Neurocognitive Effects of Risk Factors for Psychosis: Neuropsychological & Neuroimaging Studies of Immune-Related Genetic and Environmental Factors

By

Jessica Holland, B.Sc. Psychology, M.Sc.

A thesis submitted for the Degree of Doctor of Philosophy to the School of Psychology, National University of Ireland, Galway

Supervisor: Professor Gary Donohoe

School of Psychology, National University of Ireland, Galway

Submitted: September 2019
Table of Contents

Introduction ..................................................................................................................xxi

1.1 Schizophrenia ........................................................................................................1

1.1.1 An overview of SZ ......................................................................................1

1.1.2 Prevalence and Socioeconomic Burden ......................................................1

1.1.3 SZ aetiology ................................................................................................2

1.1.4 Treatment ......................................................................................................3

1.1.5 Neurodevelopmental Hypothesis ...............................................................4

1.1.6 The Immune Hypothesis ............................................................................5

1.2 Genetics of SZ ....................................................................................................6

1.2.1 Heritability of SZ .......................................................................................6

1.2.2 Candidate Gene studies ............................................................................7

1.2.3 GWAS .........................................................................................................7

1.2.4 Polygenic Risk Profiling and LD score regression .....................................8

1.3 Cognition in SZ ..................................................................................................10

1.3.1 Cognition as an Endophenotype for SZ ...............................................10

1.3.2 General Cognitive ability .........................................................................11

1.3.3 Episodic and Visual Memory .....................................................................12

1.3.4 Working Memory .......................................................................................13

1.3.5 Social Cognition .........................................................................................13

Genetic architecture of cognition in SZ ..................................................................18

1.3.6 Medication effects on cognition ...............................................................19

1.4 Brain Structure in SZ .......................................................................................20

1.5 Neuroimmunology .............................................................................................22

1.5.1 Defining the neuroimmune system .........................................................23

1.5.2 Immune System structure (Innate and Adaptive, Cross talk) ...............24

1.5.3 Immune Abnormalities in SZ- Epidemiological Findings .....................27

1.5.4 Maternal Immune Activation (MIA) .......................................................27

1.5.5 Immune markers ........................................................................................29

1.5.6 IL6 ..............................................................................................................30

1.5.7 CRP ............................................................................................................31

1.5.8 Environment, cognition and Immune function ......................................32

1.5.9 Genetic findings: MHC ............................................................................33
2. Study 1: Effects of Early Life Adversity on Immune Function and Cognitive Performance: Results from the ALSPAC cohort .................................................. 38
  2.1 Introduction .............................................................................. 40
    2.1.1 ELA and Cognition ............................................................... 40
    2.1.2 Inflammatory markers & Cognitive performance .................. 41
    2.1.3 The present study ............................................................... 42
  2.2 Methods .................................................................................. 42
    2.2.1 Participants ........................................................................ 42
    2.2.2 Childhood adversity ............................................................ 43
    2.2.3 Measure of social cognition ............................................... 45
    2.2.4 Measure of general cognition ............................................. 45
    2.2.5 Potential Confounders ......................................................... 46
    2.2.6 Measurement of IL6 and CRP ............................................ 46
    2.2.7 Analysis ............................................................................. 47
  2.3 Results ..................................................................................... 49
    2.3.1 Association between ELA score and social cognition (Measured by Triangles score) ............................................................... 51
    2.3.2 Association between ELA score and cognition (measured by IQ) .. 53
    2.3.3 Relationship between immune markers, ELAs and cognition ..... 56
    2.3.4 Does Immune response partly account for the variance in cognitive performance explained by ELA? ............................................... 60
  2.4 Discussion ............................................................................... 61
    2.4.1 Harsh parenting and cognition. ............................................ 61
    2.4.2 Harsh parenting, lower IQ, and inflammatory response. ....... 62
    2.4.3 Alternative explanations for the relationship between ELA and cognition 62
    2.4.4 Strengths and Limitations .................................................... 63
    2.4.5 Conclusion .......................................................................... 64
  2.5 Supplementary Figure ................................................................ 66
  2.6 Acknowledgements .................................................................. 67

3. Study 2. Beyond C4: Analysis Of The Complement Gene Pathway Shows Enrichment For IQ In Patients With Psychotic Disorders And Healthy Controls .... 68
  3.1 Introduction ............................................................................. 70
  3.2 Methodology .......................................................................... 72
3.2.1 Samples included in enrichment analysis .............................. 72
3.2.2 Complement gene-set ..................................................... 73
3.2.3 Enrichment analysis statistical approach .......................... 74
3.2.4 Samples included in polygenic score (PGS) analysis .......... 74
3.2.5 Cognitive assessment .................................................... 75
3.2.6 Genotyping ................................................................. 76
3.2.7 PGS score calculation ................................................... 76
3.2.8 Statistical Analysis ....................................................... 77

3.3 Results .................................................................................. 77
3.3.1 Enrichment analysis ......................................................... 77
3.3.2 PGS analysis: Clinical and demographic information .... 82
3.3.3 Association between complement-based polygenic score (PGS) and IQ 84

3.4 Discussion .............................................................................. 88
3.4.1 Enrichment analysis of complement genes ...................... 89
3.4.2 PGS analysis of complement gene-set .............................. 89
3.4.3 The complement pathway, neurodevelopment, and cognition .... 90
3.4.4 Limitations ......................................................................... 91
3.4.5 Future directions ............................................................. 93
3.4.6 Conclusion ......................................................................... 93

3.5 Acknowledgements & Funding ........................................... 94

Supplementary Tables .................................................................. 95

4. Study 3. Effects Of Complement Gene-set Polygenic Risk Score on Brain Volume and Cortical Measures in Patients with Psychotic Disorders and Healthy Controls ................................................................. 98
4.1 Introduction ............................................................................ 100
4.2 Methods ................................................................................ 102
4.2.1 Sample description ......................................................... 102
4.2.2 Cognitive assessment of memory ..................................... 102
4.2.3 Magnetic resonance imaging (MRI) ................................. 103
4.2.4 Processing of imaging data .............................................. 104
4.2.5 Genotyping ...................................................................... 105
4.2.6 C4 Expression ................................................................. 105
4.2.7 Complement gene-set ..................................................... 106
4.2.8 PRS score calculation ..................................................... 106
4.3 Results .................................................................................. 107
4.3.1 Complement gene-set PRS and memory scores ................................ 109
4.3.2 Complement PRS and hippocampal volume ................................ 109
4.3.3 Complement PRS and Cortical Thickness, Surface Area and Total Brain Volume ................................................................. 109
4.3.4 Post-hoc Analyses: sub-regions of the hippocampus .................... 111
4.3.5 C4A and structural metrics ..................................................... 113
4.4 Discussion .................................................................................. 115
  4.4.1 Strengths ................................................................................ 116
  4.4.2 Limitations .......................................................................... 117
  4.4.3 Future Directions .................................................................. 117
  4.4.4 Conclusion ........................................................................... 118
4.5 Acknowledgements & Funding .................................................... 119
4.6 Supplementary Table Study 3 ................................................... Error! Bookmark not defined.
5. Discussion .................................................................................. 121
  5.1 Summary of Main Findings ......................................................... 121
    5.1.1 Study 1 ............................................................................. 121
    5.1.2 Study 2 ............................................................................. 122
    5.1.3 Study 3 ............................................................................. 122
    5.1.4 Candidate’s Role in Studies .................................................. 124
    5.1.5 Strengths of this Thesis ......................................................... 125
    5.1.6 Limitations ......................................................................... 129
    5.1.7 Future Directions for Research ............................................. 133
    5.1.8 Clinical application of results .............................................. 135
    5.1.9 Concluding remarks ............................................................ 137
References ....................................................................................... 140
Appendix A: Publications Arising from this Thesis ............................. 163
Appendix B: Awards and Bursaries ................................................... 164
Appendix C: Presentations ................................................................ 165
Appendix D: Study Funding and Acknowledgements ......................... 166
Declaration

I declare that this thesis has not been submitted as an exercise at this or any other university. I declare that this thesis is entirely my own work, except where otherwise stated.

Signed: Jessica Holland
Statement of Contribution

For study 1, data access was granted from the ALSPAC executive committee to satisfy specific research questions outlined in a proposal (Access granted February 2016, B2448). All analysis for this study was undertaken by the author.

Studies 2 and 3 include work that forms part of a collaborative project on psychosis. The author was involved with recruitment of patients with schizophrenia, schizoaffective, and bipolar disorder for neuropsychological testing and neuroimaging components, and administering a neuropsychological test battery to participants as part of the Cognitive Genetics and Cognitive Therapy Group, NUI Galway. Data used in this thesis was collected in multiple stages. The data collected by the author was included in the latest wave of data collection that contributed towards the Resource for Psychoses Genomics in Ireland.

For studies 2 and 3, all neuropsychological test analysis and structural imaging analyses were undertaken by the author, with the help of Dr. Donna Cosgrove, Dr. Laurena Holleran and Dr. David Mothersill in reconstruction of MRI images.

All genotyping and genetics analysis reported in this thesis was carried out by Dr Derek Morris and colleagues in the Neuropsychiatric Research Group, Trinity College Dublin, and in the Cognitive Genetics and Cognitive Therapy Group, NUI Galway.

The supervisor, Graduate Research Committee, and local experts advised and provided support in conducting the research.
Funding

This thesis is principally funded by the Irish Research Council through a PhD scholarship. Recruitment and genotyping of Irish samples in Chapter 2 and 3 was supported by SFI (grants 12/IP.1359 and 08/IN.1/B1916) and the Wellcome Trust Case Control Consortium 2 project (grant 085475/B/08/Z) and the Wellcome Trust (grants 072894/Z/03/Z, 090532/Z/09/Z, and 075491/Z/04/B).
Acknowledgements

I would like to start by thanking my supervisor, Professor Gary Donohoe, who has been a reassuring and enduring supervisor throughout my PhD experience. Thank you for always being approachable, and for being so generous with your time over the past four years. I have gained invaluable experience in your lab and I feel lucky to have had you as a supervisor. I would further like to thank Dr Derek Morris for his input in all of my projects, and for always being a friendly face I could turn to when in doubt.

I am also grateful for the additional support and advice that was provided by my Graduate Research Committee: Dr. John Bogue, Dr. Mark Elliott and Dr. Padraig MacNeela.

I’d like to thank all ‘the PhD’s’ in the psychology department, past and present, who really made this experience the best it could be (G037 office in particular). The crazy conversations, laughing until we cried and celebrations (when possible) genuinely made my 4 years in Galway. Furthermore, the whole PhD process was definitely made easier by the supportive and collaborative atmosphere fostered in the department, and indeed in our research group: it has been a pleasure to work alongside the other students and researchers in the Centre for Neuroimaging and Cognitive Genomics, iRELATE and CogGene groups… I really enjoyed working with you all!

A special thanks to all my friends, especially ‘the two Sarahs’ who stuck with me through my many ups and downs. To Casey, who always provided me with clarity (and hilarity) in the face of hardship. A big thanks to my housemates Brid and Ciara- I couldn’t have asked for better people to live with during my PhD. I’m so fortunate to have met so many friends for life while in Galway.

Finally, and most importantly, I’d like to thank my family, for the warmth, love and support during my PhD. Particularly my sister Rebecca, my brother
Sam, and Aunts Josephine and Marie: thank you for everything. Thank you to Mum and Dad, who, as always, have gone above and beyond in keeping me spirited, strong and hopeful. I would like to dedicate this thesis to you.

Lastly, thank you to all of the participants who shared their time making this research possible.
Abstract

Schizophrenia (SZ) is a debilitating psychiatric disorder that afflicts approximately 1% of people worldwide. It is often characterized by hallucinations, sometimes by emotional blunting, and in a majority of cases by a decline in cognitive functioning. In understanding the aetiology of SZ (which is likely to involve both genetic and environmental factors), immune dysfunction has recently become an area of significant attention. This is in part because many recently identified genetic variants found to be associated with SZ have biological roles that are associated with immune function. This thesis sought to characterize the effects of immune-related genetic variants and immune-related environmental factors on neurocognitive outcomes associated with SZ.

My first study focused on Early Life Adversity (ELA) and immune markers, which have been associated with SZ. Immune processes, when perturbed, are thought to contribute to SZ pathophysiology, and some research suggests ELA could trigger this dysregulation. I investigated the association between ELA and cognitive deficits common in SZ, as well as the relationship between ELA and immune function, using data from the Avon Longitudinal Study of Parents And Children (ALSPAC). I created 4 ELA variables up to age 5, which were tested for association with a later measure of cognitive performance at ages 9 and 13. Counter to the hypothesis, I found no association between ELA measures and social cognition. However, ‘Harsh Parenting’, an ELA measure of physical discipline from mother to child was found to be associated with decreased IQ scores at age 9. Finally, although I found that lower IQ scores were associated with increased immune marker scores, these did not account for the relationship between ELA and IQ.

Following from this, I focused on genetic variants known to impact immune response (aside from the targeted immune markers in study 1), to further investigate how the immune system can affect cognitive deficits associated with SZ. Of various immune genes identified in recent research, the gene family associated with complement expression was chosen for further
analysis. Recent studies have suggested that structural variation within one complement gene, C4, may account for a large proportion of SZ risk implicated in many GWAS to date. Aside from C4, I wanted to investigate how other complement genes contributed to variation in IQ and cognitive domains associated with SZ risk.

For my second study I created a ‘complement’ gene set using various online databases, informed by recent research. This ‘complement’ gene set was tested for enrichment using the largest available GWAS dataset for SZ (36,989 cases, 113,075 controls). Based on the genetic overlap between SZ and cognition, gene-sets were also investigated for a genetic contribution to IQ using the largest available IQ GWAS dataset (N=269,867). Interestingly, this complement gene-set was enriched for the phenotype of IQ, but not enriched for the phenotype of SZ. Thus, a polygenic score (PGS) was created using the same ‘complement’ gene-set, for genes associated with IQ. This PGS was tested for an association with cognition in a dataset of ~1000 patients with SZ and healthy controls. The ‘complement’ PGS created was found to be positively associated with general cognitive ability, specifically premorbid IQ, whereby higher expression of complement PGS was associated with higher IQ. At a gene-set level, this may suggest that the wider complement pathway is more strongly associated with neurodevelopmental processes important to global cognitive development.

Given recent studies of complement genes and cognition in SZ (G Donohoe et al., 2018; G Donohoe, Walters, et al., 2013; C. Zhang, Lv, Fan, Tang, & Yi, 2017) a question remained about the possible relationship between SZ risk genes associated with complement and cognitive function. Complement genes have historically been related to synapse formation (Veerhuis, Nielsen, & Tenner, 2011), and more recently synaptic pruning in SZ (Sekar et al., 2016; Sellgren et al., 2019) thus brain development could be differentially impacted by complement expression. In order to further investigate this, the third study of this thesis examined the differential impact of C4 expression and a polygenic risk score (PRS) for ‘complement’ (using SZ GWAS data) on neurocognitive outcomes.
Firstly, this ‘complement’ PRS created using SZ GWAS data was seen to predict memory scores similar to a recent publication on C4 expression by our group (Donohoe, 2018). Thus, we wished to test for a differential relationship between the ‘complement’ PRS and C4 expression levels on brain volume in Irish participants. C4 was not found to be associated with any brain volume measure included. However, the ‘complement’ PRS was found to impact on hippocampal volume measures, albeit nominally. The association between complement genes and brain volume warrants further exploration in SZ patients only. On the whole, the role of the complement system aside from C4 may be relevant to the study of cognitive development as well as SZ pathology. Aspects of the immune system related to SZ incidence may be a joining point between genetics and environment, and both gene and environment influences are seen to impact cognition as well as diagnosis.
**List of Tables**

| Table 1.1: | Tests used in neuropsychological analysis in this thesis and evidence for association with SZ. | 15 |
| Table 1.2: | sMRI measures used in brain volume analysis in this thesis and evidence for association with SZ. | 22 |
| Table 2.1: | Demographic information and descriptive statistics for variables used in analysis. | 48 |
| Table 2.2: | Linear regression models for Early Life Adversity (Total [poly-victimisation] score, Harsh parenting, Domestic Abuse, Physical Abuse, Physical abuse and emotional abuse) and Theory of Mind (Triangles total) scores. | 50 |
| Table 2.3: | Linear regression models for Early life adversity (Total [poly-victimisation] scores, Harsh parenting, Domestic Abuse, Physical Abuse, Physical abuse and emotional abuse) and general cognitive ability (Full Scale IQ scores). | 52 |
| Table 2.4: | Linear regression models for ELA (Domestic Violence, Physical Abuse, Physical abuse, Emotional abuse, & Harsh parenting) predicting Immune activation (measured by log transformed CRP and IL6 levels). | 55 |
| Table 2.5: | Linear regression models for Immune activation (measured by log transformed CRP and IL6 levels) and General Cognitive Function (measured by Full scale IQ). | 57 |
| Table 3.1: | MAGMA gene-wide analysis of enrichment for Complement system genes for IQ. | 76 |
| Table 3.2: | Clinical and demographic information for the whole sample, and subsets of individuals based on diagnostic category. | 81 |
Table 3.3: Regression analysis of complement PGS IQ scores in patients and controls, using gender as a covariate for IQ analysis, age and gender as covariates in others.

Table 4.1: Demographics for whole sample of patients with psychosis and healthy controls

Table 4.2: Effect of SZ complement PGS (standardized z-score) on hippocampal volume, total brain volume, surface area and cortical thickness (age, gender and ICV as covariates)

Table 4.3: Hippocampal Subfield analysis for complement PRS score and hippocampal subregion volumes

Table 4.4: Effect of C4 Expression on hippocampal volume, total brain volume, surface area and cortical thickness (age, gender and ICV as covariates), in 150 individuals with C4 expression data in volumetric analysis
List of Supplementary Tables

**Supplementary Table 3.1.** List of 108 selected genes in the complement-related gene-set. 93

**Supplementary Table 3.2.** Analysis of associations between complement PGS and age, gender or years of education, positive/negative/disorganised symptoms or medication dose using correlational analysis. 95

**Supplementary Table 4.1.** List of 111 selected genes in the complement-related gene-set 118
List of Supplementary Figures

Figure 2.1: ALSPAC data collection timeline 64

List of Appendices

Appendix A: Publications Arising from this Thesis 154

Appendix B: List of Awards and Bursaries 155

Appendix C: List of Presentations 156

Appendix D: Study Funding and Acknowledgements 157
List of Abbreviations

C4 complement component 4
BBB blood brain barrier
CNS central nervous system
CRP C-reactive protein
DSM5 diagnostic and statistical manual of mental disorders 5th edition
ELA early life adversity
HLA human leukocyte antigen
IL6 interleukin 6
MIA maternal immune activation
MRI magnetic resonance imaging
GSA gene-set analysis
GWAS genome-wide association studies
LD linkage disequilibrium
LNS letter number sequencing
MAF minor allele frequency
MHC major histocompatibility complex
IQ intelligence quotient
OR odds ratio
PAL paired association learning
PGC psychiatric genomics consortium
PGS polygenic score
PRS polygenic risk score
QC quality control
RCT randomised control trial
SNP single-nucleotide polymorphism
SZ schizophrenia
ToM theory of mind
WAIS III Wechsler Adult Intelligence Scale 3rd Edition
WISC Wechsler Intelligence Scale for Children
WMS III Wechsler Memory Scale 3rd Edition
List of Definitions

**Adaptive immune system:** The adaptive immune system, also known as the acquired immune system or, more rarely, as the specific immune system, is a subsystem of the overall immune system that is composed of highly specialized, systemic cells and processes that eliminate pathogens or prevent their growth. Adaptive immunity creates immunological memory after an initial response to a specific pathogen, and leads to an enhanced response to subsequent encounters with that pathogen. The adaptive system includes both humoral immunity components and cell-mediated immunity components.

**Antibody:** An antibody is a large, Y-shaped protein produced mainly by plasma cells that is used by the immune system to neutralize pathogens such as pathogenic bacteria and viruses.

**Antigen:** Any substance foreign to the body that evokes an immune response either alone or after forming a complex with a larger molecule (as a protein), and that is capable of binding with a product of the immune response (such as an antibody).

**C-reactive Protein (CRP):** C-reactive protein is a protein made by the liver. CRP is classified as an acute phase reactant, which means that its levels will rise in response to inflammation. CRP levels in the blood increase when there is inflammation somewhere in the body.

**Complement:** The complement system is the major humoral mediator of innate immunity, which enhances the ability of antibodies and cells to clear microbes and damaged cells from an organism, promotes inflammation, and attacks the pathogen's cell membrane. The complement system can also be recruited and brought into action by antibodies generated by the adaptive immune system.

**Complement Component 4 (C4):** Complement component 4 (C4), in humans, is a protein involved in the intricate complement system, originating from the human leukocyte antigen (HLA) system. It serves a
number of critical functions in immunity, tolerance, and autoimmunity with the other numerous components. Furthermore, it is a crucial factor in connecting the recognition pathways of the overall system instigated by antibody-antigen complexes to the other effector proteins of the innate immune response.

**Cytokine:** Cytokines are a broad and loose category of small proteins that are important in cell signalling. Cytokines cannot cross the lipid bilayer of cells to enter the cytoplasm. Cytokines have been shown to be involved in autocrine, paracrine and endocrine signalling as immunomodulation agents.

**Haplotype:** A set of genetic determinants located on a single chromosome, inherited together from a single parent.

**Human Leukocyte Antigen (HLA):** The human leukocyte antigen system is a gene complex encoding the major histocompatibility complex (MHC) proteins in humans. HLA consists of a family of genes within the MHC on the short arm of chromosome 6 in humans. It contains over 200 genes, over 40 of which encode leukocyte proteins that distinguish self from non-self antigens.

**Innate immune system:** Innate immunity refers to nonspecific defense mechanisms that come into play immediately or within hours of an antigen's appearance in the body. These mechanisms include physical barriers such as skin, chemicals in the blood, and immune system cells that attack foreign cells in the body.

**Interleukin 6 (IL6):** Interleukin-6 (IL−6) is a molecule produced by different cells and tissues of the organism as a soluble mediator with a pleiotropic effect on inflammation, immune response, and hematopoiesis. It is involved in the production of acute phase proteins, with dysregulated continual synthesis of IL−6 associated with chronic inflammation and autoimmunity.

**Linkage disequilibrium (LD):** Linkage disequilibrium refers to the non-random association of alleles at two or more loci in a general population.
When alleles are in linkage disequilibrium, haplotypes do not occur at the expected frequencies.

**Macrophages**: A type of white blood cell of the immune system, that engulfs and digests cellular debris, foreign substances, microbes, cancer cells, and anything else that does not have the type of proteins specific to healthy body function.

**Maternal immune activation** (MIA): MIA refers to a maternal immune system triggered by infectious or infectious-like stimuli. A cascade of cytokines and immunologic alterations are transmitted to the fetus, resulting in adverse phenotypes most notably in the central nervous system.

**Major histocompatibility complex** (MHC): The major histocompatibility complex (MHC) is a set of genes that code for cell surface proteins essential for the acquired immune system to recognize foreign molecules. These cell-surface proteins are responsible for the regulation of the immune system in humans.

**Phagocytosis**: A process by which certain living cells called phagocytes ingest or engulf other cells or particles.

**Single-nucleotide polymorphism** (SNP): A single-nucleotide polymorphism (SNP, pronounced snip) is a DNA sequence variation occurring when a single nucleotide [adenine (A), thymine (T), cytosine (C), or guanine (G)] in the genome (or other shared sequence) differs between members of a species or paired chromosomes in an individual.
Chapter 1

Introduction
1.1 Schizophrenia

1.1.1 An overview of SZ

Schizophrenia (SZ) is a chronic and disabling neuropsychiatric disorder, affecting 1% of the world’s population. Susceptibility to SZ is multifactorial, with numerous causes related to polygenic risk for the disorder as well as environmental triggers (Insel, 2010). According to the Diagnostic and Statistical Manual of Mental Disorders, the following symptoms are required for a diagnosis of SZ: positive symptoms, additional sensory or perceptual changes as well as thoughts (Association, 2013), negative symptoms, such as avolition, asociality or diminished expression (Cohen, Forbes, Mann, & Blanchard, 2006) and disorganized speech or behavior. Contributing to illness severity, cognitive deficits are core to disability associated with in SZ, including deficits in IQ (Fujino et al., 2017; Woodberry, Giuliano, & Seidman, 2008), memory (Ricarte, Ros, Latorre, & Watkins, 2017), attention (K. Wang et al., 2005) and social functioning (Michael F Green, Horan, & Lee, 2015). The diagnosis of SZ requires persistent disturbance over at least a six month period, with symptoms usually first appearing in late adolescence or early adulthood (Bradford, 2009). Despite more than 100 years of SZ research, and over 100,000 scientific publications devoted to the subject, the causes of SZ remain poorly understood, and we still lack an objective diagnostic test for the disease. Without an understanding of the underlying pathophysiology, developing valid diagnostic criteria for SZ is an ongoing challenge.

1.1.2 Prevalence and Socioeconomic Burden

SZ lifetime incidence of 0.3% to 0.7% according to the DSM (McGrath, Saha, Chant, & Welham, 2008) while lifetime morbid risk of SZ has been estimated at 7.2 in 1000 people (Saha, Chant, Welham, & McGrath, 2005). Onset is typically in early adulthood, with an earlier onset in men than women by on average 3–4 years, at 20–28 years for men and 26–32 years for women (Castle, Sham, & Murray, 1998; Häfner et al., 1995). Furthermore,
the diagnosis appears to be 1.4 times more frequent in males than females (Castle et al., 1998).

While the incidence level of the disorder is relatively low, SZ is a major contributor to global burden of disease (C. J. Murray & Lopez, 1997), ranked as one of the top 25 causes of disability worldwide in 2013 (H. Y. Chong et al., 2016) This burden can be attributed to 2 factors associated with SZ: 1. The age of onset of the disorder (Insel, 2010) and 2. The chronic nature of the disorder, with up to two-thirds of individuals experiencing persistent and fluctuating symptoms (Association, 2013). The World Health Organisation estimated that costs associated with SZ account for between 1.6% and 2.6% of total healthcare expenditures in Europe (Barbato, 1998).

An Irish study of SZ estimated costs associated with SZ to be as high as 460.6 million Euro for the country in 2006 (Behan, Kennelly, & O'Callaghan, 2008), with the majority of expenditure falling into indirect costs such as loss in productivity and costs of care (H. Y. Chong et al., 2016). Individuals with SZ have a 2-3 times higher risk of death compared to the general population (McGrath et al., 2008). Furthermore, higher risk of suicide is reported in those with psychotic disorders, with depressive symptoms during FEP associated with increased longer-term risk of suicidal behaviour and self-harm (McGinty, Haque, & Upthegrove, 2018; Upthegrove et al., 2010).

1.1.3 SZ aetiology

One of the largest risk factors that predisposes people to SZ is genetic risk (Scolnick, 2017). Historically, observational studies as well twin and family studies suggest that there is a large familial component to SZ, with up to 80% of the genetic liability to SZ is inherited (Lichtenstein et al., 2009; Y. Wang et al., 2016). Heritability will be further discussed in section 1.2.1. Importantly, an inherited susceptibility to SZ due to genetic variation can be exacerbated by environmental influences (Bayer, Falkai, & Maier, 1999), with environmental effects on genetic liability to develop SZ estimated at 11% (McGue, Gottesman, & Rao, 1983; Rao, Morton, Gottesman, & Lew, 1981; P. F. Sullivan, Kendler, & Neale, 2003). Studies of SZ have found
that factors such as hypoxia, obstetric complications, low weight at birth and maternal infection were associated with an increased risk of SZ (Forsyth et al., 2012), and may contribute to a gene x environment interaction related to disease (Nicodemus et al., 2008). Other early life stressors including childhood abuse or neglect, urban upbringing or bullying were also found to contribute to SZ risk (Fisher et al., 2012; Van Os, 2004; van Os, Kenis, & Rutten, 2010), with familial risk exacerbated by early environmental factors such as parent-child relationships (Tienari et al., 1994). Possibly bridging the gap between genes and environment, altered immune response has been cited as a possible risk factor for SZ (Nimgaonkar, Prasad, Chowdari, Severance, & Yolken, 2017; Sellgren et al., 2019). This will be discussed further in section 1.1.6.

### 1.1.4 Treatment

Standard treatment of SZ consists of pharmacotherapy with dopamine receptor antagonists, using first and second generation antipsychotic drugs. Despite different classes of psychiatric drugs being among the pharmaceutical industry’s most profitable products, investment into new treatments for psychiatric disorders, including SZ, has significantly decreased. This is due to the molecular and cellular causes of SZ remaining largely unknown. However, there is growing interest in the potential utility of anti-inflammatory drugs in treating SZ (Müller, 2019).

Anti-inflammatory drugs have been used as add-on to antipsychotic treatment in patients with SZ, with some success in trials (Berk et al., 2013; Sommer et al., 2013). In a broad meta-analysis of anti-inflammatory agents Sommer et al (2013) found a significant beneficial effect on PANSS total score of treatment with aspirin. In another meta-analysis (Nitta et al., 2013), the COX-2 inhibitor Celecoxib has shown efficacy in significantly reducing PANSS positive symptoms, in particular, in early stages of disease (Müller, 2019). However, a RCT of Tocilizumab, which is a humanized IL-6 receptor monoclonal antibody, was carried out in 36 patients with residual symptoms of SZ (Girgis et al., 2018). The results showed no effect of tocilizumab on any behavioral outcome, while CRP decreased and IL-6 and
IL-8 increased. Finally minocycline, a drug thought to produce anti-inflammatory effects has also produced mixed results in research. Meta-analysis of previous research reports an overall reduction in positive symptoms for those assigned to receive minocycline (Solmi et al., 2017; Xiang et al., 2017). Although many studies suggest a reduction in positive symptoms with this treatment, large RCTs of this anti-inflammatory antibiotic such as (Deakin et al., 2019) and (Weiser et al., 2019) report no effect on the PANSS scores, positive or negative subscale or other symptoms of SZ.

One randomized controlled trial (RCT) with aspirin treatment showed stronger effects when stratifying participants on the basis of a marker of inflammation (Laan et al., 2010). No selection procedure based on elevated baseline inflammation measures was implemented in many of these trials, and the authors of several studies underlined the possibility that enriching the sample with inclusion of only individuals with an elevated CRP could have influenced the results (Girgis et al., 2018; Kroken, Sommer, Steen, Dieset, & Johnsen, 2018) This is because the immune system may have differing inflammatory signatures in SZ, caused by diverse immune involvement in the disease.

1.1.5 Neurodevelopmental Hypothesis

According to the neurodevelopmental hypothesis, the aetiology of SZ may involve aberrant brain development, caused by both genetic and environmental factors that come into effect before the brain approaches maturation. These neurodevelopmental abnormalities, developing in utero for some and thereafter for others, have been suggested to lead to the activation of maladaptive neural circuits during adolescence or young adulthood (sometimes owing to severe stress), which leads to the emergence of positive or negative symptoms. A “2-hit” model proposed by Keshavan (Keshavan, 1999) works within the framework of the neurodevelopmental theory, suggesting that maladaptive development during 2 critical time points (early brain development and adolescence) combine to produce the symptoms associated with SZ. According to this model, early
developmental insults may lead to dysfunction of specific neural networks that would account for premorbid signs and symptoms observed in individuals that later develop SZ.

Substantial evidence supports a neurodevelopmental hypothesis of SZ, as research indicates that cognitive deficits usually precede the onset of the full-blown clinical presentation of the disease, implying that early cognitive deterioration might represent an important risk factor or prodromal condition of SZ (Kahn & Keefe, 2013). These studies have yielded a consistent picture in which “approximately 50–70% of the offspring of parents with SZ manifest a range of observable difficulties including socioemotional, cognitive, neuromotor, speech-language problems, and psychopathology, and roughly 10% will develop psychosis.” (Liu, Keshavan, Tronick, & Seidman, 2015). At adolescence, excessive elimination of synapses and loss of plasticity may account for the emergence of symptoms (Insel, 2010).

