedigrees that may provide more power to find such events.

Much has been learned. Gene-disruptive and putatively protein-damaging dURVs, but not synonymous URVs (mutations not altering the amino acid sequence of a protein), are more abundant in people with schizophrenia meaning that it is only mutations that potentially impact gene function which are enriched in patient populations. Genovese et al. identified a case-associated elevation in dURVs (of about 0.25 variants per patient; 95% CI = 0.17–0.32) occurring on a background of about four dURVs per patient. To estimate the rate of de novo mutations (DNMs) that were protein damaging or

protein disruptive, they used data from a trio study of 617 affected and 1911 unaffected father-mother-offspring trios [52, 53], yielding an elevation of about 0.03 of these DNMs per exome. This estimate (0.03 per exome) was several times smaller than the elevation of dURVs in affected individuals in the population-based study (0.25 per exome). This suggests that most dURVs are inherited, even if only a few generations old [54•]. dURVs are enriched in genes overlapping regions near common variants associated with schizophrenia indicating at least some convergent genetic architecture. Mutations are also concentrated in genes implicated in X-linked intellectual disability and other developmental disorders. Such genes are known to be risk factors for syndromal forms of autism and the same may be true for schizophrenia. Looking at genes and gene sets, the mutations are concentrated in neuron-specific genes. More specifically, there is an increased rare variant burden in genes coding for members of the ARC and NMDAR, fragile X mental retardation protein (FMRP) targets, and possibly voltage-gated calcium ion channels; representing substantial overlap with the reported findings from pathway analyses of SNP and CNV data. From a recent study by Genovese and colleagues, potentially synaptic genes appear to explain more than 70% of the exome enrichment in dURVs [54•]. More extensive investigation of the protein complexes and structures at the synapse is likely to be critical in parsing out the molecular etiology or etiologies involved.

Translational Impact

Unprecedented progress has been made in understanding schizophrenia susceptibility. This has identified a spectrum of risk variation that is likely to extend from common, small genetic effects to rare, even private mutations. Much of the focus has been on common variants assayed using GWAS-arrays, but increasingly affordable whole genome sequencing will provide a more complete picture of genomic risk. There is also an emerging roadmap for the required steps in bringing together genomics, other "omics" data and environmental data to inform translational research (for examples see [55, 56]). Here, we consider how this is likely to impact on diagnostics, drug discovery, and the development of personalized medicine.

Diagnostics

Most cases of schizophrenia are likely to involve polygenic risk spread across many variants. In the PGC2 paper, the authors looked at how the transition from carrying relatively small numbers of risk variants (first decile) to almost all known risk variants (tenth decile) impacted on risk to the individual. They confirmed that odds ratios increased almost linearly across this transition in all participating populations. However, in each subsample, there were individuals in the control population who were unaffected despite harboring many common risk variants. The area under the receiver operating curve (AUC) ranged from 0.65–0.8 across samples; significantly below the threshold for clinical diagnostics utility [11•]. As more loci are identified these values will likely improve, but are unlikely to provide the basis for a simple diagnostic test. Their utility in precision medicine is discussed below.

For some people, the situation will be more nuanced. Chromosomal microarray (CMA) testing for identification of clinically significant CNVs is recommended as a first tier genetic test for patients with autism spectrum disorder, developmental delay, and intellectual disability [57, 58]. Rates of CNVs are lower in schizophrenia, but carrying such a mutation will have implications even if much is still to be done to understand the pathogenicity and penetrance of individual events. Eight rare CNV loci that surpass genome-wide significance explain <1% of schizophrenia variance in the population and are carried by <2% of patients [30•, 34]. These loci are of moderate penetrance for schizophrenia (OR = 2–30) [26, 35–39] but almost all have pleiotrophic effects conferring risk for other developmental phenotypes including ASD, ID, epilepsy, and congenital anomalies. In fact, the risk of developing any early developmental disorder (e.g., ID, ASD, developmental delay) is significantly higher (penetrance 10–88%) than the risk of developing schizophrenia (penetrance 2–18%) [38•]. Even in the absence of a psychiatric diagnosis, carriers of CNVs associated with schizophrenia have significant, but variable cognitive deficits [59, 60]. This is potentially important information for individuals, their families and treating clinicians.

