Application and integration of bioinformatic strategies towards central and peripheral proteomic profiling for diagnosis and drug discovery in schizophrenia

David Alan Cox

Robinson College

September 2017

This dissertation is submitted for the degree of Doctor of Philosophy

Acknowledgements

First and foremost, I would like to thank my supervisor Professor Sabine Bahn for giving me the opportunity to research this PhD as part of the Cambridge Centre for Neuropsychiatric Research. I am grateful to Professor Bahn for funding my PhD for four years through her Stanley Medical Research Institute grant, and for instigating a range of interesting projects.

In addition, I am deeply grateful to Dr Jordan Ramsey, Dr Michael Gottschalk, and Dr Jakub Tomasik for providing guidance on the statistical aspects of this PhD, and the many discussions over the past four years which have shaped the various projects in this thesis. I would also like to thank Dr Andrew Parnell and Dr Belinda Hernandez of University College Dublin for their advice and insights on applying Bayesian methodologies to proteomic datasets.

I would like to thank my many experimental collaborators, past and present, at the Cambridge Centre for Neuropsychiatric Research, who generated the datasets analyzed as part of this PhD. They include Dr Hendrik Wesseling, Dr Santiago Lago, Geertje van Rees, and Dr Sureyya Ozcan.

The work presented here would not have been possible without our many collaborators at institutions around the world, including Nico JM van Beveren, F
Markus Leweke, Matthias Rothermundt, Johann Steiner, Marie-Odile Krebs, Paula Suarez-Pinilla, Benedicto Crespo-Facorro, David Niebuhr, David Cowan, Thalia van der Doef, Lot de Witte, and Hemmo Drexhage.

I would like to thank my friends and housemates at Robinson College for the many enjoyable times I've experienced during my time in Cambridge.

Finally I would like to express my deep gratitude to my parents for all the support they've given me and the patience they've shown during the process of writing this thesis and the various publications resulting from this work.

Declaration and Statement of Length

This thesis is my own work, except where declared in the Acknowledgements or elsewhere in the text. This thesis has not been submitted for any degree, diploma, or other qualification at any university or other institution, be that in entirety or in part. It does not exceed the prescribed 65,000 word limit and does not contain more than 150 figures, as set by the Degree Committee of Engineering.

Cambridge, United Kingdom, September 2017

Publications arising from this work

The following publications arise directly from the work presented in this thesis:

D. Cox, M.K. Chan, S. Bahn, The potential of immune biomarkers to advance personalized medicine approaches for schizophrenia. J. Nerv. Ment. Dis. 203(5):393-9 (2015).

D. A. Cox, M.G. Gottschalk, H. Wesseling, A. Ernst, J.D. Cooper, S. Bahn. Proteomic systems evaluation of the molecular validity of preclinical psychosis models compared to schizophrenia brain pathology. Schizophr Res. 177(1-3):98-107 (2016).

D. A. Cox, M.G. Gottschalk, V. Stelzhammer, H. Wesseling, J.D. Cooper, S. Bahn, Evaluation of molecular brain changes associated with environmental stress in rodent models compared to human major depressive disorder: A proteomic systems approach. World J. Biol. Psychiatry. Nov 25:1-12 (2016).

D. A. Cox, J.M. Ramsey, M. Rothermundt, J. Steiner, F.M. Leweke, N.J. van Beveren, D.W. Niebuhr, N.S. Weber, D.N. Cowan, P. Suarez-Pinilla, B. Crespo-Facorro, C. Mam-Lam-Fook, J. Bourgin, R.J. Wenstrup, R.R. Kaldate, M.O. Krebs, S. Bahn. Comparison of statistical techniques for the classification and prediction of schizophrenia diagnosis through molecular profiling of serum. (Submitted - Under Review).

D. A. Cox, S. Ozcan, M.O. Krebs, J.D. Cooper, J.M. Ramsey, N. Rustogi, F.M. Leweke, N.J. van Beveren, H. Drexhage, D.W. Niebuhr, N.S. Weber, D.N. Cowan, S. Bahn. Targeted molecular profiling of serum for the prediction of schizophrenia before disease onset. (Submitted – Under Review)

G. F. van Rees*, S.G. Lago*, **D. A. Cox*,** J. Tomasik, N. Rustogi, K. Weigelt, S. Ozcan, J. Cooper, H. Drexhage, F.M. Leweke, S. Bahn. Evidence of microglial activation following exposure to serum from first-onset drug naive schizophrenia patients. (Submitted - Under Review)
***authors contributed equally**

G. F. van Rees*, **D. A. Cox*,** S.G. Lago, T.F. van der Doef, J. Tomasik, N. Rustogi, S. Ozcan, J. Cooper, L.D. de Witte, S. Bahn. Microglia activity in schizophrenia: linking PET data to serum analytes and intracellular microglia signalling cascades. (Manuscript in preparation)
***authors contributed equally**

The following publications are related to the work presented in this thesis:

J.A.C. Broek, Z. Lin, H.M. de Gruiter, H. van 't Spijker, E.D. Haasdijk, **D. Cox**, S. Ozcan, G.W.A. van Cappellen, A.B. Houtsmuller, R. Willemsen, C.I. de Zeeuw, S. Bahn. Synaptic vesicle dynamic changes in a model of fragile X. Mol Autism. 2016 Mar 1;7:17.

S.G. Lago, J. Tomasik, G.F. van Rees, H. Steeb, **D.A. Cox**, N. Rustogi, J.M. Ramsey, T. Petryshen, S.J. Haggarty, N.J. van Beveren, S. Bahn. A novel pipeline for drug discovery in neuropsychiatric disorders using high-content single-cell screening of signaling network responses ex vivo. (Submitted - Under Review)

Abstract

Proteomic profiling studies of schizophrenia have the potential to shed further light on this debilitating and poorly understood condition which affects up to 1% of the world's population. However, recent studies suggest that the field of proteomics in general has been hindered by poor application of bioinformatic strategies, contributing to the failure of many findings to validate. In the context of schizophrenia research, there is therefore a need for a more robust application and integration of existing statistical approaches to proteomic datasets, as well as the development of new methodologies to offer solutions to current challenges.

The aims of this thesis were multi-fold. Many studies have stipulated the need for new diagnostic and prognostic strategies to aid psychiatrists, particularly in predicting disease conversion from the prodromal phase. Proteomic data from serum samples was used to investigate the potential for statistical models based on biomarker panels to offer a new and clinically relevant approach. Models were trained based on either differential protein (chapter 3) or peptide (chapter 4) expression levels between schizophrenia patients and controls, as measured through multiplex immunoassay or targeted mass spectrometry technologies. In chapter 4, an SVM model based on 21 peptides was identified that is both highly sensitive and specific as a diagnostic and prognostic tool in symptomatic individuals.

Furthermore, in recent years, few preclinical innovations have been made in schizophrenia research in either *in vitro* or *in vivo* studies, resulting in a standstill in the development of treatments. In chapters 5 and 6 of this thesis, proteomic information from a novel cellular model of schizophrenia was analyzed. In chapter 5, cell signalling alterations *in vitro* were identified which may underpin dysfunctional microglial activation in at least a subgroup of patients, thus representing new drug targets in the CNS. Subsequent analysis identified compounds which have the potential to ameliorate the observed changes. Lastly, in chapters 7 and 8, a novel systems biology methodology was developed for the functional comparison of proteomic changes in brain tissue from existing preclinical rodent models of psychiatric disorders to those in human post-mortem samples, providing a new means of overcoming some of the translational hurdles of preclinical psychiatric research.

The application of different bioinformatic strategies to a range of proteomic datasets in this thesis has yielded a number of findings which have enhanced the understanding of schizophrenia pathophysiology and provide a platform for future studies towards the goal of improving outcomes for patients affected by this disorder.

.

Contents

Chapter 2 Methods

List of Figures

Figure 5.1. Characteristic responses at selected microglial signalling epitopes following stimulation with ligands which specifically induce either M1 or M2 phenotypes.

Figure 5.2. Stimulant responses across 62 microglial signalling epitopes for antipsychotive naive schizophrenia patient serum, control serum and positive controls calyculin and staurosporine.

Figure 5.3. Epitopes which responded differently to antipsychotic naive schizophrenia patient serum exposure relative to healthy controls are shown in the mechanistic context of adjacent proteins in Akt/mTORC1 and JAK/STAT signalling pathways.

Figure 5.4. Volcano plot illustrating the relationship between log2 fold change and statistical significance for differential peptide abundance changes between antipsychotic naive schizophrenia patients and controls as measured through multiplex reaction monitoring (MRM) mass spectrometry profiling.

Figure 5.5. Influence plot for the ten epitopes measured in the Akt/mTORC1 pathway.

Figure 5.6. Responses to minocycline and rapamycin in comparison to the vehicle condition, for each of the six microglial signalling epitopes which showed differential responses between patient and control serum exposure.

List of Tables

Abbreviations

Acute phencyclidine	aPCP
Antithrombin-III	ANT3
Apolipoprotein A-IV	APOA4
Apolipoprotein C-III	APOC3
Apolipoprotein H	APOH
Area under the curve	AUC
Brodmann area 10	BA10
Blood-brain barrier	BBB
Brain-derived neurotrophic factor	BDNF
Bayesian information criterion	BIC
Binding potential	BPND
Comprehensive Assessment of At-Risk Mental State	CAARMS
Chronic mild stress	CMS
Central nervous system	CNS
Complement C2	CO2
Complement C4-A	CO4A
Complement component C9	CO9
Chronic phencyclidine	cPCP
C Reactive Protein	CRP
Cerebrospinal fluid	CSF
Disability-adjusted life year	DALY
Department of Defence Serum Repository	DoDSR
Diagnostic and Statistical Manual of Mental Disorders	DSM
Single enzyme-linked immunosorbent assay	ELISA
Events per predictor variable	EPV
Extrapyramidal symptoms	EPS
Ficolin-3	FCN3
False discovery rate	FDR
False positive rate	FPR
Gene ontology	GO
Genome-wide association studies	GWAS
Haptoglobin	HPT
Interleukin	IL

Inter-α-trypsin inhibitor heavy chain H4	ITIH4
Least absolute shrinkage and selection operator	LASSO
Liquid chromatography mass spectrometry in expression mode	$LC\text{-}MS^{E}$
Local Linear Embedding	LLE
Liquid chromatography multiple reaction monitoring mass spectrometry	MRM
Markov Chain Monte Carlo	MCMC
Major depressive disorder	MDD
Major histocompatability complex	MHC
Magnetic resonance imaging	MRI
N-acetyl cysteine	NAC
N-methyl-D-aspartate receptor	NMDAR
Positive and Negative Syndrome Scale	PANSS
Principal component analysis	PCA
Phencyclidine	PCP
Positron-emission tomography	PET
Post-mortem interval	PMI
Prenatal stress	PNS
Protein-protein interaction	PPI
Retinol-binding protein 4	RET4
Recursive feature elimination	RFE
Receiver operating characteristic	ROC
Residual sum of squares	RSS
Social defeat	SD
Support vector machines	SVM
18 kDa translocator protein	TSPO
True positive rate	TPR
Vitronectin	VTNC
World Health Organization	WHO

Chapter 1 Introduction

1.1 The burden and challenges of schizophrenia

Schizophrenia is a major psychiatric disorder that affects up to 1% of the global population and has been listed among the world's top ten causes of disease-related disability(1). Research has shown that the disorder is a known risk factor to reduced life expectancy by 15 to 20 years in the Western world(2–4). Schizophrenia has an annual incidence rate of around 15 in every 100,000 people and the lifetime risk of developing the disease is approximately 0.7%(1). It has no obvious gender bias, affecting men and women with approximate equal frequency(5) although gender discrepancies occur regarding the onset of the disease. In males, disease onset typically occurs between 16-20 years whereas in females, disease onset is slightly later and typically between 20-30 years(5).

The symptoms of schizophrenia range from positive (hallucinations, delusions, and movement impairments) to negative (avolition, anhedonia, and poverty of speech) as summarized in **Table 1.1**(1). In addition, schizophrenia can lead to deterioration in cognitive functions such as working memory and attention which can greatly impair social and occupational functioning(6,7). Mental health studies by the World Health Organization (WHO) have found the disorder to be the 5th and 6th leading cause worldwide of years lost to disability, for men and women respectively(8). When viewed on the disability-adjusted life year (DALY) scale of assessing overall disease burden, schizophrenia is eighth for individuals between 15 and 44 years old(9). In addition studies have noted the considerable burden on family members who care for the sufferers(9). The annual cost of schizophrenia to the US healthcare system has been estimated to be in excess of 60 billion dollars due to life-long disability and the need for ongoing treatment and provision of care required by patients(10).

The field of schizophrenia research has acknowledged considerable limitations in both the diagnosis and treatment of the disorder which remain areas of ongoing research(11). The means of diagnosing of schizophrenia has not changed over the last 100 years since Emil Kraepelin first defined the disease and is still based on evaluation of signs and symptoms in clinical interviews, which means that misdiagnosis is common(12). Studies following patients from first admission for psychosis for a period of ten years or more, found that diagnoses changed in 50% of cases(13). In addition there is a pressing clinical need to develop prognostic tools to help detect the 20-30% of prodromal individuals who go on to develop schizophrenia, at an early stage, from those who do not(14) (discussed further in **1.5**).