1.1.6 The Immune Hypothesis

Building on a neurodevelopmental model of SZ, the ‘immune hypothesis’ of SZ posits that increased risk for SZ results from early environmentally-driven immune challenges (e.g. maternal viral load) and/or genetically driven suboptimal immune responsiveness, manifesting itself during gestation or the perinatal period (Fatemi & Folsom, 2009). Adding to this, immune reactions to psychological stress during early developmental periods can stimulate perturbations in neuronal activity that could otherwise be endured without long-term psychological consequence. An immune-based ‘two-hit model for psychosis’ has been proposed (Bergink, Gibney, & Drexhage, 2014; Feigenson, Kusnecov, & Silverstein, 2014). This model suggests that early stimulation of immune reactions results in cytokine-based developmental brain abnormalities that are then exacerbated by further abnormal immune activation and inflammation in adolescence and adulthood. Uncovering the precise mechanism by which altered immune function results in a developmental cascade of changes is the focus of much current research. This will be further discussed in section 1.4.
1.2 Genetics of SZ

1.2.1 Heritability of SZ

Heritability refers to the proportion of variation in a population that can be attributed to additive genetic or total genetic contribution (Visscher, Hill, & Wray, 2008). Heritability of SZ has been estimated as high as 79% (Hilker et al., 2017), with the classic twin design the most popular method for quantifying the variance of a trait attributable to genetic and environmental factors. MZ twins share approximately 100% of their genetic information, whereas DZ twins share on average 50%. For monozygotic twins where one twin has SZ, the risk for the unaffected twin is approximately 50% (Lichtenstein et al., 2009). In an overview of five twin studies, SZ concordance rates are between 41% and 65% for MZ twins and 0% to 28% for DZ twins (Cardno & Gottesman, 2000), suggesting a strong genetic origin. Children of two affected parents carry a 27% risk for SZ (Gottesman, Laursen, Bertelsen, & Mortensen, 2010). For adopted children with an affected biological parent, risk remains much higher than the general population, further supporting a role of genetics in SZ aetiology (Gottesman et al., 2010). Despite the strong evidence for genetic contribution, variation directly attributing to SZ has proved challenging to find. This could be due to several limitations of using twin studies to estimate the heritability of SZ (Rosenthal, Wender, Kety, Welner, & Schulsinger, 1971; E Fuller Torrey & Yolken, 2019). As twin researchers themselves have noted, heritability of SZ could be equally attributed to birth complications (Reveley and Reveley, 1987), with twins in general having more birth injuries, a higher mortality rate and lower birthweights (Eagles, 1994). In approximately 15% of MZ twins one twin gets more blood than the other, producing an unequal exposure to hormones, drugs, and infectious agents coming from the maternal circulation (Edwin Fuller Torrey, Bowler, & Taylor, 1994; E Fuller Torrey & Yolken, 2019). Finally, a further caveat is that the twin method assumes the effects of genes and environment do not interact (Tenesa & Haley, 2013). For SZ, however, it is assumed that such interactions do occur (E Fuller Torrey & Yolken, 2019). These caveats are
important to consider when considering heritability. Other approaches to the genetic study of SZ are indeed easier to interpret.

1.2.2 Candidate Gene studies

To try to identify candidate genes that may have a role in the aetiology of SZ, the candidate gene approach has been employed in early SZ research—using knowledge of SZ aetiology to generate hypotheses about a possible genetic origin (Tabor, Risch, & Myers, 2002). A problem with this approach is that we have insufficient understanding of the aetiology of SZ, making candidate gene selection difficult. Despite this, in 2008 the SZ Gene database (Allen, Larøi, McGuire, & Aleman, 2008) reported on 1,406 studies that have identified over 700 candidate genes for SZ. The database was updated to include 1,008 candidate genes identified in studies published before 2012 (http://www.szgene.org/). In a recent study, researchers used data from largest schizophrenia genome-wide association study (GWAS) conducted to date as input to a gene set analysis to investigate whether variants within SZ candidate genes are enriched for association with SZ (Johnson et al., 2017). This analysis found that variants in the most-studied candidate genes were no more associated with SZ than were variants in control sets of non-candidate genes.

1.2.3 GWAS

As hypothesis-driven candidate gene studies of SZ could not uncover a central ‘SZ gene’, a hypothesis-free approach to analysing genes driving SZ risk was created, called Genome Wide Association Study (GWAS) research. This approach involves mapping the entire genome and searching for individual variants associated with disease by focusing on associations between single-nucleotide polymorphisms (SNPs) and specific diagnoses or traits like SZ or IQ in large populations. Each SNP is tested for association with the phenotype of interest in a GWAS, with ability to identify risk variants depending upon several factors such as effect size, sample size and population frequency of the variant (Bergen et al., 2012). Stringent multiple correction thresholds are necessary for the amount of comparisons made in
GWAS: a p-value threshold in and around $5 \times 10^{-8}$ derived from a Bonferroni correction for about 550,000 observations is often used (G Donohoe, Deary, Glahn, Malhotra, & Burdick, 2013). The advantage of GWAS is that there is no a priori biological hypothesis required: the whole genome is examined for associations with SZ, overcoming the obstacle of unclear disease pathology (Kitsios & Zintzaras, 2009).

From advancing GWAS research it is clear that SZ has a highly complex genetic architecture, with between 23% and 50% of variance explained by thousands of independent common genetic variants (I. S. G. Consortium & 2, 2012; Lee et al., 2012). Ten years ago, the first use of big data to create a SZ GWAS found that SNPs corresponding to the major histocompatibility complex (MHC), on chromosome 6p21.3–22.1, and the 1q24.2, and 18q21.2 genomic regions were associated with SZ risk (Purcell et al., 2009; J. Shi et al., 2009; Stefansson et al., 2009). With increased power and numbers in subsequent years, a higher detection rate unveiled avenues for research by hinting at specific biological pathways. In 2014, advances in sample size available led to a ground-breaking GWAS of 36,989 cases and 113,075 controls, in which 108 SNPs were identified as related to SZ diagnosis (Ripke et al., 2014) with a SNP representing MHC remaining the top hit. Following this, the most recent and largest GWAS publication provides consistent results using an independent GWAS sample (Pardiñas et al., 2018), with a SNP representing the MHC region identified as the top associated variant. Further information on the MHC region provided in section 3.7.

1.2.4 Polygenic Risk Profiling and LD score regression

When testing for association with a phenotype such as SZ, only a small proportion of all SNPs reach genome-wide significance, but in aggregation, genes may contribute substantially to disease susceptibility. Methods such as the polygenic risk score (PRS) calculation have been employed to detect an overlap of multiple common variants associated with SZ and different phenotypes, even if variants are not individually statistically significant in GWAS (Coelewij & Curtis, 2018). Several studies have shown that such
PRS scores differ between patients and controls, thus providing a useful tool to measure genetic liability to psychosis in independent samples (Derks et al., 2012; Purcell et al., 2009). A recent review of the use of PRS in SZ research has identified 31 articles examining association with another phenotype (Mistry, Harrison, Smith, Escott-Price, & Zammit, 2018). PRS can be generated across multiple P-value thresholds and can be used to predict case-control status. PRS can also be tested for association with other phenotypes, such as cognition, in order to further understand the functionality of already identified SZ variants. Building on this, association analysis of a group of functionally related genes, called pathway-based analysis, has been proposed to boost statistical power and improve interpretability over gene-based analysis for GWAS (Wu & Pan, 2018).
1.3 Cognition in SZ

As a tractable target for intervention, lessening cognitive deficits could impact on the daily lives of those with SZ. Not only is cognitive impairment a core symptom of SZ, it is also an important correlate of functioning, with cognition strongly predicting disability in SZ (Michael Foster Green, Kern, Braff, & Mintz, 2000; Michael F Green, Kern, & Heaton, 2004). Moreover, cognitive ability predicts functional outcome in employment, independent living and maintenance of relationships in those with a SZ diagnosis (Lewandowski, Cohen, & Öngur, 2011). Individuals with SZ show widespread deficits in neurocognitive functions, including memory, attention, verbal learning, and executive functioning (Aleman, Hijman, De Haan, & Kahn, 1999; Kahn & Keefe, 2013). Furthermore, general cognitive ability is often impaired, with SZ patients scoring 1-2 standard deviations in IQ below healthy controls (Reichenberg, 2005). However, it is unclear how neurobiological mechanisms contributing to SZ also influence this cognitive decline (Danese et al., 2016).

A neurodevelopmental hypothesis of SZ is supported by the neurocognitive lag observed among children and adolescents reporting psychotic experiences (Horwood et al., 2008; Golam M Khandaker, Stochl, Zammit, Lewis, & Jones, 2014; Rossi et al., 2016; S. A. Sullivan, Thompson, Kounali, Lewis, & Zammit, 2017; Thompson et al., 2011), which might reflect the common genetically determined neurodevelopmental origin of aberrant neurocognition (Hatzimanolis et al., 2015). As well as this, non-affected siblings of probands perform on an intermediate level between patients with SZ and healthy controls across the cognitive domains of memory and attention (Egan et al., 2001). The same was found to be true with regard to Theory of Mind, a measure of social cognition (Bora & Murray, 2013; Bora & Pantelis, 2016).

1.3.1 Cognition as an Endophenotype for SZ

Historically, an endophenotype approach has been adopted in SZ research (Gottesman & Gould, 2003). Endophenotypes are heritable and quantitative traits such as cognitive function and brain structure common in a disorder.
Endophenotypes are said to be under stronger genetic rather than environmental influence, and were considered to be less complex and heterogeneous than the overarching disease (Glahn et al., 2014). The combination of genes required to produce these simpler phenotypes is potentially fewer than those required to produce the full range of symptoms included in the ‘fuzzy’ construct of SZ (Braff & Tamminga, 2017). Furthermore, the fact that endophenotypes are quantitative variables (e.g. IQ score, grey matter volume) rather than qualitative or dichotomous (e.g. ‘healthy’ vs. ‘SZ’) provides more power for statistical analysis (Glahn et al., 2014). Traits such as cognition and brain volume were thought to be more stable to measure, with affects preceding core symptoms of the disorder, and are typically assessed by laboratory-based methods rather than observation (Iacono, Malone, & Vrieze, 2017). Using this approach in research, neuroimaging and neurocognitive testing are employed to study brain activity in light of recent discoveries related to genetic risk factors in SZ patients (Insel, 2010). The cognitive endophenotypes included in this thesis include 1) general cognitive ability, 2) Episodic and visual memory, 3) working memory, and 4) social cognition.

1.3.2 General Cognitive ability

General cognitive ability- a broad definition of intelligence, refers to an individual’s ability to reason, plan and problem solve, to engage in abstract thought, and to learn quickly and from their environment (Gottfredson, 1997). Although it is plausible that each of cognitive deficit in SZ is subserved by distinct psychological and neural systems, this explanation is inconsistent with data that points to a more generalized cognitive deficit (Dickinson, Iannone, Wilk, & Gold, 2004; Dickinson, Ragland, Gold, & Gur, 2008; Nuechterlein et al., 2004). Structural equation modelling has demonstrated that a SCZ-related deficit in neuropsychological performance is largely mediated through a general ability factor that accounts for 63.3% of diagnosis-related variance (Dickinson et al., 2008). Only verbal memory (13.8%) and processing speed (9.1%) also accounted for direct effects, albeit at significantly reduced magnitude. Similarly, a meta-analysis of
cognitive deficits in SCZ indicates a generalized impairment that cuts across domains (Fioravanti, Bianchi, & Cinti, 2012). In a factor analysis of eighteen subtest scores from the WAIS and WMS III, two thirds of the overall effect of SZ diagnosis on cognitive ability was explained by a single factor. The authors concluded that a generalized cognitive deficit is a core feature of SZ, rather than specific independent domain effects (Dickinson et al., 2004). The Wechsler Adult Intelligence Scale (WAIS) is the most commonly used 18 scale to measure IQ. Version 3 (WAIS III) was developed in 1997 and includes measures of performance IQ, verbal IQ, full scale IQ as well as secondary tests measuring verbal comprehension, processing speed, and perceptual organisation (Wechsler, 1997). Four subtests of the WAIS III battery have been included in this thesis; full scale IQ, performance IQ, verbal IQ and the Wechsler Test of Adult Reading (WTAR) (Holdnack, 2001), which together, take approximately 30 minutes to complete. A shortened version of the WISC 3rd U.K. Edition, which measures intellectual ability of children from 6 to 16 years is also used in analysis for the first study (Kaufman, 1994). All cognitive outcome measures are in table 1.1.

1.3.3 Episodic and Visual Memory

Episodic and visual memory systems are both impaired in patients with SZ (Aleman et al., 1999). Episodic memory is a memory system that allows people to consciously recall past memories and experiences. The tests used in this thesis to measure immediate and delayed episodic memory include the Logical Memory (I and II) subtests from the WAIS III (Wechsler, 1997). Furthermore, PAL (paired associations learning (I and II) provide an estimation of immediate and delayed visual memory, a type of declarative memory (Wechsler, 1997). See table 1.1. For analysis in studies 2 and 3, these memory test scores were combined in an unrotated principle components to reduce multiple testing burden. This memory factor explained 72% of variance in memory scores, as described previously (G Donohoe et al., 2018).
1.3.4 Working Memory

Working memory refers to a cognitive system with a limited capacity that allows the temporary storage and manipulation of information (Baddeley, 2000). This information updating and monitoring is another component of executive function (Miyake & Shah, 1999). Patients with SZ have shorter spatial spans compared to controls, suggesting an impairment in working memory ability (Hutton et al., 1998). The Letter Number Sequencing (LNS) subtest from the WMS III (Wechsler, 1997) and the spatial working memory task from the Cambridge Neuropsychological Test Automated Battery (CANTAB SWM) (Barnett et al., 2010; J. Fray, W. Robbins, & J. Sahakian, 1996) are the measures used to assess working memory in this thesis. See table 1.1.

1.3.5 Social Cognition

Social cognition refers to the mental operations that underlie social interactions, including perceiving, interpreting, and generating responses to the intentions, dispositions, and behaviours of others (Michael F Green et al., 2015). A meta-analysis of social cognition found that deficits in social cognition predicted a larger proportion of variation in social and occupational functioning in SZ than any other aspect of cognition (Fett, Viechtbauer, Penn, van Os, & Krabbendam, 2011). Furthermore, social cognitive deficits have been found to relate to genetic risk and altered brain activity (Mothersill & Donohoe, 2016), and these deficits relate directly to environmental influences (van Os et al., 2010). It has been shown that both first episode and chronic SZ patients perform more poorly than healthy participants in tasks measuring ToM, although it is more severe in the acute phase of SZ (Bora & Pantelis, 2016). Finally, unaffected relatives perform on an intermediate level between patients with psychosis and healthy controls (Bora & Murray, 2013; Bora & Pantelis, 2016). Social cognitive test analysed in this thesis include measures of Theory of Mind (ToM), called the triangles task. See table 1.1.
Table 1.1 Tests used in neuropsychological analysis in this thesis and evidence for association with SZ.

<table>
<thead>
<tr>
<th>Cognitive Domain</th>
<th>Neuropsychological Test</th>
<th>Reference</th>
<th>Sample</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Cognitive ability (IQ)</td>
<td>WISC- IQ</td>
<td>(Fagerlund, Pagsberg, &amp; Hemmingsen, 2006)</td>
<td>• First-episode psychotic adolescents (N = 39)</td>
<td>Analyses of WISC-III factor profiles suggested that early onset schizophrenia patients may have more global IQ deficits</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Healthy control group (N = 40) between the ages 11 and 17</td>
<td></td>
</tr>
<tr>
<td>Social Cognition</td>
<td>Triangles task</td>
<td>(Russell, Reynaud, Herba, Morris, &amp; Corcoran, 2006)</td>
<td>• Patients with SZ (N=61)</td>
<td>significant differences in intentionality scores on the triangles task where healthy controls scored higher than patients with SZ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Healthy controls (N=22)</td>
<td></td>
</tr>
<tr>
<td>Study 2</td>
<td>Premorbid IQ</td>
<td>WTARR Wechsler Test of Adult Reading (Raw) (Woodberry et al., 2008)</td>
<td>Meta-analysis, 18 studies</td>
<td>Overall, schizophrenia samples demonstrated a reliable, medium-sized impairment in premorbid IQ.</td>
</tr>
<tr>
<td>Study 3</td>
<td>Episodic Memory</td>
<td>LM1 Wechsler Memory Scale III: Logical Memory 1, (immediate/delayed) (Toulopoulou, Rabe-Hesketh, King, Murray, 2006)</td>
<td></td>
<td>Patients with SZ N=36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Healthy control N=36</td>
</tr>
</tbody>
</table>

- Patients with SZ N=62
- Healthy SZ relatives N=98

Schizophrenic patients performed significantly worse than controls on nearly all measures of
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CANTAB SWM Cambridge Neuropsychological Test Automated Battery: Spatial Working Memory (errors) (Forbes, Carrick, McIntosh, &amp; Lawrie, 2009)</td>
<td></td>
<td>Overall, schizophrenia samples demonstrated a deficit in working memory.</td>
</tr>
<tr>
<td></td>
<td>The Letter Number Sequencing (LNS) subtest from the WMS III (Wechsler, 1997b) (Nuechterlein et al., 2004)</td>
<td></td>
<td>Patients with SZ N=97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Healthy controls N=87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Significant deficits in LNS in patients with SZ.</td>
</tr>
</tbody>
</table>
Genetic architecture of cognition in SZ

A number of research groups have discovered that multiple risk variants implicated in SZ are associated with variation in cognitive performance. Family studies have identified a phenotypic relationship between SZ and neurocognition, suggesting an overlap between genetic factors inducing cognitive deficits and increased SZ susceptibility. In a study of 267 twins concordant and discordant for SZ diagnosis performed by (Timothea Toulopoulou et al., 2007), a negative correlation of $r = -0.61$ was observed between intelligence and SZ. Furthermore, the shared genetic variance between intelligence and SZ was estimated to be 92% A larger study using 657 SZ cases, 674 first-degree relatives and 725 controls further supported this, estimating approximately 89% of the correlation between SZ and general cognitive ability was attributable to genetic factors (Timothea Toulopoulou et al., 2010). In addition, authors also showed that genetic factors contributing to the phenotypic correlation between SZ and immediate and delayed recall were 72% and 86% respectively (Toulopoulou et al., 2010). In a later study of 1986 MZ and 2253 DZ Swedish twins, a substantially lower phenotypic correlation between IQ and psychosis of just $r = -0.11$ was observed, although 91% of this was due to shared genetic influences, which was similar to the previous estimations (Fowler, Zammit, Owen, & Rasmussen, 2012). However, this study provided additional details showing genetic factors associated with psychosis are largely independent of IQ; only 6.8% was shared.

Due to the close relationship between neurocognitive deficits and psychiatric illness, nationwide collaborative efforts have joined forces in order to investigate the potential contribution of common genetic variation to neurocognitive performance, which may also lead to the identification of specific genetic loci involved in psychiatric (G Donohoe, Walters, et al., 2013; Lencz et al., 2014). A meta-analytical study showed high genetic overlap between cognitive phenotypes and SZ risk, also confirming high heritability of various cognitive domains, like IQ (Blokland et al., 2016). A higher PRS for SZ has been associated with poorer performance in cognitive
tasks measuring IQ and memory in cases and to a lesser extent, in controls (Mistry et al., 2018; Nicodemus et al., 2008), suggesting that the polygenic contribution to SZ risk is associated with poorer cognitive functioning. Using PRS analysis, a SZ polygenic risk score was associated with lower performance IQ and full-scale IQ in a sample of n = 11 853 (Hubbard et al., 2015). This finding was attenuated when cases were examined alone, suggesting an overall effect of SZ risk genes on cognition van (Van Scheltinga et al., 2013). Higher SZ-PRS was also associated with worse cognitive ability in people with SZ within the Cognitive Genomics Consortium (Lencz et al., 2014). This finding has been replicated in several individual samples, for instance the Lothian birth cohort, which showed a relationship between higher PRS and cognitive decline and ability at age 70 (McIntosh et al., 2013), and the ALSPAC birth cohort which tested IQ at varying ages in childhood and adolescence (Hubbard et al., 2015).

1.3.6 Medication effects on cognition

Memory and attention problems are present even in medication naïve, first-episode patients; independent of any possible medication effects (Torrey, 2002). Pharmacological treatment has even been reported to improve cognitive function in patients with severe symptoms (Mirsky et al., 2000), and medication adherent patients score higher on measures of IQ than those who are non-adherent (El-Missiry et al., 2015). Furthermore, adjuvant immunomodulatory drugs trials have reported benefits to cognition, opening new windows to tackling these deficits in SZ (Levkovitz et al., 2010). Of these, minocycline seemed to have the most promising effects (Levkovitz et al., 2010). However, in the largest double-blind, placebo-controlled trial of minocycline to date, this drug was not found to have an impact on cognition or negative symptoms (Deakin et al., 2018). Overall, evidence indicates that, at best, there is a nominal effect of antipsychotics or anti-inflammatory drugs on the improvement of cognitive symptoms, with unequivocal evidence suggesting no effect (Keefe et al., 2007).
1.4 Brain Structure in SZ

As well as the aforementioned effects of SZ-related genetic variants on cognition, widespread structural alterations in the brains of SZ patients have been identified (described in Table 1.2). Structural brain imaging studies have shown subtle, almost universal, decreases in grey matter, enlargement of ventricles, altered amygdala and hippocampal size, and additional subcortical structure alterations in individuals with SZ compared to healthy controls (Hartberg et al., 2011; Iglesias et al., 2015; R. Murray, Lewis, & Reveley, 1985; Rimol et al., 2010; Rimol et al., 2012). Research implicates excessive loss of grey matter (Cannon et al., 2002) and numbers of synapses on neurons (Cannon et al., 2015) as pathogenic mechanisms leading to these brain volume changes, which could be influenced by genetic variation (Cannon et al., 2015; Sekar et al., 2016).

Furthermore, comparable neural variations (decreased hippocampal and grey matter volume, increased ventricles) have been observed in the unaffected siblings of SZ probands (Boos, Aleman, Cahn, Pol, & Kahn, 2007). These individuals carry a similar genetic and environmental risk set compared to affected siblings; albeit at a sub-pathological threshold (Moran, Hulshoff Pol, & Gogtay, 2013). A review of genetic influences on human brain structure by (Peper, Brouwer, Boomsma, Kahn, & Hulshoff Pol, 2007) found that the heritability of the hippocampus is 40 – 69%, with cortical surface area, thickness and total intracranial volume also highly heritable (89%, 81% and 78% respectively) (Panizzon et al., 2009). Finally, Some of the grey matter abnormalities present in patients with SZ are also thought to predate the onset of psychosis (Pantelis et al., 2003), with volume reductions in amygdala and hippocampus and prefrontal cortex observed in schizotypal high risk subjects (McIntosh, Harrison, Forrester, Lawrie, & Johnstone, 2005; McIntosh et al., 2004), suggesting an underlying neurodevelopmental origin.

This thesis uses structural MRI measures as an outcome variable in the third study, thought to be differentially impacted in SZ patients compared to controls. See table 1.2.
Table 1.2 sMRI measures used in this thesis, and studies of differences in psychosis.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Citation</th>
<th>Sample</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Brain Volume</td>
<td>(Hajima et al., 2012)</td>
<td>Meta-analysis of 33 studies</td>
<td>Significant overall effect sizes were demonstrated for volume reduction of whole brain in patients with SZ.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface Area</td>
<td>(Rimol et al., 2012)</td>
<td>• 173 SZ</td>
<td>Patients with schizophrenia had a significant deficit in cortical surface size and cortical area reductions in circumscribed regions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 207 HC</td>
<td></td>
</tr>
<tr>
<td>Cortical Thickness</td>
<td>(Van Erp et al., 2018)</td>
<td>• 4474 SZ</td>
<td>There was widespread cortical thinning in schizophrenia compared with control subjects.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 5098 HC</td>
<td></td>
</tr>
<tr>
<td>Hippocampal Volume</td>
<td>(Van Erp et al., 2016)</td>
<td>• 2028 SZ</td>
<td>Patients with SZ had smaller hippocampal volume than healthy controls, among many reductions in subcortical volume.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 2540 HC</td>
<td></td>
</tr>
</tbody>
</table>

HC= Healthy Control subjects; SZ=Patients with Schizophrenia.
1.5 Neuroimmunology

The immune hypothesis rests on converging lines of evidence that indicate the presence of immune abnormalities among patients with SZ. Firstly, the largest and most robust genetic association with SZ is found at the major histocompatibility complex (MHC), a genomic region including loci that influence the immune response (Pardiñas et al., 2018; Ripke et al., 2014). Secondly, epidemiological studies have cited a link between prenatal and childhood infection leading to increased incidence (Brown, 2011; Golam M Khandaker, Zammit, Lewis, & Jones, 2014). This observation was further supported by epidemiological studies showing mothers infected with the influenza virus in pregnancy are more likely to go onto having children who develop SZ (Mednick, Machon, Huttunen, & Bonett, 1988), and children born in an urban centre, which may have more exposure to infectious agents are more likely to develop SZ. Thirdly, serological studies have shown elevated levels of immune markers in the blood of SZ patients, including C-reactive protein and Interleukin 6, are linked to SZ incidence and cognitive impairment. Fourthly, studies of immune genes informed by GWAS discovery (Ripke et al., 2014) and post-mortem studies (Fillman et al., 2013; Sekar et al., 2016; Trepanier, Hopperton, Mizrahi, Mechawar, & Bazinet, 2016) have all pointed to increased inflammatory response and differential immune reactivity associated with SZ risk. Finally, the prevalence of autoimmune disorders and atopic conditions is increased in patients with SZ and in their first-degree relatives, suggesting a shared immune aetiology between disorders (Benros et al., 2013; Eaton et al., 2006; Stringer, Kahn, de Witte, Ophoff, & Derks, 2014). This proinflammatory status is thought to result from the interaction between genetic vulnerability and environmental factors such as infections, trauma, nutrition, and stress (Fineberg & Ellman, 2013). Taken together, this evidence has led to the hypothesis that—in at least a subset of patients—SZ may be a neuroimmune disorder mediated by alterations in pro- and anti-inflammatory processes in the central nervous system (CNS) (O. Howes & McCutcheon, 2017; O. D. Howes,
Although many indirect links between abnormal immune reaction and SZ risk have been made, it is important to note that attempts to identify a specific infectious agent or an antibody directed against CNS tissue have not produced a consistently replicable finding (Kirch, 1993). Ultimately, it may be that multiple infectious agents, and even multiple autoantibodies, contribute in different ways and in different cases to the development of a SZ syndrome. The clinical heterogeneity of SZ is an inescapable fact that must be acknowledged by accepting the very real possibility of etiologic heterogeneity (Kirch, 1993). Furthermore, other serious mental illnesses, for example depression, show similar physiological alterations, therefore these findings may not be specific to psychosis or psychotic disorders (Pillinger, D’Ambrosio, McCutcheon, & Howes, 2019).

1.5.1 Defining the neuroimmune system

The central nervous system has been historically considered ‘immune-privileged’, with much research dismissing immune surveillance in the brain (Birnbaum et al., 2018). This was due to the supposition that the blood brain barrier was not permeable, and therefore the brain could not be influenced by periphery pathogens. Advances in neuroimmunology have allowed us to uncover various processes of the neuroimmune system—a system involving the biochemical and electrophysiological interactions between the nervous system and immune system which protect neurons from pathogens. These influences on the CNS are useful in homeostasis, as the brain can create anti-inflammatory responses based on input from peripheral signals. Conversely, the neuroimmunogenic architecture may work in overtime and produce excess inflammatory response to this plethora of external signals to the immune system. Its effects on monoamine metabolism, neuroendocrine function, and synaptic plasticity have all been proposed as potential pathways mediating the link between inflammation and SZ (Haroon, Raison, & Miller, 2012). Interestingly, acute infection has also been associated with psychotic relapse (Miller et al., 2013), suggesting that immediate activation of immune response exacerbates symptoms.
1.5.2 Immune System processes relevant to this thesis

‘Immunity’ refers to the global ability of the host to resist the predation of microorganisms that would otherwise destroy it (Hoebe, Janssen & Beutler, 2004). The immune system can be divided into two subcategories: ‘adaptive immunity’ and ‘innate immunity’. The innate immune system provides mechanisms for the rapid sensing and elimination of pathogens, and creates an immediate immune response (Bonilla and Oettegen, 2010). On the other hand, the adaptive immune system is composed of highly specialized cells which evolved to provide a broader and more finely tuned recognition of foreign bodies, differentiating which belong to the host, and which come from other sources (Bonilla and Oettegen, 2010). Furthermore, adaptive immunity creates immunological memory: after an initial response to a specific pathogen, the adaptive immune system creates an enhanced response to subsequent encounters with that pathogen. As explained in this section, studies of innate and adaptive immunity have suggested an immune dysregulation is associated with SZ incidence, however the mechanism by which this occurs remains unclear. Further adding to the mystery of this association is that each component of the immune system fulfils multiple functions, and mediators of immune response often engage in cross-talk between each subsystem.

The innate immune system shows signs of overactivation in SZ, as indicated by increased inflammatory response in unmedicated patients (Müller, Riedel, Ackenheim & Schwarz, 1999). To assess the presence of inflammation in a clinical setting, laboratories routinely assess the plasma concentrations of various acute phase proteins as biomarkers of inflammation (Slaats, ten Oever, van de Veerdonk, & Netea, 2016), such as C-reactive protein (CRP) which a protein produced in the liver in an innate immune response (Hurlimann, Thorbecke, & Hochwald, 1966). CRP is produced in response to inflammation as triggered by Interleukin-6 (IL-6) a molecule manufactured by different cells and tissues of the organism as a soluble mediator of innate immunity. Elevated levels of CRP in the bloodstream have been associated with SZ incidence (see section 1.5.7), and dysregulated continual synthesis of IL-6 has been associated with chronic
inflammation and autoimmunity, which often co-occur with SZ (see section 1.5.3). Both IL6 and CRP are used as measures of immune activation in study 1 of this thesis.

The complement system, which is a major humoral mediator of innate immunity, not only participates in inflammation but also acts to enhance the adaptive immune response (see figure 1) (Janeway et al., 2001). The complement system consists of roughly 30 proteins and protein fragments that are activated by inflammation and synthesized by the liver, releasing cytokines (IL6) and initiating an amplifying cascade of further complement activity (Janeway et al., 2001). Furthermore, CRP is thought to assist in complement binding to foreign and damaged cells, enhancing phagocytosis by macrophages (see definitions page xviii). The end result of this complement activation cascade is stimulation of phagocytes to clear foreign and damaged material, inflammation to attract additional phagocytes, and activation of the cell-killing membrane attack complex (Janeway et al., 2001). Complement component 4 (C4), a specific component of the complement system associated with SZ incidence, serves a number of critical functions in connecting the recognition pathways of the overall complement system instigated by antibody-antigen complexes to the other effector proteins of the innate immune response (Sekar et al., 2016) (discussed in section 1.5.10). The complement system is analysed in studies 2 and 3 of this thesis.