Pick up rates for CNVs in the general schizophrenia population are likely to be low. In autism and intellectual disability (ID), "syndromal" cases where additional phenotypes and dysmorphic features are also present and are more likely to have a genomic etiology [61]. We have started to explore whether syndromal features including co-morbid ID, autism, epilepsy, developmental delay, or learning difficulties identify a schizophrenia population at greater risk. This is challenging because these are rare events. In a pilot study of 1215 schizophrenia cases, we found significantly higher rates of the known large CNVs (>100 kb) in cases having a co-morbid neurodevelopmental disorder, developmental delay, or learning difficulty compared to the rest of the schizophrenia population. As more data becomes available from sequencing studies, we may be able to identify or refine patient subgroups for which genetic testing will be of diagnostic significance.

Drug Discovery

Knowing the molecular etiological mechanism of a disease is critical to drug discovery. We are in an era where the tools

available to dissect such etiology are rapidly advancing. We are beginning to consider the functional organization of the genome and how the regulation of transcription and spatial configuration of DNA may be important in determining the translation of genes [62]. The development of clustered regularly interspaced short palindromic repeat screening (CRISPR—allows genes to be permanently modified in living cells and organisms [63]) is making it easier to identify novel therapeutic targets and even to identify synergistic drugs to target more than one molecular pathway at a time [64]. This may be important given the extensive list of target genes emerging from gene discovery. Functional assays in model systems, including iPS cells [65, 66], will be important in trying to understand whether genetic findings converge at the level of molecular pathways. This in turn will inform the critical question for personalized medicine as to whether schizophrenia represents a single or multiple disease processes. The tools are also becoming available to build from cellular studies to more complex systems investigating neural circuitry in vivo [67].

The GWAS era and CNV studies have provided a framework identifying post-synaptic signaling complexes (ARC, NMDAR), FMRP targets, and voltage-gated calcium channels as a starting point for discovery. Common variants have been the focus of most scrutiny, but as with GWAS from other complex traits, resolving the mechanism of risk at each locus is a significant challenge. Part of the challenge is that many variants fall within regions of significant linkage disequilibrium (LD) containing multiple variants that could drive an association. A significant number fall outside the exome and do not change the protein coding sequence of genes and are presumed to have subtle regulatory effects. Initial analysis of the PGC2 loci yielded only ten instances where the association signal could be attributed to a known non-synonymous exonic polymorphism [11•].

Recently, progress has been made in elucidating the strong genetic association of schizophrenia with the MHC locus [11•, 14, 16], despite the LD-complexity of this locus [68]. Steve McCarroll's group identified association with structural alleles (variants coding for amino acid sequences) of the complement component 4 (C4) genes. They reported that the C4 alleles associated with schizophrenia tended to generate greater expression of C4A—localizing to neuronal synapses, dendrites, axons, and cell bodies. Further, they showed that C4 promoted synapse elimination during the developmentally timed maturation of neuronal circuits in mice. Reduced synaptic structures in neurons are commonly observed in schizophrenia patients [69–71], which McCarroll's group suggest may be mediated by C4 along with other components of the complement cascade to promote excessive synaptic pruning in schizophrenia. The C4 complement cascade may be a potential therapeutic target to prevent this excessive synaptic pruning.

Rare variants directly impacting gene function may be more obvious targets for drug development. Although sequencing studies have been a relatively recent addition, the identification of loss-of-function effects involving *SETD1A* indicates that histone H3 methylation may also be an important avenue for research. In previous sections, we have noted the evidence of molecular overlap between schizophrenia and other neurodevelopmental disorders. Large-scale sequencing efforts in ID and autism may identify novel therapeutic pathways relevant to schizophrenia. As an example, an exonic mutation of *CACNA1C* causes Timothy syndrome a condition characterized by autistic features and long QT syndrome. The genetic mechanism involved is understood and the subject of intense molecular research in model systems including iPSC cells from patients [72•]. Common variants at the gene are implicated in schizophrenia, and this body of research may inform understanding of its role in schizophrenia if the molecular mechanism involved can be elucidated.