Table 1.1. Positive and negative symptoms in schizophrenia(1)

Positive symptoms	Negative & Cognitive symptoms
Hallucinations (eg: visual, auditory, tactile, olfactory)	Affective flattening - reduction in range & intensity of emotional expression
Delusions (eg: paranoid delusions, delusions of reference, somatic delusions)	Alogia - lessening of speech fluency and productivity
Disorganized speech	Avolition - reduction or inability to persist in goal-driven behaviour
Grossly disorganized or catatonic behaviour	Emotional withdrawal
	Stereotyped thinking
	Social isolation
	Poor concentration & memory
	Difficulty integrating thoughts, feelings and behaviour

Development of effective treatments for schizophrenia patients has been hampered by a lack of significant innovations over the past 60 years since first-generation (typical) antipsychotics were introduced in the 1950s(15). The difficulty in developing new medications is thought to be due to both the complexity and the diversity of the underlying pathophysiological mechanisms associated with the schizophrenia disease process, which is still largely unknown. Studies have estimated that as many as 60% of schizophrenia patients do not respond at all to typical antipsychotics, or are only partially responsive(16).

Despite the introduction of second-generation (atypical) antipsychotics in the past four decades, response rates to both first and second-generation antipsychotics are low, especially after multiple treatment attempts, with many patients continuing to experience residual symptoms(17,18). Treatment with atypical antipsychotics is still not effective in up to 40% patients(19). In addition, both typical and atypical antipsychotics are associated with prominent side effects which can lead to serious secondary health conditions. In some countries, the use of typical antipsychotics has been restricted due to their propensity for inducing acute and long-term neurological side effects such as extrapyramidal symptoms (EPS)(20). While atypical antipsychotics were introduced in an attempt to deliver enhanced safety and tolerability by reducing EPS, they have their own safety concerns with side effects such as weight gain, hyperglycemia and dyslipidemia(20,21). The ongoing discovery of novel therapeutic compounds has been slowed down by the development of “me-too” or “follow-on” drugs which mainly duplicate the actions of existing antipsychotics with minor

improvements in efficacy. These developments reduce incentives for pioneering innovation and diminish valuable R&D resources(22,23).

As a result, there is a great need to improve our understanding of the complex and diverse underlying pathophysiological mechanisms associated with the schizophrenia disease process at the molecular level, both from the perspective of improving diagnosis and therapeutic options.

1.2 Proteomic methods in the study of schizophrenia

A variety of different 'omics' strategies have been used for biomarker discovery in psychiatric disorders, including genomics, proteomics, transcriptomics, metabolomics, and lipidomics(24). While this thesis focuses on proteomic profiling, this section provides a brief overview of the other molecular fingerprinting approaches which have been used to study schizophrenia.

Genomic studies have identified various gene variants which have been linked to schizophrenia through genome-wide association studies (GWAS)(25) and in recent years, two comprehensive GWAS identified 13 and 108 schizophrenia-associated risk loci that contribute to disease susceptibility and point to various functional targets such as calcium channel subunits(26,27). Transcriptomics biomarker discovery analyses typically use high-throughput microarray gene chips which can identify mRNA abundances as well as more recent studies using microRNAs(28). To date transcriptomics studies examining gene expression changes are the most common biomarker research undertakings in the literature for schizophrenia(29). Metabolomics analyses in schizophrenia are based around the hypothesis that the disorder disrupts biochemical pathways resulting in a characteristic metabolic signature. Metabolomic profiling of schizophrenia patients known to be at risk of metabolic syndrome, has shed some light on the molecular underpinnings of the metabolic disturbances experienced by at least a subset of schizophrenia patients(29,30). Lipidomics platforms have been used to study metabolic vulnerability in schizophrenia patients through measuring alterations in different lipids classes in blood plasma(31).

Proteomics techniques are being increasingly used to screen for biomarkers of schizophrenia with various studies revealing new information regarding etiology and mechanisms of the disease over the past ten years through quantitative and qualitative identification of protein patterns in a variety of tissue types(24). The commonly used platforms for identifying new protein biomarkers include multiplex immunoassay platforms,

liquid chromatography mass spectrometry in expression mode (LC-MS^E), liquid chromatography multiple reaction monitoring mass spectrometry (MRM), single enzyme-linked immunosorbent assay (ELISA) tests, and flow cytometry (see **Chapter 2** for more details). These platforms have typically been used to screen for protein alterations in post-mortem brain tissue, CSF, plasma, serum and other tissues such as the liver(32–37). Peripheral studies have the advantage of ease of access to study tissue from living patients, enabling the study of disease onset and development. Patterns observed in the periphery can mirror certain aspects of brain function. Post-mortem studies have revealed information on abnormalities in brain regions including the frontal cortex(38), thalamus(39), anterior cingulate cortex(40), corpus callosum(41) and the hippocampus(42) which will be described in more detail in **1.3**. With the well-known role of various environmental factors as potential triggers for schizophrenia(43), animal models of the disease can be created through different environmental and biological manipulations, and proteomic profiling of various tissue types can be used to study the physiological mechanisms behind certain behavioural phenotypes. Proteomics is thought to be a viable tool both for developing new diagnostic strategies, and identifying novel therapeutic drug targets for schizophrenia. These are both areas of research which will be investigated later in this thesis as while proteomics has already provided valuable information, there remains a need for the identification of more robust biomarkers to improve diagnosis and prognosis, and breakthroughs in developing and analyzing reliable preclinical models of the disease to aid the development of more efficacious treatments.

Several review studies have argued that proteomics holds a number of advantages compared to genomics and transcriptomics approaches for identifying diagnostic and therapeutic biomarkers of schizophrenia(24,43–45), reasons summarized in **Table 1.2**. While genomic studies have generated significant data, so far they have been unsuccessful in advancing diagnostic and therapeutic options in psychiatric disorders after two decades of research(44). This may be because genomic studies alone are not sufficient to fully elucidate the pathophysiology of complex mental illnesses, as variations of many genes, typically with subtle effects, are believed to be involved in the etiology of these disorders(46). While determination of gene expression and DNA variations in schizophrenia remains important, many researchers believe the study of proteins is one of the most informative reflections of biological functionality, as these molecules are involved in most physiological processes(45). One of the main limitations of genomic and transcriptomic studies is that they are unable to extrapolate to functional protein expression in healthy or disease states, as proteins undergo multiple modifications from transcription to posttranslation(43,47,48). While proteomic platforms typically detect fewer proteins than expressed genes detected through

transcriptomic approaches, protein expression is thought to present a more precise reflection of the physiological state(49,50). Current hypotheses of schizophrenia postulate that the disorder arises due to heterogenous interactions of genetic and environmental factors, the biochemical basis of which is most strongly reflected by proteomic changes in various pathways in the brain, periphery and other tissues rather than individual gene products(24).

Table 1.2. Summary of recent review studies discussing the advantages of proteomic research compared to genomics and transcriptomics.

Proteomics vs Genomics and Transcriptomics	References
Genomic studies have found involvement of many genes, typically common variants with subtle effects in the disease etiology but they haven't answered the main questions on disease pathophysiology.	Owen et al. (2010)(46) Nascimento et al. (2015)(43) Davalieva et al. (2016)(44)
Difficult to extrapolate from gene to protein expression as proteins undergo modifications from transcription → posttranslation and transcript abundance can't predict protein levels	Baloyianni et al. (2009)(47) Vogel et al. (2012)(48)
Proteomics can show global expression of proteins and is more complex than genomics as it can change from each cell type at any given time or state	Baloyianni et al. (2009)(47)
Proteomic studies detect fewer expressed proteins than expressed genes in transcriptomic studies but protein expression provides a more precise functional profile reflecting the current physiological state	Guest et al. (2014)(50) English et al. (2011)(49)
Proteomics is more likely to unravel the signal transduction pathways and complex interaction networks behind the gene-environment interactions which underpin schizophrenia	Sethi et al. (2016)(24)

1.3 Pathology

The varying hypotheses as to the etiology of schizophrenia have largely been derived from studying the pathology of the disorder and the identification of central and peripheral abnormalities which are briefly summarized in this section. The hypotheses will be discussed in further detail in **1.4**.

1.3.1 Central

Many studies have identified structural and functional brain abnormalities in schizophrenia, many of which are already present in first-onset drug naive patients by the time of diagnosis(51). A meta-analysis of structural magnetic resonance imaging (MRI) studies of regional brain volumes found global structural differences between schizophrenia patients and those without the disorder. A decrease was seen in cerebrum volume, while a marked increase was found in total ventricular volume(52). In addition, a decrease was found in amygdala, hippocampus and parahippocampus volume. The extent to which these abnormalities progress during the course of the disorder has yet to be conclusively determined. Some studies have attempted to quantify the longitudinal aspect of such structural abnormalities and have concluded that they may deteriorate over time and may be worsened by long-term antipsychotic use(53,54).

Further MRI studies comparing schizophrenia patients with controls have found significant reductions in cortical gray matter volume, white matter volume, and increases in sulcal and ventricular cerebrospinal fluid (CSF)(55). As some of these changes were additionally found in the siblings of these patients, who did not have the disorder, the causal factors are thought to be linked to particular gene variants leading to alterations in synaptic, dendritic and axonal organization.

Functional imaging studies have found evidence of abnormalities in neural connectivity in circuits involving the prefrontal cortex, hippocampus and various subcortical regions, which appears to be involved in the pathophysiology of schizophrenia(56). Evidence from functional MRI and PET studies has suggested that the dorsolateral prefrontal circuit and in particular, the dorsolateral prefrontal cortex is compromised in at least a subgroup of schizophrenia patients through decreased metabolism, abnormal cortical connectivity as a consequence of epigenetic factors and alterations in neurotransmission(57). This may be behind the cognitive deterioration experienced by some patients, as described in **1.1**.

In addition, post-mortem brain studies of schizophrenia patients have found alterations in gene and protein expression after molecular profiling analysis using microarray techniques and proteomic technologies. These alterations were found to be related to various functional pathways involving energy metabolism, oxidative stress, inflammatory processes, myelination, cell communication and synaptic transmission(58–61).

1.3.2 Peripheral

Extensive studies have been conducted examining peripheral alterations in schizophrenia patients with the aim of enhancing the quality of therapeutic interventions and detecting patient subgroups through varying pathophysiological causes. A variety of immune/inflammatory, metabolic, neurotrophic and endocrine abnormalities have been detected through proteomic studies comparing the serum of patients with controls(62–65).

Many peripheral alterations are thought to either reflect or act as a causal factor for central disturbances. Altered levels of cytokines and other inflammatory markers have been found in schizophrenia patients in the blood and CSF(62,63) which can cross the blood-brain barrier (BBB) and induce a variety of harmful consequences such as neuronal inflammation, damage, and degeneration, which have been implicated in the disorder(66). The elevated pro-inflammatory status seen in some schizophrenia patients has led to research into new therapeutic strategies exploring the potential for anti-inflammatory medications such as aspirin(67), celecoxib(68) and N-acetyl cysteine (NAC) as add-on treatments to antipsychotics. NAC acts as a strong inhibitor on cytokines such as TNF- α, IL-6 and IL-1 β, which are the main mediators of pro-inflammatory status(69) and is thought to exert beneficial regulatory effects on neurotransmission pathways that are known to be altered in schizophrenia(70).

Peripheral abnormalities have been observed which point to metabolic dysfunction in some patients. Higher insulin levels are frequently observed in first onset schizophrenia patients along with impaired glucose tolerance and elevated insulin resistance(71). In addition lipid based treatments have shown some therapeutic efficacy with omega-3 fatty acids reported as reducing conversion rates to psychosis in individuals identified as being at risk(72).

Serum studies of first-onset schizophrenia patients have suggested that increased concentrations of neuroendocrine hormones such as chromogranin A, pancreatic polypeptide, prolactin, progesterone, and cortisol may be implicated in the disease onset(64). The role of brain-derived neurotrophic factor (BDNF) levels in the periphery and central nervous system (CNS) is an area of ongoing research with some studies reporting reduced levels both in post-mortem brain tissue and in the serum and plasma of patients, both in drug naive and medicated patients(65). BDNF is involved in regulating neuron growth and survival, and synaptic plasticity and so reduced BDNF may play a role in the schizophrenia pathology by contributing to dysfunction in neuronal signalling during neurodevelopment(65).

1.4 Hypotheses

A variety of models have been hypothesized to explain the onset and development of schizophrenia but while many mechanisms have been proposed, as yet there is no exact consensus. Such is the heterogeneity of schizophrenia that there may be multiple disease mechanisms affecting different patient subpopulations, which are not mutually exclusive(73). Some of the main hypotheses for the disease pathophysiology are summarized in this section. These hypotheses are always developing, based on new pathological findings.

1.4.1 Genetic vulnerability

Schizophrenia is thought to have a high heritability component with some estimates placing the liability of genetic factors and gene-environment interactions for developing the disorder at 80-85%(1). The illness commonly runs in families with the risk of developing it increasing considerably if a close family member has been affected. One of the well-known findings among twin studies is that monozygotic twin pairs are far more likely to both develop the disorder, than dizygotic twin pairs. If one monozygotic twin has developed schizophrenia, then the risk for the other twin is 40-50%(1).

To date, no single gene locus has been identified as a major risk factor for the disorder. Instead research so far supports a highly polygenic architecture(74) with a large proportion of the variation in liability to schizophrenia thought to be spread across thousands of common SNPs in various genetic loci, over 100 of which have been discovered in sizeable GWAS studies(26,27,74–76). A 2014 meta-analysis of GWAS studies linked effects in 108 loci to schizophrenia(26). The associations found convergence upon genes expressed in tissue types likely to be relevant to the disease, for example those with important immune functions like B-lymphocyte lineages involved in acquired immunity. In addition, the associations included genes involved in glutamatergic neurotransmission and the dopamine D2 receptor, supporting previously defined hypotheses of schizophrenia(26).

However, while hundreds of genetic loci have been weakly implicated in the pathogenesis of schizophrenia, there is still much that remains unknown about the genetic architecture of the disorder(77). In addition, some studies have questioned the effectiveness of GWAS research in schizophrenia and other psychiatric disorders due to their heterogeneity. They argue that because each GWAS is conducted on a subset of a diverse overall patient population, and the loci identified in each study are very diverse, there are concerns that as more and more patients are studied, loci will ultimately be implicated across the entire genome, thus rendering the findings increasingly uninformative(77). One of the reasons for this is the

genetic heterogeneity of schizophrenia across ethnic populations. This was illustrated by a 2017 study of 128 schizophrenia associated loci across five ethic populations of patients which found significant heterogeneity between ethnic groups in the majority of these loci(78).