**Figure 1**: Schematic of immune reactivity relevant to this thesis: Interleukins (IL6) as humoral mediators of innate immune response to inflammation, which trigger CRP and complement synthesis in the liver. These proteins then signal the complement cascade and phagocytosis as well as cell-to-cell immune reactions. Just as resistance to disease can be innate or acquired, the mechanisms mediating it can be correspondingly divided into innate and adaptive, each composed of both cellular and humoral elements (i.e. free in serum or body fluids). Adaptive mechanisms, more recently evolved, perform many of their functions by interacting with the older innate ones (e.g. complement).
INNATE IMMUNE SYSTEM

IL6

C-reactive protein

Phagocytosis

Liver

Pro-inflammatory Cytokines

INFLAMMATION

ADAPTIVE IMMUNE SYSTEM

Complement activation/regulation

Complement Cascade

C4

1.5.3 Immune Abnormalities in SZ- Epidemiological Findings

Epidemiological studies of SZ have highlighted an overlap between incidences of SZ and other autoimmune and immune-related disorders. While there are discrepancies among studies regarding which autoimmune diseases are most strongly correlated with SZ risk (Benros et al., 2011; Benros et al., 2014; Benros et al., 2013; Eaton et al., 2006; J. Pouget et al., 2017; J. G. Pouget et al., 2016; Stringer et al., 2014), there is converging evidence that SZ and autoimmune diseases co-occur at a higher rate than is expected by chance. A large epidemiological study performed in Denmark, including 7704 patients with SZ and 192,590 age and gender matched controls without a psychiatric record, showed that SZ patients have a 45% higher lifetime incidence of one or more immune disorders which were present prior to the onset of SZ compared to matched controls (Eaton et al., 2006). Furthermore, the incidence of immune disorders was also higher in parents of SZ patients than among parents of comparison subjects (Stringer et al., 2014). In the largest retrospective cohort study to date (N = 39,364 SZ cases and < 3 million healthy controls) individuals a prior hospital contact because of autoimmune disease increased the risk of a subsequent mood disorder diagnosis by 45% (Benros et al. 2013). Importantly, the type of infectious agent does not seem to matter (Golam M Khandaker et al., 2015)-increased risk of SZ is observed following viral (e.g. influenza (Brown, 2011), rubella (Brown et al., 2001), bacterial (e.g. meningitis (Abrahao, Focaccia, & Gattaz, 2005), and parasitic (e.g. Toxoplasma gondii infections (E Fuller Torrey & Yolken, 2003). Furthermore, the number of hospitalizations for infection and psychosis risk have been reported in a dose-response relationship (Benros et al., 2011).

1.5.4 Maternal Immune Activation (MIA)

For the past 40 years, significant support for the neurodevelopmental hypothesis of SZ has derived from evidence that risk is increased by maternal infection during pregnancy, and postnatal infection (Golam M Khandaker et al., 2015). As with obstetric complications that cause hypoxic damage, these environmental risk factors are associated with an increased
inflammatory response in the developing brain. In SZ, and in other neurodevelopmental disorders, this inflammatory response has in turn been associated with abnormalities in the location, number and function of microglia - innate immune cells that develop during early gestation and whose longevity make them particularly vulnerable to developmental disturbance. As reviewed by Prinz & Priller (Prinz & Priller, 2014), this vulnerability is critically important for the developing brain given the role of microglia in synaptic pruning and remodelling during development and adulthood. More recently, this microglial activation has been linked to complement activity (Sellgren et al., 2019). Complement will be further discussed in section 1.6.11.

The first proposal that maternal infections increased rates of SZ in offspring followed observation that the risk of SZ was higher among individuals who were born in winter months, and in observation that those were in utero during the 1957 influenza epidemic were also more likely to develop SZ (Bradbury & Miller, 1985). Subsequent human studies have found that children with a viral infection of the CNS in childhood were 1.7 times more likely to develop a psychosis in later life in a meta-analysis of 2,424 cases and over 2 million healthy controls (Golam M Khandaker, Zimbron, Dalman, Lewis, & Jones, 2012). Another study of 9,596 individuals exposed to in utero infection and 13,808 of their siblings (of whom 36 and 35 developed SZ, respectively) by Clarke et al. reported a synergy between genetic risk for SZ (measured by positive family history) and infections with respect to disease susceptibility, suggesting a potential gene-environment interaction whereby infections increase the risk of SZ in genetically vulnerable individuals (Clarke, Tanskanen, Huttunen, Whittaker, & Cannon, 2009). Taken together, these findings suggest that individuals with SZ have been exposed to both genetic and environmental factors (e.g. infections) causing immune disturbances that predispose them to autoimmune disease as well as SZ risk (Benros et al., 2013).

Meyer et al (Meyer, Feldon, & Yee, 2008) have proposed a ‘Prenatal cytokine hypothesis’ based on evidence from animal models that abnormal
cytokine function prenatally leads to abnormalities in cognition and behaviour into adulthood, similar to those reported in SZ. Since cytokines have known roles in the development, maturation, and plasticity of cortical connections (Deverman & Patterson, 2009), these observed changes in cytokine levels during the period of synaptogenesis, dendritic spine formation, and activity-dependent plasticity may lead to chronic changes in cortical connectivity and altered behaviours in offspring (Estes & McAllister, 2016). In rodent models, MIA offspring show dynamic oscillating alterations in cytokine receptors during sensitive periods of neural growth and synaptogenesis (Estes et al., 2018). This is one of a number of such models of immune reaction proposed during the last 20 years (Anderson & Maes, 2013; Monji, Kato, & Kanba, 2009; Smith, Li, Garbett, Mirnics, & Patterson, 2007; E Fuller Torrey & Yolken, 2003).

1.5.5 Immune markers

As regards immune response to a pathogen, cytokines and reactive proteins are cellular communicators of the immune system, and used largely as markers of systemic inflammation in research (Monji et al., 2009). At a neural systems level, cytokines may play a role in synaptic plasticity, synaptic transmission and neurogenesis, and appear to be tempered differently in SZ (Di Nicola et al., 2013). Furthermore, elevated C-Reactive Protein (CRP) has been reported as a risk factor for SZ (Z. Wang et al., 2017). Cytokines and CRP are said to be involved in neurodevelopment, as increased immune markers in the CNS could cause disruption in the blood-brain barrier (BBB), weakening it and increasing immune secretions which permeate the BBB. Interleukin-6 (IL-6) and CRP, immune markers of inflammatory response, are used as immune markers in this thesis (section 1.6.5 and 1.6.6).

In serological studies of individuals with SZ and matched healthy controls, elevated levels of inflammatory markers have been consistently reported (Bergink et al., 2014; Di Nicola et al., 2013; Mondelli & Howes, 2014; Potvin et al., 2008; Upthegrove, Manzanares-Teson, & Barnes, 2014). As well an overall increase, a number of these markers including IL-6 appeared
to be responsive to anti-psychotic treatment (Mondelli & Howes, 2014; Tourjman et al., 2013). Recent findings link inflammatory markers (i.e. CRP and IL-6) to cognitive performance in SZ, suggesting that inflammation is associated with poorer cognitive performance both during and outside of acute illness phases. This could be caused by a spike in inflammation related to disease onset, a response to medication, or indeed a genetically programmed difference in immune activation and immune markers.

1.5.6 IL6

Among the aforementioned cytokines, IL-6 has been the target of much investigation, as it is involved in numerous fundamental processes of the CNS. IL6’s role in brain function was first identified when it was observed to be up-regulated in neurodegenerative disorders, such as in Alzheimer’s disease, Parkinson’s disease and Multiple Sclerosis (Spooren et al., 2011). Furthermore, Khandaker et al. found that high concentrations of circulating IL6 can be detected in childhood as predicting psychosis later in development, indicating a differential immune response long before onset of illness (Khandaker et al., 2014). Finally, maternal serum and placental IL-6 was found to increase exponentially during the acute phase of infection, sufficient to cause SZ-relevant behaviours and neuropathologies in offspring using animal models of MIA (Buka, Cannon, Torrey, Yolken, & Disorders, 2008; Choi et al., 2016; Dorrington et al., 2014; Estes & McAllister, 2016; Smith et al., 2007; Spinrad et al., 2012; Walshe et al., 2011; Zammit et al., 2009).

1.5.6.1 IL6 and Cognition

In a recent meta-analysis by Misiak et al. of several studies examining the relationship between inflammation and cognition, five studies of IL6 and cognition using a SZ sample were analysed (Fillman et al., 2013; Frydecka et al., 2015; Hope et al., 2015; Hori et al., 2017; Misiak et al., 2018; X. Y. Zhang et al., 2016). Only the study by Frydecka et al. (2015a) reported higher IL-6 levels were associated with worse cognitive performance on visual attention, visuomotor processing speed, semantic and working memory, task-switching ability and executive control function. No
significant correlations between IL-6 levels and cognitive performance were reported in other studies included in this analysis. However, Fillman et al. (2016) revealed that IL-6 levels contributed to the effects of all measured cytokines on cognitive performance in patients with SZ.

1.5.7 CRP

The study of neuroimmunology in SZ has often involved serological studies of CRP, or C-reactive protein, as a hallmark of systemic inflammation (Okada et al., 2010). CRP is a gun-shot protein produced by the liver, involved in direct immune response to a pathogen. In SZ patients, CRP levels are abnormally high, thus demonstrating presence of inflammation (Frydecka, 2015a). Recent meta-analyses suggest elevated levels of circulating CRP in patients with SZ and Bipolar disorder (Dargel et al., 2015; Fernandes et al., 2016), with birth cohort study also demonstrating that elevated CRP at the age of 15–16 years could predict subsequent development of SZ-spectrum disorders later in life (Metcalf et al., 2017). In contrast to IL6 and other cytokines, it has been found that elevated CRP levels in SZ occur regardless of antipsychotic treatment (Fernandes et al., 2016).

1.5.7.1 CRP and cognition

In the meta-analysis mentioned above by Misiak et al., the association between CRP levels and cognitive performance was tested in seven studies on SZ patients (Bulzacka et al., 2016; Dickerson et al., 2014; Dickerson, Stallings, Origoni, Boronow, & Yolken, 2007; Dickerson et al., 2012; Joseph et al., 2015; Micoulaud-Franchi et al., 2015) in which abnormal CRP levels were associated with worse cognitive performance on general intellectual ability, abstract reasoning, memory, working memory, semantic memory, learning abilities, attention, verbal learning and memory, mental flexibility and processing speed (Bulzacka et al., 2016) as well as global cognition in patients with SZ (Misiak, 2018, Dickerson et al., 2007b, Frydecka et al. 2015a). Another study revealed no significant correlations between CRP levels and cognitive performance in patients with SZ or schizoaffective disorder (executive functioning); however, in healthy
controls CRP levels negatively correlated with performance of executive functioning (Joseph et al., 2015).

1.5.8 Environment, cognition and Immune function

One intriguing aspect of the immune hypothesis is that the effects of both immune challenge and response may be mediated by environment. Early life adversity (ELA), including physical and emotional abuse/neglect, is a significant risk factor for a range of mental health disorders including psychosis (Varese et al., 2012). Immune markers seen to be elevated in SZ can be rapidly upregulated in response to inflammatory states, but also in response to psychosocial experiences such as stress (Garcia-Oscos et al., 2012). Increased risk for psychotic experiences have been observed at age 12 following maltreatment by an adult in childhood (relative risk=3.16) or bullying by peers (2.47) (Fisher et al., 2012). Furthermore, increased incidence of ELA can negatively impacts general cognitive ability and social cognition later in life, in both SZ populations and in healthy controls (Dauvermann & Donohoe, 2019; Rokita, Dauvermann, & Donohoe, 2018).

A recent meta-analysis by Baumeister et al. (Baumeister, Akhtar, Ciufolini, Pariante, & Mondelli, 2016) suggests that ELA results in significant changes in immune response that vary according to the trauma experienced. In a 2018 study, exposure to childhood victimization was associated with higher CRP levels at age 18 in females only, independent of latent genetic influence, as well as childhood socioeconomic status, and waist-hip ratio and body temperature at the time of CRP assessment (Baldwin et al., 2018).

Another study found that patients with SZ and a history of childhood trauma were found to have a pro-inflammatory phenotype in response to stress (Dennison, McKernan, Cryan and Dinan, 2012). As such, changes in immune response may represent at least one mechanism by which the negative effects of ELA on psychiatric illness risk is mediated. Thus, physiological stress due to emotional reactivity in childhood may have a multiplicative effect on genes related to SZ risk. A study by Tienari et al. (2004) illustrates the importance of this gene-environment interaction in SZ incidence. In this study, adopted children with a high genetic risk for
developing SZ are seen to be more likely to develop SZ in adoptive homes with a maladaptive pattern of upbringing. In contrast, children with high genetic risk who were brought up in harmonious homes were found to be protected from their biological vulnerabilities (Tienari et al., 2004). This supports a ‘developmental programming’ theory of disorder, whereby permanent alterations of psychological and physiological pathways in the brain can be altered by adversity, disturbing a key ‘developmental window’ in cognition (Khandaker and Dantzer, 2016; Collip et al., 2008).

Recognition is growing that it is necessary to tease out these types of adverse childhood experience, that carry risk for later psychosis symptoms (Upthegrove, 2015).

1.5.9 Genetic findings: MHC

Along with the clinical and epidemiological data already presented, genetic data have been interpreted to support an immunological cause of SZ (Stefansson et al. 2009, Debnath et al. 2012, Horvath and Mirnics 2014). Analyses of SZ GWASs suggested a role for immune pathways, largely because SNPs in the MHC region, which encodes many proteins critical to immunity, have shown the strongest association with SZ in every GWAS conducted to date (Woo et al., 2019, Purcell et al. 2009, Shi et al. 2009, Stefansson et al. 2009, Ripke et al. 2011, Irish SZ Genomics Consortium and Wellcome Trust Case Control Consortium 2012, Ripke et al. 2013, SZ Working Group of the PGC 2014; Pardiñas et al., 2018). Representing ~0.1% of the human genome, the MHC is a four-million base pair region encoding for a large gene family (~250 genes) of classical and transplantation HLA genes but also many other immune and non-immune genes (Corvin & Morris, 2014). The MHC is complex, with extreme sequence diversity, substantial linkage disequilibrium (LD) and high gene density, which makes it the most speculated upon gene region in recent SZ research. MHC variants represent both the strongest signal and the largest class of variants to be associated with SZ risk (Corvin & Morris, 2014). The function of the MHC variant in SZ has been related to early differentiation of neural circuitry, synaptic function, and excitatory and inhibitory
neurotransmitter imbalance (McAllister, 2014). In another study, a SNP representing the MHC region had an association with delayed episodic memory and decreased hippocampal volume (Walters, 2013). In combination, the synaptic processes thought to be impaired by MHC risk variants may relate to differential brain development in SZ (Sekar 2016).

1.5.10 C4 expression in SZ

Due to the large size and diversity of the MHC region, risk factors could relate to SZ pathogenesis though the disruption of many neural processes (Corvin & Morris, 2014). Recent research has tried to eliminate ambiguity surrounding MHC by targeting the ‘Complement component 4’ or C4 gene, encoding for complement component 4 within the MHC region. C4 is part of the complement system, an innate immune system pathway comprising a large number of plasma proteins that help antibodies and phagocytes clear pathogens (Janeway et al., 2001), including the removal of damaged cells, autoantigens and environmentally derived antigens (Nimgaonkar et al., 2017). Compellingly, recent genome-wide association studies suggest repeat polymorphisms incorporating the complement 4A (C4A) and 4B (C4B) genes as risk factors for SZ (Nimgaonkar et al., 2017). C4 genetic variation has been associated with different RNA expression levels across post-mortem brain samples, as well as increased risk for SZ in a large study of 28,799 patients and 35,986 controls (Sekar et al., 2016). In the same study, high risk C4 genetic variation was associated with increasing brain C4 expression, and in a separate analysis, C4 expression was 1.4 times higher in post-mortem brain samples from SZ patients compared to controls (Sekar et al., 2016). The C4A/C4B genetic associations have re-ignited interest not only in inflammation-related models for SZ pathogenesis, but also in neurodevelopmental theories, because rodent models indicate a role for complement proteins in synaptic pruning and neurodevelopment. Excessive C4 activity is hypothesized to result in increased synapse elimination, which involves destruction of inappropriate synaptic connections, and has been proposed to be dysfunctional in SZ in the brains of SZ patients (Mayilyan, Weinberger & Sim, 2008, Boksa, 2012; Sekar et al., 2016, Boyajyan, Khoyetsyan & Chavushyan, 2010). Furthermore, in a proteomic analysis of
the complement signalling pathway using longitudinal population-based data, individuals who went on to experience psychotic-like experience and psychotic disorders were found to have upregulation of multiple complement proteins in childhood (English et al., 2017, Föcking et al., 2019). More efforts are needed to clarify whether and how synaptic pruning plays a role in SZ, and if so, how the complement system is involved in this process. Thus, the complement system could be used as one of the ‘staging posts’ for a variety of focused studies of SZ pathogenesis.

1.5.11 Complement and cognition

Following up the seminal paper suggesting C4 is linked to SZ, we (Donohoe et al. 2018) found that higher predicted C4 RNA expression was associated with poorer performance on measures of memory function in both patients with SZ and healthy participants, and with reduced cortical activation during visual task performance in healthy participants. In our own group and others, the complement-related gene CSMD1, encoding a regulator of complement, has been associated with decreased neurocognitive ability in general cognitive measures and poorer episodic memory performance in patients with SZ and healthy controls (Donohoe 2013, Athanasiu, 2017). Furthermore, in a study of healthy males the CSMD1 risk variant associated with SZ risk was associated with poorer performance on measures of general cognitive ability, strategy formation, spatial and visual working memory, set shifting, target detection and planning for problem solving (Koiliari, Roussos, Pasparakis & Lencz, 2014). In further studies from our group, SNPs tagging variation at or within the MHC region (which contains complement genes) have been association with variation in episodic memory and hippocampal volume in a sample of SZ patients and healthy adults (Walters et al., 2013). As regards other genes associated with complement, in a study of patients with mild cognitive impairment and Alzheimer’s disease (AD) altered levels of complement factor H, a gene involved in the complement cascade were reported SZ patients compared to age and gender matched healthy controls (Gezen-Ak, 2013). This, along with evidence that other risk-related genes within the complement system...
are also associated with variation in cognition as well as cortical thickness (Alleswede et al.), led us to consider whether variation within genes that encode the complement system could as a whole influence cognition.

1.6 Rationale and hypotheses

It is known that immune components such as MHC regulate brain development and adult neural plasticity (McAllister et al., 2014) and more recently, C4 and the complement system have been implicated in SZ risk and cognition (Sekar et al., 2016, Donohoe et al., 2018). Exposure to the wrong level of an immune factor at the wrong time may consequently disrupt brain development and adult neural functioning, as supported by in utero immune activation studies, as well as serological studies of inflammatory markers for immune function in SZ. Many SZ patients show hallmarks of immune disease—such as prior infection, co-occurring autoimmunity, and inflammation—support the idea that immune disturbances may play a role in SZ by disrupting brain development and/or adult neural function. In this thesis three studies have been performed in order to further characterize genetic and environmental factors related to immune function, that may disrupt normal neural and cognitive functioning in SZ.

The first study took advantage of a wide range of environmental and immune marker variables available in a longitudinal birth cohort, ALSPAC. This study hypothesized that ELA, characterized by parental maltreatment in the form of hitting, emotional abuse, physical abuse and domestic abuse may impact on immune markers in the blood. This hypothesis was based on previous research linking high levels of immune markers in the blood to PE in the same sample (Khandaker et al., 2012), and other research from separate samples linking ELA to elevated immune response (Baumeister et al., 2016), as well as a long established connection between ELA and SZ (Fisher et al., 2012).

As the MHC signal is difficult to disentangle in terms of multiple genetic variants in high linkage disequilibrium, the second study investigated the
The genetic influence of immune reaction on cognition in SZ. This interplay of SZ risk genes is important to study, given the polygenic nature of the disorder, as well as the polygenic nature of cognition. Using a list of identified complement genes in this network, a polygenic score (PGS) was constructed. This PGS of genes associated with the complement system was identified as enriched for the phenotype of IQ. It was hypothesised that a higher PGS would be linked to the cognitive deficits commonly seen in SZ patients.

The third study in this thesis was based on the results of the second, in which an association between complement PGS and general cognition was observed. To further characterise the effect of the complement system, a PRS score was created for complement genes associated with SZ risk. The association between the score and structural neuroimaging measures linked with memory function were investigated. It was hypothesised that a higher PRS would associate with deficits in measures of hippocampal volume, brain volume, cortical thickness and surface area. This PRS was used in analysis to compare effects of the whole complement PRS on brain structures with C4 expression levels alone.
2. Study 1: Effects of Early Life Adversity on Immune Function and Cognitive Performance: Results from the ALSPAC cohort

Jessica F Holland¹, Golam M Khandaker²,³, Maria MR Dauverman¹, Derek Morris¹, Stanley Zammit⁴,⁵, Gary Donohoe¹,

1. Cognitive Genetics & Cognitive Therapy Group, The Centre for Neuroimaging, Cognition and Genomics (NICOG), School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Ireland.
2. Department of Psychiatry, University of Cambridge, UK
3. Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, UK
4. Population Health Sciences, Bristol Medical School, University of Bristol, UK.
5. MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, UK
Abstract

Background Early Life Adversity (ELA) is a significant risk factor for mental health disorders. One hypothesised mechanism by which this occurs is via an effect on immune response. In this analysis of epidemiological data, we tested whether ELA was associated with cognitive performance, and if so, whether these effects were influenced by immune function.

Methods We investigated the longitudinal relationship between ELA, inflammatory markers and cognition in data from Avon Longitudinal Study of Parents And Children (ALSPAC; n~5,000). ELA was defined in terms of physical/emotional abuse, harsh parenting, or domestic violence before 5 years. Social cognition was measured in terms of theory of mind, and general cognitive ability was measured using IQ. Inflammatory markers included serum C-reactive protein and interleukin-6 levels.

Results A significant association was observed between IQ and harsh parenting, whereby children who were physically disciplined had lower IQ scores (accounting for relevant social factors). Both immune markers were associated with variation in cognition; however, neither accounted for the effects of ELA on cognition.

Discussion This study highlights the impact of ELA on cognition. In the absence of evidence that these effects are explained by inflammation, other mechanisms by which the effects of ELA are mediated are discussed.
2.1 Introduction

Human cognitive development is an open system, where children are in constant interaction with their environment (Bornstein et al., 2006). Although early life adversity (ELA) is known to widely impact developmental outcomes, the mechanism by which this occurs is only beginning to be elucidated, with many pathways to pathology still underexplored (Shonkoff et al., 2012). Negative impacts of ELA include disrupted cognitive (Ehler, 2013) and neural (Mothersill & Donohoe, 2016) function during childhood, and increased risk for psychosis and other psychiatric disorders in adulthood (Alemany et al., 2015; Varese et al., 2012). One mechanism by which this change in functioning may occur is via elevated immune response. A meta-analysis by Baumeister and colleagues suggests that ELA is associated with low-grade inflammation as reflected by increased levels of a number of circulating inflammatory markers, such as C-reactive protein (CRP) and interleukin 6 (IL-6), which are still apparent in adulthood (Baumeister et al., 2016). As such, changes in immune response may represent at least one mechanism by which the negative effects of ELA on cognitive function (and psychiatric illness risk) is mediated.

2.1.1 ELA and Cognition

Studies during the last 30 years have consistently documented impaired cognitive abilities and poor academic achievement in maltreated young people (Slopen, Kubzansky, McLaughlin, & Koenen, 2013). Various traumas and ELAs have been studied in relation to receptive and expressive language skills (Fox, 1993), IQ (Delaney-Black et al., 2002; Slopen, Kubzansky, McLaughlin, et al., 2013), and deficits in social cognition (Germine, Dunn, McLaughlin, & Smoller, 2015; Polanczyk et al., 2010). Some studies have indicated that social cognition is specifically affected by exposure to severe physical adversity in childhood; i.e. physical abuse or domestic violence (Germine et al., 2015). This could be because children who have experienced physical or emotional abuse are frequently reported to engage in general maladaptive thinking patterns and negative self-schema (Appiah-Kusi et al., 2017). Studies of ELA and cognition have also reported that children who
experienced physical abuse have deficits in verbal memory and general cognition (Friedrich, Einbender, & Leucke, 1983) and that the degree of these effects may relate to adverse events occurring at particular developmental stages (Pollak & Kistler, 2002). Furthermore, ‘poly-victimization’ or non-specific multiple abusive experience has been associated with higher rates of mental illness later in development (Ehlert, 2013; Hardy et al., 2016).

Conversely, there may be alternative pathways to cognitive difficulties which involve genetic and environmental confounds, making certain children more susceptible to experiencing ELA [8]. When examining 2 large cohort studies of ELA and IQ, Danese and colleagues (Danese et al., 2016) observed that relationships between childhood victimization and IQ were largely accounted for when pre-existing cognitive ability (prior to ELA) was included in analysis. This finding challenges developmental programming theories that ELA causes later deficits in cognition, suggesting instead that lower cognitive function may be a risk factor for ELA.

### 2.1.2 Inflammatory markers & Cognitive performance

C-reactive protein (CRP), an acute-phase reactant, and interleukin-6 (IL6), a pro-inflammatory cytokine, have been examined in many studies as hallmarks of systemic inflammation (Okada et al., 2010). Population-based longitudinal studies have reported an association between elevated levels of IL-6 and CRP in childhood and increased risks of symptoms/diagnosis of SZ and depression in adulthood (Golam Mohammed Khandaker et al., 2017; Metcalf et al., 2017). While the relationship between elevated levels of childhood CRP and IL6 and mental health outcomes in adulthood has received some attention (Gimeno et al., 2009; Zalli, Jovanova, Hoogendijk, Tiemeier, & Carvalho, 2016), the longitudinal effects of elevated IL6 and CRP on cognition during childhood development remains unexplored (Slopen, Koenen, & Kubzansky, 2012, 2014; Slopen, Kubzansky, & Koenen, 2013; Slopen, Kubzansky, McLaughlin, et al., 2013). Abnormal levels of CRP, elevated above 3mg/l in adults was found to be associated with lower IQ scores in a large sample of patients with SZ (Fond et al., 2018). Higher IL6 blood concentrations were also found to be negatively associated with performance on a range of
memory tests in a study of 151 patients with SZ (Frydecka et al., 2015). These findings suggest that that low-grade inflammation could be a precursor to illness (Gimeno et al., 2009; G. M. Khandaker, R. M. Pearson, S. Zammit, G. Lewis, & P. B. Jones, 2014; Zalli et al., 2016). The mechanism by which inflammatory response may affect cognition remains to be elucidated, and this may shed light on the association with psychotic disorders.

2.1.3 The present study

The aim of this study was to examine the relationship between ELA and social cognitive ability, while exploring whether any observed relationship was accounted for by variation in inflammatory markers. Based on available variables from ALSPAC, we defined ELA in terms of a range of negative early parenting experiences, including physical discipline by a parent (‘harsh parenting’), physical abuse by a parent, emotional abuse by a parent, and/or presence of domestic violence between parents before the age of five years. In ALSPAC literature, this is considered ‘preschool’ age (Hay, Heron, & Ness, 2005), thus accounting for ELA in the home and not including factors outside the home like bullying or later adversities. Immune response was measured using CRP and IL-6 levels taken from peripheral blood at 9 years. Social cognition was measured using a computerized Theory of Mind task administered at age 13 years. General cognitive ability was assessed in terms of IQ at age 8 years, used in secondary analysis. Based on these data, we tested the following hypotheses: (1) That ELA at the age of 5 years was associated with poorer performance in and social cognition (theory of mind) at 13 years, and secondly with general cognitive ability (IQ) at 8 years; (2) That a relationship exists between elevated levels of immune response and cognitive function; (3) That the relationship between immune response (IL-6 and CRP) and cognition partially accounts for variation in cognition that is explained by ELA.

2.2 Methods

2.2.1 Participants

The current sample was drawn from the Avon Longitudinal Study of Parents and Children (ALSPAC), a birth cohort based on a recruited 14,541 pregnant
women residing in the county of Avon, with expected delivery dates between April 1991 and December 1992. This is the initial number of pregnancies for which the mother enrolled in the ALSPAC study and had either returned at least one questionnaire or attended a “Children in Focus” clinic by 19/07/99 (www.alspac.bris.ac.uk). Parents completed regular postage questionnaires about many aspects of their child’s development and health since birth. This resource includes a wide range of phenotypic and environmental measures in addition to biological samples. The area of Avon includes both urban and rural areas, with the population representative of all children in the UK (Boyd et al., 2013). Please note that the study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/).

The current study is based on up to 5709 participants with information regarding adverse childhood events up to age 5 years, and relevant experiential and cognitive assessments completed at ages 8.5 years or 13 years. Sample sizes vary for each analysis depending on availability of data. Choice of variables was driven by availability of data. Ethical approval for the study was obtained from ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

### 2.2.2 Childhood adversity

#### 2.2.2.1 Domestic violence

Mothers reported via a series of postal questionnaires on a range of victimization experiences that they and their children had been exposed to since birth. At interview time-points of 8 months, 1.5, 2.5, 4 and 5 years old, mothers were asked whether they experienced emotional or physical abuse from their partner: “Your husband/partner was physically cruel to you” and “Your husband/partner was emotionally cruel to you.” This measure was then split into binary terms- A positive response to either of these two questions at any of the time points was considered to be evidence of domestic violence (see (Bowen, Heron, Waylen, Wolke, & Team, 2005)). Use of a binary domestic violence measure has been reported
on previously as part of a composite victimization score (Fisher et al., 2012).

2.2.2.2 Physical cruelty

Physical cruelty was assessed based on mother’s self-report at interview time-points of when the child was ages 8 months, 1.75yrs, 2.75yrs, 4yrs and 5 years. Again, a binary yes/no variable was calculated to reflect the presence or absence of any daily/weekly physical cruelty up to age 5 was considered a ‘yes’ and recorded (Plant, Jones, Pariante, & Pawlby, 2017).

2.2.2.3 Emotional Cruelty

A measure of emotional cruelty was calculated by asking mothers to assess whether they had been emotionally cruel to their child at ages 8 months, 1.75yrs, 2.75yrs, 4yrs and 5 years interview time-points. Again, a binary yes/no variable was computed in which any indication of emotional cruelty up to age 5 was considered a ‘yes’, and recorded, similar to emotional abuse measures calculated elsewhere (Plant et al., 2017).

2.2.2.4 Harsh parenting

Harsh parenting was measured based on responses made by mothers to the question “When you are at home with your child how often do you slap him’ or ‘do you smack your child when they are naughty?’ at the interview time-points for when the child was aged 2 or 3.5 years. The answer of ‘daily’ or ‘weekly’ was coded as indicating the presence of Harsh parenting, with answers of ‘once a month’, ‘never’ or ‘rarely’ coded as absence of Harsh parenting. In secondary analysis, the same group was divided into a scaled score based on frequency of hitting. Measures of maternal maladaptive parenting in ALSPAC have been reported on previously using these measures as part of composite scores on victimization (Plant et al., 2017) (Fisher et al., 2012)
Poly-victimisation has been operationalized as the total number of victimization types that a child experiences (Germine et al., 2015). Our poly-victimization variable was derived by summing all ELA experiences that each child experienced on a regular basis prior to the age of 5. The score ranged from 0-4, with 4 representing presence of all the above ELAs. As some victimization measures correlated, this poly-victimization score is used for primary analysis of effects. Secondary analysis will include individual ELAs.

2.2.3 Measure of social cognition

The ‘triangles’ social cognition assessment was completed at the age 13.5 years visit (mean subject age: 13.75 years (SD= 71 days). This test involves the use of computerized abstract animations to measure the participant’s ability to attribute an emotional mental state to non-human animate entities. The animations are used to test the participant’s ability to use motion cues, such as speed and trajectory of movement, and movement in relation to others, to infer emotions (Barona, Kothari, Skuse, & Micali, 2015). Participants are shown 5-s animations of a circle and a triangle on a computer screen. In some of the animations, the triangle moves in a self-propelled manner designed to evoke a particular emotion: angry, happy, sad, or scared. In the other animations, it moves in a manner designed to make it appear “non-living.” Participants are asked (a) whether the triangle is living, and if so, how living (measured on a Likert scale 0–5), or (b) whether the triangle has a particular emotion (happy, sad, angry, or scared). We calculated the total score by adding the score for all the emotional items. To avoid negative scores, we added 40 to the total score, giving the score a range from 0–80, similar to other studies of Triangles data (Warrier & Baron-Cohen, 2018).