Personalized Treatment

The extent to which gene discovery will inform personalized treatment in schizophrenia is unclear. We are beginning to understand a fraction of the molecular etiology involved. It is premature to speculate on whether the disorder captures one or more disease processes. However, empowered by the information we have, we can begin to ask important questions about the onset of the disorder, severity, and possibly treatment response. There is a significant history of pharmacogenetics research in the field investigating the genetic contribution to variance in drug response, genetic risk of developing side effects and aiming to improve treatment efficacy. Malhotra et al. conducted a meta-analysis on the association between a variant (−141C Ins/Del SNP) in the promoter region of the DRD2 gene and antipsychotic drug response in almost 700 patients. There was a significant difference in response rate (clinical response defined as a 50% reduction of either the Brief Psychiatric Rating Scale total score or Positive and Negative Syndrome Scale total score at approximately 8 weeks of follow-up evaluation) between deletion carriers and Ins/Ins genotypes (pooled odds ratio = 0.65, 95% CI = 0.43~0.97, $p = 0.03$), indicating that patients who carry one or two deletion alleles tend to have less favorable antipsychotic drug responses than patients with the Ins/Ins genotype. Put another way, patients with the Ins/Ins genotype were 54% more likely to respond to antipsychotic drugs than those with at least one copy of the Del allele [73].

Zhang et al. have reviewed pharmacogenetic findings for two important antipsychotic adverse effects: antipsychotic weight gain and clozapine-induced agranulocytosis (CIA) [74]. Variations in the human leukocyte antigens (HLA) genomic region have been associated with clozapine-induced agranulocytosis (CIA) [75, 76]. Association of one variant

of HLA-DQB1 with CIA (OR 16.9, $p < 0.001$) was identified by Athanasiou and colleagues, with high specificity but low sensitivity [77]. The low sensitivity indicates that other genetic variants likely contribute to risk of CIA and further studies with large sample sizes will be required to identify these other genetic risks [74]. The pharmacogenetic literature has been the subject of further substantial reviews elsewhere (e.g., [74, 78]), but this body of work highlights the point that the genetic architectures underlying illness severity, treatment response, or side effects are likely to be complex and require similar genomics, with statistically rigorous methods, to those being applied to understand pathophysiology.

For example, the PRS method is being applied to look at common genetic risk in aggregate. A recent study has shown that PRS-SCZ predicts distinct types of prepubertal developmental impairments at different stages of childhood; higher risk score is associated with lower performance IQ, poorer speech intelligibility, and fluency, and more "headstrong" behavior at age from 7 to 9 years or with social difficulties and behavior problems as early as 4 years but the effect sizes were small (OR 1.07–1.13) [79]. Furthermore, PRS-SCZ has a greater power to explain the heritability of childhood-onset schizophrenia, a severe form of psychotic symptoms beginning before age 13, compared with that of late-onset schizophrenia [80]. PRS-SCZ may help to detect individuals who potentially would benefit from attending high-risk clinics or for preventative strategies, but further work in larger samples numbers is required to provide better effect size estimates.

The second question is if we can use PRS-SCZ to predict patient's response to drug or treatment. Around 30% of schizophrenia patients are known to be treatment-resistant [6, 81]. The atypical antipsychotic medication clozapine is associated with severe adverse events and hence only prescribed to the treatment-resistant patients [82]. Because a delay in providing proper treatments and a lack of efficacy for the initial treatment lead to a poorer prognosis, there is an urgent need for new approaches to the identification of patients who will require treatment with clozapine. The PRS-SCZ is being used to explore whether genetic information can help identify this patient group early as treatment-resistant patients tend to have higher risk scores than those who can respond to standard medications [83]. Follow-up studies with a large set of samples are necessary to enable genetic testing for the early intervention with clozapine.