1.4.2 Immune system dysfunction

Evidence from neurobiological, genetic and environmental studies suggests that alterations in immune function may be associated with schizophrenia pathogenesis in at least a subgroup of patients(62). Schizophrenia patients and their relatives are often found to have a higher or lower than expected prevalence of auto-immune disorders such as multiple sclerosis, autoimmune encephalitis, rheumatoid arthritis and celiac disease(79).

Genetics has provided considerable evidence that immune system dysfunction plays a key role in the pathophysiology of schizophrenia(80). A 2013 study identified 144 genes relating to inflammation and immune response in the hippocampus of schizophrenia patients, which were differentially expressed compared with controls(81). In addition one of the most replicated genetic findings is the association between schizophrenia and the immune response regulating major histocompatability complex (MHC) region which is thought to play a role in the pathogenesis of multiple sclerosis. The genetic-inflammatory-vascular hypothesis of schizophrenia suggests that a genetic predisposition to an over-expressed inflammatory response could damage the microvascular system in the brain, with such immune disturbances thereby triggering metabolic dysfunction(82).

Analyses of blood levels in first onset and relapsed patients have found elevated concentrations of numerous inflammatory cytokines in serum and plasma(62,63,83). One hypothesis is that increased psychological stress on susceptible individuals can trigger an exaggerated inflammatory response via pathways involving defective glucocorticoid-mediated feedback inhibition and exaggerated sympathetic nervous system mediated activation of immune responses(62). The peripheral levels of C Reactive Protein (CRP), a protein mainly generated in the liver, are also a good indicator of chronic inflammation. Recent data have shown that it is increased in around 30% of schizophrenia patients(84,85).

Studies have pointed to the potential harmful role of activated inflammatory processes in schizophrenia on the CNS(86) which will be discussed further in **1.6.1**. The release of cytokines and free radicals through activation of the brain's immune cells, the microglia, is central to the microglia hypothesis of schizophrenia which will be discussed in **1.6.2**.

1.4.3 Infectious disease

Exposure to pathogenic microbes, during either the prenatal period or childhood, has been suggested as being a contributing factor to the etiology of schizophrenia by causing disruptions in brain development(87). This is a potential component of the two-hit hypothesis of schizophrenia, discussed further in **1.6.3**. Studies investigating the role of specific infectious agents in schizophrenia have suggested that the risk is mainly posed by viruses and parasites with a tendency to invade the CNS.

A number of viral traces have been identified in first episode drug naive schizophrenia patients and over the past decade, studies applying bioassays to measure anti-microbial antibodies in archived maternal serum samples have yielded some notable results. They suggest that infections posing a particular risk factor for schizophrenia include influenza A virus, polio virus, and rubella virus(88–90). In particular, influenza A virus has been suggested to increase the risk of developing schizophrenia by sevenfold when mothers are exposed to the virus during the first trimester of pregnancy(88).

In addition, there is evidence that viral CNS infections during childhood increase the risk of developing schizophrenia as an adult by at least two-fold(91,92) and schizophrenia patients who have suffered a viral CNS infection earlier in life, tend to have an earlier disease onset(92). Post-mortem brain tissue studies have also yielded evidence of viral traces, with one study finding retroviral RNA to be present in the frontal cortices of all patient samples(93).

Considerable research has been undertaken on the parasite Toxoplasma (T.) gondii and its potential contribution to the etiopathogenesis of schizophrenia(94,95). As of 2013, thirty-eight studies have shown positive correlations between T. Gondii antibody titers and schizophrenia(96), and links have been observed between exposure in adulthood and schizophrenia in studies of the US military(97) and ultra-high risk individuals(98).

However although associations have been made between schizophrenia and infections years to decades before diagnosis, investigations into how infectious agents contribute to the pathomechanisms of the disorder, after a lengthy latency period, are still ongoing.

1.4.4 Metabolic dysfunction

A major contributing factor to this increased risk of mortality from schizophrenia is believed to be the greater susceptibility of patients to metabolic syndrome, cardiovascular disease and diabetes compared to the general population(31,99–101). For example, a large UK

study of 46,000 patients with either schizophrenia-spectrum or bipolar disorder and 300,000 controls reported that hazard ratios for cardiovascular disease mortality were threefold for schizophrenia patients aged 18-49 years and twofold for patients aged 50–75 years(102).

Metabolic dysfunctions include insulin resistance, hyperinsulinemia, dyslipidemia, hyperglycemia, central obesity and hypertension(103). While metabolic dysfunctions can be side effects of antipsychotics, studies showing a high prevalence of metabolic syndrome in drug naïve patients imply that metabolic dysfunction is associated with the schizophrenia disease process(104–106) .

The exact role which metabolic dysfunction plays in the disease process of schizophrenia is still little understood but proteomic serum findings suggest that it represents a vulnerability factor to developing the disorder. Increased levels of molecules such as insulin, C-peptide and proinsulin have been found in both patients(107,108) and unaffected first degree relatives(107), suggesting there is a shared genetic and environmental disposition in families to conditions such as hyperinsulinemia.

1.4.5 Neurodevelopmental

The neurodevelopmental hypothesis of schizophrenia posits that brain development in the early stages of gestation is altered, and the onset of psychosis occurs at a later stage as the disease progresses, due to abnormal cortical development(109). The hypothesis was initially formed following epidemiological studies of an influenza epidemic in 1957. It was found that children whose mothers were pregnant during the epidemic, had an 88% higher likelihood of later developing schizophrenia compared to those born in the periods before and after the epidemic(109,110). The hypothesis is based around the fact that the typical age-range for the onset of schizophrenia is between 18 and 36, a period where the prefrontal cortex is still developing, indicating involvement of cortex development in the disorder(111).

The neurodevelopmental hypothesis has been further supported by associations to schizophrenia with genetic and early risk factors that affect development, along with imaging studies which have identified structural brain abnormalities in patients which were already present before disease onset(112,113). Similar alterations have been found to be present to a lesser degree in prodromal individuals(114).

The neurodevelopmental hypothesis is linked to the two-hit hypothesis which will be discussed further in **1.6.3.**

1.4.6 Neurotransmitter dysfunction

Multiple neurotransmitter theories have been proposed for schizophrenia, which have formed the basis of all antipsychotics developed so far. Each of these theories implies a circuit-based pathology for the disorder involving multiple interconnected brain regions.

It is thought that there is a dopaminergic abnormality underlying the pathogenesis of schizophrenia(115). The current dopamine hypothesis is that presynaptic dopamine availability and dopamine release are increased in schizophrenia patients, and in particular sub-cortical dopamine dysfunction plays a contributory role to both psychosis and the cognitive and negative symptoms of the disorder(116–118). This theory is given additional support by studies of prodromal individuals which have found that dopamine synthesis capacity appears to specifically increase in those who develop the disorder(119–121). However other studies have suggested that there may be 'hyperdopaminergic' and 'normodopaminergic' subtypes of schizophrenia, as for a significant number of patients, the mechanisms underlying their symptoms appear unrelated to dopamine dysregulation(122).

N-methyl-D-aspartate receptor (NMDAR) hypofunction has been hypothesized as being involved in schizophrenia, through development of earlier theories for the involvement of glutamatergic mechanisms in the disorder(123). NMDAR abnormalities have been observed in post-mortem studies(124), genetic studies(26) and in preclinical studies(125) using NMDAR antagonists as a model system for schizophrenia, though NMDAR hypofunction abnormalities may be present in only a subset of patients, and possibly only at a particular phase of the disorder(126). While recent studies have attempted to integrate the dopamine and NMDAR hypofunction hypotheses, it may be that they each underpin different subtypes of schizophrenia. The NMDAR hypofunction hypothesis provides a clearer way of accounting for the negative and cognitive symptoms of the disorder, and studies examining the effectiveness of antipsychotics have found that different individuals show treatment resistance to drugs which target dopaminergic systems to those which target glutamatergic systems(127).

Other neurotransmitter hypotheses include serotonin and the cholinergic system. The latter has been linked to schizophrenia through environmental associations such as the high prevalence of patients who smoke(128). Serotonergic dysfunction in the cerebral cortex is thought to be a cause of abnormalities in glutamate signalling and cortical atrophy in schizophrenia(129).

1.5 Diagnosis

1.5.1 Diagnosis through serum molecular profiling

Over the past decade, there have been attempts to investigate novel diagnostic strategies for schizophrenia through the identification of biomarkers in transcriptomic(130,131), proteomic(132), epigenetic(133) and neuroimaging studies(134). Particular attention has been given to identifying serum-based diagnostic biomarkers due to the ease of access to samples, and the potential for translation to clinical utility at relatively low cost(29).

1.5.2 Current diagnostic strategies and limitations

Currently, the diagnosis of schizophrenia is entirely based on the evaluation of signs and symptoms in clinical interviews, a procedure which has changed little in the past century(135). **Table 1.3** shows the current clinical diagnostic criteria for schizophrenia according to the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5), produced by the American Psychiatric Association(136).

Previous studies have stated that one of the key problems afflicting schizophrenia patients is the time taken to reach a correct diagnosis(137). As mentioned in **1.1**, one study found 50% patient diagnoses changed over the course of ten years, following initial admission for psychosis(13). A large component of this is thought to be the reliance on clinical interviews as the sole diagnostic technique. The problems of just using clinical interviews for diagnosis were highlighted in a landmark study in 1973 in which healthy individuals faking hallucinations were diagnosed with psychiatric disorders and admitted to hospital(138). More than 40 years on, misdiagnosis through clinical interviews remains common across all psychiatric disorders, with one study reporting that approximately one third of bipolar patients are mistakenly diagnosed with schizophrenia or other psychotic disorders(139). Various reasons have been postulated for why schizophrenia patients in particular are misdiagnosed, for example the refusal of some patients to acknowledge symptoms of psychosis such as hallucinations or delusions, the overlap of symptoms between different psychiatric disorders(140), the occurrence of similar symptoms in schizophrenia and mood and personality disorders(141), frequently occurring confounding factors such as substance abuse, and symptoms of other medical conditions(142). The heterogeneous nature of schizophrenia itself can make diagnoses difficult and unreliable as each patient may manifest a different subset of symptoms(143).

Table 1.3. DSM-5 criteria for the diagnosis of schizophrenia(136)

A	Two or more of the following, each present for a significant portion of time during a 1 month period: a. Delusions b. Hallucinations c. Disorganized speech d. Grossly disorganized or catatonic behaviour e. Negative symptoms, i.e., affective flattening, alogia or avolition Note: Only one Criterion A symptom is required if delusions are bizarre or hallucinations consist of a voice consisting of a running commentary on the person's behaviour or thoughts, or two or more voices conversing with each other.
B	For a significant portion of the time since the onset of the disturbance, one or more major areas of functioning, such as work, interpersonal relations, or self-care, are markedly below the level achieved prior to the onset (or when the onset is in childhood or adolescence, failure to achieve expected level of interpersonal, academic, or occupational achievement).
C	Continuous signs of the disturbance persist for at least 6 months. This 6-month period must include at least 1 month of symptoms (or less if successfully treated) that meet Criterion A (i.e., active-phase symptoms) and may include periods of prodromal or residual symptoms. During these prodromal or residual periods, the signs of the disturbance may be manifested by only negative symptoms or by two or more symptoms listed in Criterion A present in an attenuated form (e.g., odd beliefs, unusual perceptual experiences).
D	Schizoaffective disorder and depressive or bipolar disorder with psychotic features have been ruled out because either (1) no major depressive or manic episodes have occurred concurrently with the active phase symptoms; or (2) if mood episodes have occurred during active-phase symptoms, their total duration has been brief relative to the duration of the active and residual periods
E	Substance/general medical condition exclusion: The disturbance is not attributed to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or another medical condition.
F	If there is a history of autism spectrum disorder, the additional diagnosis of schizophrenia is made only if prominent delusions or hallucinations are also present for at least 1 month (or less if successfully treated).

In addition, there are currently no tools available to aid psychiatrists either with diagnosis, or prognosis for those in the prodromal phase to detect which individuals are more likely to develop schizophrenia. The latter is a particularly critical area of research in psychiatry as studies have shown that the earlier treatment intervention occurs, the more likely it is to be effective(144).

1.5.3 Serum biomarker panels

The identification of serum biomarker panels as the basis for a blood test which could be used as a clinical diagnostic tool, has long been an objective across the healthcare spectrum. Considerable investment has been made in biomarker studies in cancer(145), tuberculosis(146), Alzheimer's(147) and many other diseases(148). The WHO has encouraged biomarker research, stating that millions of patients could be saved from premature deaths through more accurate diagnostics(149).

If such a biomarker panel could be identified and established to work reproducibly and efficaciously in schizophrenia, it could potentially yield considerable clinical benefits. The current DSM-5 based diagnostic criteria, described in **1.5.1**, requires a 6 month duration of continuous symptoms before diagnosis is confirmed(132). The use of a biomarker panel could establish a diagnosis more rapidly, resulting in less delay before treatment and therefore better patient outcomes(132). Studies have shown that schizophrenia patients with a shorter duration of untreated psychosis are typically more responsive to antipsychotic treatment as well as being less symptomatic later on, resulting in lower maintenance doses of medication(144). One of the minimum requirements of a biomarker panel would be for it to be highly sensitive and specific. From a diagnostic perspective, the minimum sensitivity of clinical interviews in correctly diagnosing schizophrenia patients has been estimated to be 75%, hence it is felt that any additional diagnostic tool should have a sensitivity and specificity of at least 80%(150).