2.2.4 Measure of general cognition

IQ assessment was completed during the age 8 years visit (mean subject age: 8 years, 8 months SD=3.1 months), using a shortened version of the WISC 3rd U.K. Edition, administered by trained psychologists. Use of the shortened version reduced the length of assessment, so the children were less likely to
fatigue (Mackinnon, Zammit, Lewis, Jones, & Khandaker, 2018). IQ data obtained using this method have been shown to be valid (G. M. Khandaker, J. Stochl, et al., 2014), shown robust correlations with neurodevelopmental disorders, and correlates with other concurrent neurocognitive measures such as short-term memory, working memory and socio-economic factors (G. M. Khandaker, R. M. Pearson, et al., 2014; Shonkoff et al., 2012). While IQ was measured in the year prior to serum inflammatory markers, this measure shows robust test–retest reliability during childhood (6–13 years), and was considered unlikely to have changed during this period (Fraser et al., 2012). For the purposes of this analysis, total IQ scores were used.

2.2.5 Potential Confounders
Maternal education at pregnancy was found to be the highest predictor of children’s educational attainment at ages 16 and 18 years in previous studies of ALSPAC (Morris, Dorling, & Davey Smith, 2016), even when accounting for children’s cognitive ability. This variable has been strongly and robustly linked to multiple later life outcomes (Slopen, Kubzansky, McLaughlin, et al., 2013), and is likely to represent an important environmental confound. Therefore, maternal educational attainment was examined as a possible confounder in the analysis of the relationship between ELA and IQ. Gender, BMI and maternal education were entered as covariates in analysis, as these factors have been found to impact on immune response and/or cognition in previous ALSPAC studies (G. M. Khandaker, R. M. Pearson, et al., 2014). Finally, paternal social class and ethnicity were included as potential confounders (G. M. Khandaker, S. Zammit, et al., 2014).

2.2.6 Measurement of IL6 and CRP
Serum Interleukin-6 and C-reactive protein levels, measured at 9 years, were examined in the current study. Non- fasting blood samples were collected from participants in assessment clinic using standard procedure. The samples were immediately spun and frozen at -80°C. Measurements were assayed in 2008 after a median of 7.5 years in storage. There was no evidence of freeze-thaw during this period. An enzyme-linked immunosorbent assay (ELISA) was used to measure IL-6 levels (R&D systems, Abingdon, UK), and CRP
was measured by automated particle-enhanced immunoturbidimetric assay (Roche UK, Welwyn Garden City, UK). These methods have been reported elsewhere (G. M. Khandaker, R. M. Pearson, et al., 2014). Participants with infection were excluded from analysis, as this known to transiently increase levels of inflammatory markers (G. M. Khandaker, R. M. Pearson, et al., 2014).

2.2.7 Analysis

2.2.7.1 Primary Analysis: Poly-victimization score

To test the associations between childhood victimization (independent variable) and cognition (dependent variable), Linear Regressions were carried out in SPSS (version 23). Covariates were included in the first step (Model 1 no covariates, model 2 gender, model 3 covariates of gender and maternal education) prior to inclusion of independent variables. ‘Triangles Score’, a factor created using a combination score of 4 Triangles task subtests was used as the outcome variable. Secondly, Full scale IQ was analysed as potentially effected by ELA. Poly-victimization score was used as the predictor in this model, ranging from 0-4. Therefore, 2 analyses are run in the primary: one for poly-victimization and social cognition; and subsequently an analysis of poly-victimization and General Cognition (IQ).

2.2.7.2 Secondary Analysis: Differential impact of specific ELAs

We looked at each ELA separately as predictors of cognitive deficits (Social cognition at 13 years and IQ at 8 years). This allowed us to take advantage of the richness of the dataset, and extra participants available in different measures of ELA. The aim of secondary analysis is to identify specific associations in the data which may be differential depending on the type of adversity experienced. These effects would have been diluted if a composite score of adversity was to be used only. Gender was entered as a possible confounder of results in each analysis, along with maternal education.

1.8.3 Are ELAs associated with CRP/IL-6?
Linear Regressions were analysed to test the relationship between elevated levels of IL6 or CRP and cognition. Log transformed values of IL-6 and CRP were used as the dependent variables in analysis. ELA was the predictor as a binary variable. All those with infection at the time of blood draw were removed from analysis. Covariates of age, gender, BMI, maternal education, ethnicity and paternal class were added to the model, as these have been found to predict immune response in previous studies (G. Khandaker, R. Pearson, S. Zammit, G. Lewis, & P. Jones, 2014; Golam M Khandaker & Dantzer, 2016; G. M. Khandaker, R. M. Pearson, et al., 2014; Golam Mohammed Khandaker et al., 2017; G. M. Khandaker, J. Stochl, et al., 2014).

**1.8.4 Does the association between CRP/IL-6 and cognition partly explain the relationship between ELA and cognition – regression analysis**

As a mediating role of IL6 and CRP cannot be observed due to temporality of variables examined, IL6 and CRP respectively were added as covariates in a regression model, to test whether inclusion of inflammatory markers in the model changes the relationship between ELA and cognition. Interaction terms representing the combined effect of inflammatory markers and ELA were also tested as predicting cognition. Gender, age, BMI, maternal education, ethnicity and paternal class were entered as covariates in the model (G. M. Khandaker, R. M. Pearson, et al., 2014; Golam Mohammed Khandaker et al., 2017; G. M. Khandaker, S. Zammit, et al., 2014).
2.3 Results

Descriptive information for all variables analysed are presented in Table 2.1. For each regression analysis, assumptions of linear regression were checked using P-P plots and scatterplots. Multi-collinearity was not observed between independent variables, with all predictor variables yielding a variance inflation factor (VIF) of <1.15. Correlations between ELA measurements were either low or non-significant (Pearsons R less than .3).
Table 2.1: Demographic information and descriptive statistics for variables used in analysis

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>ELA</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domestic Abuse</td>
<td>Up to 5 years</td>
<td>+</td>
<td>654</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>2068</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical cruelty</td>
<td>Up to 5 years</td>
<td>+</td>
<td>107</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>2630</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional cruelty</td>
<td>Up to 5 years</td>
<td>+</td>
<td>290</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>2431</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harsh parenting</td>
<td>Up to 3 years</td>
<td>+</td>
<td>2108</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>2558</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poly Victimization</td>
<td>Up to 5 years</td>
<td>+</td>
<td>1594</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>1074</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL6 Level</td>
<td>9.8 years</td>
<td></td>
<td>4143</td>
<td>1.26</td>
<td>1.55</td>
</tr>
<tr>
<td>CRP Levels</td>
<td>9.8 years</td>
<td></td>
<td>4152</td>
<td>0.76</td>
<td>2.55</td>
</tr>
<tr>
<td>IQ Score</td>
<td>8.5 years</td>
<td></td>
<td>5787</td>
<td>105.27</td>
<td>16.31</td>
</tr>
<tr>
<td>Triangles Task</td>
<td>13.75 years</td>
<td></td>
<td>5220</td>
<td>48.40</td>
<td>3.74</td>
</tr>
</tbody>
</table>

*ELA* = Early Life Adversity

*N* = Number of participants with information

*M* = Mean

*SD* = Standard Deviation
2.3.1 Association between ELA score and social cognition (Measured by Triangles score)

To test for an association between cognition and ‘poly victimization’, linear regression was used based on the 3 models described in the methods section. No association between effects of ELA ‘poly-victimization’ were observed using the ‘Triangles’ theory of mind task scores (Table 2.2). Similarly, when examining each ELA separately no relationship between ELA and social cognition was observed (Table 2.3).
Table 2.2: Linear regression models for Early life adversity (Total [poly-victimisation] score, Harsh parenting, Domestic Abuse, Physical Abuse, Physical abuse and emotional abuse) and Theory of Mind (Triangles total) scores.

<table>
<thead>
<tr>
<th>ELA Score</th>
<th>N</th>
<th>B</th>
<th>B1</th>
<th>p</th>
<th>R² change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly-victimization</td>
<td>2192</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.59</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Domestic Abuse</td>
<td>2612</td>
<td>-0.04</td>
<td>-0.02</td>
<td>0.39</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Physical Abuse</td>
<td>2626</td>
<td>0.15</td>
<td>0.03</td>
<td>0.14</td>
<td>0.001</td>
</tr>
<tr>
<td>Emotional Abuse</td>
<td>2609</td>
<td>0.12</td>
<td>0.04</td>
<td>0.05</td>
<td>0.001</td>
</tr>
<tr>
<td>Harsh Parenting</td>
<td>4679</td>
<td>-0.03</td>
<td>-0.02</td>
<td>0.29</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Analysis adjusted for covariates of age, gender and maternal education

*B = unstandardized beta value; B1 = standardized beta value.*
2.3.2 Association between ELA score and cognition (measured by IQ)

An association was observed between poly-victimization scores and lower general IQ at 8 years (Table 2.3). When the effects of adversity variables were considered separately, a significant association between ‘harsh parenting’ and Full scale IQ was observed that survived correction for multiple testing (Bonferroni; 4 ELA types, p=0.0125). Furthermore, while maternal education was, as expected, strongly associated with child IQ (b=4.612; p=7.22E-117), the association between ‘Harsh Parenting’ and IQ remained significant after co-varying for this and all other variables of interest (see Table 2.3).

As a post hoc analysis, when harsh parenting was re-classified according to frequency - ‘never’, ‘sometimes’ (less than once per week), or ‘often’ (more than once per week), a dose-response effect was observed: children who were disciplined frequently had significantly lower IQ scores than those disciplined less frequently (Tukey HSD=-4.10, p=3.08x10^-7), and those hit less frequently had lower scores than those never hit (Tukey HSD=-2.37; p=0.00001).
Table 2.3: Linear regression models for Early life adversity (Total [poly-victimisation] scores, Harsh parenting, Domestic Abuse, Physical Abuse, Physical abuse and emotional abuse) and general cognitive ability (Full Scale IQ scores).

<table>
<thead>
<tr>
<th>ELA Score</th>
<th>B</th>
<th>B1</th>
<th>p</th>
<th>R² change</th>
<th>B</th>
<th>B1</th>
<th>p</th>
<th>R² change</th>
<th>B</th>
<th>B1</th>
<th>p</th>
<th>R² change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pol-y-victimization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n= 2147</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1.9</td>
<td>-0.068</td>
<td>0.002</td>
<td>0.005</td>
<td>-1.92</td>
<td>-0.069</td>
<td>0.001</td>
<td>&lt;0.001</td>
<td>-1.093</td>
<td>-0.039</td>
<td>0.057</td>
<td>0.003</td>
</tr>
<tr>
<td>2</td>
<td>-0.346</td>
<td>-0.011</td>
<td>0.617</td>
<td></td>
<td>0.455</td>
<td>0.012</td>
<td>0.55</td>
<td>&lt;0.001</td>
<td>0.166</td>
<td>0.004</td>
<td>0.816</td>
<td>0.119</td>
</tr>
<tr>
<td>3</td>
<td>0.14</td>
<td>0.004</td>
<td>0.829</td>
<td></td>
<td>0.118</td>
<td>0.004</td>
<td>0.855</td>
<td></td>
<td>0.016</td>
<td>0</td>
<td>0.979</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.759</td>
<td>0.346</td>
<td>&lt;0.001</td>
<td></td>
<td>4.713</td>
<td>0.343</td>
<td>&lt;0.001</td>
<td></td>
<td>4.742</td>
<td>0.345</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Domestic Abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=2533</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.453</td>
<td>0.012</td>
<td>0.552</td>
<td>&lt;0.001</td>
<td>0.455</td>
<td>0.012</td>
<td>0.55</td>
<td>&lt;0.001</td>
<td>0.166</td>
<td>0.004</td>
<td>0.816</td>
<td>0.119</td>
</tr>
<tr>
<td>2</td>
<td>0.14</td>
<td>0.004</td>
<td>0.829</td>
<td></td>
<td>0.118</td>
<td>0.004</td>
<td>0.855</td>
<td></td>
<td>0.016</td>
<td>0</td>
<td>0.979</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4.759</td>
<td>0.346</td>
<td>&lt;0.001</td>
<td></td>
<td>4.713</td>
<td>0.343</td>
<td>&lt;0.001</td>
<td></td>
<td>4.742</td>
<td>0.345</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Physical Abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n= 2548</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.817</td>
<td>0.034</td>
<td>0.085</td>
<td>0.001</td>
<td>2.809</td>
<td>0.034</td>
<td>0.086</td>
<td>&lt;0.001</td>
<td>1.947</td>
<td>0.024</td>
<td>0.206</td>
<td>0.117</td>
</tr>
<tr>
<td>2</td>
<td>0.118</td>
<td>0.004</td>
<td>0.855</td>
<td></td>
<td>0.016</td>
<td>0</td>
<td>0.979</td>
<td></td>
<td>4.713</td>
<td>0.343</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4.742</td>
<td>0.345</td>
<td>&lt;0.001</td>
<td></td>
<td>4.742</td>
<td>0.345</td>
<td>&lt;0.001</td>
<td></td>
<td>4.742</td>
<td>0.345</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Emotional Abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n= 2531</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.288</td>
<td>0.024</td>
<td>0.22</td>
<td>0.001</td>
<td>1.292</td>
<td>0.024</td>
<td>0.218</td>
<td>&lt;0.001</td>
<td>1.311</td>
<td>0.025</td>
<td>0.183</td>
<td>0.119</td>
</tr>
<tr>
<td>2</td>
<td>0.178</td>
<td>0.005</td>
<td>0.784</td>
<td></td>
<td>0.036</td>
<td>0.001</td>
<td>0.953</td>
<td></td>
<td>4.713</td>
<td>0.343</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4.742</td>
<td>0.345</td>
<td>&lt;0.001</td>
<td></td>
<td>4.742</td>
<td>0.345</td>
<td>&lt;0.001</td>
<td></td>
<td>4.742</td>
<td>0.345</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Harsh Parenting</td>
<td>1</td>
<td>-3.378</td>
<td>-0.105</td>
<td>1.07E-12</td>
<td>0.011</td>
<td>-3.45</td>
<td>-0.107</td>
<td>4.29E-13</td>
<td>0.001</td>
<td>-1.913</td>
<td>-0.059</td>
<td>0.000024</td>
</tr>
<tr>
<td>----------------</td>
<td>---</td>
<td>---------</td>
<td>--------</td>
<td>----------</td>
<td>--------</td>
<td>--------</td>
<td>---------</td>
<td>----------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.694</td>
<td>-0.022</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>4.612</td>
<td>0.333</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model 1: Analysis unadjusted for other variables; Model 2: Gender included as a covariate; Model 3: Gender and maternal education included as covariates.

*B=unstandardized coefficients; BI = standardized coefficients*
2.3.3 Relationship between immune markers, ELAs and cognition

To test for significant associations between ELA and either IL6 or CRP levels at age 9, Linear regression was carried out with each ELA as the independent variable, and IL6 or CRP as the dependent variable including covariates of age, gender maternal education, paternal class and ethnicity in each analysis. Based on these analyses, no significant association between any ELA and levels of either IL6 or CRP levels was observed (see Table 2.4).
Table 2.4: Linear regression models for ELA (Domestic Violence, Physical Abuse, Physical abuse, Emotional abuse, & Harsh parenting) predicting Immune activation (measured by CRP and IL6 levels).

<table>
<thead>
<tr>
<th></th>
<th>IL6 N</th>
<th>B</th>
<th>B1</th>
<th>P</th>
<th>CRP N</th>
<th>B</th>
<th>B1</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly-victimization</td>
<td>1580</td>
<td>0.008</td>
<td>.004</td>
<td>0.877</td>
<td>1582</td>
<td>-0.041</td>
<td>-0.025</td>
<td>0.303</td>
</tr>
<tr>
<td>Domestic abuse</td>
<td>1920</td>
<td>0.023</td>
<td>0.009</td>
<td>0.389</td>
<td>1922</td>
<td>-0.030</td>
<td>-0.013</td>
<td>0.540</td>
</tr>
<tr>
<td>Physical Abuse</td>
<td>1930</td>
<td>-0.045</td>
<td>-0.008</td>
<td>0.728</td>
<td>1932</td>
<td>-0.138</td>
<td>-0.028</td>
<td>0.192</td>
</tr>
<tr>
<td>Emotional Abuse</td>
<td>1916</td>
<td>-0.098</td>
<td>-0.027</td>
<td>0.227</td>
<td>1918</td>
<td>-0.051</td>
<td>-0.016</td>
<td>0.447</td>
</tr>
<tr>
<td>Harsh Parenting</td>
<td>3441</td>
<td>0.028</td>
<td>0.012</td>
<td>0.463</td>
<td>3449</td>
<td>0.015</td>
<td>0.008</td>
<td>0.636</td>
</tr>
</tbody>
</table>

Analysis adjusted for covariates of age, gender, BMI and maternal education;

*B* = unstandardized beta value; *B1* = standardized beta value.
As Table 5 indicates, significant negative relationships between both IL-6 or CRP and IQ were observed, such that higher inflammatory response (measured using either marker) was associated with lower IQ scores. With the inclusion of BMI, maternal education, paternal class and ethnicity in the model, the significance of these results was diminished (See Table 2.5). Finally, no association between either IL6 or CRP, and ‘Triangles’ task scores were observed.
### Table 2.5: Linear regression models for Immune activation (measured by CRP and IL6 levels) and General Cognitive Function (measured by Full scale IQ).

| Model | Variable | IL6 | N   | B    | B1   | P   | CRP | N   | B    | B1   | P    |
|-------|----------|-----|-----|------|------|-----|-----|-----|------|------|------|------|
| 1     | IQ       |     | 3940| -0.576| -0.039| 0.014|     | 3948| -1.109| -0.065| 0.00005|
|       | Gender   |     |     | 0.208| 0.007| 0.684|     | 0.221| 0.007| 0.666|
|       | Age      |     |     | -0.012| -0.08| <0.001|     | -0.012| -0.083| <0.001|
|       | BMI      |     |     | -0.093| -0.016| 0.338|     | -0.058| -0.01| 0.563|
|       | Mother’s education |     | 4.667| 0.339| <0.001|     | 4.654| 0.339| <0.001|
|       | Paternal class |     | -0.694| -0.108| <0.001|     | -0.693| -0.108| <0.001|
|       | ethnicity |     | 0.305| 0.012| 0.446|     | 0.3| 0.012| 0.453|

Model 1: Analysis unadjusted for other variables; Model 2: Gender, age, maternal education, BMI, ethnicity and paternal class included as covariates.

*B* = unstandardized beta value; *B1* = standardized beta value; *R*\(^2\) change per model.
2.3.4 Does Immune response partly account for the variance in cognitive performance explained by ELA?

Given the association between IL6, CRP and IQ, an analysis was carried out in which IQ was the dependent variable, Harsh Parenting was the independent variable, and each of IL6 or CRP (taken in turn) were tested as covariates. This analysis examined whether the variation in IQ explained by Harsh parenting was reduced (hence, partially accounted for) when immune function was considered. Neither immune marker (IL6 or CRP) was observed to explain the relationship between harsh parenting and IQ; After immune response was covaried for, Harsh Parenting continued to explain a comparable percentage of variation in IQ when either marker was included in the analysis (IL6: \( b = -3.209, p = 2.4356 \times 10^{-9} \); CRP: \( b = 3.204, p = 2.5535 \times 10^{-9} \)). To further rule out an interaction between immune markers and ELA, interaction terms were created using mean centered values for IL6 and CRP multiplied by scores of Harsh Parenting. These interaction terms were tested for association with IQ in linear regression. When added to the model, neither interaction term predicted IQ scores (IL6*harsh parenting: \( b = -0.462, p = 0.392, N = 3501 \); CRP*Harsh Parenting \( b = 0.158, p = 0.443, N = 3511 \)).
2.4 Discussion

This paper sought to characterise the effects of ELA on cognitive ability in children, and to examine whether any observed effects were partly accounted for by immune response. We hypothesised that ELA would negatively impact on children’s cognitive development, specially disrupting social cognition. Unfortunately, a relationship was not observed between poly-victimization and social cognition at 13 years. Following a nominal association between total number of ELAs experienced (poly-victimisation score) and general cognitive ability (IQ) at age 8, we observed a significant association between ‘Harsh Parenting’ before age 5 (measured in terms of maternal physical discipline) and IQ at age 8 years in secondary analysis. This association survived correction and remained strongly significant after accounting for other relevant covariates, including maternal education and gender. Finally, although elevated levels of IL-6 and CRP were significantly associated with lower IQ scores, these inflammatory markers were not observed to account for the association between ELA and IQ.

2.4.1 Harsh parenting and cognition.

Our analyses support the view that ‘harsh parenting’ in early childhood is associated with lower general cognitive ability in middle childhood. Specifically, the experience of harsh parenting prior to the age of 5 was associated with an average reduction of 3 scaled score IQ points compared to children without this experience. In a post-hoc analysis, we further observed a dose-response in this relationship, such that increased frequency of harsh parenting was associated with larger effects on IQ. Furthermore, covarying for maternal educational attainment in the analysis, as a proxy for maternal IQ, did not appear to explain this relationship. This association may reflect a causal relationship, or alternatively the combined effect of other biological and/or environmental factors jointly affecting both parenting and general cognitive ability. The suggestion of a causal relationship is supported by the temporality of variables, as ELA was measured before age 5 and IQ measured later at age 8. However, in a large study of ELA and IQ, cognitive deficits in victimized individuals were largely
explained by cognitive deficits that predated childhood victimization and by confounding genetic and environmental risks (Danese et al., 2016). As there was only one time-point for IQ available for the current analysis, an alternative explanation is that those with lower IQ may be more susceptible to harsh parenting.

2.4.2 Harsh parenting, lower IQ, and inflammatory response.

Our hypothesis that the effects of ELA (such as harsh parenting) on cognition might be partially explained by inflammatory response was not supported. While elevated levels of inflammatory markers IL-6 and CRP were both found to be negatively correlated with IQ scores, neither IL-6 nor CRP significantly accounted for variation in IQ explained by ELA. A recent meta-analysis of ELA and inflammatory response found that physical abuse was associated with higher IL6 scores but not CRP, emotional abuse was not associated with either marker, while Tumour Necrosis Factor a (TNFa) was strongly influenced by early trauma (Baumeister et al., 2016). The unavailability of additional immune markers, such as TNFa, precludes us from drawing firm conclusions about the potential effects of these markers on the relationship between ELA and IQ in this dataset.

2.4.3 Alternative explanations for the relationship between ELA and cognition

As an alternative to immune-based accounts for the effects of ELA on cognition and social cognition, cognitive developmental accounts of ELA focus on how a child’s representation of the world is thwarted due to past experiences in a ‘cascade of increasingly deviant development’ (Bramon, Kelly, Van Os, & Murray, 2001). If children are deprived of consistent, sensitive caregiving and a responsive loving exchange early in life, it has been argued that this can have a profound effect on cognitive development (Fox, 1993). Prior to developing coping strategies for stress, children withdraw and isolate themselves following ELA, as their understanding of the world is biased due to perceived threat at home (Fox, 1993). Avoiding adult social interactions slows down the process of social learning which might otherwise take place – for example, learning language, appropriate behaviours and
communication of ideas (May-Chahal & Cawson, 2005). The relevance of parent-child interactions to cognitive development has already been supported by Chong et al. (S. Y. Chong et al., 2016) in a previous study of the ALSPAC dataset. In this study differences in IQ at age 8 were associated with ‘parental warmth’ and ‘parental control’, measured up to 47 months. They found that higher parental control was associated with lower IQ scores (S. Y. Chong et al., 2016). While physical reprimanding was not assessed in the Chong et al study (S. Y. Chong et al., 2016), and parenting styles were not a subject of the current study, it is not difficult to imagine that these variables may be correlated. Future research should investigate whether and how these variables are related regarding cognitive development.

2.4.4 Strengths and Limitations

A strength of the study was availability of a large epidemiological dataset consisting of longitudinal data, where the effects of ELA could be associated with later cognitive function. Often in psychiatric studies ELA is assessed retrospectively considering a diagnosis, whereas here ELA data was collected contemporaneously on a year to year basis. Another strength of this study was the availability of IL6 and CRP blood markers for analysis in longitudinal data. A limitation of the study, however, is the unavailability of other markers of immune response: it is possible that a measure of TNFa may have been more relevant to ELA analysis (Baumeister et al., 2016). Similarly, the fact that these measures were taken on average one year after assessment of IQ precluded a more direct analysis of the mediating role of inflammatory response between ELA and IQ.

A limitation to this study is the lack of clear-cut, well validated ELA definitions in the field. While we studied here the presence or absence of individual ELAs, other ALSPAC-based studies have used a cumulative score of childhood adverse events not specific to parent-child relationship (Fisher et al., 2012) (Slopen, Kubzansky, McLaughlin, et al., 2013). Sexual Abuse was not discussed in the current study, as there was only a small number of children with experience of this ELA (n=27), making the analysis low powered in comparison to others. Furthermore, there was a substantial
amount of missing data for each of the ELA measures computed. Missingsness may represent a bias in the data, as many of those in abusive homes may not attend follow-up visits, and so the data available may be based on more functional homes. The ‘harsh parenting’ variable is based on time-points of 2 and year 3.5 years, thus captures a more representative sample than other measures of ELA. This score was derived from composite ELA measures in previous ALSPAC research (Fisher et al., 2012). A limitation of the use of this variable is a lack of corroborative evidence from witnesses of ELA, such as evidence from a partner and the child themselves.

Finally, no family and child medical history was used in this analysis, nor were medications used as covariates. From previous studies of the ALSPAC data it is clear that various atopic disorders can impact on inflammatory markers (G. M. Khandaker, S. Zammit, et al., 2014), as can various medications and conditions. Furthermore, environmental factors like sibling number, birth difficulties and genetics could be also included to further understand developmental context and inflammatory response to stress. Although beyond the scope of the current study, further data on depression scores or evidence of psychotic disorders at later time-points could have been included in analysis, especially as these have already been cited as impacting cognition and immune response in ALSPAC (G. M. Khandaker, R. M. Pearson, et al., 2014; Golam Mohammed Khandaker et al., 2017). Strategies for coping including internalizing and externalizing indicators could also be taken into consideration in analysis on childhood development whereby an underlying vulnerability can predict psychological and cognitive outcomes (McElroy, Shevlin, & Murphy, 2017). Emotional regulation, which was not included in this analysis, may be integral to understanding cognitive outcomes, as distress has been cited as a key factor in cognitive development (Thewissen et al., 2011).

2.4.5 Conclusion

This study, based on a longitudinal epidemiological cohort, provides evidence that harsh parenting during the early years of development is associated with lower cognitive performance in middle childhood by
comparison with children who have not had this experience. This study also suggested a dose-response in this relationship, whereby more frequent smacking/slapping was associated with lower cognitive performance than less frequent physical discipline. These findings may have implications for public health interventions aimed at supporting caregivers, particularly in emphasising the importance of adopting alternative parental disciplining methods. Finally, our study suggested that altered immune function, as measured by IL6 and CRP, is unlikely to explain a large proportion of the variation in IQ explained by ELA. While other immune markers remain to be examined, our study highlights the need to consider alternative biological and cognitive pathways for explaining the relationship between ELA and cognition.
2.5 Supplementary Figure

Timeline for ALSPAC data collection
2.6 Acknowledgements

Jessica Holland’s research is supported by a PhD scholarship from the Irish Research Council. GD’s research is supported by generous funding from the European Research Council (grant no.677467) and Science Foundation Ireland grant no.16/ERCS/3787).

Dr Khandaker acknowledges funding support from the Wellcome Trust (Intermediate Clinical Fellowship; grant no: 201486/Z/16/Z), MRC (grant no: MC_PC_17213), and MQ: Transforming Mental Health (grant no: MQDS17/40).

Prof Zammit acknowledges funding support from the National Institute for Health Research Biomedical Research Centre at the University Hospitals Bristol National Health Service Foundation Trust and the University of Bristol.

Access to the ALSPAC dataset was granted by University of Bristol ALSPAC committee. We are grateful to all families who took part in the ALSPAC study, midwives for their help in recruitment, and the whole ALSPAC team, including interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

Conflicts of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.
3. Study 2. Beyond C4: Analysis Of The Complement Gene Pathway Shows Enrichment For IQ In Patients With Psychotic Disorders And Healthy Controls

Jessica F. Holland¹, Donna Cosgrove¹, Laura Whitton¹, Denise Harold²,³, Aiden Corvin², Michael Gill², David O. Mothersill¹, Derek W. Morris¹, Gary Donohoe¹.

1. Cognitive Genetics & Cognitive Therapy Group, The Centre for Neuroimaging, Cognition and Genomics (NICOG), School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Ireland.

2. Neuropsychiatric Genetics Research Group, Department of Psychiatry, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland

3. School of Biotechnology, Dublin City University, Dublin, Ireland
Abstract

Introduction Variation in cognitive performance, which strongly predicts functional outcome in SZ (SZ), has been associated with multiple immune-relevant genetic loci. These loci include complement component 4 (C4A), structural variation at which was recently associated with SZ risk and synaptic pruning during neurodevelopment and cognitive function. Here, we test whether this genetic association with cognition and SZ risk is specific to C4A, or extends more broadly to genes related to the complement system.

Methods Using a gene-set with an identified role in ‘complement’ function (excluding C4A), we used MAGMA to test if this gene-set was enriched for genes associated with human intelligence and SZ risk, using genome-wide association summary statistics (IQ: N=269,867, SZ; N=105,318). We followed up this gene-set analysis with a complement gene-set polygenic score (PGS) regression analysis in an independent dataset of patients with psychotic disorders and healthy participants with cognitive and genomic data (N=1000).

Results and conclusions Enrichment analysis suggested that genes within the complement pathway were significantly enriched for genes associated with IQ, but not SZ. In a gene-based analysis of 90 genes, SERPING1 was the most enriched gene for the phenotype of IQ. In a PGS regression analysis, we found that a complement pathway PGS associated with IQ GWAS statistics also predicted variation in IQ in our independent sample. This association (observed across both patients and controls) remained significant after controlling for the relationship between C4A and cognition. These results suggest a robust association between the complement system and cognitive function, extending beyond structural variation at C4A.
3.1 Introduction

SZ is a complex heritable neurodevelopmental disorder, in which level of disability is strongly predicted by deficits in cognitive function (Sekar et al., 2016). Genome-wide association studies (GWAS) have consistently found evidence of association between SZ and variation at or within the major histocompatibility complex (MHC) locus on chromosome 6 (McAllister, 2014). While both the size and high degree of linkage disequilibrium within the region has made it difficult to disentangle the contribution of this region to SZ risk, Sekar et al. (Sekar et al., 2016) demonstrated that part of the association signal at this locus could be explained by structural variation mapping to C4A. C4A is part of the complement system, an innate immune system pathway comprising a large number of plasma proteins that help antibodies and phagocytes to clear pathogens (Sekar et al., 2016). This study found that structural variation at C4 was associated with (a) variation in RNA expression levels in post-mortem brain samples, (b) increased risk for SZ, and (c) that increased RNA expression predicted by C4 structural variation influenced development of the neural system via a pattern of altered synaptic pruning. Following up these findings, we recently showed that higher predicted C4 RNA expression was associated with poorer performance on measures of memory function in both patients with SZ and healthy participants, and with reduced cortical activation during visual task performance in healthy participants (G Donohoe et al., 2018). This, along with evidence that other risk-related genes and proteins within the complement system are also associated with variation in both cognition and brain structure in both patients and healthy individuals (Allswede et al., 2018; Athanasiu et al., 2017; Bralten et al., 2011; Chung et al., 2014; G Donohoe, Walters, et al., 2013; Nettiksimmons, Tranah, Evans, Yokoyama, & Yaffe, 2016; Rose et al., 2013; Song et al., 2014; C. Zhang et al., 2017), led us to consider whether variation within genes that encode the complement system could, as a whole, influence cognition.