Genomic research uncovers a wide spectrum of allele frequency associated with schizophrenia; risk alleles have the range from common alleles of weak effect to rare alleles or CNVs of relatively large effect. However, a question remains about whether the integration of copy number, rare, and common variants can provide a better estimate of polygenic risk and increase the power in the patient stratification. A recent study has made a comparison of PRS-SCZ between carriers of a known pathogenic CNV from cases and controls, as well as between the cases carrying the CNV and controls. The CNV carrier cases show higher PRS-SCZ than controls, suggesting that common variants contribute to the susceptibility in the presence of a CNV. Even within the CNV carriers, PRS-SCZ is able to statistically differentiate between cases and controls. Since carriers of a pathogenic CNV are prone to have other developmental disorders or cognitive deficits, the PRS-SCZ calculation could be useful for the detection of CNV carriers at increased risk for schizophrenia [39].

Although we discussed case studies that illustrate potential applications of genomic findings for translational medicine, the practical use of PRS-SCZ has several caveats [22]. First, since the number of susceptible SNPs that can be detected from GWAS is limited, a choice of SNPs to be included in the calculation has a critical effect on the estimate. Second, the use of only SNPs above genome-wide significance does not always show the best performance, suggesting that a threshold for the inclusion/exclusion of SNPs has to be optimized in empirical ways, such as using independent sets of samples or cross-validation. Third, given the highly polygenic architecture of schizophrenia, the precision of the PRS-SCZ model strongly depends on the sample size used for GWAS; a huge sample size is required to significantly increase the accuracy of PRS-SCZ.

Conclusion

Despite substantial progress, much of the heritability of schizophrenia is yet to be explained. The advent of genome sequencing is likely to lead to greater understanding. A substantial contribution of common risk variants across all populations, even controls, is an important message in destigmatizing this disorder but also suggests limitations for molecular diagnostics. Whether a subgroup of patients with other neurodevelopmental pathologies will benefit from molecular diagnostics is a focus of active research. As yet, it is unclear whether the schizophrenia syndrome captures one or more disease states, but the evidence is converging on a number of molecular pathways at a time when neuroscience is developing more sophisticated tools to understand how neural circuits function in vivo. This convergence represents the start of an exciting new era for drug discovery in this devastating disorder.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as
• Of importance