It is thought to be possible to identify CNS disorders relative to healthy controls through a fingerprint of alterations in peripheral protein biomarkers(132). The peripheral alterations described in **1.3.2** are believed to be a consequence of CNS dysfunction, and thus can be used as a tool for diagnosis. For example, as discussed in **1.4.2 and 1.4.4**, two of the main hypotheses for the onset of schizophrenia are immune and metabolic system dysfunction. These peripheral pathways can be altered via crosstalk with the CNS, mediated by various hormonal signalling networks, molecular signalling pathways and neuroanatomical networks(151). Various proteomic profiling studies conducted on the serum of schizophrenia patients have identified significant alterations in proteins relating to immune and metabolic pathways, as described in **1.3.2**. However, research remains ongoing into whether such alterations provide consistent predictive power across the spectrum of heterogeneity within the schizophrenia patient population, and whether their predictive ability remains robust to confounding factors which can affect the levels of serum proteins such as antipsychotic medication(152).

As of yet, no diagnostic biomarker tests for schizophrenia and other psychiatric disorders are clinically available. A variety of gene expression and microRNA profiling studies have been carried out to identify blood-based biomarkers for schizophrenia with a range of diagnostic accuracies obtained by plotting receiver operating characteristic (ROC) curves. Classification performances, based on the area under the curve (AUC), of 0.69-0.85 have been reported between patients and controls using these biomarkers (153–155). A 2010 proteomics study reported the identification of a diagnostic signature of 51 inflammatory, oxidative stress and hypothalamic–pituitary–adrenal signalling serum proteins altered in first onset schizophrenia patients compared to controls with 83% accuracy(156). This 51-analyte assay was incorporated into a clinical blood test but was subsequently withdrawn(152). While clinicians agreed that the test could be helpful in aiding diagnosis schizophrenia, measuring 51 proteins meant that it cost $2500 per patient, which was considered prohibitive(152,157).

In addition, subsequent studies have found that psychiatrists consider one of the most pressing clinical needs is a biomarker test which can predict psychosis or schizophrenia conversion in prodromal individuals(152). Overall, the biomarker field for schizophrenia remains in its early stages. The major limitation has been that it is difficult and expensive to recruit patients, and therefore many discovery studies are underpowered statistically, and lack validation cohorts. As the field develops, and more data becomes available, more meta analyses and validation will be possible(158).

1.5.4 Predicting disease conversion from the prodromal phase

The need for a prognostic tool based on biomarkers which can predict conversion to psychosis or schizophrenia in prodromal individuals, arises from evidence that one of the major causes of the delay in reaching a correct diagnosis of the disorder is the insidious disease onset. As a result many psychiatrists prescribe antipsychotic medication to individuals thought to be at high risk of schizophrenia purely on the basis of family history and early functional decline, even though a schizophrenia diagnosis has not been reached(159). Hence over the past two decades much focus has been placed on the prodromal phase of schizophrenia, such that the recent revision of the DSM-5 has seen prodromal syndrome listed as a ‘condition for further systematic study’ and a potential future diagnostic category(160). While prodromal symptoms such as disturbances in perception, thought processing, language and attention can resolve(161) and hence do not guarantee transition to an initial psychotic episode, they can indicate an increased risk of this transition(162). Studies have found that around 20-35% of individuals in the prodromal

phase go on to develop schizophrenia over a 2-3 year period(14). As a result prodromal schizophrenia is often referred to as ultra-high-risk syndrome(12).

A biomarker panel which could perform with high efficacy as both a diagnostic and a prognostic tool would be an important breakthrough in helping psychiatrists to identify particularly vulnerable patients early in the disease process, allowing for earlier or even preventative therapeutic intervention(163). Some studies have hypothesized that if it were possible to predict disease onset with near certainty, then it may be possible to reduce the duration of untreated psychosis to zero(152).

However, a biomarker panel intended for potential prognostic use in the clinic would need to be highly robust due to the potential consequences of false-positive predictions. The majority, approximately 70%, of individuals with prodromal symptoms do not develop schizophrenia, and a false diagnosis may have devastating consequences due to the stigma associated with the disorder(164). In addition, the consequences of unwarranted treatment could be serious as studies have shown that when antipsychotics are used to treat individuals with ultra-high-risk syndrome, there is an increased risk factor for motor side effects for example akathesia and EPS with first generation antipsychotics and metabolic side effects such as cardiac disease and diabetes with second generation antipsychotics(165).

So far few biomarker studies have been conducted in prodromal schizophrenia cohorts. A 2015 study identified a panel of 15 analytes in plasma capable of predicting conversions from individuals with ultra-high-risk syndrome to psychosis, but to date, this has not been taken forward to the clinic(166).

1.5.5 Current proteomic biomarker study limitations

There is a need for novel studies for the identification of diagnostic and prognostic biomarker panels for schizophrenia which avoid some of the limitations of existing proteomic biomarker studies across the field of medicine. Many of these limitations have been observed in studies attempting to identify biomarkers for different psychiatric disorders. This section looks to summarize these limitations, and address methodologies for tackling them.

Between 2008 and 2013, over 3000 publications were listed in PubMed for ‘biomarker discovery.’(167) Across these studies, a wide variety of bioinformatic strategies were used to try and identify panels of proteomic biomarkers which were sensitive and specific enough to be used as diagnostic tools. Various statistical and machine learning based classification methods such as support vector machines (SVM)(168), random forest(169,170), artificial

neural networks(171), and linear discriminant analysis(172) have been used to try and identify global trends in discovery datasets. Alternative approaches include the use of dimensionality reduction techniques such as principal component analysis (PCA) to identify important influences in data which can be taken as biomarkers(167). Newer studies have begun to explore the potential of Bayesian methods in proteomic analysis due to their capacity for handling complex, noisy and incomplete data(173,174). However, despite many model comparison biomarker studies using classification criteria including ROC based performance metrics, as yet no one model has been identified as being optimal for biomarker discovery(175–177). Instead the "correct" solution often hinges on a variety of factors ranging from the characteristics of the discovery dataset to the interpretability of the model for clinicians(178).

Regardless of the model employed, many proteomic biomarker discovery studies have suffered from limitations based on poor statistical design(179,180). This has seen the field come under scrutiny in recent years because few published candidate biomarkers have 'survived' validation and reached clinical utility, despite significant government and industry investment(181,182). Among the few positive case studies has been the identification of the HER-2/neu, estrogen receptor, and progesterone receptor biomarkers which are all used clinically for diagnosis and prognosis(183). However numerous other protein biomarkers which were initially thought to be highly discriminatory in breast cancer, and other diseases, for example in prostate, pancreatic, and ovarian cancers, have all failed to validate across multiple follow-up studies. Those studies revealed that the markers, having appeared promising, do not provide sufficient sensitivity, specificity and prognostic value for clinical decision making across a range of independent sample cohorts(184,185). These setbacks have illustrated that the challenge of finding a sensitive and specific panel of protein biomarkers is more complex than previously thought, particularly problematic in a highly heterogeneous disorder such as schizophrenia.

1.5.5.1 Statistical Shortcomings

Reviews seeking to address why so many discovery studies have ultimately failed to achieve their desired goals have pinpointed a variety of statistical downfalls(186). These are summarized as follows:

(a) Overfitting: Overfitting refers to the chosen statistical model fitting the discovery dataset to the extent that it picks up and learns random noise in the data as concepts, therefore negatively impacting its ability to generalize(187). Because the model does not reflect a general trend, its performance on independent validation datasets is negatively impacted.

Studies have cited overfitting as one of the main reasons why many proteomic biomarkers have failed to reach clinical utility, as they have been identified as strong predictors in the initial discovery study, but subsequent validation studies have proved disappointing(188).

Overfitting is exacerbated by small sample sizes, which is a particular problem in psychiatric research where substantial patient cohorts are difficult to obtain and expensive, particularly for serum profiling. Many proteomic discovery experiments involve the use of multiplex immunoassay technology for the measurement of protein levels or LC-MS^E for peptide identification and quantification, measuring sometimes thousands of variables (referred to as p) on a fairly small number of samples (referred to as n). The typical statistical guideline for discovery cohort dimensions is the number of events per predictor variable (EPV) where an 'event' is defined as the clinical group with the lower frequency, and the predictor variables are the proteins or peptides being measured(189). Studies suggest that the minimum EPV should be at least 5(189), but many discovery studies in psychiatry have an EPV of less than 1 because there are far fewer clinical samples than proteins/peptides being measured (190,191). For example, in *Perkins et al.*(166) the discovery cohort has an EPV of just 0.27, as 117 proteins are measured across 32 psychosis converters (the clinical group with the lower frequency). One of the consequences of a low EPV is that the statistical models trained on this discovery cohort may not have sufficient degrees of freedom to accurately estimate the full model. This is typically referred to as the "small n, large p" problem, and can affect the performance of the model on validation cohorts(187).

(b) No validation cohorts: To obtain an unbiased estimate of the performance of the statistical model trained on the discovery data, it is necessary to validate it on independent cohorts. This is required to evaluate whether the model is overfitting. However many study designs do not include validation cohorts, and only use the discovery data to estimate the model performance(188). This is again commonly seen in psychiatric research due to the difficulties in obtaining samples. For example, *Perkins et al*. identified a panel of 15 biomarkers for predicting conversion to psychosis from prodromal individuals with an AUC of 0.88 but the model was trained and tested on the same data(190). If the model is overfitting, the performance estimate will be overly optimistic compared to its true classification ability, as the biomarker coefficients are already strongly associated with the response variable(192,193).

(c) Inclusion criteria: An additional problem for biomarker discovery studies is that discovery datasets can be unintentionally pre-biased through selection of subjects(167,168). This is particularly pronounced in psychiatric disorders due to the high disease heterogeneity stemming from the various underlying pathophysiological causes. Therefore, biomarkers

identified in one particular set of subjects may not be robust in another dataset obtained through different selection criteria. This problem is especially pronounced when discovery datasets are small, but it can be negated somewhat by conducting biomarker discovery through meta-analysis of data from multiple independent studies(167,168). However, in psychiatric disorders, this problem is especially challenging because there are so many confounding variables associated with these disorders for example age, sex, and ethnicity, as well as environmental factors such as substance abuse, and life traumas(191).

(d) Erroneous application of models: In studies which do have validation cohorts, inflated performances estimates can be obtained through erroneous application of statistical methodologies(188). For example in *Chan et al*. 26 biomarkers were obtained through applying a logistic regression model to the discovery dataset(12). Subsequently, a new regression model was fitted to each of the validation cohorts using those 26 biomarkers, resulting in biased performance estimates of AUC = 0.9 and greater. The only way to ensure unbiased and accurate evaluation of the performance of a biomarker panel is to use the same model for discovery and validation.

1.5.5.2 Strategies for negating statistical shortcomings

Reviews seeking to improve the quality of future proteomic biomarker studies have proposed various strategies which can be used to address the above mentioned shortcomings, in an attempt to identify more robust and clinically relevant biomarkers(167).

(a) Meta-analysis: In addition to addressing the inclusion criteria problem, meta-analysis of multiple independent datasets to create a larger discovery dataset can be used to reduce the overfitting problem, by increasing the EPV(186). However due to sample limitations in psychiatric studies, it may still be difficult to reach the recommended 5 EPV.

(b) Reducing the model space: With p predictor variables, the total model space is 2^p. One of the strategies which can be used to reduce overfitting, is to prefilter the number of predictors in the discovery dataset through running a series of t-tests or linear regression model(186). Thus the dataset is stratified by eliminating proteins which were not significant ($p < 0.05$) according to the t-test/regression model, and then training a classification model on this filtered data.

(c) Utilizing prior information: If relevant prior information exists on the predictor variables in the discovery dataset, either through independent datasets or previously reported findings in the literature, this can be utilized through applying Bayesian models with an informed prior(194). This is thought to result in more stable and informed models through the twin

benefits of filtering spurious proteins, and potentially identifying higher levels of abstraction through groups of proteins acting in concert in a signalling pathway to influence biological response. This is thought to be a potential improvement on classification models such as Random Forest(167,194). Such methods have already been incorporated in other fields of medicine, for example the integration of existing data to predict response to an anti-cancer agent in breast cancer studies(195).

1.6 Preclinical models for drug discovery

Despite the findings from proteomic studies over the past decade regarding schizophrenia pathology, as summarized earlier in this chapter, there remain difficulties in translating findings into improved treatment options. This is felt to be partly due to a lack of effective preclinical models of schizophrenia which has hindered the ability to screen for novel compounds against targets of interest(196). There is a need for novel preclinical models of the disease which can be utilized to explore the affected functional pathways relating to changes observed in the periphery between control and disease states. To date, the majority of preclinical models for developing novel drugs for schizophrenia have been animal models, but it has been commonly observed that a key driver in the high drug attrition rate is the lack of a methodology for translating pathophysiologies identified in these models to the human disease (described further in **1.6.2**)(196).

1.6.1 Cellular models of schizophrenia

Over the past six years, cellular models of schizophrenia have formed a new class of preclinical models for investigating various hypotheses of the disorder(197). Peripheral blood cells, which express signalling pathways of interest found in the brain, for example those involving neurotransmitter and cytokine receptors(198), have already been used as a cellular model to identify known disease signatures in schizophrenia patients, as well as novel functional changes relating to impaired energy metabolism(197). Subsequent investigations have sought to use this model to screen various compounds against identified targets. A different cellular model, obtained directly from patients, has used neuron-like cells from reprogrammed fibroblasts to investigate the neurodevelopmental hypothesis of schizophrenia, reproducing abnormalities found in patients with the disorder such as reduced neuronal connectivity and post-synaptic density(199).