Examples of other complement function related genes that have been associated with both SZ risk and cognition include Cub and Sushi Multiple
Domains-1 (CSMD1), complement Factor H (CFH), and complement C3b/C4b receptor 1 (CR1) (G Donohoe et al., 2018; Dunkelberger & Song, 2010; Föcking et al., 2019; Gigante et al., 2011). In patients with SZ and health participants, we previously reported that a GWAS-identified SZ risk variant within CSMD1, which encodes for a regulator of complement, was associated with poorer general cognitive ability and episodic memory function (G Donohoe, Walters, et al., 2013). Other variants within CSMD1 gene have also been associated with poorer cognitive performance in large samples of healthy participants (Athanasiu et al., 2017). In a sample of 1,783 patients with SZ and healthy controls, expression of complement factor H (CFH) in the hippocampus was found to be associated with both increased SZ risk and poorer memory function (C. Zhang et al., 2017). The CR1 gene has also been associated with cognitive impairment in a study of AD (Chung et al., 2014), and entorhinal cortex volume in young healthy adults (Bralten et al., 2011). In addition to these examples, complement gene expression more broadly has been linked to superior frontal cortex thickness in healthy humans (Allswede et al., 2018), which has been linked to general intelligence in previous studies (Jung & Haier, 2007).

Furthermore, in a proteomic analysis of the complement signalling pathway using longitudinal population-based data, individuals who went on to experience psychotic like experience were found to have upregulation of multiple complement proteins. (English et al., 2017; Föcking et al., 2019).

Given these lines of evidence, the present study aimed to systematically characterise, for the first time, the association between genetic variation within genes related to complement function and both SZ risk and cognition. To do this, we performed gene-set analysis (GSA) of complement pathway genes, based on prior work in curating complement gene-sets based on publically available data by (Birnbaum et al., 2018; Qian et al., 2019). The resulting complement gene-set was tested for enrichment using IQ GWAS summary data from Savage et al. (Savage et al., 2018), and SZ GWAS summary data from the Psychiatric Genomics Consortium (PGC) (Ripke et al., 2014). We tested the hypotheses that 1) the complement pathway gene-set (excluding C4) would show significant
enrichment for genes associated with both IQ and with SZ. Depending on whether these hypotheses were supported, we sought to further test whether 2) a polygenic score for complement pathway genes based on IQ or SZ GWAS summary statistics would explain variation in cognitive performance in a sample of patients with psychotic disorders and healthy controls.

3.2 Methodology

3.2.1 Samples included in enrichment analysis

IQ GWAS dataset (Savage et al 2018; 18): This dataset included GWAS summary statistics on 269,867 individuals from 14 cohorts: UK Biobank (UKB), the Cognitive Genomics Consortium (COGENT), the Rotterdam Study (RS), the Generation R Study (GENR), the Swedish Twin Registry (STR), Spit for Science (S4S), the High-IQ/Health and Retirement Study (HiQ/HRS), the Twins Early Development Study (TEDS), the Danish Twin Registry (DTR), IMAGEN: reinforcement-related behaviour in normal brain function and psychopathology (Schumann et al., 2010), the Brisbane Longitudinal Twin Study (BLTS), the Netherlands Study of Cognition, Environment, and Genes (NESCOG), Genes for Good (GfG), and the Swedish Twin Studies of Aging (STSA) (Savage et al., 2018). Participants ranged from children to older adults, with older samples being screened for cognitive decline to exclude the possibility of dementia affecting performance on neurocognitive tests. This analysis identified 205 genomic loci (harbouring 1,016 genes) related to human intelligence. In each cohort, varying measures of intelligence were subject to a principal components analysis to create a common latent g factor underlying multiple dimensions of cognition (Savage et al., 2018). For each single nucleotide polymorphism (SNP), computed beta value represented the likelihood of increased ‘g’ factor scores based on the SNP of interest. The combined summary statistics of these GWAS in relation to the IQ phenotype were used for enrichment analysis.

SZ GWAS dataset (Pardiñas et al., 2018): This recent GWAS for SZ combined existing genome wide association data (Ripke et al., 2014) with genome-wide genotype information for SZ cases from the UK (the
CLOZUK sample), amalgamated with control datasets obtained from public repositories to form a meta-analysis of 40,675 SZ cases and 64,643 healthy controls (Pardiñas et al., 2018), excluding overlapping samples. This meta-analysis found 145 loci related to SZ diagnosis, with 50 of these novel discoveries. An independent sample total of 5,762 cases and 154,224 controls was used in analysis to replicate meta-analytic findings. Details of samples can be found in supplementary information on the original paper (Pardiñas et al., 2018; Ripke et al., 2014). Participants in the SZ sample were diagnosed based on clinicians report or using a specific inclusion criteria and research-based assessment. Healthy controls were screened for other psychiatric illness and were within normal IQ range. Summary statistics of this GWAS provided an odds ratio score for likelihood of developing SZ based on risk alleles of each SNP.

### 3.2.2 Complement gene-set

To test for enrichment of genes related to complement function in IQ and SZ, we based our gene-set list on recent publications targeting complement-related immunity genes (Birnbaum et al., 2018; Qian et al., 2019). The Birnbaum paper (Birnbaum et al., 2018) provided a gene-set related to complement in SZ by collating information from neuroimmunology and general immunology literature, as well as pathway annotation programs, cross-referenced with multiple gene expression databases (eg. GO, KEGG, IMPORT, IPA, and IMMUNOME) to assemble 34 complement-related genes (Dunkelberger & Song, 2010; Orsini, De Blasio, Zangari, Zanier, & De Simoni, 2014; Ricklin, Hajishengallis, Yang, & Lambris, 2010; Sarma & Ward, 2011; Veerhuis et al., 2011). The second publication (Qian et al., 2019) included 32 of the original 34 genes with additional gene lists included from Molecular Signatures Database, the Human Biological Pathway Unification Database and the HUGO Gene Nomenclature Committee (https://pathcards.genecards.org, http://software.broadinstitute.org/gsea/msigdb/index.jsp, https://www.genenames.org) searching the keyword “complement”. After removing the duplicated genes and genes directly encoding for C4 (C4A, C4B, C4BPA, C4BPB, C4_B), 108 genes remained as the candidate genes
for further analysis (Supplementary Table 1). As 90 of these genes were available to be tested for enrichment using MAGMA, these 90 genes were brought forward for analysis (listed in Table 1).

### 3.2.3 Enrichment analysis statistical approach

The statistical toolbox MAGMA (https://ctg.cncr.nl/software/magma; (de Leeuw, Mooij, Heskes, & Posthuma, 2015) was used to test for enrichment of the complement gene-set for genes associated with the IQ and SZ phenotypes. Enrichment analysis based on MAGMA consisted of three steps: first, an annotation step to map SNPs onto genes; second, a gene analysis step to compute gene p-values; and three, a gene-level analysis step, carried out both on a gene-wide and a gene-set-wide level. By doing so, MAGMA tests whether multiple genetic markers in a given gene or gene-set (e.g. complement genes) are more strongly associated with SZ or IQ than other genes in the genome.

### 3.2.4 Samples included in polygenic score (PGS) analysis

In total, 808 cases and 192 healthy participants completed a full neuropsychological assessment battery and had full genome-wide data available. Cases consisted of n=585 clinically stable patients with a diagnosis of SZ and schizoaffective disorder (SZA), and an additional n=223 patients diagnosed with bipolar disorder with psychotic features, major depressive disorder with psychotic features, delusional disorder, or psychosis not otherwise specified, as described elsewhere (D Cosgrove et al., 2017; G Donohoe et al., 2018; Whitton et al., 2016). Participants were recruited from five sites across Ireland, and ethics approval was obtained from local ethics committees. Written informed consent was obtained from all participants. Inclusion criteria required participants to be clinically stable at the time of cognitive assessment, aged between 18–65 years, had no history of comorbid psychiatric disorder, history of seizures, substance abuse in the preceding 6 months, or prior head injury with loss of consciousness. All participants had Irish ancestry (all four grandparents born in Ireland). Patients were diagnosed by trained psychiatrists using the Structured Clinical Interview for DSM-IV Axis I Diagnosis (First, Spitzer,
Gibbon, & Williams, 2002). Additional diagnostic and clinical information ascertained at time of interview, including symptom severity (Andreasen, 1984) and medication dosage were also recorded for analysis.

The healthy control sample was recruited on the basis of response to local media advertisements. Control participants were included if they were aged between 18 and 65 and met the criteria of having no history of major mental health problems, intellectual disability or acquired brain injury based on clinical interview, and no history of substance misuse in the preceding 6 months based on self-report. Control participants were also excluded from the study if they reported having a first-degree relative with a history of psychosis. Both patient and control clinical assessments were conducted in accordance with the relevant ethics committees' approval from each participating site. All participants provided written informed consent.

### 3.2.5 Cognitive assessment

General cognitive functioning (IQ) was measured in the sample using selected subtests (Vocabulary, Similarities, Block Design and Matrix Reasoning) from the Wechsler Adult Intelligence Scale, third edition (WAIS-III; (Wechsler, 1997)), deriving a full scale, verbal and performance IQ for each participant. Premorbid IQ was measured using the Wechsler test of adult reading (WTAR; (Holdnack, 2001). Episodic memory was assessed in the Irish samples using the logical memory subtests (immediate and delayed conditions) from the Wechsler Memory Scale, third edition (WMS-III) (Wechsler, 1997), and the paired associations learning task (PAL; stages completed and total errors) from the Cambridge Automated Neuropsychological Test Battery (CANTAB; (Robbins et al., 1994). As in our previous study of C4, an unrotated principle components analysis based on the four available episodic memory tests was carried out to reduce multiple testing burden. This memory factor explained 72% of variance in memory scores, as described previously (G Donohoe et al., 2018). Spatial working memory was assessed in the samples using the spatial working memory task from the Cambridge Automated Neuropsychological Test Battery (CANTAB SWM; (Robbins et al., 1994).
### 3.2.6 Genotyping

Genotyping was conducted on DNA extracted from whole blood or saliva. Full GWAS data were available for all samples. A proportion of samples (n=575) were genotyped with an Affymetrix 6.0 chip (Santa Clara, CA, USA; as part of the WTCCC2 (I. S. G. Consortium & 2, 2012) and the remainder on the Illumina HumanCoreExome chip (San Diego, CA, USA), granting comparisons between platforms indicate high agreement (Barnes, Freudenberg, Thompson, Aronow, & Pavlidis, 2005). SNPs were excluded on the basis of MAF (minor allele frequency) <0.1% (i.e rate of second most common allele occurs at a frequency of 10% or greater in the population), SNP missingness $\leq 2\%$, and Hardy–Weinberg equilibrium $\leq \text{P}^{10^{-6}}$. Imputation was carried out on these data sets separately using 1000 Genomes Phase I integrated haplotypes (Dec 2013 release) and IMPUTE2 to give ~10 million SNPs genome-wide per sample.

### 3.2.7 PGS score calculation

Based on the complement set of 90 genes described above, we began by identifying all SNPS within (+/−20 kb) of these gene loci, and then extracting the genotype values for our samples, which had been genotyped using either Affymetrix 6.0 or Illumina HumanCoreExome. PRSice software (https://choishingwan.github.io/PRSice/(Euesden, Lewis, & O’reilly, 2014)) was used to perform quality control on data from each sample and SNPs were excluded from further analysis if there was >10% missing genotype data, HardyWeinberg equilibrium $<1 \times 10^{-5}$, or MAF $<1\%$. Next, the lists of SNPs that passed QC for each sample were joined, providing a SNP list common to each sample (n = 14237 SNPs). Based on this list, an effect-size weighted PGS was generated for each individual using PRSice based on a threshold of p<0.05 for SNPs in the IQ and SZ GWAS sets. While multiple p value thresholds can be used, we have previously found a threshold of p = 0.05 to be the most informative in terms of signal-to-noise ratio (D Cosgrove et al., 2017).
3.2.8 Statistical Analysis

To estimate the association between computed complement PGS and performance on cognitive tasks, multiple regression analyses were carried out on the whole sample and on the patient sample only using IBM SPSS Statistics Version 25.0 (IBM Corp, Armonk, NY, 2017). Three cognitive domains of IQ, episodic memory and working memory were tested. PGS scores for complement were used as the independent variable, and age and gender were entered as covariates of no interest. For IQ analysis, only gender was entered as a covariate, as scaled scores were used. To maximize power to detect differences, we carried out our analysis on the full dataset of cases and controls (n=1000). To test the possibility of a particular phenotype population driving results, analysis was also carried out in healthy controls and patients separately.

3.3 Results

3.3.1 Enrichment analysis

Using MAGMA to test for the enrichment of the complement gene-set in IQ and SZ GWAS data, the complement gene-set showed enrichment for IQ (\( \beta = 0.29, p = 0.010; \alpha = 0.05 / 2 = 0.025 \) after bonferroni correction for the two phenotypes tested). By comparison, no enrichment was observed using the SZ data (\( \beta = -0.07, p = 0.26 \)). Following up the complement gene-set analysis with a gene-based association analysis, 32 of the 90 genes included showed nominal association for the phenotype of IQ, with 12 of these genes surviving correction for multiple testing, using bonferroni correction for 90 genes (\( \alpha = 0.05 / 90 = 0.00055 \)) of which SERPING1 was the highest associated (see Table 1).
Table 3.1: MAGMA gene-wide analysis of enrichment for Complement system genes for IQ.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Entrez ID</th>
<th>CHR</th>
<th>start</th>
<th>stop</th>
<th>#SNPS</th>
<th>ZSTAT</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SERPING1</td>
<td>710</td>
<td>11</td>
<td>57345027</td>
<td>57402326</td>
<td>145</td>
<td>5.7697</td>
<td>3.97E-09</td>
</tr>
<tr>
<td>PROS1</td>
<td>5627</td>
<td>3</td>
<td>93571881</td>
<td>93712934</td>
<td>233</td>
<td>5.0005</td>
<td>2.86E-07</td>
</tr>
<tr>
<td>F2</td>
<td>2147</td>
<td>11</td>
<td>46720743</td>
<td>46781056</td>
<td>145</td>
<td>4.8382</td>
<td>6.55E-07</td>
</tr>
<tr>
<td>CFH</td>
<td>3075</td>
<td>1</td>
<td>196601008</td>
<td>196736634</td>
<td>445</td>
<td>4.7923</td>
<td>8.24E-07</td>
</tr>
<tr>
<td>CD46</td>
<td>4179</td>
<td>1</td>
<td>207905383</td>
<td>207988861</td>
<td>211</td>
<td>4.673</td>
<td>1.48E-06</td>
</tr>
<tr>
<td>PROC</td>
<td>5624</td>
<td>2</td>
<td>128155996</td>
<td>128206822</td>
<td>158</td>
<td>3.841</td>
<td>6.13E-05</td>
</tr>
<tr>
<td>F7</td>
<td>2155</td>
<td>13</td>
<td>113740102</td>
<td>113794995</td>
<td>162</td>
<td>3.652</td>
<td>0.00013</td>
</tr>
<tr>
<td>PROC1</td>
<td>10544</td>
<td>20</td>
<td>33739740</td>
<td>33823771</td>
<td>190</td>
<td>3.604</td>
<td>0.000157</td>
</tr>
<tr>
<td>CD59</td>
<td>966</td>
<td>11</td>
<td>33704556</td>
<td>33778025</td>
<td>245</td>
<td>3.5001</td>
<td>0.000233</td>
</tr>
<tr>
<td>CR1L</td>
<td>1379</td>
<td>1</td>
<td>207798458</td>
<td>207917036</td>
<td>483</td>
<td>3.48</td>
<td>0.000251</td>
</tr>
<tr>
<td>CR1</td>
<td>1378</td>
<td>1</td>
<td>207649473</td>
<td>207835110</td>
<td>531</td>
<td>3.4712</td>
<td>0.000259</td>
</tr>
<tr>
<td>C1QB</td>
<td>713</td>
<td>1</td>
<td>22959682</td>
<td>23008130</td>
<td>165</td>
<td>3.355</td>
<td>0.000397</td>
</tr>
<tr>
<td>F12</td>
<td>2161</td>
<td>5</td>
<td>176809139</td>
<td>176856577</td>
<td>68</td>
<td>3.2023</td>
<td>0.000682</td>
</tr>
<tr>
<td>F10</td>
<td>2159</td>
<td>13</td>
<td>113757113</td>
<td>113823843</td>
<td>140</td>
<td>3.1311</td>
<td>0.000871</td>
</tr>
<tr>
<td>C1QC</td>
<td>714</td>
<td>1</td>
<td>22949969</td>
<td>22994603</td>
<td>161</td>
<td>3.1088</td>
<td>0.000939</td>
</tr>
<tr>
<td>C1QA</td>
<td>712</td>
<td>1</td>
<td>22943118</td>
<td>22986175</td>
<td>155</td>
<td>2.9931</td>
<td>0.001381</td>
</tr>
<tr>
<td>CFHR2</td>
<td>3080</td>
<td>1</td>
<td>196892888</td>
<td>196948356</td>
<td>135</td>
<td>2.8875</td>
<td>0.001942</td>
</tr>
<tr>
<td>THBD</td>
<td>7056</td>
<td>20</td>
<td>23006270</td>
<td>23050301</td>
<td>168</td>
<td>2.8522</td>
<td>0.002171</td>
</tr>
<tr>
<td>C8A</td>
<td>731</td>
<td>1</td>
<td>57300443</td>
<td>57403894</td>
<td>453</td>
<td>2.805</td>
<td>0.002516</td>
</tr>
<tr>
<td>C2</td>
<td>717</td>
<td>6</td>
<td>31845562</td>
<td>31933449</td>
<td>320</td>
<td>2.63</td>
<td>0.00427</td>
</tr>
<tr>
<td>CFI</td>
<td>3426</td>
<td>4</td>
<td>110641848</td>
<td>110743381</td>
<td>355</td>
<td>2.593</td>
<td>0.004757</td>
</tr>
<tr>
<td>Gene</td>
<td>ID</td>
<td>Chromosome</td>
<td>Start</td>
<td>End</td>
<td>Length</td>
<td>Log2 Fold Change</td>
<td>p-value</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>------------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-----------------</td>
<td>---------</td>
</tr>
<tr>
<td>PLAU</td>
<td>5328</td>
<td>10</td>
<td>75650862</td>
<td>75697259</td>
<td>136</td>
<td>2.4843</td>
<td>0.00649</td>
</tr>
<tr>
<td>CFHR5</td>
<td>81494</td>
<td>1</td>
<td>196926667</td>
<td>196998804</td>
<td>182</td>
<td>2.3663</td>
<td>0.008983</td>
</tr>
<tr>
<td>SERPINC1</td>
<td>462</td>
<td>1</td>
<td>173852938</td>
<td>173906516</td>
<td>139</td>
<td>2.3036</td>
<td>0.010623</td>
</tr>
<tr>
<td>CFHR4</td>
<td>10877</td>
<td>1</td>
<td>196837144</td>
<td>196907843</td>
<td>183</td>
<td>2.2494</td>
<td>0.012244</td>
</tr>
<tr>
<td>SERPINA5</td>
<td>5104</td>
<td>14</td>
<td>95027706</td>
<td>95079457</td>
<td>300</td>
<td>2.1918</td>
<td>0.014197</td>
</tr>
<tr>
<td>CFB</td>
<td>629</td>
<td>6</td>
<td>31893721</td>
<td>31939861</td>
<td>178</td>
<td>2.1439</td>
<td>0.016021</td>
</tr>
<tr>
<td>PLAUR</td>
<td>5329</td>
<td>19</td>
<td>44130247</td>
<td>44194498</td>
<td>206</td>
<td>2.0839</td>
<td>0.018585</td>
</tr>
<tr>
<td>F5</td>
<td>2153</td>
<td>1</td>
<td>169461192</td>
<td>169575769</td>
<td>510</td>
<td>2.0443</td>
<td>0.020461</td>
</tr>
<tr>
<td>CFHR1</td>
<td>3078</td>
<td>1</td>
<td>196768861</td>
<td>196821319</td>
<td>27</td>
<td>1.9089</td>
<td>0.028136</td>
</tr>
<tr>
<td>F2RL2</td>
<td>2151</td>
<td>5</td>
<td>75891307</td>
<td>75939259</td>
<td>245</td>
<td>1.8605</td>
<td>0.031407</td>
</tr>
<tr>
<td>CLTC</td>
<td>1213</td>
<td>17</td>
<td>57677050</td>
<td>57794317</td>
<td>210</td>
<td>1.6987</td>
<td>0.04469</td>
</tr>
<tr>
<td>CLU</td>
<td>1191</td>
<td>8</td>
<td>27434434</td>
<td>27492328</td>
<td>185</td>
<td>1.5628</td>
<td>0.059049</td>
</tr>
<tr>
<td>PLAT</td>
<td>5327</td>
<td>8</td>
<td>42012236</td>
<td>42085242</td>
<td>187</td>
<td>1.5332</td>
<td>0.062618</td>
</tr>
<tr>
<td>COLEC10</td>
<td>10584</td>
<td>8</td>
<td>119944527</td>
<td>120139207</td>
<td>849</td>
<td>1.4206</td>
<td>0.077709</td>
</tr>
<tr>
<td>CR2</td>
<td>1380</td>
<td>1</td>
<td>207607645</td>
<td>207683240</td>
<td>244</td>
<td>1.3591</td>
<td>0.087065</td>
</tr>
<tr>
<td>CFHR3</td>
<td>10878</td>
<td>1</td>
<td>196723925</td>
<td>196783203</td>
<td>33</td>
<td>1.3096</td>
<td>0.095166</td>
</tr>
<tr>
<td>FCN1</td>
<td>2219</td>
<td>9</td>
<td>137777711</td>
<td>137829849</td>
<td>238</td>
<td>1.3075</td>
<td>0.095516</td>
</tr>
<tr>
<td>F2R</td>
<td>2149</td>
<td>5</td>
<td>75991868</td>
<td>76051595</td>
<td>238</td>
<td>1.2813</td>
<td>0.10004</td>
</tr>
<tr>
<td>C3</td>
<td>718</td>
<td>19</td>
<td>6657846</td>
<td>6740662</td>
<td>317</td>
<td>1.2771</td>
<td>0.10079</td>
</tr>
<tr>
<td>LMAN1</td>
<td>3998</td>
<td>18</td>
<td>56975055</td>
<td>57046508</td>
<td>245</td>
<td>1.1553</td>
<td>0.12399</td>
</tr>
<tr>
<td>SERPINA1</td>
<td>5265</td>
<td>14</td>
<td>94823084</td>
<td>94877029</td>
<td>269</td>
<td>1.0904</td>
<td>0.13777</td>
</tr>
<tr>
<td>FCN2</td>
<td>2220</td>
<td>9</td>
<td>137752658</td>
<td>137799366</td>
<td>242</td>
<td>1.0838</td>
<td>0.13924</td>
</tr>
<tr>
<td>BDKRB2</td>
<td>624</td>
<td>14</td>
<td>96651016</td>
<td>96730666</td>
<td>332</td>
<td>1.0752</td>
<td>0.14115</td>
</tr>
<tr>
<td>ITGAX</td>
<td>3687</td>
<td>16</td>
<td>31346488</td>
<td>31414320</td>
<td>141</td>
<td>0.9437</td>
<td>0.17266</td>
</tr>
<tr>
<td>SERPINF2</td>
<td>5345</td>
<td>17</td>
<td>1626015</td>
<td>1678562</td>
<td>191</td>
<td>0.93225</td>
<td>0.1756</td>
</tr>
<tr>
<td>Gene</td>
<td>Score</td>
<td>Rank</td>
<td>Fold</td>
<td>F1</td>
<td>Fold</td>
<td>Score</td>
<td>Rank</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>C1QBP</td>
<td>708</td>
<td>17</td>
<td>5316099</td>
<td>5362471</td>
<td>197</td>
<td>0.86574</td>
<td>0.19332</td>
</tr>
<tr>
<td>CFD</td>
<td>1675</td>
<td>19</td>
<td>839665</td>
<td>883624</td>
<td>109</td>
<td>0.76449</td>
<td>0.22229</td>
</tr>
<tr>
<td>C8G</td>
<td>733</td>
<td>9</td>
<td>139819333</td>
<td>139861426</td>
<td>160</td>
<td>0.74116</td>
<td>0.2293</td>
</tr>
<tr>
<td>MASP2</td>
<td>10747</td>
<td>1</td>
<td>11066580</td>
<td>11127296</td>
<td>152</td>
<td>0.72154</td>
<td>0.23529</td>
</tr>
<tr>
<td>SERPINE1</td>
<td>5054</td>
<td>7</td>
<td>100750370</td>
<td>100802547</td>
<td>180</td>
<td>0.69982</td>
<td>0.24202</td>
</tr>
<tr>
<td>F13B</td>
<td>2165</td>
<td>1</td>
<td>196987877</td>
<td>197056568</td>
<td>176</td>
<td>0.65701</td>
<td>0.25559</td>
</tr>
<tr>
<td>VTN</td>
<td>7448</td>
<td>17</td>
<td>26674298</td>
<td>26717373</td>
<td>83</td>
<td>0.61589</td>
<td>0.26898</td>
</tr>
<tr>
<td>C6</td>
<td>729</td>
<td>5</td>
<td>41122248</td>
<td>41281588</td>
<td>532</td>
<td>0.61434</td>
<td>0.2695</td>
</tr>
<tr>
<td>C8B</td>
<td>732</td>
<td>1</td>
<td>57374883</td>
<td>57453213</td>
<td>403</td>
<td>0.6037</td>
<td>0.27302</td>
</tr>
<tr>
<td>CRP</td>
<td>1401</td>
<td>1</td>
<td>159662079</td>
<td>159704396</td>
<td>146</td>
<td>0.55353</td>
<td>0.28995</td>
</tr>
<tr>
<td>EMR1</td>
<td>2015</td>
<td>19</td>
<td>6867560</td>
<td>6960464</td>
<td>517</td>
<td>0.52641</td>
<td>0.2993</td>
</tr>
<tr>
<td>VWF</td>
<td>7450</td>
<td>12</td>
<td>6038040</td>
<td>6253836</td>
<td>812</td>
<td>0.52451</td>
<td>0.29996</td>
</tr>
<tr>
<td>CD55</td>
<td>1604</td>
<td>1</td>
<td>207474817</td>
<td>207554311</td>
<td>191</td>
<td>0.521</td>
<td>0.30118</td>
</tr>
<tr>
<td>C5AR2</td>
<td>27202</td>
<td>19</td>
<td>47815404</td>
<td>47865272</td>
<td>165</td>
<td>0.51465</td>
<td>0.3034</td>
</tr>
<tr>
<td>BDKRB1</td>
<td>623</td>
<td>14</td>
<td>96701641</td>
<td>96755304</td>
<td>217</td>
<td>0.47856</td>
<td>0.31613</td>
</tr>
<tr>
<td>ITGB2</td>
<td>3689</td>
<td>21</td>
<td>46285864</td>
<td>46368753</td>
<td>401</td>
<td>0.42917</td>
<td>0.3339</td>
</tr>
<tr>
<td>IGL</td>
<td>3535</td>
<td>22</td>
<td>22360474</td>
<td>23285085</td>
<td>3273</td>
<td>0.3135</td>
<td>0.37695</td>
</tr>
<tr>
<td>TFPI</td>
<td>7035</td>
<td>2</td>
<td>188308957</td>
<td>188439219</td>
<td>359</td>
<td>0.30263</td>
<td>0.38109</td>
</tr>
<tr>
<td>ITGAM</td>
<td>3684</td>
<td>16</td>
<td>31251288</td>
<td>31364213</td>
<td>236</td>
<td>0.28119</td>
<td>0.38928</td>
</tr>
<tr>
<td>FGA</td>
<td>2243</td>
<td>4</td>
<td>155484280</td>
<td>155531897</td>
<td>103</td>
<td>0.25514</td>
<td>0.39931</td>
</tr>
<tr>
<td>CD93</td>
<td>22918</td>
<td>20</td>
<td>23039993</td>
<td>23086977</td>
<td>168</td>
<td>0.24433</td>
<td>0.40349</td>
</tr>
<tr>
<td>C9</td>
<td>735</td>
<td>5</td>
<td>39264377</td>
<td>39384655</td>
<td>473</td>
<td>0.22934</td>
<td>0.4093</td>
</tr>
<tr>
<td>GZMM</td>
<td>3004</td>
<td>19</td>
<td>524034</td>
<td>569920</td>
<td>207</td>
<td>0.094176</td>
<td>0.46248</td>
</tr>
<tr>
<td>FCN3</td>
<td>8547</td>
<td>1</td>
<td>27675601</td>
<td>27721315</td>
<td>105</td>
<td>0.091124</td>
<td>0.4637</td>
</tr>
<tr>
<td>FGB</td>
<td>2244</td>
<td>4</td>
<td>155464132</td>
<td>155513915</td>
<td>152</td>
<td>0.059684</td>
<td>0.4762</td>
</tr>
<tr>
<td>Gene</td>
<td>Chrm</td>
<td>Start</td>
<td>End</td>
<td>Length</td>
<td>Genomic Start</td>
<td>Genomic End</td>
<td>Fold Change</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>---------------</td>
<td>---------------</td>
<td>--------------</td>
</tr>
<tr>
<td>IGH</td>
<td></td>
<td>3492</td>
<td>14</td>
<td>106012614</td>
<td>107308051</td>
<td>2383</td>
<td>0.043655</td>
</tr>
<tr>
<td>CPB2</td>
<td></td>
<td>1361</td>
<td>13</td>
<td>46607201</td>
<td>46715318</td>
<td>324</td>
<td>-0.00138</td>
</tr>
<tr>
<td>FGG</td>
<td></td>
<td>2266</td>
<td>4</td>
<td>155505286</td>
<td>155553902</td>
<td>98</td>
<td>-0.01297</td>
</tr>
<tr>
<td>C5AR1</td>
<td></td>
<td>728</td>
<td>19</td>
<td>47793104</td>
<td>47845327</td>
<td>175</td>
<td>-0.01506</td>
</tr>
<tr>
<td>F13A1</td>
<td></td>
<td>2162</td>
<td>6</td>
<td>6124311</td>
<td>6340924</td>
<td>843</td>
<td>-0.04834</td>
</tr>
<tr>
<td>COLEC11</td>
<td></td>
<td>78989</td>
<td>2</td>
<td>3622422</td>
<td>3712234</td>
<td>455</td>
<td>-0.106</td>
</tr>
<tr>
<td>SERPINB2</td>
<td></td>
<td>5055</td>
<td>18</td>
<td>61534939</td>
<td>61591124</td>
<td>266</td>
<td>-0.2191</td>
</tr>
<tr>
<td>SERPIND1</td>
<td></td>
<td>3053</td>
<td>22</td>
<td>21108383</td>
<td>21162008</td>
<td>190</td>
<td>-0.27605</td>
</tr>
<tr>
<td>C7</td>
<td></td>
<td>730</td>
<td>5</td>
<td>40889599</td>
<td>41003041</td>
<td>511</td>
<td>-0.35549</td>
</tr>
<tr>
<td>C5</td>
<td></td>
<td>727</td>
<td>9</td>
<td>123694613</td>
<td>123832554</td>
<td>394</td>
<td>-0.36952</td>
</tr>
<tr>
<td>PLG</td>
<td></td>
<td>5340</td>
<td>6</td>
<td>161103225</td>
<td>161195086</td>
<td>346</td>
<td>-0.5789</td>
</tr>
<tr>
<td>KNG1</td>
<td></td>
<td>3827</td>
<td>3</td>
<td>186415098</td>
<td>186482199</td>
<td>381</td>
<td>-0.6265</td>
</tr>
<tr>
<td>MBL2</td>
<td></td>
<td>4153</td>
<td>10</td>
<td>54505140</td>
<td>54552578</td>
<td>230</td>
<td>-0.78353</td>
</tr>
<tr>
<td>F3</td>
<td></td>
<td>2152</td>
<td>1</td>
<td>94974732</td>
<td>95027413</td>
<td>128</td>
<td>-0.88538</td>
</tr>
<tr>
<td>F2RL3</td>
<td></td>
<td>9002</td>
<td>19</td>
<td>16979826</td>
<td>17022830</td>
<td>124</td>
<td>-0.91454</td>
</tr>
<tr>
<td>F11</td>
<td></td>
<td>2160</td>
<td>4</td>
<td>187167099</td>
<td>187230835</td>
<td>229</td>
<td>-0.93805</td>
</tr>
<tr>
<td>MASP1</td>
<td></td>
<td>5648</td>
<td>3</td>
<td>186913873</td>
<td>187029810</td>
<td>445</td>
<td>-0.96466</td>
</tr>
<tr>
<td>CFP</td>
<td></td>
<td>5199</td>
<td>X</td>
<td>47463612</td>
<td>47509704</td>
<td>8</td>
<td>-1.1077</td>
</tr>
<tr>
<td>KLKB1</td>
<td></td>
<td>3818</td>
<td>4</td>
<td>187128502</td>
<td>187199625</td>
<td>256</td>
<td>-1.1082</td>
</tr>
</tbody>
</table>
3.3.2 PGS analysis: Clinical and demographic information

Demographic and clinical information for all participants included in the PGS analysis is outlined in table 2. No association between complement PGS and either age, gender, or years of education were observed, see supplementary table 2. Similarly, based on a principle components analysis described previously by the group (O’Gráda et al., 2009), no association was found between complement PGS and positive or negative symptom factor scores. No association between medication dosage and complement PGS was observed.
Table 3.2: Clinical and demographic information for the whole sample, and subsets of individuals based on diagnostic category.