1. Hegarty JD, Baldessarini RJ, Tohen M, Waternaux C, Oepen G. One hundred years of schizophrenia: a meta-analysis of the outcome literature. Am J Psychiatry. 1994;151(10):1409–16.
2. Matheson S, Shepherd A, Pinchbeck R, Laurens K, Carr V. Childhood adversity in schizophrenia: a systematic meta-analysis. Psychol Med. 2013;43(2):225.
3. Cantor-Graae E, Selten J-P. Schizophrenia and migration: a meta-analysis and review. Am J Psychiatr. 2005;162(1):12–24.
4. Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, "just the facts": what we know in 2008 part 1: overview. Schizophr Res. 2008;100(1–3):4–19. doi:10.1016/j.schres.2008.01.022.
5. Miyamoto S, Miyake N, Jarskog LF, Fleischhacker WW, Lieberman JA. Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. Mol Psychiatry. 2012;17(12):1206–27. doi:10.1038/mp.2012.47.
6. Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet (London, England). 2013;382(9896):951–62. doi:10.1016/s0140-6736(13)60733-3.
7. Tiihonen J, Lonnqvist J, Wahlbeck K, Klaukka T, Niskanen L, Tanskanen A, et al. 11-year follow-up of mortality in patients with schizophrenia: a population-based cohort study (FIN11 study). Lancet. 2009;374(9690):620–7. doi:10.1016/s0140-6736(09)60742-x.
8. Kamb A, Harper S, Stefansson K. Human genetics as a foundation for innovative drug development. Nat Biotechnol. 2013;31(11):975–8. doi:10.1038/nbt.2732.
9.• Nioi P, Sigurdsson A, Thorleifsson G, Helgason H, Agustsdottir AB, Norddahl GL, et al. Variant ASGR1 associated with a reduced risk of coronary artery disease. N Engl J Med. 2016;374(22):2131–41. doi:10.1056/NEJMoa1508419. **This whole-genome sequencing study identified a deletion leading to inactivation of ASGR1, associated with reduced levels of non-HDL cholesterol and a resultant reduced risk of coronary artery disease.**
10. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014;42(Database issue):6.
11.• Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511(7510):421–7. doi:10.1038/nature13595. **This paper reported on an extensive GWAS analysis on schizophrenia, including data on up to 36,989 cases and 113,075 controls. 108 independent genomic risk loci associated with schizophrenia were identified. http://www.nature.com/nature/journal/v511/n7510/abs/nature13595.html#supplementary-information**
12. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci. 2015;18(2):199–209.
13. Corvin A, Morris DW. Genome-wide association studies: findings at the major histocompatibility complex locus in psychosis. Biol Psychiatry. 2014;75(4):276–83. doi:10.1016/j.biopsych.2013.09.018.
14. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, et al. Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature. 2009;460(7256):753–7. doi:10.1038/nature08192.
15. Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. Biol Psychiatry. 2012;72(8):620–8. doi:10.1016/j.biopsych.2012.05.035.
16. Genome-wide association study identifies five new schizophrenia loci. Nature genetics. 2011;43(10):969–76. doi:10.1038/ng.940.
17. Morris DW, Pearson RD, Cormican P, Kenny EM, O'Dushlaine CT, Lemieux Perreault L-P, et al. An inherited duplication at the gene p21 protein-activated kinase 7 (PAK7) is a risk factor for psychosis. Hum Mol Genet. 2014; doi:10.1093/hmg/ddu025.
18.• Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. Nature. 2016;530(7589):177–83. doi:10.1038/nature16549. **This study examined the association of schizophrenia with variation in the MHC locus, identifying a key role of excessive complement component 4 activity, possibly resulting in the synapse loss observed in schizophrenia**
19. Loh PR, Bhatia G, Gusev A, Finucane HK, Bulik-Sullivan BK, Pollack SJ, et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. Nat Genet. 2015;47(12):1385–92. doi:10.1038/ng.3431.
20. Gratten J, Wray NR, Keller MC, Visscher PM. Large-scale genomics unveils the genetic architecture of psychiatric disorders. Nat Neurosci. 2014;17(6):782–90. doi:10.1038/nn.3708.
21. Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet. 2002;31(1):33–6. doi:10.1038/ng853.
22. Chatterjee N, Shi J, Garcia-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. Nat Rev Genet. 2016;17(7):392–406. doi:10.1038/nrg.2016.27.
23. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42(7):565–9.
24. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nat Genet. 2004;36(9):949–51. doi:10.1038/ng1416.
25. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al. Large-scale copy number polymorphism in the human genome. Science (New York, NY). 2004;305(5683):525–8. doi:10.1126/science.1098918.
26. Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell. 2012;148(6):1223–41. doi:10.1016/j.cell.2012.02.039.
27. Pulver AE, Nestadt G, Goldberg R, Shprintzen RJ, Lamacz M, Wolyniec PS, et al. Psychotic illness in patients diagnosed with velo-cardio-facial syndrome and their relatives. J Nerv Ment Dis. 1994;182(8):476–8.
28. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science (New York, NY). 2008;320(5875):539–43. doi:10.1126/science.1155174.
29. Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature. 2008;455(7210):237–41. doi:http://www.nature.com/nature/journal/v455/n7210/suppinfo/nature07239_S1.html.
30.• Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet. 2017;49(1):27–35. doi:10.1038/ng.3725. This study found genome-wide significant evidence for eight CNV loci contributing to schizophrenia risk with suggestive support for a further eight candidate susceptibility and protective loci

31. Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S, et al. Large recurrent microdeletions associated with schizophrenia. Nature. 2008;455(7210):232–6. doi:10.1038/nature07229.
32. Rees E, Kendall K, Pardinas AF, Legge SE, Pocklington A, Escott-Price V, et al. Analysis of intellectual disability copy number variants for association with schizophrenia. JAMA psychiatry. 2016;73(9):963–9. doi:10.1001/jamapsychiatry.2016.1831.
33. Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Mol Psychiatry. 2012;17(2):142–53. doi:10.1038/mp.2011.154.
34. Kirov G. CNVs in neuropsychiatric disorders. Hum Mol Genet. 2015;24(R1):R45–9. doi:10.1093/hmg/ddv253.
35. Morris DW, Pearson RD, Cormican P, Kenny EM, O'Dushlaine CT, Perreault LP, et al. An inherited duplication at the gene p21 protein-activated kinase 7 (PAK7) is a risk factor for psychosis. Hum Mol Genet. 2014;23(12):3316–26. doi:10.1093/hmg/ddu025.
36. Rujescu D, Ingason A, Cichon S, Pietilainen OP, Barnes MR, Toulopoulou T, et al. Disruption of the neurexin 1 gene is associated with schizophrenia. Hum Mol Genet. 2009;18(5):988–96. doi:10.1093/hmg/ddn351.
37. Rees E, Walters JT, Chambert KD, O'Dushlaine C, Szatkiewicz J, Richards AL, et al. CNV analysis in a large schizophrenia sample implicates deletions at 16p12.1 and SLC1A1 and duplications at 1p36.33 and CGNL1. Hum Mol Genet. 2014;23(6):1669–76. doi:10.1093/hmg/ddt540.
38.• Kirov G, Rees E, Walters JT, Escott-Price V, Georgieva L, Richards AL, et al. The penetrance of copy number variations for schizophrenia and developmental delay. Biol Psychiatry. 2014;75(5):378–85. doi:10.1016/j.biopsych.2013.07.022. **This study estimated total penetrance of thirteen schizophrenia-associated CNVs for neurodevelopmental disorders including SCZ, DD, ASD or a congenital malformation, to be high (ranging between 10.6%-100%), with significant implications for genetic counselling**
39. Tansey KE, Rees E, Linden DE, Ripke S, Chambert KD, Moran JL, et al. Common alleles contribute to schizophrenia in CNV carriers. Mol Psychiatry. 2016;21(8):1085–9. doi:10.1038/mp.2015.143.
40. Liu P, Carvalho CM, Hastings PJ, Lupski JR. Mechanisms for recurrent and complex human genomic rearrangements. Curr Opin Genet Dev. 2012;22(3):211–20. doi:10.1016/j.gde.2012.02.012.
41. MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, et al. A systematic survey of loss-of-function variants in human protein-coding genes. Science (New York, NY). 2012;335(6070):823–8. doi:10.1126/science.1215040.
42. Mägi R, Asimit JL, Day-Williams AG, Zeggini E, Morris AP. Genome-wide association analysis of imputed rare variants: application to seven common complex diseases. Genet Epidemiol. 2012;36(8):785–96. doi:10.1002/gepi.21675.
43. de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, et al. Diagnostic exome sequencing in persons with severe intellectual disability. N Engl J Med. 2012;367(20):1921–9. doi:10.1056/NEJMoa1206524.
44. Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet. 2012;380(9854):1674–82. doi:10.1016/s0140-6736(12)61480-9.
45. O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science (New York, NY). 2012;338(6114):1619–22. doi:10.1126/science.1227764.
46. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Ercument Cicek A, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature. 2014;515(7526):209–15. doi:10.1038/nature13772.