This section describes a novel cellular model of schizophrenia which is investigated later in this thesis, exposing the brain's immune cells, the microglia, to serum from schizophrenia

patients and controls *in vitro*. As explained in **1.6.1.2**, abnormal activation of brain microglial cells has been widely implicated in the pathogenesis of schizophrenia, a process which GWAS studies have indicated may be linked tothe presence of altered circulatory proteins with microglial activation propensity(25,200–202). Many hypotheses of schizophrenia, including immune system dysfunction (**1.4.2**), infectious disease (**1.4.3**) and neurodevelopmental dysfunction (**1.4.5**), are based around dysfunctional interaction between the periphery and CNS, which will be described in more detail in the next section. However, as yet, this hypothesized effect of circulatory protein abnormalities on microglial activation status has not been explored. Therefore analysis of data generated from this model may point towards new cell signalling targets, and provide a means of testing compounds with the potential to normalize the activation phenotype, on a human cell line.

1.6.1.1 Peripheral and central inflammation in schizophrenia

CNS immune function is driven by separate systems to that of the periphery, and thus the characteristics of neuroinflammation differ somewhat to inflammatory processes in other tissues. In particular, the brain contains few of the immune system cells found elsewhere in the body. However, there is a continuous crosstalk between the CNS and the periphery, reflected by an exchange of signalling proteins which can influence central immune system processes(203).

As such, there are a variety of mechanisms by which peripheral inflammation has been hypothesized to either mirror or cause CNS dysfunction in schizophrenia. As mentioned in **1.3.2** and **1.4.2**, many molecular profiling studies have identified increased blood concentrations of inflammatory cytokines in schizophrenia patients. Studies have suggested that high concentrations of cytokines may affect brain function through interaction with the vagus nerve(204), or through direct or indirect modulation of dopaminergic or glutamatergic neurotransmission(205–208).

The amount of crosstalk between the periphery and the CNS is typically restricted in healthy individuals due to the presence of the BBB. However high cytokine concentrations in the periphery can cause disruption, lowering the BBB and inducing abnormal trafficking of leukocytes and large inflammatory molecules between the periphery and the brain(203). Damage to the BBB impairs its ability to modulate which cells and molecules enter the brain, enhancing neuroinflammation and preventing normal brain function, leading to tissue damage(203). In addition, high concentrations of peripheral cytokines can stimulate cytokine production in the cells which form the BBB, allowing cytokines to reach the brain through the circumventricular organs, structures in the brain which lack a BBB(204).

1.6.1.2 The role of microglia in schizophrenia

Microglia are the brain's immune cells and provide a similar protective function in the CNS against insults and foreign invaders, to that of the peripheral immune cells in the rest of the body(209). In healthy individuals, microglia survey the entire CNS for pathogens and debris, and are involved in a variety of homeostatic functions including synaptic pruning and maintenance, developmental apoptosis, neuronal survival, trafficking of neurotransmitters, and phagocytosis(203,210,211). However, dysfunctional microglia-mediated inflammatory processes have been theorized as one of the mechanisms behind several of the hypotheses for schizophrenia, resulting in many of the structural brain changes known to be associated with the disorder(212,213). Microglia have different phenotypical activation states ranging from an M1 pro-inflammatory phenotype which arises in response to neuroinflammation and leads to the release of glutamate and inflammatory cytokines such as interleukin (IL)-6 and TNF-alpha, to an M2 neurotrophic phenotype which resolves the inflammatory response through the release of anti-inflammatory cytokines such as IL-4 and IL-13(214). The balance between the M1 and M2 states is known to be crucial to aid the brain in repairing itself from acute injury(215,216). However, while microglia naturally activate in response to pathogens, it is thought that microglial dysfunction can accentuate the schizophrenia disease process by subjecting the brain to an inflammatory assault causing disturbances in both grey and white matter which underpin the disorder(217–221).

The release of pro-inflammatory cytokines by M1-activated microglia is thought to have the capability to disrupt the BBB, allowing cytokines from the periphery to reach the brain, creating a build-up of neuroinflammation(203). The cytotoxic consequences of persistently activated microglia are hypothesized to potentially result in secondary neuronal degeneration, decreased neurogenesis and synaptic dysfunction(212). Little is known about the consequences for the CNS of microglial dysfunction in the M2 phenotype, particularly regarding schizophrenia. In addition, microglial dysfunction is thought to be associated with the neurodevelopmental hypothesis of schizophrenia, through neurotoxic mechanisms other than cytokine production. In healthy individuals, the interaction between microglia and complement activation is thought to be necessary for brain wiring and development in the postnatal period(222). However, when this process goes awry or is recapitulated during adulthood, the overactivation of the complement cascade is thought to lead to neurodegenerative processes resulting in destabilization of neuronal circuits and synapse elimination(223,224).

As of yet there is no direct empirical evidence proving the relationship between dysfunctional microglial activation and schizophrenia, but a number of studies have pointed towards a

causal link. Some post-mortem studies of the disorder have identified increased microglia density in several brain regions(225–228), and various animal studies based on pre or postnatal infection paradigms stimulating inflammation, have pointed towards a link between increased microglial density or activation and schizophrenia endophenotypes(229–232). Moreover, clinical and preclinical studies on agents such as minocycline which are known to be capable of crossing the BBB and inhibiting microglial activation, have shown benefits in ameliorating symptoms of schizophrenia(233–237).

1.6.1.3 The two-hit hypothesis

As mentioned in **1.4.3** and **1.4.5**, the theories of dysfunctional microglial activation as one of the mechanisms for schizophrenia neuropathology, are linked to the infectious disease and neurodevelopmental hypotheses for the disorder through the so-called “two-hit” hypothesis of schizophrenia(238). This postulates an involvement of genetic liabilities to schizophrenia through various risk factors such as those touched upon in **1.4.1**, combined with an environmental insult during the key early phases in neurodevelopment(239). This comprises the first hit, affecting CNS development and creating an abnormal signalling network(240). The first hit primes the individual for an event later in life, the second hit, which results in the onset of the disorder(239). This second hit typically occurs in adolescence or post-adolescence and can be one of a range of incidents from viral infection to immune system dysfunction or environmental stress(241). The two-hit hypothesis is summarized in **Figure 1.1**.

Various environmental insults have been suggested as potential contributory factors to the first hit. Past studies have pointed towards factors such as extreme stress causing an immune insult during gestation and affecting the developing nervous system(242,243). Evidence comes, for example, from epidemiological evidence of disease incidence following the Dutch famine of 1944-1945 which resulted in a two-fold increased in schizophrenia prevalence for offspring whose second trimester occurred during this period(244,245). Other studies have suggested that maternal infection during the prenatal phase could be a key component of the first hit as discussed in **1.4.3**, and that the resulting maternal immune activation could interfere with fetal brain development(241).

The role of microglia in the two-hit hypothesis is thought to be through early microglial activation via the first hit, subsequently sensitizing them for later activation(246). Animal studies have provided evidence for this theory, suggesting that infections or severe stress in the prenatal or early life phases can result in primed microglia which are then more easily activated(247–249). From a molecular perspective, this priming may result in abnormal

development of microglial cells themselves, possibly in terms of their glutamate receptor composition, leading to increased inflammatory cytokine release upon activation, which in turn could result in excitotoxic neuronal death(250).

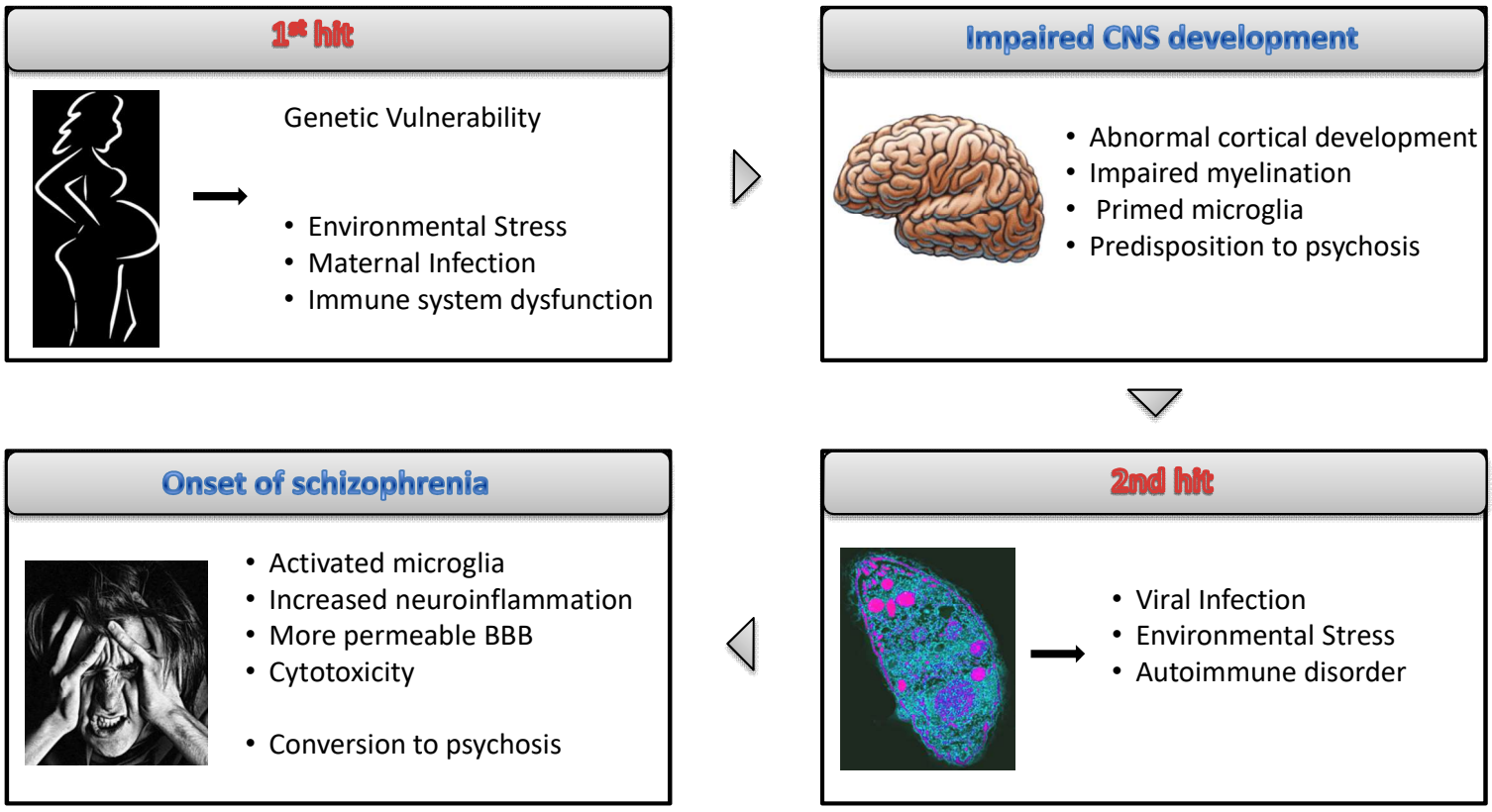


Figure 1.1. The two-hit hypothesis of schizophrenia. During key early phases in neurodevelopment, environmental insults combined with genetic vulnerability can prime microglia for a further event later in life, triggering microglial activation and neuroinflammatory processes which lead to the onset of schizophrenia. Abbreviations: BBB, blood brain barrier.

1.6.1.4 Bioinformatic strategies for investigating data from cellular models of schizophrenia

As mentioned earlier in this section, while cellular models of schizophrenia represent a novel area of preclinical research into the disorder, thus far no studies have investigated the functional effects of circulatory protein abnormalities on microglial activation status. One of the proteomic platforms used to analyze such cellular models is flow cytometry (described in **Chapter 2**) which typically yields median fluorescent intensities for different cellular epitopes in response to a stimulant. Experiments are often run in triplicate, and thus a linear mixed model provides a means of modeling any experimental variation between measurements. In addition, the molecular pathways in such models are typically driven by modules of proteins

working in concert rather than the expression of individual molecules(251). Such is the likely complexity of the signaling networks driving microglial phenotypes in schizophrenia, approaches such as Goeman's global test will be a necessary addition to univariate analysis to order to examine expression changes across the differentially regulated pathways. Through analyses which are based not just on single proteins, but affected pathways or signaling chains, a more complete interpretation can be reached(251).

1.6.2 Animal models of schizophrenia

As described in **1.1**, there has been a lack of significant breakthroughs in developing new treatments for schizophrenia over the past half century. This has had a significant impact on patients suffering from the disorder as while positive symptoms can be ameliorated to a certain extent by first and second generation antipsychotics, negative and cognitive symptoms remain resistant(252–254). As such there is a pressing need to develop novel compounds which display efficacy against these aetiopathological facets of the disorder.

1.6.2.1 The need for animal models of schizophrenia

Animal models represent the first phase in the drug discovery pipeline and are regarded as valuable preclinical tools to investigate the mechanisms behind disease progression and molecular pathways forming the basis of a complex disorder such as schizophrenia. Such models can provide a more rapid platform for monitoring disease progression, and provide a means of testing novel therapeutics which would not be feasible in humans(255). In addition, animal models of psychiatric disorders enable various environmental or genetic risk factors to examined under controlled conditions, and it has been hoped that models which recapitulate certain aspects of the human pathology in response to environmental or genetic manipulations can facilitate identification of novel targets for further validation.

It is currently estimated that there are over 20 different animal models of "schizophrenia", most of which produce a phenotype replicating certain aspects of the positive symptoms of schizophrenia although some also demonstrate aspects of impairments in cognition(22). These models are typically selected based on a spectrum of behavioural and biochemical abnormalities such as vulnerability to stress, abnormal response to reward, limbic dopamine dysregulation and cortical glutamatergic hypofunction.