<table>
<thead>
<tr>
<th></th>
<th>Whole Group</th>
<th>Cases</th>
<th>Healthy Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>1000</td>
<td>808</td>
<td>192</td>
</tr>
<tr>
<td>Age, mean (s.d.)</td>
<td>41.40±12.74</td>
<td>42.88±12.46</td>
<td>35.20±12.04</td>
</tr>
<tr>
<td>Female sex, %</td>
<td>37.7</td>
<td>34.4</td>
<td>51.6</td>
</tr>
<tr>
<td>WAIS FSIQ, mean (s.d.)</td>
<td>99.03±22.21</td>
<td>92.17±19.49</td>
<td>121.20±14.68</td>
</tr>
<tr>
<td>CPZ, mean (s.d.)</td>
<td>—</td>
<td>445.81±467.35</td>
<td>—</td>
</tr>
<tr>
<td>SAPS mean (s.d.)</td>
<td>—</td>
<td>19.81±19.28</td>
<td>—</td>
</tr>
<tr>
<td>SANS mean (s.d.)</td>
<td>—</td>
<td>23.46±19.91</td>
<td>—</td>
</tr>
<tr>
<td>Years in education</td>
<td>—</td>
<td>12.67±2.54</td>
<td>—</td>
</tr>
</tbody>
</table>

Abbreviations: M, mean; SD, standard deviation; FSIQ, Full scale intelligence quotient; CPZ, chlorpromazine equivalents; SAPS, Scale for the Assessment of Positive Symptoms; SANS, Scale for the Assessment of Negative Symptoms.
3.3.3 Association between complement-based polygenic score (PGS) and IQ

As the complement pathway was enriched for IQ-associated genes based on GWAS summary statistics (Savage et al., 2018), polygenic scores were created for the complement gene-set to test for an association with cognition in an independent sample of cases and healthy participants. Based on a regression analysis on the total sample (cases and healthy participants combined), complement PGS significantly predicted variation in measures of Premorbid IQ and verbal IQ (see table 3). The observed direction of effect was that increased complement-based PGS was associated with an increase in IQ in the sample. These findings remained significant when gender was covaried for (premorbid IQ: $F$ change=5.938; Std Beta=0.087, df = 2,772; $p = 0.015$; verbal IQ: $F$ change=4.037, Std Beta=0.067, df=2,903; $p=0.045$). If corrected for 3 domains of cognition analysed, however, premorbid IQ only survives multiple testing correction ($\alpha = 0.05 / 3= 0.016$). When the sample was sub-divided into patients and control groups (considered separately) as a post-hoc analysis, the association between complement PGS and IQ was no longer significant in cases. However, complement PGS was found to influence Full Scale IQ scores in healthy controls only ($F$ change=4.195; Std Beta= 0.147, df = 2,182; $p =0.042$). Notably, a significant difference in complement PGS between patients and healthy participants was observed, such that patients had lower PGS than healthy participants ($F=1.82$, df 1,953; $p=0.018$). No significant associations were observed between complement PGS and memory scores (see table 3).

To determine whether the association between complement PGS and IQ remained significant once variations in the reported C4A structural haplotype were accounted for, we re-ran the analysis by including predicted C4A RNA expression scores as a covariate (score previously used by group (G Donohoe et al., 2018), seen to impact on cognition). After the effects of C4A had been accounted for in this way, complement PGS continued to significantly predict variation in Premorbid IQ ($F$ change=6.057; Std Beta= 0.088, df = 3,771; $p = 0.014$), and Verbal IQ ($F$ change=4.016, Std Beta=0.067, df=3,612; $p =0.045$).
0.067, df=3,902; p=0.045). For healthy controls only, complement PGS continued to predict variation in full scale IQ scores (F change=4.153; Std Beta= 0.147, df = 3,181; p = 0.043).
Table 3: Regression analysis of complement PGS IQ scores in patients and controls, using gender as a covariate for IQ analysis, age and gender as covariates in others.

<table>
<thead>
<tr>
<th></th>
<th>Whole group</th>
<th>Cases Only</th>
<th>Controls only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>F change</td>
<td>R2 change</td>
</tr>
<tr>
<td>Premorbid IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>774</td>
<td>5.938</td>
<td>0.008</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>905</td>
<td>4.037</td>
<td>0.004</td>
</tr>
<tr>
<td>Performance IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>750</td>
<td>1.44</td>
<td>0.002</td>
</tr>
<tr>
<td>Full scale IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>743</td>
<td>2.96</td>
<td>0.004</td>
</tr>
<tr>
<td>Memory PCA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>631</td>
<td>0.928</td>
<td>0.001</td>
</tr>
<tr>
<td>SWM errors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>696</td>
<td>2.656</td>
<td>0.003</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Value</td>
<td>p-value</td>
<td>Beta</td>
</tr>
<tr>
<td>--------------</td>
<td>-------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>LNS</td>
<td>911</td>
<td>3.204</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Abbreviations: IQ, intelligence quotient; Memory PCA-score created in Principle components analysis of memory scores; SWM, spatial working memory; LNS, letter-number sequencing, B= Standardised beta coefficient
3.4 Discussion

The purpose of this study was to characterise the association between genetic variation within the complement system and performance on measures of cognitive function related to disability in SZ. Following earlier studies from our group and others, demonstrating that multiple individual loci within complement, including C4A, are associated with cognition, we sought to determine whether variation in the complement gene-set as a whole was associated with cognitive performance. To do this, we carried out a series of enrichment analyses, and PGS analysis in independent datasets. We found that (1) the complement gene-set was enriched for association with cognitive function (as measured by IQ) but not SZ risk, with multiple individual genes being associated with IQ, of which SERPING1 was the strongest associated. We further found, based on independent samples of patients and controls, that (2) a complement-based polygenic score for IQ predicted variation on multiple IQ measures, even when the effects of C4A was accounted for. Importantly, (3) these cognitive effects did not appear to differ between patient and healthy participant samples, but in post-hoc analysis of healthy controls only the complement PGS still affected full scale IQ.

Collectively, these findings suggest that whereas individual genetic components of the complement pathway (including C4A as analysed in (G Donohoe et al., 2018) and closely related gene CSMD1 (G Donohoe, Walters, et al., 2013) maybe associated with SZ risk, the complement pathway as a whole is associated with neurodevelopmental processes related to cognition. This is evident in the effects of PGS on IQ in the healthy control population for full scale IQ, and in the whole group irrespective of diagnosis. At a gene-set level, this may suggest that the wider complement pathway is more strongly associated with neurodevelopmental processes important to global cognitive development than with illness processes per se. In the context of the present study and previous studies on individual complement genes carried out by our group (G Donohoe et al., 2018; G
Donohoe, Walters, et al., 2013), while the role of an individual gene in neurodevelopment may make it particularly important for illness risk (e.g. C4A and SZ), the pathway’s broader neurodevelopmental role is largely independent of SZ risk.

3.4.1 Enrichment analysis of complement genes

In addition to observing enrichment to the overall complement pathway for IQ, we found that 32 of the available 90 genes listed in the complement pathway showed gene-based association with IQ. 12 of these genes remained significant after correction for multiple testing, with SERPING1 the highest associated. SERPING1, also known as C1 inhibitor, has been suggested to participate in neuronal stem cell proliferation, and play a role in the three activation arms of the complement system (Gorelik, Sapir, Woodruff, & Reiner, 2017). Knockdown of circulating C1 inhibitor induces neurovascular impairment, glial cell activation, neuroinflammation, and behavioral deficits (Farfara et al., 2019). Using a shRNA mediated knock-out of SERPING1 to silence the gene expression, a recent study suggests knock-out of this gene impairs neuronal migration and affects brain development (Gorelik et al., 2017). Preliminary evidence in a study of complement genes in twins with SZ implicated peripheral expression of C5 and SERPING1 to be associated with cortical thinning in the superior frontal region (Allswede et al., 2018); SERPING1 has also been found to be overexpressed in the amygdala of patients with SZ (Chang et al., 2017). In animal studies, C1 inhibitor created by SERPING1 attenuated acute neurobehavioral deficits, ischemic volume, and neurodegeneration (Longhi et al., 2009). Therefore, SERPING1 is a gene pivotal to inflammatory response regulation, neural development, and a key contributor to complement cascade activation.

3.4.2 PGS analysis of complement gene-set

In regression analysis in a sample of 1000 individuals with psychosis and healthy controls, increased PGS (i.e. carrying a higher proportion of IQ-associated alleles) was associated with increased scores across several IQ measures (Pre-morbid IQ as measured by the WTAR, Verbal and Full-scale
IQ from the WAIS) in the whole sample and healthy controls only. However, we did not have a specific hypothesis around this aspect of cognitive performance, and instead interpret the data to indicate a broad association with each of the measures available, depending of the sensitivity of each measure used. The direction of this association is interesting given multiple previous reports citing upregulation of complement genes (theoretically driving increased neuroinflammation and synaptic pruning) to be associated with cognitive decline (Gigante et al., 2011; Heyer et al., 2013; Q. Shi et al., 2015; Song et al., 2014). In enrichment analysis, genes related to the phenotype of IQ (SERPING1, CD46) tended to be regulators of the complement system, thus our PGS may represent regulatory activities of the complement cascade, inhibiting inappropriate activation of complement and inflammatory processes (Farfara et al., 2019; Cardone et al., 2011; Winston et al., 2019). This finding was in the overall sample of healthy participants and patients, showing an association between complement and cognition irrespective of diagnosis. This is consistent with MAGMA analysis, as complement genes were enriched for the phenotype of IQ in a GWAS of healthy participants, but no enrichment of these genes for the phenotype of SZ was observed. When taking cases and healthy controls separately, effects on these measures of IQ were no longer significant, but full scale IQ was found to be affected, explaining 2% of variance, a higher proportion than initial findings in patients.

### 3.4.3 The complement pathway, neurodevelopment, and cognition

The current results are consistent with several studies linking genetic variation within complement to cognitive processes. In human studies of complement genes, CFH- a gene significantly enriched for IQ in MAGMA analysis- has been seen to influence cognitive decline in several different cohorts, including healthy controls, patients with SZ (C. Zhang et al., 2017), patients with AD (Gezen-Ak et al., 2013), and post-operative patients with heart conditions (Gigante et al., 2011; Heyer et al., 2013). In another study of post-operative cognitive decline, the complement gene CD59 (which was also significantly enriched for IQ in our analysis) was found to be upregulated in patients with neurocognitive decline as measured by
assessments of memory, executive function, attention, language, and global cognition (Ramlawi et al., 2007), suggesting that several elements of the complement system influence this cognitive phenotype. Again, this cognitive decline may be the result of upregulated neural inflammation, attributed to various complement components.

In addition to C4, other complement genes have been associated with variation in cognition via biological processes related to synaptic pruning. Chibnik et al. (Chibnik et al., 2011) found an association between variation in the CR1 gene and age related changes in cognition. The CR1 gene is associated with general loss of cognitive function, and with the development of AD, possibly due to alternations in neural circuits via excessive synaptic pruning (Chibnik et al., 2011; Chung et al., 2014). In a sample of healthy controls only, CR1 was also found to contribute to a reduction in entorhinal cortex volume (Bralten et al., 2011). This further supports a theory that complement gene expression plays a role in neural processes throughout development, such as synaptic pruning, which may be particularly disrupted in the development of SZ (Shatz, 2009).

3.4.4 Limitations

The ‘complement’ gene-set which formed the basis of this study was curated by Birnbaum (Birnbaum et al., 2018), based on multiple independent publicly available databases (GO, KEGG, IMPORT, IPA, and IMMUNOME) combined with a complement gene-set curated by Qian et al (Qian et al., 2019). While this list is an empirically well validated, as with any list, some gene inclusions/exclusions will make it imperfect. For example, CSMD1, a regulator of C4 and a multiple-domain regulator of the complement system was not included in this gene-set (Kraus et al., 2006). Genetic variation at a locus corresponding to CSMD1 has previously been associated with both increased SZ risk, cognitive performance (G Donohoe, Walters, et al., 2013), and variation in brain structure and function (Rose et al., 2013).

Another limitation of the gene-set chosen is that some components of the complement pathway are neuroprotective, and others
neurodegenerative, with very diverse functions. The effects of individual complement genes on cognition are likely to be both nuanced and interactive and this may be masked when viewed as a set. For instance, increased C5aR (which is expressed in neural stem cells) is said to increase neurogenesis (Hernandez et al., 2017), whereas increased CR1 is said to eliminate neural connections (Chibnik et al., 2011). In other words, some aspects of the complement system may disrupt neural functioning when expressed at a high level, while others may disrupt outcomes when expressed at a low level. Therefore, having one score for complement components irrespective of direction of affects could cancel out the full extent that complement genes affect cognition. Assuming a linear relationship between genes in the PGS and cognition may be an oversight, as some biological processes could be non-linear, with too little or too much complement activity impairing outcome (Veerhuis et al., 2011). Evidence from proteomic profiling of blood plasma samples in children indicated that a majority of complement proteins were upregulated in those who go on to develop psychotic disorders, although some downregulated complement proteins were also associated with disease (English et al., 2017). Further investigation will be required to understand the direction of the association observed here- for example, by subdividing variants into those associated with up-regulation or down-regulation of complement expression using proteomic analysis. More studies establishing the effect each individual gene has on cognition could shed light on the functional and phenotypic effects of up-regulation or down-regulation, and therefore future gene-sets could be weighted based on effect, or reversed scores could be used. A final potential limitation is that the complement PGS was created using weightings from a GWAS of IQ, a phenotype whose genetic architecture is likely to differ from that of the memory phenotype used in our cognitive analysis. Using other, more refined GWAS information to create PGS for specific cognitive measures/domains will be important to shed further light on the specific role of complement in cognition.
3.4.5 Future directions

Given that an immune response is principally a response to environmental factors (infection, stress), the inclusion of environmental data relevant to immune function would enhance analysis of immune relevant genetic variation. The multi-factorial polygenic threshold model of SZ posits that a large number of genetic risk factors with distinct, aggregate, small effects and environmental factors that can interact with established genetic risk factors (Nimgaonkar et al., 2017). Similarly, we and others have shown that environmental processes associated with immune function have a significant impact on cognitive development (Rokita et al., 2018). Future studies of SZ risk, cognitive function, and genetic variation related to immune function may benefit from the inclusion of information about exposure to relevant environmental factors (maternal infection, early social adversity) likely to mediate the relationship between these variables. Finally, structural and functional brain imaging studies of immune genetics related to SZ may also help shed light on how these immune and complement genetics affect neural structure and function.

3.4.6 Conclusion

This study contributes to the growing literature on the role of complement in brain function, SZ pathology and cognition. An association between a previously curated complement gene-set and cognitive function was supported based on both enrichment analysis of GWAS data and polygenic score analysis. This was in the absence of a similar association between complement and SZ risk, despite some of the individual genes within this gene-set previously showing genome wide association with SZ. In terms of the broader genetic architecture of SZ, these findings suggest that while SZ risk variants are likely to be found at neurodevelopmentally relevant loci (genes involved in synaptogenesis). When complement gene-sets are taken as a whole, their relevance is to neurodevelopment, not illness.
3.5 Acknowledgements & Funding

This publication is the work of the authors and Jessica Holland and Gary Donohoe will serve as guarantors for the contents of this paper. This research was funded through an Irish Research Council PhD scholarship grant to JH, and European Research Council (grant no.677467) and Science Foundation Ireland (grant no.16/ERCS/3787) funding to GD.
### Supplementary Tables

#### Supplementary table 3.1. List of 108 selected genes in the complement-related gene-set

<table>
<thead>
<tr>
<th>Publication</th>
<th>Dataset</th>
<th>Name of pathway(^a)</th>
<th>Selected genes(^b)</th>
<th>Number of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genecards</td>
<td></td>
<td>Complement Pathway</td>
<td>C1QA, C1QB, C1QBP, C1QC, C1R, C1S, C2, C3, C5, C5AR1, C6, C7, C8A, C8B, C8G, C9,</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CD46, CD55, CD59, CFB, CFD, CFH, CR1, CR2, MASP1, MASP2, MBL2, SERPING1</td>
<td></td>
</tr>
<tr>
<td>Qian et al., 2019</td>
<td></td>
<td></td>
<td>A2M, BDKRB1, BDKRB2, C3AR1, CFI, CLTC, CLU, CPB2, F10, F11, F12, F13A1, F13B, F2,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F2R, F2RL2, F2RL3, F3, F5, F7, FGA, FGB, FGG, ITGAM, ITGAX, ITGB2, KLKB1, KNG1,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LMAN1, MIR6843, PLAT, PLAUR, PLG, PROC, PROCR, PROS1, SERPINA1, SERPINA5, SERPINB2,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SERPINC1, SERPIND1, SERPINF2, TFPI, THBD, VTN, VWF</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complement and coagulation cascades</td>
<td>A2M, BDKRB1, BDKRB2, C3AR1, CFI, CLTC, CLU, CPB2, F10, F11, F12, F13A1, F13B, F2,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F2R, F2RL2, F2RL3, F3, F5, F7, FGA, FGB, FGG, ITGAM, ITGAX, ITGB2, KLKB1, KNG1,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LMAN1, MIR6843, PLAT, PLAUR, PLG, PROC, PROCR, PROS1, SERPINA1, SERPINA5, SERPINB2,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SERPINC1, SERPIND1, SERPINF2, TFPI, THBD, VTN, VWF</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CD93, CFHR3, CRP, FCN2, FCN3, IGH, IGHG1, IGHM, IGHV3-23, IGK, IGKC, IGL</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C5AR2, COLEC10, COLEC11, FCN1, GZMM, IGHG2, IGHG3, IGHG4, IGHV3D-11, IGHV4-1, IGHV5-2,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Immune response Lectin induced complement pathway</td>
<td>CD93, CFHR3, CRP, FCN2, FCN3, IGH, IGHG1, IGHM, IGHV3-23, IGK, IGKC, IGL</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C5AR2, COLEC10, COLEC11, FCN1, GZMM, IGHG2, IGHG3, IGHG4, IGHV3D-11, IGHV4-1, IGHV5-2,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IGLC1, IGLC6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Creation of C4 and C2 activators</td>
<td>C5AR2, COLEC10, COLEC11, FCN1, GZMM, IGHG2, IGHG3, IGHG4, IGHV3D-11, IGHV4-1, IGHV5-2,</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IGLC1, IGLC6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HGNC</td>
<td>Complement system</td>
<td><strong>CFHR1</strong>, <strong>CFHR2</strong>, <strong>CFHR4</strong>, <strong>CFHR5</strong>, <strong>CRIL</strong></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>------------</td>
<td>---------------------------</td>
<td>--------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td><em>Birnbaum et al., 2018</em></td>
<td>From publications</td>
<td>Complement system</td>
<td>EMR1, CFP</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td>108</td>
</tr>
</tbody>
</table>

*a* Genes were selected based on online datasets and literatures;

*b* Duplicated genes had been removed

*c* Genes underlined unavailable in MAGMA analysis
**Supplementary Table 3.2.** Analysis of associations between complement PGS and age, gender or years of education, positive/negative/disorganised symptoms or medication dose using correlational analysis.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Gender</th>
<th>Education Level</th>
<th>SAPS</th>
<th>SANS</th>
<th>CPZ equivalents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson</td>
<td>.000</td>
<td>.015</td>
<td>.070</td>
<td>.046</td>
<td>.031</td>
<td>.025</td>
</tr>
<tr>
<td>P-value</td>
<td>.999</td>
<td>.643</td>
<td>.132</td>
<td>.249</td>
<td>.439</td>
<td>.541</td>
</tr>
<tr>
<td>N</td>
<td>954</td>
<td>955</td>
<td>462</td>
<td>627</td>
<td>623</td>
<td>592</td>
</tr>
</tbody>
</table>

SAPS, Scale for the Assessment of Positive Symptoms; SANS, Scale for the Assessment of Negative Symptoms; CPZ, chlorpromazine equivalents.
4. Study 3. Effects Of Complement Gene-set Polygenic Risk Score on Brain Volume and Cortical Measures in Patients with Psychotic Disorders and Healthy Controls

Authors: Jessica F. Holland¹, Donna Cosgrove¹, Laura Whitton¹, Denise Harold²,³, Aiden Corvin², Michael Gill², David Mothershill¹, Derek W. Morris¹, Gary Donohoe¹.

1. Cognitive Genetics & Cognitive Therapy Group, The Centre for Neuroimaging, Cognition and Genomics (NICOG), School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Ireland. Neuropsychiatric Genetics Research Group
2. Department of Psychiatry, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland
3. School of Biotechnology, Dublin City University, Dublin, Ireland
Abstract

Multiple genome-wide association studies of SZ have reported associations between genetic variants within the MHC region and disease risk, an association that has been partially accounted for by alleles of the complement component 4 (C4) gene. Following on previous findings of association between both C4 and other complement-related variants and memory function, we tested the hypothesis that polygenic scores calculated based on identified SZ risk alleles within the ‘complement’ system would be broadly associated with memory function and associated brain structure. We tested this using a polygenic risk score (PRS) calculated for complement genes, but excluding C4 variants. Higher complement-based PRS scores were observed to be associated with lower memory scores for the sample as a whole (N=620, F change= 8.25; p = 0.004). A significant association between higher PRS and lower hippocampal volume was also observed (R2 change= 0.016, Beta =−0.13, p= 0.015). However, after correcting for further testing of association with the more general indices of cortical thickness, surface area or total brain volume, none of which were associated with complement, the association with hippocampal volume became non-significant. A post-hoc analysis of hippocampal subfields suggested an association between complement PRS and several hippocampal subfields, findings that appeared to be particularly driven by the patient sample. In conclusion, our study yielded suggestive evidence of association between complement-based SZ PRS and variation in memory function and hippocampal volume.
4.1 Introduction

Evidence from Genome-Wide Association Studies (GWAS) of schizophrenia (SZ) have repeatedly pointed to an association with the major histocompatibility complex (MHC), contributing to the hypothesis that genes relevant to immune function may influence illness risk (Bergen et al., 2012; Moss & Khanna 1999; McAllister, 2014; Pardiñas et al., 2018; Walters et al., 2013). Located on chromosome 6, the MHC is a complex region that includes over 200 genes in high linkage disequilibrium (Moss & Khanna, 1999). In a seminal study of the MHC region and SZ, Sekar et al. found that part of this signal could be explained by structural variation mapping to complement component 4 (C4) gene expression (Sekar et al., 2016). In that study, the authors found that structural variation at C4A was associated with increased risk for SZ using genetic data from the Psychiatric Genomics Consortium (Ripke et al., 2014) and in an animal model, increased RNA expression predicted by C4 structural variation influenced development of the neural system via a pattern of altered synaptic pruning (Sekar et al., 2016).

Following this study, we recently reported that higher predicted C4 expression was associated with poorer performance on measures of memory function (Donohoe et al., 2018). These findings mirrored several studies that have reported association between other genetic variants functionally involved in the complement system and variation in memory-related outcomes and brain structure (Allswede et al., 2018; Athanasiu et al., 2017; Bralten et al., 2011; Chung et al., 2014; Donohoe et al., 2013; Nettiksimmons et al., 2016; Rose et al., 2013; Zhang et al., 2017). In studies from our group, for example, SNPs tagging variation at or within the MHC region have been associated with variation in episodic memory and hippocampal volume in a sample of SZ patients and healthy adults (Walters et al., 2013). Furthermore, a SZ risk variant within CSMD1, encoding a regulator of complement, was also associated with poorer episodic memory performance (Donohoe et al., 2013), and in large independent samples of
patients and controls (n=1783), expression of complement factor H in the hippocampus was found to be associated with both increased SZ risk and poorer memory function (Zhang et al., 2017). At a gene-set level, we’ve recently shown that a complement polygenic score calculated using summary data from a GWAS of IQ (Savage et al., 2018) was associated with variation in general cognitive function (Holland et al., 2019).

Furthermore, in a proteomic analysis of the complement signalling pathway using longitudinal population-based data, individuals who went on to experience psychotic-like experience and psychotic disorders were found to have upregulation of multiple complement proteins in childhood (English et al., 2017, Föcking et al., 2019).

Given the above evidence that variation within genes encoding for complement associated with SZ could affect memory and brain structure, we hypothesised that (1) A ‘complement’ gene-set PRS created using SZ GWAS summary statistics (Pardiñas et al., 2018) explained variation in behavioural measures of memory function in a sample of patients with psychotic disorders and healthy controls, using a gene-set curated by previous studies (Birnbaum et al., 2018; Qian et al., 2019), excluding variants from C4 in our PRS calculation. Depending on whether this hypothesis was supported, we further sought to test whether (2) The same ‘complement’ PRS based on SZ GWAS summary statistics (excluding C4) would explain variation in memory-related brain structures (i.e. the hippocampus). Finally, because a study of the relationship between C4A and brain structural measures was not undertaken in our previous study of C4A expression and memory (Donohoe et al., 2018), we further tested the hypothesis that (3) C4A predicted gene expression would be associated with variation in the same memory-focused and general metrics of brain structure as hypothesised to be affected by ‘complement’ PRS.
4.2 Methods

4.2.1 Sample description

In total, 479 cases and 141 healthy participants completed a full neuropsychological assessment battery of memory function and had full genome-wide SNP data available on the basis of which complement PRS levels could be calculated (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Patients were diagnosed by trained psychiatrists using the Structured Clinical Interview for DSM-IV Axis I Diagnosis (First, 2005). Cases consisted of n=359 clinically stable patients with a diagnosis of SZ and schizoaffective disorder (SZA), and an additional n=120 patients diagnosed with bipolar disorder with psychotic features, major depressive disorder with psychotic features, delusional disorder, or psychosis not otherwise specified, as described elsewhere (Cosgrove et al., 2017, Donohoe et al., 2018, Whitton et al., 2016). These patients were recruited from five sites across Ireland. Inclusion criteria required participants to be clinically stable at the time of cognitive assessment, aged between 18–65 years, no history of co-morbid psychiatric disorder, no substance abuse in the preceding 6 months, no prior head injury with loss of consciousness, no history of seizures and with Irish ancestry (all four grandparents born in Ireland). Symptom severity was measured using the SAPS and SANS scores as previously described by us (Donohoe et al., 2008).

Healthy participants were recruited from the general population through local media advertisements. All were aged between 18 and 65 years and had Irish-born paternal and maternal grandparents, and satisfied, on the basis of clinical interview, the criteria of having no history of major mental health problems, intellectual disability or acquired brain injury, and no substance abuse in the preceding 6 months. Exclusion criteria also included having a first-degree relative with a history of psychosis. All assessments were conducted in accordance with the relevant ethics committees’ approval from each participating site, and all participants provided written informed consent. In this study, healthy participants did not represent a control group as no direct phenotypic comparison are made with patients; instead healthy
participants are included both to establish whether comparable effects of predicted PRS levels were observed in both groups and, in a subset of these samples, to test for cortical effects using MRI. 620 individuals were included in cognitive analysis, with complete data on diagnosis, age, sex, PRS, and memory scores.

4.2.2 Cognitive assessment of memory

Episodic memory was assessed in the sample using the logical memory subtests (immediate and delayed conditions) from the Wechsler Memory Scale, third edition (WMS-III) (Wechsler, 1997), and the paired associations learning task (PAL; stages completed and total errors) from the Cambridge Automated Neuropsychological Test Battery (CANTAB; (Robbins et al., 1994)). As in our previous study of C4, an unrotated principle components analysis based on the four available episodic memory tests was carried out to reduce multiple testing burden. This memory factor explained 72% of variance in memory scores, as described previously (Donohoe et al., 2018).

4.2.3 Magnetic resonance imaging (MRI)

A subset of our total cognitive sample also underwent structural MRI (35 patients with SZ, 8 with schizoaffective disorder, and 2 with bipolar disorder). Inclusion criteria for healthy participants (n= 171) was being aged 18–65 years, no history of psychiatric illness, drug abuse or head injury. Healthy participant sampling at the Dublin site included individuals involved in the Trinity College Biobank project, described in (Rose et al., 2012). All patients (n= 45) were chronic, but stable and medicated outpatients, with a confirmed diagnosis according to DSM-IV. Demographics presented for this sample in Table 1.

Structural MRI sequences were acquired on a Philips Intera Achieva 3TMR system, with whole-brain imaging consisting of a T1-weight image (180 slices; duration 6 min) using a TFE gradient echo pulse sequence, with a slice thickness of 0.9 mm, and 230 × 230 FOV (Donna Cosgrove et al., 2018).
4.2.4 Processing of imaging data

Images were processed using the most recent version of FreeSurfer (v6.0), a software which creates virtual 3D reconstruction of the human brain, stacking slices of MR images in 3D space. This software automatically detects the grey matter and white matter volumes of all brain structures, by segmenting subcortical structures by affine registration to Talairach space according to the differences in their voxel intensities, as described elsewhere (Akudjedu et al., 2018; Dale et al., 1999; Fischl et al., 2002). Each image underwent motion correction, intensity normalization, transformation to Talairach space and skull stripping (Cosgrove et al., 2018). Mathematical outliers were detected by using R software as per the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) protocol (http://enigma.loni.ucla.edu/protocols/). In addition to this, each output was visually inspected for T1 quality (artefacts, contrast, resolution, intensity) and to ensure adequate reconstruction. Images underwent a thorough visual inspection of each image slice in the coronal, sagittal, and axial planes using Freeview. Manual edits to pial surfaces and white matter (WM) intensity were carried out where necessary by trained researchers (DC, JH). Using the ENIGMA protocol for the analysis of mean cortical thickness and surface area data within FreeSurfer ROIs, values for surface area and thickness were extracted.