47. Allen AS, Berkovic SF, Cossette P, Delanty N, Dlugos D, Eichler EE, et al. De novo mutations in epileptic encephalopathies. Nature. 2013;501(7466):217–21. doi:10.1038/nature12439.
48. Singh T, Kurki MI, Curtis D, Purcell SM, Crooks L, McRae J, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. Nat Neurosci. 2016;19(4):571–7. doi:10.1038/nn.4267.
49. Zuk O, Schaffner SF, Samocha K, Do R, Hechter E, Kathiresan S, et al. Searching for missing heritability: designing rare variant association studies. Proc Natl Acad Sci U S A. 2014;111(4):E455–64. doi:10.1073/pnas.1322563111.
50.• Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, et al. A polygenic burden of rare disruptive mutations in schizophrenia. Nature. 2014;506(7487):185–90. doi:10.1038/nature12975. **This paper introduces the polygene score method and its application to schizophrenia**
51. Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, et al. De novo mutations in schizophrenia implicate synaptic networks. Nature. 2014;506(7487):179–84. doi:10.1038/nature12929.
52. Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, et al. De novo mutations in schizophrenia implicate synaptic networks. Nature. 2014;506(7487):179–84. doi:10.1038/nature12929. http://www.nature.com/nature/journal/v506/n7487/abs/nature12929.html#supplementary-information
53. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, et al. The contribution of de novo coding mutations to autism spectrum disorder. Nature. 2014;515(7526):216–21.
54.• Genovese G, Fromer M, Stahl EA, Ruderfer DM, Chambert K, Landen M, et al. Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. Nat Neurosci. 2016;19(11):1433–41. doi:10.1038/nn.4402. **This study explored the role of rare coding mutations in schizophrenia and identified that most dURVs contributing to the disorder appeared to be inherited and enriched in brain-specific genes. The findings provide evidence of a general synaptic dysfunction in the pathogenesis of schizophrenia**
55. Akil H, Brenner S, Kandel E, Kendler KS, King MC, Scolnick E, et al. Medicine. The future of psychiatric research: genomes and neural circuits. Science (New York, NY). 2010;327(5973):1580–1. doi:10.1126/science.1188654.
56. Corvin A, Sullivan PF. What next in schizophrenia genetics for the psychiatric genomics consortium? Schizophr Bull. 2016;42(3):538–41.
57. O'Byrne JJ, Lynch SA, Treacy EP, King MD, Betts DR, Mayne PD, et al. Unexplained developmental delay/learning disability: guidelines for best practice protocol for first line assessment and genetic/metabolic/radiological investigations. Ir J Med Sci. 2016;185(1):241–8. doi:10.1007/s11845-015-1284-7.
58. Schaefer GB, Mendelsohn NJ. Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. Genet Med. 2013;15(5):399–407.
59. Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, et al. Cognitive performance among carriers of pathogenic copy number variants: analysis of 152,000 UK biobank subjects. Biol Psychiatry. 2016; doi:10.1016/j.biopsych.2016.08.014.
60. Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature. 2014;505(7483):361–6. doi:10.1038/nature12818.
61. Miles JH, Takahashi TN, Bagby S, Sahota PK, Vaslow DF, Wang CH, et al. Essential versus complex autism: definition of fundamental prognostic subtypes. Am J Med Genet A. 2005;135(2):171–80. doi:10.1002/ajmg.a.30590.