1.6.2.2 Glutamatergic animal models of schizophrenia

As discussed in **1.4.6**, all currently available treatments for schizophrenia have been developed in response to various neurotransmitter theories of the disorder. The majority of preclinical research has focused on hypotheses based on dopaminergic system dysfunction but over the past two decades, there has been growing interest in glutamatergic models of schizophrenia based on evidence pointing to NMDAR abnormalities as one of the disease mechanisms(256).

NMDA antagonist models of schizophrenia are the most studied glutamatergic models of the disease(256). These models are based on the systemic injection of NMDA receptor antagonists such as the dissociative anethetic phencyclidine (PCP), ketamine, AP5 or MK801 in laboratory animals, predominantly rodents. These models have various advantages over alternative preclinical pharmacological models of the disorder such as those administering amphetamine or cannabinoids, most notably that they have a direct clinical parallel. Various clinical trials exposing healthy individuals to a single injection of ketamine or PCP have been shown to induce temporary schizophrenia-like symptomatology including positive symptoms such as paranoia and auditory hallucinations, as well as negative and cognitive symptoms such as social withdrawal and impaired working memory(257–262). In addition, both clinical and recreational use of PCP has been found to precipitate episodes of psychosis which are unresponsive to conventional antipsychotic treatment(258,263). PCP and ketamine preclinical models produce a phenotype which bears some similarities to the observed clinical effects of these agents in humans. This phenotype has been repeatedly shown to be characterised by hyperactivity, stereotypic behaviour, sensory gating deficits and impaired working memory performance(264–267).

In addition to pharmacologically based models of schizophrenia, in recent years mutant models have been developed based on studies which have identified genetic abnormalities in schizophrenia patients relating to the glutamatergic system. These models include neuregulin 1 hypomorphs, a gene which regulates expression of glutamate receptor subunits(268). Another mutant model of interest the NR1 partial knock-down model which reduces NMDAR expression and has been seen to produce a phenotype characterized by exaggerated locomotion, stereotypy and reduced social interaction(269).

1.6.2.3 The need for novel bioinformatic strategies to assessing translational relevance between animal models of psychiatric disorders and human pathology

As mentioned in **1.6**, for all animal models of psychiatric disorders, finding a means of assessing the translational relevance in a quantitative fashion is an ongoing challenge, particularly in terms of identifying which model represents a particular aspect of the human pathology most closely. *Nestler et al.* defined construct, face and predictive validity as a means of characterising preclinical models, as summarized in **Figure 1.2**(270)**.** Construct validity concerns whether a model replicates implicated aetiological factors of the human disorder. Face validity concerns whether a model shows symptomatic homology to the disorder, particularly in terms of known behavioural and biochemical aspects of the condition. Predictive validity concerns whether the observed face validity can be subsequently ameliorated following treatment interventions.

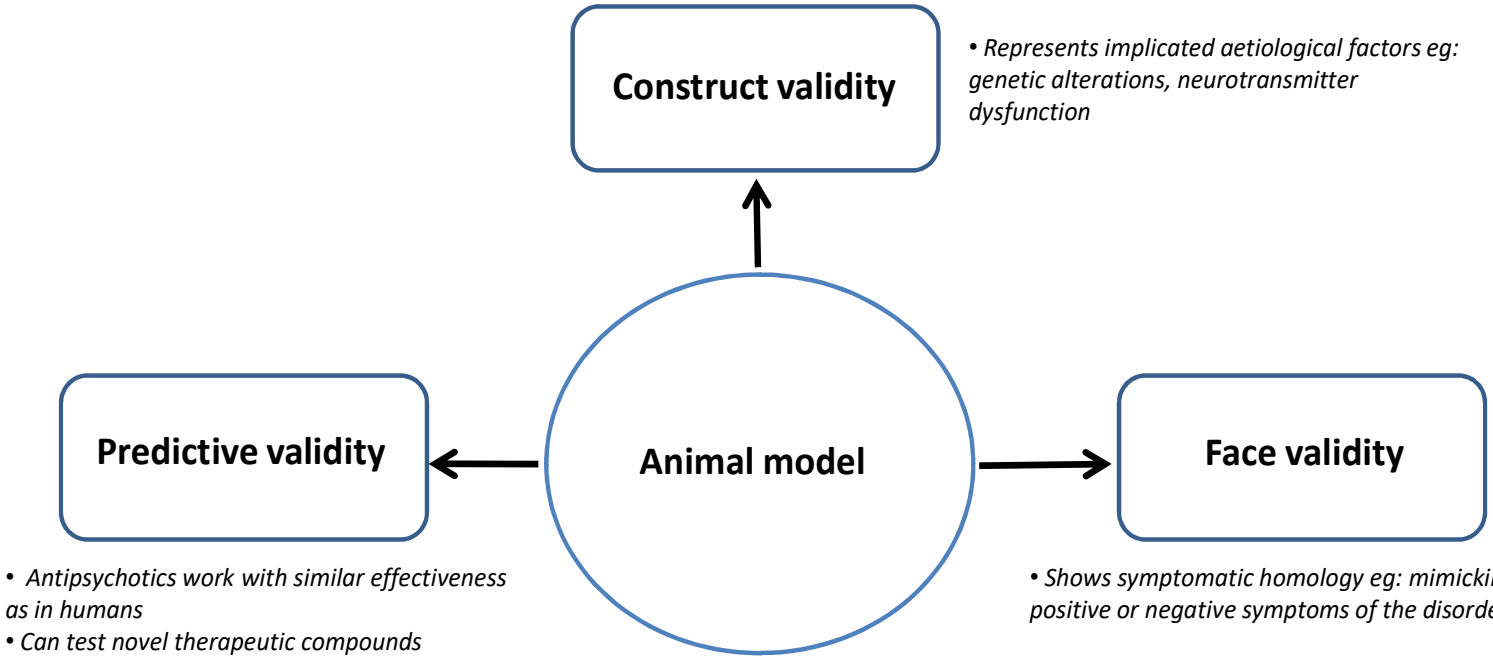


Figure 1.2. The three aspects of validity for assessing the translational relevance of an animal model of a psychiatric disorder to the human disease pathology; construct, face and predictive validity.

Despite these characterizations, many studies have found that behavioural observations associated with a particular animal model can be hard to directly quantify, even if correlated

with disease symptomatology, making it difficult to assess uniquely human behavioural symptoms in animals(255,271). Attempts to profile molecular biomarkers underpinning behavioural endophenotypes in animal models of psychiatric disorders have led to difficulties in interpretation due to the limited overlap in proteome coverage between human and rodent tissue (272) As a result, progress in developing new animal models for psychiatric disorders has slowed to a standstill, creating a pressing need for new bioinformatic methodologies to aid the comparison of these models to the disease state(270).

However, one emerging area of bioinformatics for the comparison of large scale proteomic data, such as that obtained from profiling analysis of tissue or serum from an animal model, is network biology and in particular protein-protein interaction (PPI) networks(273). These strategies can be used to obtain functional information represented by observed molecular changes and thus perhaps link functional alterations in animal models to the human disease on an ontology level, providing a new way of interrogating translational preclinical validity(274,275).

1.7 Thesis aims and outline

Despite considerable research, the aetiology and pathophysiology of schizophrenia remains little understood. Proteomic research can yield further insights into the pathophysiological mechanisms of the disorder through different profiling studies, with the aim of improving diagnosis and treatment of the disorder.

As such, the aims of this thesis are multi-faceted. This work (a) explores the potential for serum protein biomarkers to provide a new diagnostic and prognostic tool for schizophrenia; (b) analyzes data from a novel cellular model of schizophrenia to examine alterations in signalling pathways underpinning dysfunctional microglial activation in both antipsychotic naive and antipsychotic treated patients. These studies aim to identify new potential targets in the CNS and quantify the potential of various compounds for normalizing the observed changes; (c) develops a novel methodology to conduct a functional comparison between proteomic changes observed in animal models of a psychiatric disorder to those observed in human post-mortem brain tissue. The aim of this methodology is to enable more targeted preclinical studies in future, and thus aid identification of novel targets relating to particular neuropathological facets of a psychiatric disorder.

Below is an outline of the remaining chapters of this thesis:

- **Chapter 2** discusses the experimental platforms used to generate the data analyzed in this thesis and the statistical methodologies used to obtain the results presented in **Chapters 3-8**.
- **Chapter 3** uses a multiplex immunoassay platform to measure the concentrations of proteins across various inflammatory, metabolic and hormonal pathways in 11 cohorts of antipsychotic naive patients and controls, antipsychotic treated patients and controls, pre-symptomatic patients and controls, and prodromal converters and non converters. Multiple statistical models are trained on a meta-cohort of antipsychotic naive patients and controls. These models are subsequently tested on the remaining independent cohorts, assessing classification performance and thus their potential as a diagnostic and/or prognostic test. Variable selection methods are then used to identify important subsets of analytes in the training data relating to each model. New statistical models are subsequently trained on each of these sets of analytes, and classification performance on the independent cohorts is computed using these models to assess whether sufficient performance for clinical utility can be achieved with a smaller and more cost-effective set of proteins.
- **Chapter 4** applies the same methodologies as in **Chapter 3** except it utilizes targeted mass spectrometry to measure the abundances of panel of peptides linked to psychiatric disorders in 5 cohorts of antipsychotic naive patients and controls, antipsychotic treated patients and controls, pre-symptomatic patients and controls, and prodromal converters and non converters. As in **Chapter 3**, the aim is to identify a model which achieves both high diagnostic and prognostic performance when tested in independent cohorts, while remaining financially viable in terms of the number of analytes required.
- **Chapters 5 & 6** analyze data from a novel preclinical cellular model of schizophrenia to investigate dysfunctional microglial activation. **Chapter 5** investigates microglial disturbances in response to serum exposure from antipsychotic naive schizophrenia patients, and examines whether these signalling alterations can be normalized using known microglial inhibitors. **Chapter 6** investigates microglial activation in antipsychotic treated patients through analyzing data from the aforementioned cellular model, in conjunction with PET imaging data from these patients. The chapter subsequently examines whether antipsychotics may be involved in driving the observed signalling alterations in patients.
- **Chapter 7** develops a novel methodology for quantifying the molecular similarity of protein alterations observed in animal models of psychiatric disorders with the human

disease pathology, using data from proteomic profiling of brain tissue. The methodology is applied to identify common neuropathological domains between three environmental stress animal models and post-mortem tissue of major depressive disorder (MDD) patients, and quantify which model represents each neuropathological domain of MDD most closely.

- **Chapter 8** applies the novel methodology outlined in **Chapter 7** to conduct a functional comparison between four glutamatergic animal models of schizophrenia and post-mortem brain tissue from schizophrenia patients.
- Lastly, **Chapter 9** provides an overall summary of the findings from **Chapters 3-8**, and the implications of this work. In addition, this chapter discusses the limitations of the research presented and considers future studies which would build on these findings.

Chapter 2 Methods

2.1 Proteomic data generation

The proteomic data analyzed in this thesis is generated using four separate experimental platforms relating to the biological samples in question; multiplex immunoassay for the acquisition of serum protein concentrations (**Chapters 3 and 5**), MRM to quantify peptide abundances in serum (**Chapters 4 and 5**), flow cytometry for the measurement of fluorescent intensities from cellular epitopes in pathways relating to CNS pathology (**Chapter 5**) and LC-MS^E for the measurement of protein abundance in post-mortem brain tissue samples (**Chapters 6 and 7**). In **Chapter 5**, epitope measurements for schizophrenia patient and control samples are correlated with positron-emission tomography (PET) imaging data for several brain regions, an experimental platform will be very briefly summarized in this section. Details on how the data generated from each of these platforms was pre-processed and analyzed will be provided in each individual chapter.

2.1.1 Multiplex immunoassay

The multiplex immunoassay platform employed in **Chapter 3** has been utilised for the purpose of biomarker detection across a range of psychiatric illnesses including MDD(276,277) and Alzheimer's(278,279), as well as cancer(280). This particular immunoassay platform combines ELISAs with flow cytometry. The platform measures the concentrations of a pre-specified panel of proteins in serum using antibody based protein detection. This is achieved through the use of polystyrene microsphere beads coated with antibodies targeting the defined proteins. Each antibody can be individually identified by an internal dye consisting of a unique combination of red and infrared fluorophores. Following incubation with serum samples and protein binding, detection antibodies labelled with fluorescent dyes are added. Subsequently, the beads are fed into a flow cytometer where the internal and fluorescent detection dyes are utilized for identification and quantification of each protein. A red laser excites the internal dye, thus identifying the bead, while a blue laser excites the fluorescent detection dye and the amount of fluorescence is used to quantify the protein concentration levels. A similar platform is used in **Chapter 5** to simultaneously measure specific cytokines and growth factors. The technology is illustrated in **Figure 2.1**.