After the delineation of all brain structures, the subroutine of hippocampal subfield detection (Iglesias et al., 2015) was performed. This updated version of software includes a tool which generates an automated segmentation of the hippocampal subfields based on a statistical atlas built primarily upon ultra-high resolution (~0.1 mm isotropic) ex vivo MRI data (Iglesias et al., 2015). Enigma hippocampal subfield protocol was followed (Iglesias et al., 2015), and MATLAB was used to plot segmentations directly on each participant's scan and collate snap-shots of these into a webpage for visual inspection (the internal surface QC method) which was carried out by trained researchers (JH, LH). From the subfield routine, we
selected the volumes of subiculum, CA1, CA3, CA4, dentate gyrus (DG), and whole hippocampus for analysis.

4.2.5 Genotyping

Genotyping was conducted on DNA extracted from whole blood or saliva. Full GWAS data were available for all samples. A proportion of samples were genotyped with an Affymetrix 6.0 chip (Santa Clara, CA, USA; conducted as part of the Wellcome Trust Case Control Consortium 2 [Irish Schizophrenia Genomics Consortium and the Wellcome Trust Case Control Consortium 2, 2012] and the remainder on the Illumina HumanCoreExome chip (San Diego, CA, USA). SNPs were excluded on the basis of MAF \( \leq 1\% \), SNP missingness \( \geq 2\% \), and Hardy–Weinberg equilibrium \( \leq P_{10^{-6}} \). Imputation was carried out on these data sets separately using 1,000 Genomes Phase I integrated haplotypes (Dec2013 release) and IMPUTE2 to give \( \sim 10 \) million SNPs genome-wide per sample.

4.2.6 C4 Expression

Direct genotypes for SNPs in the region of 23–35 Mb on chromosome 6 from the Affymetrix (n = 3657 SNPs) and Illumina (n = 3712) data were used to impute C4 structural alleles and predicted expression (G Donohoe et al., 2018). This analysis of our data was undertaken by a member of the McCarroll group using the same methods described previously by them (Sekar et al., 2016), acknowledged in our original C4 and cognition paper (Donohoe et al., 2018). In brief, this involved imputation of C4 structural alleles in the study populations using a 222 haplotype integrated SNP and C4 reference panel. Imputed structural alleles were used to determine copy number of C4 structural elements (C4A, C4B, C4L and C4S and their co-occurrence) in each individual, and expected expression of C4A and C4B in the brain was inferred based on the previously determined relationship of copy number of C4 structural elements to gene expression in human brain samples. This resulted in a normally distributed range of predicted C4 expression scores of between 0 and 1.87 (mean 1.23, S.D. 0.45).
4.2.7 Complement gene-set

To test a gene-set related to complement function in SZ, we based our gene list on recent publications targeting complement-related immunity genes in SZ and other disorders (Birnbaum et al., 2018, Qian et al., 2019). Both papers compiled a gene-set related to ‘complement’ in SZ by collating information from neuroimmunology and general immunology literature, cross-referenced with multiple gene expression and biology databases (e.g. GO, KEGG, IMPORT, IPA, and IMMUNOME, Molecular Signatures Database, the Human Biological Pathway Unification Database and the HUGO Gene Nomenclature Committee (https://pathcards.genecards.org, http://software.broadinstitute.org/gsea/msigdb/index.jsp, https://www.genenames.org)) to assemble a list of complement-related genes (Dunkelberger & Song, 2010, Orsini et al., 2014, Ricklin et al., 2010, Sarma & Ward, 2011, Veerhuis et al., 2011). After removing the duplicated genes and genes directly encoding for C4 (C4A, C4B, C4BPA, C4BPB, C4_B), and taking into account our previous paper on a complement gene-set (Holland et al., 2019), 90 genes were brought forward for analysis, which have been shown to relate to cognition in our previous work (Holland et al., 2019).

4.2.8 PRS score calculation

Based on the complement set of 90 genes described above, we began by identifying all SNPS within (+/-20 kb) of these gene loci, and then extracting the genotype values for our samples, which had been genotyped using either Affymetrix 6.0 or Illumina HumanCoreExome. PRSice software (https://choishingwan.github.io/PRSice/) (Euesden et al., 2014) was used to perform quality control on data from each sample and SNPs were excluded from further analysis if there was >=2% missing genotype data, HardyWeinberg equilibrium <1 × 10−5, or MAF <1%. Next, the lists of SNPs that passed QC for each sample were joined, providing a SNP list common to each sample (n = 14237 SNPs). Based on this list, an effect-size weighted PRS was generated for each individual using PRSice based on a threshold of p<0.05 for SNPs in the SZ GWAS set, similar to previous
studies in on the same dataset utilizing PRS calculation methods (Cosgrove et al., 2017).

*Memory score analysis:* PRS scores for ‘complement’ were used as the independent variable, while age and gender were entered as covariates of no interest in regression analysis. The memory score created in principle components analysis was entered as the outcome measure. To maximize power to detect differences, we carried out our analysis on the full dataset of cases and controls (n=620, of which the structural MRI sample is a subset) for whom we had information on PRS and memory factor score, as per our previous findings on C4 expression and memory (G Donohoe et al., 2018).

*Imaging analysis:* Age, gender, and intracranial volume were used as covariates in all analyses. Analysis of ‘complement’ PRS and C4A expression effect on total brain volume, surface area, cortical thickness and hippocampal volume was performed in SPSS version 25 (IBM Corp, Armonk, NY, 2017) using regression ‘enter’ method. To test the possibility of a particular phenotype population driving any statistically significant results, analysis was also carried out in healthy controls and patients separately.

### 4.3 Results

Participant demographics, brain volume measures and complement gene-set PRS values are displayed in Table 1. Individuals with psychosis were observed to show lower hippocampal volume ($F(215)= 26.98, p=0.001$) and brain volume ($F(215)= 39.99, p=0.001$) compared to healthy participants, when covariates of age and gender are accounted for in analysis.
Table 4.1: Demographics for whole sample of patients with psychosis and healthy controls

<table>
<thead>
<tr>
<th></th>
<th>Healthy Participants</th>
<th>Patients with psychotic disorders</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>171</td>
<td>45</td>
</tr>
<tr>
<td>Age</td>
<td>28.8 (10.27)</td>
<td>41.73 (9.77)</td>
</tr>
<tr>
<td>Female %</td>
<td>56.7%</td>
<td>28.9%</td>
</tr>
<tr>
<td>Complement SZ PRS (z)</td>
<td>-.06 (1.00)</td>
<td>.22 (.97)</td>
</tr>
<tr>
<td>Brain Volume (mm³)</td>
<td>1189435.18 (126636.80)</td>
<td>1144009.82 (127555.99)</td>
</tr>
<tr>
<td>Hippocampal Volume (mm³)</td>
<td>6936.74 (707.50)</td>
<td>6610.68 (745.18)</td>
</tr>
</tbody>
</table>
4.3.1 Complement gene-set PRS and memory scores

After the effects of age and sex were accounted for (as covariates of no interest), computed ‘complement’ PRS significantly predicted variation in memory performance (N=620, F change = 8.25; p = 0.004), explaining 1.2% of total variation in memory factor scores. In the sample of N=620 individuals with genetic and cognitive data, C4 expression levels were found to correlate with Complement PRS (N=620, r=0.408, p=<0.001). When C4 expression levels were included as a covariate in analysis, an association between complement PRS and memory scores remained (N=620, Beta=-0.079, p=0.05), while C4 expression was observed to no longer significantly predict memory scores (M=620, Beta=-0.070, p=0.08).

4.3.2 Complement PRS and hippocampal volume

In the full sample of patients and controls with MRI data (N=216), an association between complement gene-set PRS and hippocampal volume was observed (R2= 0.016, Beta =−0.13, p=0.015, Table 2). When including diagnosis as a further covariate in the model, the association between complement PRS and hippocampal volume remained significant (R2= 0.012, Beta =−0.11, p=0.027). However, this finding was no longer significant following correction for the analyses of global brain metrics undertaken (total volume, cortical thickness surface area; see below).

4.3.3 Complement PRS and Cortical Thickness, Surface Area and Total Brain Volume

In the full sample of patients and controls with MRI data (N=216), no association of PRS on total brain volume, cortical thickness or surface area was observed (see Table 2), using age, sex and Intracranial Volume (ICV) as covariates.
Table 4.2: Effect of SZ complement PGS (standardized z-score) on hippocampal volume, total brain volume, surface area and cortical thickness (age, gender and ICV as covariates)

<table>
<thead>
<tr>
<th>Measure</th>
<th>B</th>
<th>SE B</th>
<th>Beta</th>
<th>R² change</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hippocampal volume (cm³)</td>
<td>-93.35</td>
<td>38.18</td>
<td>-0.13</td>
<td>0.016</td>
<td>0.015</td>
</tr>
<tr>
<td>Brain volume</td>
<td>-3776.15</td>
<td>5217.60</td>
<td>-0.03</td>
<td>0.001</td>
<td>0.47</td>
</tr>
<tr>
<td>Cortical thickness</td>
<td>&lt;0.001</td>
<td>0.012</td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>0.99</td>
</tr>
<tr>
<td>Surface area</td>
<td>-338.59</td>
<td>778.98</td>
<td>-0.019</td>
<td>&lt;0.001</td>
<td>0.66</td>
</tr>
</tbody>
</table>

ICV= Intracranial Volume.
4.3.4 Post-hoc Analyses: sub-regions of the hippocampus

Analyses of sub-regions of the hippocampus were carried out in post-hoc analyses, following a nominally significant finding between complement PRS and hippocampal volume. From the automatic segmentation of hippocampal subfields values for right and left CA1, CA3, CA4, subiculum and DG were extracted. Widespread volume loss in hippocampal subfields was observed in the whole group, and when divided into cases and controls the effect was observed to be driven by the sample of patients with psychosis. Patients with psychosis had lower volume in subiculum, Right CA1, Right DG, and Right CA4 associated with increased complement PRS, while controls exhibited a difference in Left subiculum only (see table 3). Following correction for multiple testing (0.05/ 10 subfield tests= 0.005) only the left subiculum was observed to be associated with complement PRS, in the whole sample. When diagnosis was included as a covariate in this analysis, this finding remained significant (N=216, R2change=0.025, Beta=-0.16, p=0.0045), surviving multiple testing correction. When including C4A expression as a covariate in analysis of the left subiculum along with covariates of age, sex and ICV, this association remained significant (N=88, R2change=0.057, Beta=-0.29, p=0.009), although this was in a subset of individuals with C4 expression scores and MRI data.
Table 4.3: Hippocampal Subfield analysis for complement PRS score and hippocampal subregion volume.

<table>
<thead>
<tr>
<th>Subfield</th>
<th>N</th>
<th>F change</th>
<th>B change</th>
<th>R2 change</th>
<th>p</th>
<th>N</th>
<th>F change</th>
<th>B change</th>
<th>R2 change</th>
<th>p</th>
<th>N</th>
<th>F change</th>
<th>B change</th>
<th>R2 change</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hippocampal volume (cm³)</td>
<td>216</td>
<td>5.98</td>
<td>-0.21</td>
<td>0.016</td>
<td>0.02*</td>
<td>45</td>
<td>3.72</td>
<td>-0.39</td>
<td>0.05</td>
<td>0.06</td>
<td>171</td>
<td>2.27</td>
<td>-0.15</td>
<td>0.008</td>
<td>0.13</td>
</tr>
<tr>
<td>R whole Hippocampus</td>
<td>5.16</td>
<td>-0.2</td>
<td>0.014</td>
<td>0.02*</td>
<td></td>
<td>4.91</td>
<td>-0.46</td>
<td>0.06</td>
<td>0.03*</td>
<td></td>
<td>1.529</td>
<td>-0.12</td>
<td>0.005</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>L whole Hippocampus</td>
<td>6.08</td>
<td>-0.21</td>
<td>0.017</td>
<td>0.01*</td>
<td></td>
<td>2.47</td>
<td>-0.31</td>
<td>0.04</td>
<td>0.12</td>
<td></td>
<td>2.847</td>
<td>-0.17</td>
<td>0.01</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>R Subiculum</td>
<td>7.70</td>
<td>-0.02</td>
<td>0.022</td>
<td>0.006*</td>
<td></td>
<td>4.36</td>
<td>-0.42</td>
<td>0.06</td>
<td>0.04*</td>
<td></td>
<td>3.681</td>
<td>-0.19</td>
<td>0.013</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>L Subiculum</td>
<td>9.29</td>
<td>-0.24</td>
<td>0.029</td>
<td>0.003**</td>
<td></td>
<td>4.87</td>
<td>-0.40</td>
<td>0.07</td>
<td>0.03*</td>
<td></td>
<td>4.448</td>
<td>-0.20</td>
<td>0.017</td>
<td>0.04*</td>
<td></td>
</tr>
<tr>
<td>R CA1</td>
<td>4.33</td>
<td>-0.17</td>
<td>0.013</td>
<td>0.04*</td>
<td></td>
<td>5.72</td>
<td>-0.46</td>
<td>0.07</td>
<td>0.02*</td>
<td></td>
<td>0.865</td>
<td>-0.08</td>
<td>0.003</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>LCA1</td>
<td>3.41</td>
<td>-0.16</td>
<td>0.01</td>
<td>0.07</td>
<td></td>
<td>0.95</td>
<td>-0.18</td>
<td>0.02</td>
<td>0.33</td>
<td></td>
<td>1.59</td>
<td>-0.13</td>
<td>0.005</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>R GC-ML-DG</td>
<td>6.30</td>
<td>-0.21</td>
<td>0.019</td>
<td>0.01*</td>
<td></td>
<td>4.55</td>
<td>-0.39</td>
<td>0.07</td>
<td>0.04*</td>
<td></td>
<td>2.198</td>
<td>-0.14</td>
<td>0.008</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>L GC-ML-DG</td>
<td>4.41</td>
<td>-0.18</td>
<td>0.013</td>
<td>0.04*</td>
<td></td>
<td>3.60</td>
<td>-0.35</td>
<td>0.05</td>
<td>0.07</td>
<td></td>
<td>1.125</td>
<td>-0.11</td>
<td>0.004</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>R CA3</td>
<td>3.36</td>
<td>-0.14</td>
<td>0.012</td>
<td>0.07</td>
<td></td>
<td>3.23</td>
<td>-0.32</td>
<td>0.06</td>
<td>0.08</td>
<td></td>
<td>0.846</td>
<td>-0.08</td>
<td>0.004</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>LCA3</td>
<td>0.16</td>
<td>-0.03</td>
<td>0.001</td>
<td>0.69</td>
<td></td>
<td>0.54</td>
<td>-0.14</td>
<td>0.01</td>
<td>0.46</td>
<td></td>
<td>0.062</td>
<td>0.02</td>
<td>&lt;0.001</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>RCA4</td>
<td>6.10</td>
<td>-0.20</td>
<td>0.019</td>
<td>0.01*</td>
<td></td>
<td>4.27</td>
<td>-0.39</td>
<td>0.06</td>
<td>0.045*</td>
<td></td>
<td>2.217</td>
<td>-0.14</td>
<td>0.009</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>L CA4</td>
<td>4.19</td>
<td>-0.17</td>
<td>0.013</td>
<td>0.04*</td>
<td></td>
<td>3.30</td>
<td>-0.34</td>
<td>0.05</td>
<td>0.08</td>
<td></td>
<td>1.14</td>
<td>-0.10</td>
<td>0.004</td>
<td>0.29</td>
<td></td>
</tr>
</tbody>
</table>

*p significant at p=0.05; **significant at p=0.0035, corrected for 14 tests. Age, gender and ICV as covariates in analysis B=standardised Beta coefficients
4.3.5  C4A and structural metrics

In the available sample of patients and controls with MRI data and C4 expression scores (N=88), no association was observed between predicted C4A expression levels and hippocampal volume, total brain volume, total cortical thickness or total surface area; using age, sex and ICV as covariates (see Table 4).
Table 4: Effect of C4 Expression on hippocampal volume, total brain volume, surface area and cortical thickness (age, gender and ICV as covariates), in 88 individuals with C4 expression data in volumetric analysis

<table>
<thead>
<tr>
<th>Measure</th>
<th>B</th>
<th>SE B</th>
<th>Beta</th>
<th>R² change</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hippocampal volume (cm³)</td>
<td>-27.730</td>
<td>92.856</td>
<td>-0.019</td>
<td>&lt;0.001</td>
<td>0.766</td>
</tr>
<tr>
<td>Brain volume</td>
<td>947.341</td>
<td>12568.548</td>
<td>0.004</td>
<td>&lt;0.001</td>
<td>0.940</td>
</tr>
<tr>
<td>Cortical thickness</td>
<td>-.039</td>
<td>0.027</td>
<td>-0.087</td>
<td>0.007</td>
<td>0.153</td>
</tr>
<tr>
<td>Surface area</td>
<td>1205.037</td>
<td>1793.710</td>
<td>0.034</td>
<td>0.001</td>
<td>0.503</td>
</tr>
</tbody>
</table>
4.4 Discussion

It has become increasingly clear that immune molecules, in addition to their roles in classical immunological pathways, contribute to brain development and cognitive function (Shatz, 2009). With many findings indicating a link between genetic variation related to SZ risk, complement function and memory (Athanasiu et al., 2017; G Donohoe et al., 2018; G Donohoe, Walters, et al., 2013; Walters et al., 2013), this study aimed to address a gap in literature which investigates whether complement function as a whole (using a PRS created with SZ GWAS data) can better predict memory and brain volume, rather than studying individual genes associated with complement function, or indeed C4A expression levels alone. This study sought to determine the effect of a ‘complement’ PRS, which omitted the already associated C4 structural variant, on measures of memory, hippocampal volume, cortical thickness and total brain volume. Given the previous findings suggesting higher C4A expression levels impact on memory scores (G Donohoe et al., 2018), and findings suggesting higher complement PRS related to IQ affects general cognitive function using the same sample (Holland et al., under review) we hypothesised that higher complement PRS associated with SZ would be associated with lower scores on measures of memory, and that either higher C4 expression or higher ‘complement’ PRS would be associated with lower hippocampal volume, cortical thickness, surface area, and total brain volume.

We found that the higher complement PRS was associated with reduced memory task performance scores. The strength of association between ‘complement’ PRS and memory scores was comparable to previous findings on C4 expression (G Donohoe et al., 2018) in the same sample using the same principal component derived memory scores. As noted, this association was observed in the absence of the influence of C4 (which was not included in the PRS score), thus minimising the possibility that this gene-set wide finding was driven by structural allelic variation at the C4A locus. This finding supports the idea that the association between memory function and variation within the complement system extends beyond C4A. Similar to previous studies, the effect of complement PRS was observed
across our full sample of both patient and healthy participants and not specific to SZ. Similarly, an association between hippocampal volume and complement PRS was observed at a nominally significant level across the whole sample of 216 individuals, with this association was no longer significant following correction for the number of brain volumes tested. In post-hoc analyses of hippocampal subfields, the volumetric difference in hippocampus seemed to be driven by differences in the patient sample. This indication of a significant association between ‘complement’ gene-set PRS and hippocampal volume may be of interest for further exploration in a patient-only sample.

4.4.1 Strengths

To our knowledge, this is the first study to examine the impact of C4A expression on hippocampal volume, and the first to use a complement PRS in volume analysis of brain areas related to memory. A strength of the study is the comparison of C4A expression versus complement PRS associated with SZ, excluding C4. The initial analysis of C4A and hippocampal volume was non-significant, which may challenge previous findings which link C4A expression to memory impairment and over-representation of these proteins in the hippocampus (G Donohoe et al., 2018; Sekar et al., 2016). Computation of a PRS not including C4 variants allowed for a considered analysis of other genes involved in the complement system. This analysis suggests a role for the wider complement system influencing hippocampal volume, particularly left subiculum volume.

A major strength of this analysis is the use of Freesurfer 6.0 for segmentation of hippocampal subfields. This allowed us to further explore the effects of complement PRS on specific areas within the hippocampus, in which we observed multiple nominal findings for sub-regions of the hippocampus contributing to effect. As subfields in the hippocampus have been shown to be influenced by different gene expression (Strange, Witter, Lein, & Moser, 2014), it is interesting to further analyse subsections of the hippocampus rather than whole hippocampus volume readings related to PRS. This allows for a more specific study of complement genes, and more
accurate targets for future studies and treatments involving complement and SZ.

4.4.2 Limitations

One limitation of this study is that the patient sample is significantly older than the healthy control group. The effect of PRS on hippocampal volume may be more readily detected in the patient group due to natural aging, or indeed because SZ patients may have been exposed to a higher number of life stressors, which is said to impact memory and hippocampal volume (Lupien et al., 1997; McEwen, 1999). Future analysis could be carried out in a better age and gender matched sample, and could record significant stressful life events to use as a covariate in analysis. Furthermore, as the patient sample accounts for less than a quarter of the sample, a potential diagnosis-specific effect may not have been detected due to power issues in MRI analysis; a larger patient sample would be needed for this specific study. When analysing cases and controls separately, the effect seems to be driven by the group of participants with a psychiatric diagnosis, which may be representative of a more pronounced effect due to reduction of hippocampal and other brain volume associated with illness (Steen et al., 2006), although we have controlled for factors such as intracranial volume in this study. As individuals with SZ tend to have higher proportion of complement ‘risk’ alleles related to SZ, this difference in populations could also drive this effect.

4.4.3 Future Directions

As this is the first study to characterise the effect of a complement polygene score on brain structure measures, replication in a different sample is required to further support these findings, preferably in a larger patient population in order to determine whether or not there are specific disease-related consequences of carrying a higher PRS score for the complement pathway. Evidence from proteomic profiling of blood plasma samples in children indicated that a majority of complement proteins were upregulated in those who go on to develop psychotic disorders, there are also complement proteins that are downregulated associated with disease
Further analysis will be required to understand the direction of the association; for example by subdividing variants into those associated with up-regulation or down-regulation of complement expression using proteomic analysis. In addition, a potential limitation in many imaging studies is that different types of software are used to reconstruct T1 images that can estimate brain volumes differently based on the algorithms used, as well as different QC methodologies. Therefore, using a defined protocol such as that of hippocampal subfield analysis used in the current study (Iglesias et al., 2015) would be of use to future studies of hippocampal volume.

4.4.4 Conclusion

This study contributes to the growing literature suggesting a role of the complement system in brain function, SZ pathology and cognition. An association between a previously curated complement gene-set and hippocampal volume was supported based on polygenic score analysis, specifically in measures of left subiculum volume. This was in the absence of a similar association between C4 expression and hippocampal volume, despite C4 expression previously showing association with deficits in memory in our wider sample (G Donohoe et al., 2018). Further studies with larger sample sizes that include more SZ patients would help clarify the dissociable effects of this complement PRS on measures of cognitive function compared to effects in controls. Additional investigations into a complement function pathway may provide new lines of evidence to explore in terms of developing a biomarker or therapeutic strategy for SZ.
4.5 Acknowledgements & Funding

This publication is the work of the authors and Jessica Holland and Gary Donohoe will serve as guarantors for the contents of this paper. This research was funded through an Irish Research Council PhD scholarship grant to JH, and European Research Council (grant no.677467) and Science Foundation Ireland (grant no.16/ERCS/3787) funding to GD.
### 4.5.1 Supplementary Table Study 3: List of 90 selected genes in the complement-related gene-set

<table>
<thead>
<tr>
<th>Publication</th>
<th>Dataset</th>
<th>Name of pathway</th>
<th>Selected genes</th>
<th>Number of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qian et al., 2019</td>
<td>Genecards</td>
<td>Complement Pathway</td>
<td><em>CIQA, CIQB, CIQBP, C1QC, C2, C3, C5, C5AR1, C6, C7, C8A, C8B, C8G, C9, CD46, CD55, CD59, CFB, CFD, CFH, CR1, CR2, MASP1, MASP2, MBL2, SERPING1 BDKRB1, BDKRB2, CFI, CLTC, CLU, CPB2, F10, F11, F12, F13A1, F13B, F2, F2R, F2RL2, F2RL3, F3, F5, F7, FGA, FGB, FGG, ITGAM, ITGAX, ITGB2, KLKB1, KNG1, LMAN1, PLAT, PLAUR, PLG, PROC, PROCR, PROS1, SERPINA1, SERPINA5, SERPINB2, SERPINC1, SERPIND1, SERPINE1, SERPINF2, TFPI, THBD, VTN, VWF</em></td>
<td>24</td>
</tr>
<tr>
<td>Genecards</td>
<td>Complement and coagulation cascades</td>
<td></td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Genecards</td>
<td>Immune response Lectin induced complement pathway</td>
<td></td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Genecards</td>
<td>Creation of C4 and C2 activators</td>
<td><em>C5AR2, COLEC10, COLEC11, FCN1, GZMM</em></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>HGNC</td>
<td>Complement system</td>
<td><em>CFHR1, CFHR2, CFHR4, CFHR5, CR1</em></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Birnbaum et al., 2018</td>
<td>From publications</td>
<td>Complement system</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>90</strong></td>
<td></td>
</tr>
</tbody>
</table>

*a* Genes were selected based on online datasets and literatures;  
*b* Duplicated genes had been removed
5. Discussion

The hypothesis that SZ may be an immunological disorder is longstanding. As cognitive deficits associated with SZ are among the most disabling of symptoms, this thesis sought to characterise the influence of immune function on cognition, and explore possible genetic and environmental factors which could influence this affect. Thus, using large collections of genetic data and measures of immune function, we examined the immune hypothesis of SZ related to cognition. Investigating the effect of immune-related variants and immune markers on neuropsychological performance may point towards biological mechanisms behind deficit, potentially leading to a better understanding and treatment of SZ. This chapter will provide a summary of three studies in this thesis which demonstrate findings adding to this field, along with strengths, limitations and suggestions for future work.

5.1 Summary of Main Findings

5.1.1 Study 1

Using data from a large epidemiological and longitudinal cohort from the UK, we found that altered immune function (measured by immune markers IL6 and CRP) explained a proportion of variation in IQ scores in the general population. However, contrary to what was hypothesized, early life adversity (ELA) did not predict alteration in immune markers, nor did immune markers influence the relationship between ELA and cognition.

A nominal association between total number of ELAs experienced (poly-victimisation score) and general cognitive ability (IQ) at age 8 was also observed. In a post-hoc analysis, we observed a significant association between the specific ELA ‘Harsh Parenting’ before age 5 (measured in terms of maternal physical discipline) and IQ at age 8 years. This association survived correction, and remained significant after accounting for other relevant covariates. Finally, although elevated levels of inflammatory markers IL-6 and CRP were significantly associated with lower IQ scores, neither inflammatory marker was observed to account for the association between ELA and IQ. While some hypotheses surrounding
the influence of early environment on cognition were supported, immune function did not appear to be the mechanism through which this occurred. This negative finding is of importance to the study of immune function in SZ, as it contradicts previous findings suggesting that ELA predicts inflammatory response alterations (Slopen et al., 2010, Danese et al., 2007).

5.1.2 Study 2

In a sample of Irish participants with genetic and cognitive data (independent from the cohort investigated in Study 1), we explored the impact of a ‘complement’ immune-related polygenic score (PGS) on cognitive outcomes shown to be relevant to immune function. The purpose of this study was to characterise the association between genetic variation within the complement system and performance on measures of cognitive function related to disability in SZ. To do this, we carried out a series of enrichment analyses, and PGS analysis in independent datasets. We found that (1) the complement gene-set was enriched for association with cognitive function (as measured by IQ) but not SZ risk, with multiple individual genes being enriched for an IQ phenotype. We further found, based on independent samples of patients and controls, that (2) a complement-based polygenic score for IQ predicted variation on multiple IQ measures, even when the effects of C4A was accounted for. Importantly, (3) these cognitive effects were comparable between patient and healthy participant samples. Collectively, these findings suggest that whereas individual genetic components of the complement pathway - including C4A as analysed in (Donohoe et al., 2018) and the closely related gene CSMD1 (Donohoe et al., 2013) - maybe associated with SZ risk, the complement pathway as a whole was associated with more general neurodevelopmental processes related to cognition.

5.1.3 Study 3

From the context of recent studies of complement genes and cognition in SZ (Donohoe et al., 2018; Donohoe et al., 2013, Zhang et al., Athanasiu et al., 2017), a question remained about the possible relationship between SZ risk genes associated with complement and cognitive function. Complement
genes have historically been related to synapse formation (Veerhuis, 2011), and more recently synaptic pruning in SZ (Sellegren, 2019; Sekar, 2016), thus brain development could be differentially impacted by complement expression. Given the previous findings suggesting higher C4A expression levels impact on memory scores (Donohoe et al., 2018), and findings suggesting higher complement PGS was associated with IQ affects general cognitive function using the same sample (Holland et al., 2019; study 2) the third study of this thesis examined the differential impact of C4 expression and a polygenic risk score (PRS) for ‘complement’ (using SZ GWAS data) on neurocognitive outcomes. To our knowledge, this was the first comprehensive investigation of the effect of a ‘complement’ pathway polygenetic risk score related to SZ risk on brain volume following seminal studies on the involvement of complement component 4 in SZ (Sekar et al., 2016). We hypothesised that higher complement PRS associated with SZ would be associated with lower scores on measures of memory, and that either higher C4 expression or higher ‘complement’ PRS would be associated with lower hippocampal volume, cortical thickness, surface area, and/or total brain volume.

Firstly, we found that the higher complement-based SZ PRS was associated with lower memory scores, comparable to the influence of C4 on memory scores previously examined in the same sample (Donohoe et al., 2018). As noted, this association was observed in the absence of the influence of C4 (which was not included in the PRS score), thus minimising the possibility that this gene-set wide finding was driven by structural allelic variation at the C4A locus. Similar to previous studies, the effect of complement PRS was observed across our full sample of both patient and healthy participants; not specific to SZ.

Next, we wished to test for a differential relationship between the ‘complement’ PRS and C4 expression levels on brain volume. Using the latest software, we tested an association between C4 expression brain volume measures in Irish participants. C4 was not found to be associated with any brain volume measure included. However, an increased SZ-related
‘complement PRS’ (not C4 expression) was associated with a decrease in hippocampal volume in the subset of participants for whom volumetric data was available (N=216, n=45 patients with SZ). This association was no longer significant following correction for the number of brain volumes tested. Finally, we observed no association between complement PRS and measures of cortical thickness, surface area, or total brain volume.

In post-hoc analyses of hippocampal subfields, the volumetric difference in hippocampus seemed to be driven by differences in the patient sample. This indication of a significant association between ‘complement’ gene-set PRS and hippocampal volume may be of interest for further exploration in a patient sample only. As a whole, the role of the complement system aside from C4 may be relevant to the study of cognitive development as well as SZ pathology.

5.1.4 Candidate’s Role in Studies

The first study was carried out using data from the Avon Longitudinal Study of Parents and Children (ALSPAC) which follows children from mother’s pregnancy, through childhood and adolescence into young adulthood. Data is currently available on the trajectories of young people from birth to age 21 and there is extensive detail on life events, parent and offspring’s mental health, cognitive, immune and environmental variables including early life adversity (ELA) (Boyd et al., 2013). Access to the ALSPAC dataset was granted to the candidate by University of Bristol ALSPAC committee.

Secondly, this thesis includes collaborative studies on psychosis involving multiple authors for studies 2 and 3, and this data was collected over many years by several members of the group. GWAS summary statistics were also used for enrichment analysis and creation of polygene scores, available via online summary statistics databases (eg. PGC consortium). This section will clarify the PhD candidate’s contributions to each study.

The following tasks were performed by the PhD candidate:
• **Study 1**: Data access process undertaken in collaboration with the University of Bristol. Selection of variables available for extraction from the ALSPAC database. Cleaning of data including identification of outliers, log transformation of immune marker scores, creation of ELA measures, main analysis using SPSS.