62. Rajarajan P, Gil SE, Brennand KJ, Akbarian S. Spatial genome organization and cognition. Nat Rev Neurosci. 2016;17(11):681–91. doi:10.1038/nrn.2016.124.
63. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science (New York, NY). 2013;339(6121):819–23. doi:10.1126/science.1231143.
64. Han K, Jeng EE, Hess GT, Morgens DW, Li A, Bassik MC. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nat Biotechnol. 2017; doi:10.1038/nbt.3834.
65. Lee IS, Carvalho CM, Douvaras P, Ho S-M, Hartley BJ, Zuccherato LW, et al. Characterization of molecular and cellular phenotypes associated with a heterozygous CNTNAP2 deletion using patient-derived hiPSC neural cells. npj Schizophrenia. 2015;1:15019.
66. Wen Z, Nguyen HN, Guo Z, Lalli MA, Wang X, Su Y, et al. Synaptic dysregulation in a human iPS cell model of mental disorders. Nature. 2014;515(7527):414–8. doi:10.1038/nature13716. http://www.nature.com/nature/journal/v515/n7527/abs/nature13716.html#supplementary-information
67. Kim CK, Adhikari A, Deisseroth K. Integration of optogenetics with complementary methodologies in systems neuroscience. Nat Rev Neurosci. 2017;18(4):222–35.
68. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. Nature. 2016; doi:10.1038/nature16549.
69. Garey LJ, Ong WY, Patel TS, Kanani M, Davis A, Mortimer AM, et al. Reduced dendritic spine density on cerebral cortical pyramidal neurons in schizophrenia. J Neurol Neurosurg Psychiatry. 1998;65(4):446–53.
70. Glantz LA, Lewis DA. Decreased dendritic spine density on prefrontal cortical pyramidal neurons in schizophrenia. Arch Gen Psychiatry. 2000;57(1):65–73.
71. Glausier JR, Lewis DA. Dendritic spine pathology in schizophrenia. Neuroscience. 2013;251:90–107. doi:10.1016/j.neuroscience.2012.04.044.
72.• Krey JF, Pasca SP, Shcheglovitov A, Yazawa M, Schwemberger R, Rasmusson R, et al. Timothy syndrome is associated with activity-dependent dendritic retraction in rodent and human neurons. Nat Neurosci. 2013;16(2):201–9. doi:10.1038/nn.3307. **This study examined the point mutation in the calcium channel CACNA1C causing Timothy syndrome (a neurodevelopmental disorder) in iPSC-derived cortical neurons and identified activity-dependent dendritic retraction in neurons, providing insights into the cellular basis of Timothy syndrome and potentially other neurodevelopmental disorders**
73. Zhang JP, Lencz T, Malhotra AK. D2 receptor genetic variation and clinical response to antipsychotic drug treatment: a meta-analysis. Am J Psychiatry. 2010;167(7):763–72. doi:10.1176/appi.ajp.2009.09040598.
74. Zhang J-P, Malhotra AK. Pharmacogenetics of antipsychotics: recent progress and methodological issues. Expert Opin Drug Metab Toxicol. 2013;9(2):183–91.
75. Lieberman JA, Yunis J, Egea E, Canoso RT, Kane JM, Yunis EJ. HLA-B38, DR4, DQw3 and clozapine-induced agranulocytosis in Jewish patients with schizophrenia. Arch Gen Psychiatry. 1990;47(10):945–8.
76. Yunis JJ, Corzo D, Salazar M, Lieberman JA, Howard A, Yunis EJ. HLA associations in clozapine-induced agranulocytosis. Blood. 1995;86(3):1177–83.
77. Athanasiou MC, Dettling M, Cascorbi I, Mosyagin I, Salisbury BA, Pierz KA, et al. Candidate gene analysis identifies a polymorphism in HLA-DQB1 associated with clozapine-induced agranulocytosis. The Journal of clinical psychiatry. 2011;72(4):458–63.
78. Zhang J-P, Malhotra AK. Pharmacogenetics and antipsychotics: therapeutic efficacy and side effects prediction. Expert Opin Drug Metab Toxicol. 2011;7(1):9–37. doi:10.1517/17425255.2011.532787.
79. Riglin L, Collishaw S, Richards A, Thapar AK, Maughan B, O'Donovan MC, et al. Schizophrenia risk alleles and neurodevelopmental outcomes in childhood: a population-based cohort study. The lancet Psychiatry. 2017;4(1):57–62. doi:10.1016/s2215-0366(16)30406-0.
80. Ahn K, An SS, Shugart YY, Rapoport JL. Common polygenic variation and risk for childhood-onset schizophrenia. Mol Psychiatry. 2016;21(1):94–6. doi:10.1038/mp.2014.158.
81. Meltzer HY. Treatment-resistant schizophrenia—the role of clozapine. Curr Med Res Opin. 1997;14(1):1–20. doi:10.1185/03007999709113338.
82. Nielsen J, Nielsen RE, Correll CU. Predictors of clozapine response in patients with treatment-refractory schizophrenia: results from a Danish register study. J Clin Psychopharmacol. 2012;32(5):678–83. doi:10.1097/JCP.0b013e318267b3cd.
83. Frank J, Lang M, Witt SH, Strohmaier J, Rujescu D, Cichon S, et al. Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients. Mol Psychiatry. 2015;20(2):150–1. doi:10.1038/mp.2014.56.