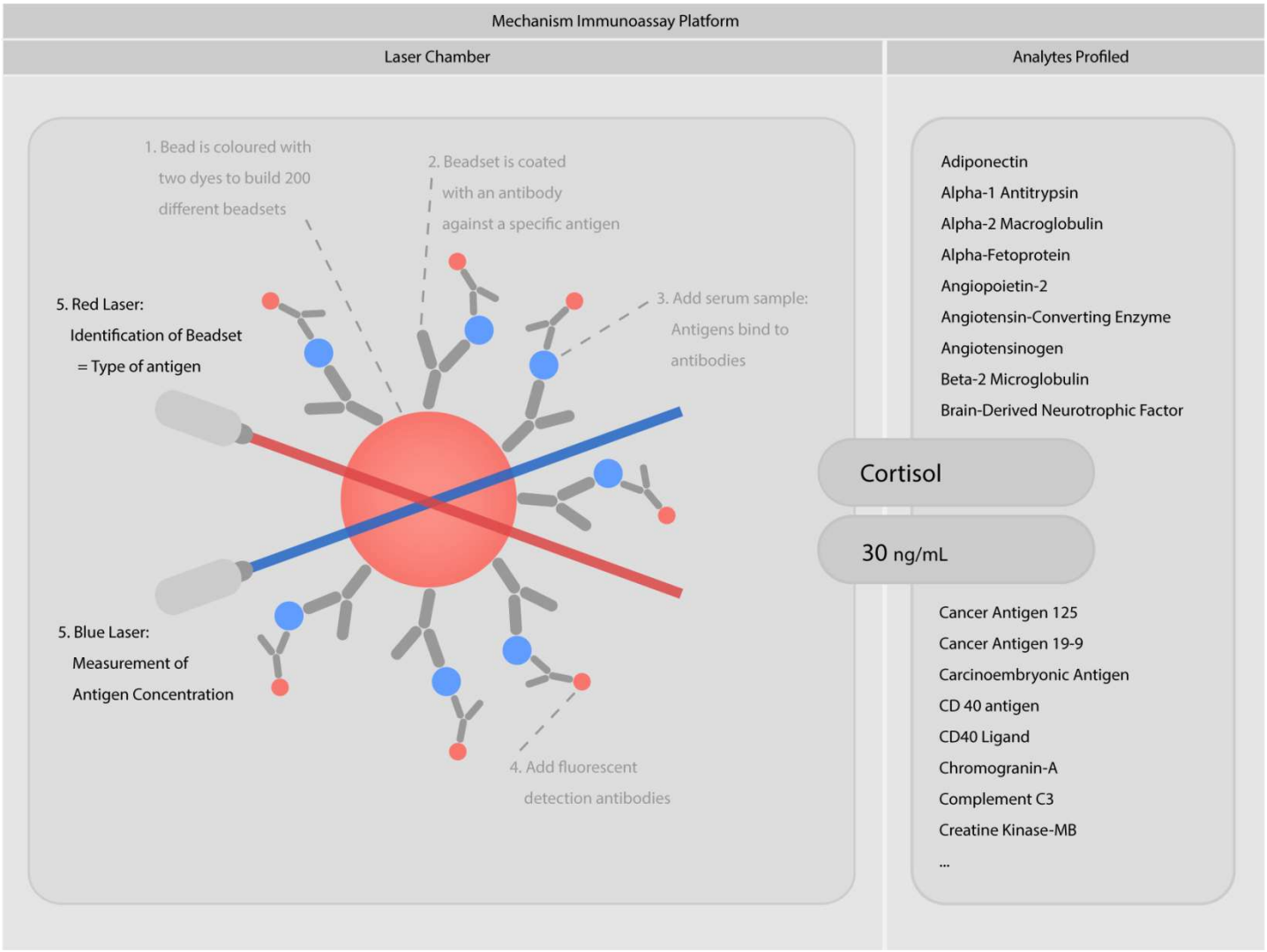


Figure 2.1 Overview of multiplex immunoassay technology used to measure protein concentrations in serum. *Image courtesy of Dr. Wolfgang Kluge, reprinted with permission*.

2.1.1.1 Advantages

In general, multiplex immunoassay platforms hold several advantages. The platform used in **Chapter 3** enables the simultaneous measurement of the concentrations of over 200 proteins in serum, namely high and low abundance proteins involved in various hormonal, immune and inflammatory, metabolic and neurotrophic pathways. The high sample throughput of these platforms enables the analysis of hundreds of samples within a short timeframe, reducing measurement variability and sample deterioration which can result from lengthy experiments. It is possible to detect proteins over a wide dynamic range of concentrations, enabling the simultaneous measurement of high and low abundance proteins. In particular, multiplex immunoassays are very sensitive and specific for proteins typically present in low serum concentrations such as hormones or cytokines which is one of the biggest advantages of the technology(281,282).

2.1.1.2 Disadvantages

The main drawback of these platforms is that there are a limited number of commercially available assays, and the proteins measured are biased towards a selected few biological pathways(283). Their reliability can also be compromised by cross reactions between proteins. Incorporating new assays is expensive as this can require generation of new antibodies, and extensive testing for interactions, for the protein of interest. In addition, they have relatively large sample volume requirements, are costly, and they can suffer from variation in concentration measurements due to the effects of batch to batch antibody variation(284,285).

2.1.2 Mass spectrometry

Two different mass spectrometry approaches are used in this thesis, which are summarized as follows.

2.1.2.1 Liquid chromatography mass spectrometry in expression mode (LC-MS^E)

In **Chapters 7 and 8**, LC-MS^E is used to measure peptide abundances in post-mortem brain tissue of patients with psychiatric disorders and animal models of those conditions. In these studies, the protocol applied involved protein digestion using trypsin, followed by peptide separation using liquid chromatography(286). Peptides are ionized as they are introduced into the mass spectrometer, before being accelerated through an electric field, meaning that the velocity of the ions depends on their molecular mass or mass to charge ratio (m/z). The accelerated ions are subsequently fragmented in the collision cell chamber by helium molecules. The peptides and thus proteins to which each of the ion fragments belong are identified by protein database algorithms which match them to corresponding precursor peptide ions through retention time, mass accuracy and other physiochemical properties. The ion fragments are fed into the time of flight analyser where they are accelerated by an electric field and the analyser measures the time taken by ions of different molecular mass values to reach the detector. The overall abundance quantification assumes that the mass spectrometry signal of a particular peptide is linearly related to its quantity(287). The data is subsequently exported for statistical analysis.

2.1.2.2 Advantages

Hundreds or even thousands of peptides can be identified in a single non-hypothesis driven discovery analysis, known as 'shotgun proteomics' which allows for more complex functional analyses such as the creation of protein-protein interaction (PPI) networks as presented in **Chapters 7 and 8**(288).

2.1.2.3 Disadvantages

Unlike multiplex immunoassays, LC-MS^E biomarker studies have been unable to detect lower abundance proteins such as cytokines, hormones and growth factors. In addition, LC-MS^E is thought to have poor reproducibility, which has contributed towards the failure of many proteomic biomarkers to validate(187,283).

2.1.2.4 Liquid chromatography multiple reaction monitoring mass spectrometry (MRM)

MRM is a highly sensitive and specific mass spectrometry technique which is used in **Chapters 4, 5 and 6** to measure the abundances of a panel of peptides known to be associated with psychiatric disorders in the serum of patients and controls. MRM is a targeted mass-spectrometry approach which is utilized for quantification of pre-selected biomarker panels based upon an existing hypothesis(289). As described in the LC-MS^E approach, the experimental protocol involved trypsin digestion of proteins, peptide separation, ionization and filtering, collision cell chamber fragmentation, and subsequent quantification. In addition, to these endogenous peptides, synthetically generated isotopically-labelled peptides with identical chromatographic, ionization and fragmentation properties, are added to the clinical samples in a 1:10 ratio to the endogenous peptides. This step is important to improve the specificity and reproducibility of the data(290). The data is subsequently exported for statistical analysis using a data pre-processing methodology outlined in **Chapter 4**.

2.1.2.5 Advantages

MRM has a number of advantages including low development costs and multiplexing capability(291). It has a higher dynamic range than traditional hypothesis-free mass spectrometry approaches like LC-MS^E allowing it to detect lower abundance proteins and is thought to be far more reproducible, allowing for sensitive and robust quantification(292). It requires lower sample volume than immunoassays which is useful from a practical

perspective as schizophrenia patient samples can be expensive and difficult to obtain(284). In addition, it does not suffer from batch to batch antibody variation resulting in fewer false positives and false negatives, and MRM experiments are not restricted to the commercial assays currently available.

2.1.2.6 Disadvantages

One of the main disadvantages of MRM is the time required to select representative peptide candidates for each protein of interest, synthesize isotopically labelled peptides for each endogenous peptide, and optimize the mass spectrometer in order to measure the particular set of peptides in question(293,294). In addition, because the intensities measured are those of peptide ion fragments (transitions) and the intensities of different transitions for a single peptide can differ substantially, manual validation is necessary to select the most abundant transition to represent that particular peptide(294). MRM also has a lower sample throughput compared to other technologies.

2.1.3 Flow cytometry

In **Chapters 5 and 6**, a flow cytometry platform incorporating fluorescent cell barcoding, is used for parallel detection of serum responses from patient and control populations across multiple intracellular signalling epitopes. This platform has been optimized across previous studies(196,295,296) to enable characterization of downstream signalling mechanisms in cellular models of schizophrenia. For the studies described in **Chapters 5 and 6**, a protocol is applied in which microglia cells are exposed to serum samples for thirty minutes before being fixed to stop signalling and phosphorylation events, and permeabilized so that antibodies for each epitope can enter the cells. Cells are subsequently multiplexed through the use of fluorescent cell barcoding technology to enable the simultaneous analysis of single-cell signalling responses of up to 139 serum samples, and then stained with antibodies specific to the phosphorylated form of the signalling epitopes of interest. Serum responses at individual epitopes are quantified through phosphoflow cytometry. This measures changes in phosphorylation status for each epitope in terms of the average of the fluorescent intensities of each antibody binding to each cell(297). These changes in phosphorylation status allows for the direct quantification of the activation or deactivation status of specific proteins(297). Protein phosphorylation is a reversible process which occurs through the addition of phosphatase to a protein by enzymes called kinases, resulting in the amino acids within the protein being phosphorylated or dephosphorylated, and thus regulating protein function(298). Phosphorylation of specific phosphoproteins is the main

mechanism through which extracellular stimuli such as cytokines, neurotrophic factors and hormones, modulate physiological processes(298).

2.1.3.1 Advantages

There are several advantages of the above flow cytometry platform in the context of proteomic data generation. Through measuring protein phosphorylation at different cellular epitopes, this approach provides more functional information than immunoassays, as it is measuring the quantity of activated or inhibited protein rather than protein concentrations(299). Secondly, many highly functional molecules such as protein kinases, have a fairly low level of expression, and thus other proteomic technologies do not detect them, focusing proteomic analyses towards more abundant proteins(300). Flow cytometry approaches make it possible to accurately measure signalling cascades downstream of these low abundance molecules. Thus thirdly, the use of flow cytometry and phospho-specific state antibodies makes it possible to interrogate *in vitro* a number of signalling processes for which function in microglial cells is hitherto unknown in the context of schizophrenia. Fourthly, this technology makes it possible to screen a variety of compounds across the same signalling epitopes previously stimulated by serum samples(295). This allows for the identification of potential clinically relevant compounds with the abilities to normalize dysregulated signalling mechanisms seen in the disease state.

2.1.3.2 Disadvantages

It is not possible to utilize the flow cytometry platform described in this thesis, to quantify causal relationships between epitopes in the microglial signalling network through correlation analysis(301). As such, in **Chapter 5 and 6** these relationships can only be speculated upon through the results of univariate and multivariate analyses of epitopes in different signalling pathways. This is because the platform measures three epitope responses at the same time in parallel, meaning that not all epitopes are measured in the same cell. Alternative technologies such as mass cytometry are capable of multiplexing 50 epitopes at a time(196), therefore offering the possibility of determining casual relationships between epitopes in the same signalling cascade across thousands of single cells.

2.1.4 Positron-emission tomography imaging

In **Chapter 6**, PET imaging is used to measure binding potential (BPND) of the PET tracer (R)-[11C]PK11195 in total gray matter, and five gray matter regions of interest in recent onset schizophrenia patients and controls. The BPND of this tracer was used to measure levels of 18 kDa translocator protein (TSPO). Elevated levels of this protein is associated with microglial activation. The protocol followed was the same as described in previous studies(302). Following tracer administration, a three-dimensional emission scan was obtained, using a head immobilization device to minimize subject motion. Binding potential information was extracted from the image data using a supervised cluster analysis algorithm described in previous studies(302).

2.2 Statistical methodologies

This section provides brief summaries of the main statistical techniques used throughout this thesis. Details of how each method is applied to each particular dataset are provided in those chapters. The data pre-processing and analyses in **Chapters 3-8** are carried out using the statistical program R (v3.1.3)(303), and the Bioconductor statistical package(304). Network analysis in **Chapters 7 and 8** is carried out using the software package Cytoscape (v3.3.0)(305).

2.2.1 Principal component analysis

PCA is a non-parametric, multivariate technique used to reduce the dimensionality of complex datasets to facilitate visualization of trends or artefacts in the data such as batch effects(306). PCA projects correlated variables to uncorrelated linear combinations of these variables called principal components. It makes the assumptions that the dataset consists of linear combination of the variables, that large variances have meaningful dynamics, and that principal components are orthogonal(306). The first principal component is obtained by identifying a vector known as the loading vector, along which the greatest variation occurs in the data in the original variable space. Subsequent principal components are obtained by identifying vectors in the direction of maximum variance, providing that they are orthogonal to all principal components previously identified. The first few principal components account for the majority of the variation in the data. PCA is used in **Chapters 3-8** to assess data artefacts such as batch effects from differing experimental run times between sample groups, or instrumental variation. The function prcomp in the R ggfortify package is used for PCA with default settings(303).

2.2.2 ComBat

Batch effects due to non-biological variation have long been known as a potential hindrance in gene expression microarray experiments, making samples in each 'batch' not directly comparable(307). Sources of such variation range from batches of reagent used to practical limitations forcing samples to be run on different well plates or different days(308). Such batch effects are not limited to microarrays. They can be seen in proteomic experiments, and when a meta-analysis of multiple studies is conducted(309). ComBat is an algorithm developed to eliminate such batch effects, which is designed to work well on relatively small sample sizes. The algorithm assumes that batch effects typically affect many analytes in similar ways, such as higher variability and increased concentration. It works by borrowing information across analytes in each batch to estimate the parameters that represent the batch effects, then shrinking these parameter estimates towards the overall mean of the estimates across all analytes in the data, then using these estimates to adjust the data. ComBat is used to normalize experimental variation in **Chapters 3, 4, 5 and 6**, using the ComBat function in the R sva package with default settings(310).