• **Study 2**: Identification of the complement function gene variants of interest from previous literature. Extraction of genetic data relevant to complement variants from master file (using PLINK), and creation of a PGS using genetic software (PRSice). Enrichment analysis of the complement gene-set was performed using MAGMA, statistical analysis of effects of PGS on cognitive performance was performed in SPSS version 24.

• **Study 3**: Reconstruction of T1 images using FreeSurfer performed by the candidate, as well as quality control and visual inspection of FreeSurfer images using updated Freesurfer reconstruction (v6.0). Creation of complement PRS related to SZ (again using PLINK and PRSice), statistical analysis of the effects of complement PRS on brain structural measures using SPSS software.

• **All studies**: writing the first draft of the paper and performing subsequent edits based on advice from other authors and reviewers.

5.1.5 **Strengths of this Thesis**

5.1.5.1 **Novelty**

One major strength of this thesis is the use of large, well-powered samples to ask a series of novel questions. In study 1, the ALSPAC sample provided us with a unique opportunity to test for associations between ELA variables and immune functioning on cognition, using data from 3 separate time points. This is of interest, as the majority of previously published studies examining ELA and immune function in SZ use retrospective measurements of ELA and furthermore, very few longitudinal studies have immune marker data available. Adding to this, different types of ELA were examined for specific influence on IQ and immune markers. ELA score calculation was informed by previous research in the ALSPAC sample (Fisher, 2012), and
use of these individual measures provides further insight into specific
effects of each ELA.

In study 2, enrichment analysis was carried out whereby the complement
gene set could be tested for enrichment in IQ GWAS and SZ GWAS data in
MAGMA, a tool for gene-set enrichment analysis. This research led to
finding an association between complement genes and cognitive function
using IQ GWAS summary statistics.

Using different GWAS summary statistics to calculate gene scores in
studies 2 and 3 helped to highlight differential functioning of complement
genes related to cognition, with the IQ complement PGS improving
cognitive outcomes, and SZ complement PRS decreasing memory scores.
The approach adopted here was to look at the additive effects of a group of
SZ-associated and IQ-associated variants by generating a PRS for
‘complement’ functioning, a method employed to investigate SZ-associated
variants in previous studies (Hargreaves et al., 2013; Hubbard et al., 2015;
Nicodemus et al., 2014; Riglin et al., 2017), but which has not previously
been used to create a ‘complement’ PGS gene-set analyses using large
GWAS datasets.

5.1.5.2 Up-to-Date analysis and reliable outcome data

In study 2 and study 3, the latest GWAS for IQ and SZ data was used to
create polygenic scores for complement, and rich data on phenotypic
measures of cognition were available to test for associations in an Irish
sample. The neurocognitive tests employed in both studies have been widely
used by psychologists and researchers and their results are reliable and
reproducible (Strauss, Sherman, & Spreen, 2006). The neuroimaging
analyses presented in this thesis were performed to build on these observed
cognitive effects. Structural MRI images which measure metrics linked to
SZ or cognition such as volume, surface area and thickness, enabled
exploration of further associations with SZ risk variants. Finally, the effect
this complement PRS on cognitive performance and brain structure had not
previously been investigated, and enrichment analysis of complement genes
was also a novel approach not previously explored.
5.1.5.3 **Sample Characteristics**

Both the ALSPAC sample and our Irish sample of patients and healthy participants are extensively phenotyped on multiple domains of cognitive ability, which provided numerous measures with which to test genetic variant and immune marker associations. Furthermore, the sample sizes in each study provide suitable power to detect effects. The ALSPAC sample provides a vast number of cases with phenotypic, environmental and genetic gathered from different cohort groups over the past 25 years, with the overall sample used in study 1 roughly N=5000. The Irish sample used in studies 2 and 3 consisted of approximately N=1000. This is among the largest single research group datasets of this kind available. To maximize power to detect differences in study 2, we carried out our analysis on the full dataset of all cases and controls (n=1000). We then followed any significant results in the patient only groups (both the broad psychosis group (n=808) and narrow psychosis group (patients with SZ and schizoaffective disorder only (n=585) to confirm the direction of effects in these groups. It was estimated in a previous study by our group suing PGS and cognitive outcomes that an n=988 has 0.88 power to detect a very conservative polygene score effect of r²=0.01, with α=0.05. Study 3 had 216 participants with full MRI data and genetic information.

5.1.5.4 **Comparison of C4 and Complement PRS**

To our knowledge, the third study is the first of it’s kind to examine the impact of C4A expression on hippocampal volume, and the first to use a complement PRS in volume analysis of brain areas related to memory. A strength of the study is the comparison of C4A expression versus complement PRS associated with SZ, excluding C4. The initial analysis of C4A and hippocampal volume was non-significant, which may challenge previous findings which link C4A expression to memory impairment and over-representation of these proteins in the hippocampus (Donohoe et al., 2018; Sekar et al., 2016). Computation of a PRS not including C4 variants allowed for a considered analysis of other genes involved in the complement
This analysis suggests a role for the wider complement system influencing hippocampal volume.

5.1.5.5 Neuroimaging data

Finally, neuroimaging analysis was performed using Freesurfer 6.0 in study 3, which provided segmentation of hippocampal subfields. This allowed us to further explore the effects of complement PRS on specific areas within the hippocampus, in which we observed multiple nominal findings for sub-regions of the hippocampus contributing to effect. Volumes of hippocampal subfields were generated according to an advanced probabilistic atlas (Zhu et al., 2017). This new programmed algorithm provided by FreeSurfer 6.0 was based on a computational atlas built upon a combination of ex vivo MRI data (manual delineation of the hippocampal substructures from 15 subjects using ultra-high resolution scanner) and in vivo MRI data (manual annotation of the adjacent extra hippocampal structures from a separate dataset of 39 subjects) (Zhu et al., 2017). FreeSurfer 6.0 has been utilized in recent studies to detect (a) an early reduction of CA1 volume in SZ patients (Ho et al., 2016); (b) smaller volumes of the whole hippocampus, CA1, CA4, DG, and the molecular layer of the hippocampus in elderly subjects with subjective memory complaints; and (c) a positive correlation between the volume of the dentate gyrus and memory performance in elderly subjects (Cantero, Iglesias, Van Leemput, & Atienza, 2016). Recent research provided strong evidence for the reliability of these hippocampal subfields' indices estimated using FreeSurfer 6.0 (Whelan et al., 2016). As subfields in the hippocampus have been shown to be influenced by different gene expression (Strange et al., 2014), it is interesting to further analyse subsections of the hippocampus rather than whole hippocampus volume readings related to PRS. This allows for a more specific study of complement genes, and more accurate targets for future studies and treatments involving complement and SZ.
5.1.6 Limitations

5.1.6.1 Missing Data
ALSPAC follows a general population sample of over 14,000 families with a child born in 1990–92. Data is currently available on the trajectories of young people from birth to age 21 and there is extensive detail on life events, parent and offspring’s mental health, cognition and immune markers for this sample (Boyd et al., 2013). Albeit extensive, the ALSPAC sample is not representative of the UK population as a whole in some aspects, because of the non-random nature of the missing data (Fraser et al., 2013). The most disadvantaged families are underrepresented in ALSPAC (Boyd, 2013), and this is due to a high degree of non-response to questionnaires which appear to be ‘non-random’, meaning respondents engage in some questionnaires but not others (Stevens, 2018). Those who completed all questionnaires in the study were more likely to be white, be an owner-occupier of their home, and were more likely to be married parents. Furthermore, mothers who completed all questionnaires were likely to have higher educational attainment scores than the national average, and those who were lost to follow-up had lower attainment on average (Stevens, 2017). ALSPAC attrition, then, is systematic and not random, being more common in lower social classes (Wolke et al., 2009), with a direct relationship between socioeconomic status and the number of questionnaires returned (Boyd et al., 2013).

5.1.6.2 Validity of ELA measures
A possible limitation of the ALSPAC dataset is the lack of clear-cut measures of ELA. While we studied here the presence of ELAs in parental relationships, other ALSPAC-based studies have used a cumulative score of childhood adverse events not specific to parent-child relationship (Slopen, Kubzansky, McLaughlin & Koenen, 2013; Fisher et al., 2012), with no set ‘ELA’ variable created in ALSPAC. A limitation the ELA variable used in this thesis is a lack of corroborative evidence from witnesses of ELA, such as evidence from a partner, from police reports or the child themselves. Furthermore, there were no reports of abuse recorded from a father’s
perspective of domestic abuse carried out by mothers, thus under-representing the incidence of ELA in the population.

5.1.6.3 Participants and Diagnosis

It is possible that patients who participated in studies 2 and 3 are not entirely representative of the general population of people with SZ in terms of cognitive performance and functioning, and may represent a more cognitively intact cohort. Participants who were 1) unwilling and unable to complete tests, 2) those admitted as inpatients at the time of recruitment, and 3) those who were unable to remain in a MRI scanner due to high anxiety did not provide data, thus the potentially lower scores of neuropsychological performances and MRI measures from poorer functioning patients with SZ were not captured.

A limitation of the third study is that the patient sample is significantly older than the healthy control group. The effect of PRS on hippocampal volume may be more readily detected in the patient group due to natural aging, or indeed because SZ patients may have been exposed to a larger number of life stressors, which is said to impact memory and hippocampal volume (Lupien et al., 1997; McEwen, 1999). Furthermore, as the patient sample accounts for less than a quarter of the sample, a potential diagnosis-specific effect may not have been detected due to power issues in MRI analysis.

5.1.6.4 GWAS limitations

A general limitation to the use of GWAS data is the difficulty around fine mapping and functional annotation of GWAS variants. Identifying casual variants in GWAS is not straightforward because of linkage disequilibrium (LD) across the genome. Large blocks of LD results in several SNPs being equally associated with the GWAS phenotype. This risk signal can also extend across multiple genes at a given locus, making SNP to gene annotation challenging. This also means potential confounding factors in gene-set enrichment analyses.

GWAS data from different ancestral populations would also help solve the fine-mapping problem. However, a second general limitation to GWAS is
the lack of data based on samples with ancestries other than European. However, GWAS of SZ and cognitive traits in different populations are now becoming available. A recent GWAS of SZ based on a population of Chinese ancestry reported on a large overlap (75%) of risk SNPs found in European samples (Li et al., 2017b).

Population stratification occurs when allele frequency differs due to systematic ancestry differences. This can cause spurious associations in disease studies (Price et al., 2006). Given the high heritability of C4 expression, future investigations should include a measure of genetic variation in the population, such as a metric of principal components based on ancestral variation, as a covariate in the statistical model: this not only minimises spurious associations, but maximises power to detect true associations. The method of constructing the PRS in this thesis was based on previous studies that did not include PCA for population stratification assessment (Hargreaves et al., 2013; Nicodemus et al., 2014).

5.1.6.5 Variability/Reliability of complement gene sets

The effects of individual complement genes on cognition are likely to be both nuanced and interactive, and this may be masked when viewed as a set. Some components of the complement pathway are neuroprotective, and others neurodegenerative, with very diverse functions. In other words, some aspects of the complement system may disrupt neural functioning when expressed at a high level, while others may disrupt outcomes when expressed at a low level. Evidence from proteomic profiling of blood plasma samples in children indicated that a majority of complement proteins were upregulated in those who go on to develop psychotic disorders, there are also complement proteins that are downregulated associated with disease (English et al., 2017). Therefore, having one score for complement components irrespective of direction of affects could cancel out the full extent that complement genes affect cognition. Assuming a linear relationship between genes in the PGS and cognition may be an oversight, as some biological processes could be non-linear, with too little or too much complement activity impairing outcome (Veerhuis et al., 2011).
Availability of Further Immune information

A limitation of Study 1 is the unavailability of other markers of immune response. It is possible that a measure of TNFa may have been more relevant to ELA analysis based on results of a meta-analysis of ELA and immune function (Baumeister, Akhtar, Ciufolini, Pariante, & Mondelli, 2016). Furthermore, no family and child medical history was used in this analysis, nor were medications used as covariates. From previous studies of the ALSPAC data it is clear that various atopic disorders can impact on inflammatory markers (Khandaker, Zammit, Lewis, & Jones 2014), as can various medications and conditions (Khandaker et al., 2015; Van Mierlo, Schot, Boks & de Witte, 2019). Although beyond the scope of the current study, further data on depression scores or evidence of psychotic disorders at later time-points could have been included in analysis, especially as these have already been cited as impacting cognition and immune response in ALSPAC (Khandaker, Pearson, Zammit, Lewis & Jones, 2014; Khandaker et al., 2017).

Finally, for studies 2 and 3, further information on past or present immune disorders, use of anti-inflammatory medication, and record of birth complications were not available. This information is paramount to the study of immunity and immune genetics (Pouget, 2017; Khandaker et al., 2015; Torrey and Yolken, 2019), as these variables could potentiate observed effects between immune genes and cognition or brain volume. It is possible that many gene X environment interactions have gone unnoticed in past studies of SZ, due to an oversight of the importance of immune reactions and previous birth complications or infections in genetic studies of SZ (Yolken and Torrey, 2019).
5.1.7 Future Directions for Research

5.1.7.1 Gene x Environment interaction: Analysis of complement genes and environment

While study 3 suggests components of the complement system contribute to cognitive decline in SZ, this finding also builds on previous studies which identified the complement system as an important source of gene-environmental interactions affecting brain development (Nimgaonkar et al., 2017). Given that an immune response is principally a response to environmental factors (infection, stress), the inclusion of environmental data relevant to immune function would enhance analysis of immune relevant genetic variation. The multi-factorial polygenic threshold model of SZ posits that a large number of genetic risk factors with distinct, aggregate, small effects exist, and that environmental factors can interact with established genetic risk factors (Nimgaonkar et al., 2017).

We have shown that environmental processes associated with immune function have a significant impact on cognitive development (Rokita et al., 2018). Future studies of SZ risk, cognitive function, and genetic variation related to immune function may benefit from the inclusion of information about exposure to relevant environmental factors (maternal infection, early social adversity) likely to interact with the relationship between these variables. Stepniak et al. (2014) studied the effect of environmental risk on age of SZ onset by creating an environmental risk score for each participant from information about perinatal brain insults, cannabis use, neurotrauma, psychotrauma, urbanicity, and migration. Including this type of metric as a interaction term in this analysis could identify additional or indeed multiplicative genetic and environmental risk factors leading to cognitive decline in SZ. Furthermore, inclusion of more information on prior or present infection, such as toxoplasma gondii exposure (Torrey et al., 2012) can further elucidate the contribution of genes to SZ as opposed to infections with periphery processes that affect brain function (Torrey & Yolken, 2019).
Future studies could investigate the utility of combining genetic risk scores and environmental risk scores to better predict those patients with the poorest cognitive function and who would most benefit from early intervention. In future, using Genome-wise environmental interaction studies (GWEIS) could shed light on how environmental risk factors may interact with genetic risk factors related to the immune system in SZ.

5.1.7.2 Effects of Genes throughout Development

Studies have identified a new role for complement 4 (C4) in synaptic pruning (Melle, 2019; Sekar et al., 2016; Sellgren et al., 2019), but further work into the mechanism of complement activity needed. Synaptic pruning peaks during adolescence, and is essential for refinement of the CNS and maturation of cognitive abilities (Insel, 2010). Structurally different variants of C4 genes are associated with differences in C4 expression and with the risk of SZ, supporting the notion that elevated complement activity could lead to increased synaptic pruning is a risk factor (Sekar et al., 2016). The stability of transcriptional regulation and expression of genes can alter across cell types and throughout development, therefore The neurodevelopmental stage during which the variants investigated here may contribute to cognitive deficits associated with SZ is unclear. The effect of the SZ-associated variants may have already had substantial influence in embryonic neurodevelopmental stages, influencing downstream developmental processes and manifesting as the adult phenotype; Establishing when (developmentally) and where these variants exert their effects on a gene is important (Bhat et al., 2012; Bray & Hill, 2016). Future studies of complement activity and SZ could be delineated based on the time point at which they are measured; for instance the developmental stage, and indeed the stage of illness progression need to be established in analysis.

5.1.7.3 Complement genes and synaptic pruning in patients with SZ

A recent study using patient-derived induced pluripotent stem cells found abnormalities in microglia-like cells and synaptic structures, in addition to increased synaptic pruning in the neuronal cultures (Sellgren et al., 2019).
Risk-associated variants of the C4 genes were linked to increased complement uptake these in synapses (Sellgren et al., 2019). In line with this, our group has found indications of poorer memory function linked to increased predicted C4A expression, across patients with SZ and healthy controls (Donohoe et al., 2018). Future studies could incorporate microglial activity, measured in tandem with complement activity to account for possible interaction effects leading to increased synaptic pruning (Sellgren et al., 2019).

5.1.7.4 Differential effects of complement genes

Further analysis will be required to understand the direction of the association between complement and cognition; for example by subdividing variants into those associated with up-regulation or down-regulation of complement expression using proteomic analysis. Further studies similar to that of Sellegren (2019) could further characterise complement activity by assessing the contribution of complement variants aside from C4 in synaptic pruning. More studies establishing the effect each individual gene has on cognition could shed light on the functional and phenotypic effects of up-regulation or down-regulation, and therefore future gene-sets could be weighted based on effect.

5.1.8 Clinical application of results

Throughout this thesis, a particular focus was given to the immune hypothesis of SZ, a longstanding theory proposing that disturbances in the immune system contribute to the development of the disease in at least a proportion of patients. If demonstrated to hold true, the immune hypothesis could significantly alter current prevention and management strategies. Possible interventions that may be indicated include infection control programs for primary prevention, screening and routine monitoring of the immune profile among high-risk patients or those with active disease, and adjunct immunotherapy in active disease. However, the results of the studies comprising this thesis suggest that, as a group, complement genes do not appear to play a major role in SZ susceptibility, but rather contribute to cognitive functioning in the general population.
5.1.8.1 Genetic variation to predict treatment response

A significant increase in risk SNP discovery via GWAS has occurred over the past decade. This opens up opportunities to detect genetic biomarkers for SZ and improve patient outcome by using personalised therapeutic interventions. Interestingly, one South African study investigated the role of variants in a gene involved in synaptic plasticity (MMP9), with treatment outcome in SZ considering the severity of childhood trauma as an interacting variable. The cohort comprised 103 previously medication-naïve first episode SZ patients treated with a single injectable antipsychotic, flupenthixol decanoate, monitored over 12 months. Relationships between novel and previously described variants, and haplotypes, with antipsychotic treatment response were found to be modified when considering childhood trauma as an interacting variable. This study provides the first evidence for the involvement of polymorphisms and childhood trauma in antipsychotic treatment response, and warrants further investigation into the role gene-environment interactions may play in the betterment of antipsychotic treatment strategies (McGregor et al., 2018). Although treatment response was not measured in this thesis, it would be interesting to evaluate whether the observed immune changes reflect a certain subgroup of patients whom could be stratified for treatment. If this would be the case, stratification of patients for the presence of these immune changes would be a potential application of research, to increase efficacy of cognitive interventions, possibly in combination with measures of ELA.

5.1.8.2 Immune-modulating drugs

One of the important reasons to focus on the role of immune changes in SZ is the fact that this research area may reveal promising targets for new forms of pharmacological therapy. For the treatment of autoimmune disorders, infectious diseases, allergies and cancer, numerous types of immune-modulating agents have been developed (van Mierlo et al., in press). Establishing the translational value of these agents is a priority for the field. However, since most of these drugs can induce severe side effects, further replication and validation of the evidence of their relevance for SZ will be
necessary. Thus far, broadly acting anti-inflammatory drugs, such as aspirin, have been shown to have some positive effect with even more pronounced effects when patients are stratified depending on their levels of CRP prior to adjunct treatment (Kroken et al., 2018). One randomized controlled trial (RCT) with aspirin treatment showed stronger effects when stratifying participants on the basis of a marker of inflammation (Laan et al., 2010). No selection procedure based on elevated baseline inflammation measures was implemented in any trials to date, and the authors of several studies underlined the possibility that enriching the sample with inclusion of only individuals with an elevated CRP could have influenced the results (Kroken et al., 2018, Girgis et al., 2017). Determining which patients respond and which do not in terms of immune activation could represent a subtype of individual for therapy, specific to a particular immune marker or indeed overall immune response. These patients may be more suitable candidates for interventions. However, as indicated in findings of this thesis, the reality is that ‘the immune system is a complicated network or circuit board, made up of many different cellular components that each have multiple functions, and inflammation is a downstream indicator of what is going on across many individual parts of the network’ (Pouget, 2016). Thus, further study of the intricacies of immune reaction in SZ is justified.

5.1.9 Concluding remarks

SZ is a complex, highly polygenic disorder, associated with the disruption of multiple immune processes and environmental risk factors. Although it is a major global health problem, pathophysiology remains largely unknown. Many epidemiological studies have sited a relationship between SZ and immune function, and between immune function and cognition.

The first study in this thesis, based on a longitudinal epidemiological cohort, provides evidence that harsh parenting during the early years of development is associated with lower cognitive performance in childhood by comparison with children who have not had this experience. These findings may have implications for public health interventions aimed at supporting caregivers, particularly in emphasising the importance of
adopting alternative parental disciplining methods. However, our study suggested that altered immune function, as measured by IL6 and CRP, is unlikely to explain a large proportion of the variation in IQ explained by ELA. While other immune markers remain to be examined, our study highlights the need to consider alternative biological and cognitive pathways for explaining the relationship between ELA and cognition.

The second and third studies in this thesis examine the impact of two immune-related ‘complement pathway’ gene-sets on cognitive performance and brain volume. An association between a previously curated complement gene-set PRS using IQ GWAS and cognitive function was supported based on enrichment analysis of and polygenic score analysis. Furthermore, an association between the complement gene-set and brain volume was observed using a SZ-related PRS score. This was in the absence of a similar association between C4 expression and hippocampal volume, despite C4 expression previously showing association with deficits in memory in our wider sample (Donohoe et al., 2018). In terms of the broader genetic architecture of SZ, these findings suggest that while SZ risk variants are likely to be found at neurodevelopmentally relevant loci (eg. genes involved in synaptogenesis), with complement gene-sets possibly relevant to neurodevelopment. Investigating the effect of genetic variants on neuropsychological performance may point towards biological mechanisms behind this deficit, potentially leading to a better understanding and treatment of the disorder.

This thesis is moving beyond the state of the art because: (1) it seeks to refocus our efforts to understand SZ away from clinical symptoms and towards deficits in cognition as a cause of enduring disability in SZ; (2) it combines this focus with a biological focus on immune response rather than the ‘traditional’ biological focus on neurotransmitters (the dopamine model) as a means of understanding genetically predisposed risk; and (3) in doing this it seeks to integrate new knowledge about the genetic basis of SZ risk with current hypotheses about immune and environmental factors.

Incorporating data and methodologies from three different disciplines
(psychiatric genetics, immunology, and social cognitive neuroscience) the core concept of this project is to develop and test an immune based translational model of disability in SZ. Further studies with larger sample sizes that include more SZ patients would help clarify the effects of complement on measures of cognitive function and brain volume. Furthermore, additional investigations into a complement function pathway could provide new lines of evidence to explore in terms of developing a biomarker or therapeutic strategy for SZ. Understanding the relationship between immune function and neurocognition in SZ may aid in future elucidation of the pathophysiology of the disorder, eventually leading to better treatments and patient outcomes.
References


Barnes, M., Freudenberg, J., Thompson, S., Aronow, B., & Pavlidis, P. (2005). Experimental comparison and cross-validation of the
Affymetrix and Illumina gene expression analysis platforms. *Nucleic acids research*, 33(18), 5914-5923.


neuropsychological measures in schizophrenia and nonpsychiatric populations: a systematic review and meta-analysis. *Schizophrenia bulletin*, 43(4), 788-800.


virus and risk of psychosis among adult offspring. *Biological psychiatry, 63*(8), 809-815.


Consortium, I. S. G., & 2, W. T. C. C. C. (2012). Genome-wide association study implicates HLA-C* 01: 02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biological psychiatry, 72*(8), 620-628.


the inhibition/excitation ratio in the rat temporal cortex via trans-
signaling. *Biological psychiatry, 71*(7), 574-582.

Childhood adversity is associated with adult theory of mind and
social affiliation, but not face processing. *PloS one, 10*(6),
e0129612.

Gezen-Ak, D., Dursun, E., Hanağası, H., Bilgiç, B., Lohman, E., Araz, Ö.
S., et al. (2013). BDNF, TNFα, HSP90, CFH, and IL-10 serum
levels in patients with early or late onset Alzheimer's disease or mild
cognitive impairment. *Journal of Alzheimer's Disease, 37*(1), 185-
195.

Gigante, P. R., Kotchetkov, I. S., Kellner, C. P., Haque, R., Ducruet, A. F.,
component 3 (C3F) and complement factor H (Y402H) increase
the risk of postoperative neurocognitive dysfunction following carotid
endarterectomy. *Journal of Neurology, Neurosurgery & Psychiatry,
82*(3), 247-253.

Gimeno, D., Kivimäki, M., Brunner, E. J., Eloainio, M., De Vogli, R.,
Steptoe, A., et al. (2009). Associations of C-reactive protein and
interleukin-6 with cognitive symptoms of depression: 12-year
follow-up of the Whitehall II study. *Psychological medicine, 39*(3),
413-423.

Girgis, R. R., Ciarleglio, A., Choo, T., Haynes, G., Bathon, J. M., Cremers,
clinical trial of tocilizumab, an interleukin-6 receptor antibody, for
residual symptoms in schizophrenia. *Neuropsychopharmacology, 43*(6),
1317.

Glahn, D. C., Knowles, E. E., McKay, D. R., Sprooten, E., Raventós, H.,
Blangero, J., et al. (2014). Arguments for the sake of endophenotypes: examining common misconceptions about the use
of endophenotypes in psychiatric genetics. *American Journal of
Medical Genetics Part B: Neuropsychiatric Genetics, 165*(2), 122-
130.

inhibitor affects cortical development in a cell autonomous and non-
cell autonomous manner. *Frontiers in cellular neuroscience, 11*,
169.

psychiatry: etymology and strategic intentions. *American Journal of

Severe mental disorders in offspring with 2 psychiatrically ill
parents. *Archives of general psychiatry, 67*(3), 252-257.

*Intelligence, 24*(1), 79-132.

schizophrenia. *Nature Reviews Neuroscience, 16*(10), 620.

deficits and functional outcome in schizophrenia: are we measuring
the “right stuff”? *Schizophrenia bulletin, 26*(1), 119-136.
Green, M. F., Kern, R. S., & Heaton, R. K. (2004). Longitudinal studies of
cognition and functional outcome in schizophrenia: implications for
MATRICS. *Schizophrenia research, 72*(1), 41-51.

Häfner, H., Maurer, K., Löffler, W., Bustamante, S., Van der Heiden, W.,
schizophrenia *Search for the Causes of Schizophrenia* (pp. 43-66):
Springer.

Hajima, S. V., Van Haren, N., Cahn, W., Koolschijn, P. C. M., Hulshoff
Pol, H. E., & Kahn, R. S. (2012). Brain volumes in schizophrenia: a
meta-analysis in over 18 000 subjects. *Schizophrenia bulletin, 39*(5),
1129-1138.

Hardy, A., Emsley, R., Freeman, D., Bebbington, P., Garety, P. A., Kuipers,
E. E., et al. (2016). Psychological mechanisms mediating effects
between trauma and psychotic symptoms: the role of affect
regulation, intrusive trauma memory, beliefs, and depression.
*Schizophrenia bulletin, 42*(suppl_1), S34-S43.

meets neuropsychopharmacology: translational implications of the
impact of inflammation on behavior. *Neuropsychopharmacology, 37*(1),
137.

Hartberg, C. B., Sundet, K., Rimol, L. M., Haukvik, U. K., Lange, E. H.,
Nesvåg, R., et al. (2011). Subcortical brain volumes relate to
neurocognition in schizophrenia and bipolar disorder and healthy
controls. *Progress in Neuro-Psychopharmacology and Biological

Hatzimanolis, A., Bhatnagar, P., Moes, A., Wang, R., Roussos, P., Bitsios,
P., et al. (2015). Common genetic variation and schizophrenia
polygenic risk influence neurocognitive performance in young

Hay, A. D., Heron, J., & Ness, A. (2005). The prevalence of symptoms and
consultations in pre-school children in the Avon Longitudinal Study

inflammatory polarization, favors clearance pathways and

Heyer, E. J., Kellner, C. P., Malone, H. R., Bruce, S. S., Mergeche, J. L.,
Ward, J. T., et al. (2013). Complement polymorphisms and cognitive

Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge,
T. M., et al. (2017). Is an early age at illness onset in schizophrenia
associated with increased genetic susceptibility? Analysis of data
from the Nationwide Danish Twin Register. *EBioMedicine, 18*, 320-
326.


Kahn, R. S., & Keefe, R. S. (2013). Schizophrenia is a cognitive illness: time for a change in focus. *JAMA psychiatry, 70*(10), 1107-1112.


Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., et al. (2012). Childhood adversities increase the risk of


Appendix A: Publications Arising from this Thesis

1. Cosgrove, D; Mothersill, D; Whitton, L; Harold, D; Kelly, S; Holleran, L; **Holland, J**; Anney, R; Richards, A; Mantripragada, K; Owen, M; O'Donovan, M; Gill, M; Corvin, A; Morris, D; Donohoe, G. (2018) "Effects of miR-137 Genetic Risk Score on Brain Volume and Cortical Measures in Patients with Schizophrenia and Controls" American Journal of Medical Genetics Part B: Neuropsychiatric Genetics.


Appendix B: Awards and Bursaries

- ENCP young scientist award and Travel Grant 2017. Bursary for workshop attendance with 100 young scientists. Nice, France (3 days).
- YITP young scientist award and Travel Grant, 2016. Bursary to attend FENS conference with 2 week lab visit to NRU, Neurobiology Research Unit, Copenhagen University Hospital, Denmark (3 Weeks).
- College of Arts, Social Sciences, and Celtic Studies Research Travel Bursary, April 2016. Bursary to attend FreeSurfer software course in Tours, France (4 days).
- Shortlisted for Cambridge Cognition CANTAB Bursary Award, 2016.
- Awarded Irish Research council (IRC) PhD scholarship: 2016-present.
- PSI EGG conference best poster presentation 2015- runner up
Appendix C: Presentations

Oral Presentations

- Annual School of Psychology Research Day, National University of Ireland, Galway, April 2018, May 2017.
- Immune Function in Psychosis (iPsychosis) Meeting, NUI Galway, June 2017.
- PSI EGG student congress, NUI Galway, April 2017.

Poster Presentations

- BNA (British Neuroscience Association) 2019 Festival of Neuroscience, Dublin, April 2019.
- Irish Society of Human Genetics, Dublin, September 2018
- Cognomics Conference, Radboud, NL, September 2017.
- World Congress of Psychiatric Genetics, Orlando FL, October 2017
- Irish Society of Human Genetics, Dublin, September 2017.
- Annual School of Psychology Research Day, May 2017, National University of Ireland, Galway
- ECNP workshop for young scientists and Young scientist award (Bursary for attendance), March 2017.
- FENs forum of European Neuroscience. Poster Presentation and Young scientist award (Bursary for attendance), July 2016.
- 5th Annual NUIG-UL Graduate Research Day, NUIG, April 2015
Appendix D: Study Funding and Acknowledgements

Recruitment and genotyping was supported by Science Foundation Ireland (SFI) (grants 12/IP.1359 and 08/IN.1/B1916) and the Wellcome Trust Case Control Consortium 2 project (grant 085475/B/08/Z) and the Wellcome Trust (grants 072894/Z/03/Z, 090532/Z/09/Z, and 075491/Z/04/B).

We wish to thank all patients and their support staff, and all healthy volunteers for participating in the data collection on which this thesis is based.