2.2.3 Stepwise variable selection

Stepwise variable selection is used for biomarker selection in conjunction with the logistic regression model in **Chapters 3 and 4**, and to adjust for confounding variables in univariate analyses in **Chapters 4-8** using the R function step in the stats package, with the arguments direction = "both" and k=log(n)(303). The Stepwise approach used in this thesis was mixed stepwise selection (chosen by selecting the direction="both" option). This considers either dropping or adding a variable at each stage, depending on whichever option minimizes the Bayesian information criterion (BIC). Setting k=log(n) in the step function ensures that the BIC is used as the criterion of choice. BIC penalizes larger models more severely than other criteria, and is asymptotically consistent(311). It provides an estimate of the test error of the model by adjusting the training error to account for the number of model parameters.

2.2.4 Wilcoxon rank-sum test

The Wilcoxon rank-sum test is used to compute epitope expression in response to serum/compound exposure relative to the vehicle condition for the cellular model investigated in **Chapters 5 and 6** using the R function wilcox_test in the coin package, with distribution = “exact” (312). It is a non-parametric test for comparing samples of observations n_A and n_B from two populations *A* and *B*. The test assumes that the two samples are independent of each other, and that the populations have equal variance(313). Applying the test in the case of unequal variance would be inaccurate due to lack of statistical power. The aim of the Wilcoxon rank-sum test is to examine the null hypothesis H_0 that the distributions of populations *A* and *B* are the same, as displayed below:

$$H_0: A = B$$

$$H_1: A \neq B$$

Hypothesis H_1 is as such unless there is a strong prior reason for hypothesising a shift in a particular direction.

The Wilcoxon rank-sum test tests H_0 through the following algorithm.

1. Combine the $n_A + n_B$ observations into one group *N* and rank them from smallest to largest. Compute the observed Wilcoxon rank-sum test statistic, W_A, whereby W_A = sum of the ranks for the observations from population *A*.
2. Permute population class labels across the ranks in *N* (10,000 permutations are used in this thesis). For each permutation, compute *W*, where *W* is the sum of the ranks now in population *A*.
3. Determine the *p-value* through:

$$P = \frac{\sum(W \neq W_A)}{n_{permute}}$$

2.2.5 Linear regression

Linear regression is used in this thesis to conduct exploratory univariate analyses, modelling the relationship between a response variable (clinical status) and one or more explanatory variables (typically protein concentration measurements and covariates such as age and gender). Algebraically, the model can be represented as(314):

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1 \ldots n$$

Where y_i represents the response variable for observation i, x_{i1} to x_{ik} represents the values of each of the k independent variables for observation i, β_i to β_k are the regression coefficients for each of the independent variables, n represents the number of observations, and ε_i is an unobserved error term. The regression coefficient β_k for an independent variable x_{ik} represents the predicted change in the response variable y_i for a unit change in x_{ik}, while keeping all other independent variables constant(314).

In order for linear regression estimates to be valid, several assumptions must be satisfied(315–317). These include:

- The response variable is a linear function of the independent variables and the error term.
- The samples are drawn randomly from the population.
- The expected value of the mean of the errors in the regression should be zero given the values of the independent variables, denoted algebraically as $E(\varepsilon/X) = 0$. This means that there is no relationship between errors and independent variables.
- There is homoskedasticity, meaning that the variance of the regression errors is constant. Heteroskedasticity can result in inaccurate standard error calculations.
- Error terms should be normally distributed, conditional on the independent variables.
- No multi-collinearity between independent variables which could result in unstable coefficients and inflate model variance.

Linear regression was used in **Chapter 4** using the lm function with default settings in the R stats package(303).

2.2.6 Logistic regression

Logistic regression is used in **Chapters 3 and 4** for classification purposes. Logistic regression models the probability of a binary event given the independent variables. Algebraically, the logistic regression model is represented as(318):

$$\log\left(\frac{\pi(x_i)}{1-\pi(x_i)}\right) = \beta_1 x_{i1} + \ldots. + \beta_k x_{ik} \quad , i = 1 \ldots. n$$

Where $\log\left(\frac{\pi(x_i)}{1-\pi(x_i)}\right)$ is referred to as the logit or the log odds of an event, x_{i1} to x_{ik} represents the values of each of the k independent variables for observation i, β_i to β_k are the regression coefficients for each of the independent variables, and n represents the number of observations. The regression coefficient β_k for an independent variable x_{ik} represents the

change in log odds ratio for a unit change in x_{ik}, while keeping all other independent variables constant.

In order for logistic regression estimates to be valid, several assumptions must be satisfied(318,319). These include:

- The response variable should be binomially distributed.
- The logit and the independent variables should be linearly related
- No multi-collinearity between independent variables which could result in unstable coefficients and inflate model variance.
- Error terms should be independent of each other

Logistic regression is implemented in R in this thesis using the glm function with family = "binomial" in the stats package(303).

2.2.7 Ridge regression and LASSO

Ridge regression and least absolute shrinkage and selection operator (LASSO) are forms of penalized regression applied in **Chapters 3 and 4** using the glmnet function in the R glmnet package with family = "binomial" (320). As is customary, ridge regression was applied by setting alpha = 0 in the function, LASSO was applied by setting alpha = 1. The tuning parameter λ was obtained through cross-validation.

Ridge regression is used with a binomial distribution for binary classification and LASSO is used for variable selection. The ridge regression method is designed to negate overfitting which may occur due to problems associated with multicollinearity which are inherent to many datasets. Multicollinearity causes the variability of parameter coefficients to be large, meaning they are far from the true value. Ridge regression reduces the standard errors through adding a degree of bias by penalizing large coefficients through the L2 norm (the Euclidean length) of the parameter coefficient vector. Algebraically, ridge regression looks to minimize the log likelihood function (L) added to the L2 norm:

$$L + \lambda \sum_{i=1}^{n} \beta_j^2$$

Where λ is the tuning parameter, n is the number of coefficients, and β_i are the estimated coefficients.

LASSO also penalizes large coefficients but it does so through the L1 norm (the sum of the absolute values of the parameter coefficient vector) which enables it to shrink some parameter coefficients to zero, meaning it can be used as a variable selection technique. Algebraically, LASSO looks to minimize the log likelihood function added to the L1 norm(321):

$$\mathrm{L} + \lambda \sum_{i=1}^{n} |\beta_i|$$

Where λ is the tuning parameter, n is the number of coefficients, and β_i are the estimated coefficients.

Elastic net regularization, a compromise between the LASSO and Ridge approaches which selects an optimum value of the alpha parameter between 0 and 1 (typically through cross-validation), was not considered for fitting classification models in the work conducted in **Chapters 3 and 4**, due to the model comparison nature of these studies. Ridge regression was used for the classification stages of these studies, whilst Elastic net would not have been suitable in this case, because it would have selected a subset of the initial predictors. To enable model comparison, it was important that the classification methods chosen (Logistic regression, Ridge regression, Random forest, SVM, Bayesian LASSO and BART), were not implementing any variable selection as part of the model fitting. LASSO was used at the 'Biomarker Identification' stages of these studies to select the most important variables from the initial sets of predictors. Elastic net would have been a viable alternative, but LASSO was selected as it has been used for this purpose in a number of previous proteomics studies in this field(322,323).

2.2.8 Random forest

The Random Forest classification algorithm is an example of a classification tree based method. Many classification trees (the exact number is defined as a model parameter) are constructed, and test data is classified based on the majority decision of these trees. Each tree in the forest is grown from the training data, beginning with all training samples in the root node. Every time this node is split into two child nodes (which then become the new root nodes), the search for the best variable (in **Chapter 3** the variables are protein concentrations, and in **Chapter 4** the variables are peptide abundances) to split on is limited to a subset of m variables (the model parameter mtry) which is randomly drawn (with replacement) from the total number of variables n, in the training data(324). The variable chosen to split the node sample pool on is determined by the maximum decrease in the Gini impurity criterion, a measure of node class heterogeneity. The algorithm terminates when for

each tree, all samples belonging to the same node are from the same clinical group, and the Gini impurity criterion is equal to 0(325).

For binary classification problems such as those examined in this thesis, the Gini impurity criterion I_g for a particular node, is defined as(314):

$$\mathrm{I_g} = 1 - \sum_{\mathrm{i=0}}^{1} \mathrm{P_i^2}$$

Where p_i is the proportion of data points with label i in the node sample pool.

Random Forest is used in **Chapters 3 and 4**, using the train function in the R caret package with method="rf"(326). Model parameters number of trees and mtry are obtained through cross-validation.

2.2.9 Support vector machines

SVM is a machine learning method which is used for classification purposes in **Chapter 3** and **Chapter 4**. When applied to the training data, the algorithm attempts to find an optimal hyperplane which will separate the data into two clinical groups with minimal error(327). However, as a linear hyperplane is typically insufficient for stratifying the data in this way, a nonlinear hyperplane is required. This hyperplane is constructed throughout mapping the training data to points in a higher dimensional space, called the feature space, through the use of an appropriate kernel function(327). The closest equidistant data points to the hyperplane in the feature space are known as support vectors.

In **Chapter 3** and **Chapter 4**, a polynomial kernel is used. For prediction on a test data, the independent samples are projected into the feature space and assigned a class, based upon the side of the margin that they fall. Algebraically, the SVM obtains predictions through(328):

$$\mathrm{f(x)} = \mathrm{sign}(\sum_{\mathrm{i=1}}^{\mathrm{n_{sv}}} \alpha_\mathrm{i}\, \mathrm{y_i K(x, x_i)} + \mathrm{b})$$

where $f(x)$ represents the predicted classes for each of the test data samples, x_i is the training data samples, y_i is the training data classes, n_{sv} is the number of support vectors, and α and b are obtained by solving the classifier optimization problem.

The polynomial kernel is defined as(327):

$$K(x, x_i) = (\gamma(x^T x_i) + b)^d$$

Where *d* is the degree parameter, *γ* is a scale parameter, and *b* is an offset parameter.

SVM is applied in **Chapters 3 and 4**, using the train function in the R caret package with method='svmPoly'(326). The cost, degree and scale parameters are obtained through cross-validation.

2.2.10 Bayesian modelling

In **Chapters 3 and 4**, two different Bayesian modelling approaches are used to train a classifier and test on an independent dataset. The key to all Bayesian approaches is Bayes' theorem, summarized as *'the posterior is proportional to the prior times likelihood'* or algebraically(187):

$$P(\theta|Y) \propto P(\theta) \times P(Y|\theta)$$

Where *Y* refers to the experimental data, and ϑ is the unknown parameter variables. $P(\theta|Y)$is the posterior distribution, or the joint probability distribution of the unknown parameters given the experimental data. $P(\theta)$is the prior distribution, or the existing knowledge regarding the unknown parameters before any data has been measured, in the form of a probability distribution. $P(Y|\theta)$ is the likelihood or the conditional probability distribution of the experimental data, given the unknown parameters(187).

2.2.10.1 Bayesian LASSO

One of the Bayesian methodologies used in this thesis is the Bayesian LASSO which has been utilized in a variety of GWAS datasets with sparse parameter spaces(329). The Bayesian LASSO places a Laplace prior on the parameter vector $P(\beta|\sigma^2)$ This can be expressed algebraically as(330):

$$P(\beta|\sigma^2) = \prod_{j=1}^{p} \frac{\lambda}{2\sqrt{\sigma^2}} e^{-\lambda|\beta_j|\sqrt{\sigma^2}}$$

Where *β* refers to the parameter coefficients and σ^2 is the model variance. λ is the LASSO tuning parameter.*j* is the number of parameters. The median of the Bayesian LASSO posterior distribution estimates corresponds to the standard LASSO point estimates. The Bayesian LASSO approach differs from LASSO in that it automatically provides credible interval estimates for all parameters in the model.

In **Chapters 3 and 4**, the Bayesian LASSO is implemented using the reglogit function in the reglogit R package with 10,000 iterations and normalize=FALSE. This is a Markov Chain Monte Carlo (MCMC) implementation of the model algorithm that is equivalent to a logistic regression with double exponential priors(331).

2.2.10.2 BART

BART is a non-parametric, classification tree based method, which uses sums of regression trees. BART and other Bayesian classification and regression decision tree (CART) approaches are popular in classification problems due to their ability to capture interactions, non-linearities and additive effects in datasets(332). Unlike Random Forest, BART takes a set of trees, $T_0 \ldots . T_k$ (the number k is defined as a model parameter) and updates them again and again through MCMC methods, known as a stochastic search, throughout the model training process.

Each tree has a prior distribution, and a prior on the tree terminal node outputs which represents the probability of that node splitting (thus an informative prior can be incorporated into the model providing information from independent datasets on which node is likely to split)(333). At each iteration *k* a new tree is proposed through either growing or swapping the nodes of the most recently accepted tree. The new tree is either accepted or rejected based on how well it matches the training data and prior distribution. The algorithm continues to iteratively sample trees until the model parameter estimates are stable, and convergence is reached(333).

Posterior predictions are then obtained by adding the MCMC samples across all trees, thus combining the prior distributions across each tree with the tree model likelihood. The BART model can be summarized algebraically for binary classification problems as(327):

$$P(Y = 1|X) = \varphi(T_1^M(X) + T_2^M(X) + \cdots + T_m^M(X))$$

Where *Y* is the response variable, *X* is the dataset of parameter variables, φ is the cumulative density function of the standard normal distribution, *m* is the number of trees and *T* is the total tree structure. *M* is the parameters at each root node in the tree. T^M denotes an entire tree with its structure and all node parameters.

In **Chapters 3 and 4**, BART is implemented using the bartMachineCV function in the bartMachine R package with serialize=TRUE(334). Model hyperparameters and number of trees are obtained through cross-validation.

2.2.11 Mixed effects models

While a linear regression model, contains only fixed effects, mixed effects models have both fixed and random effects(335). This can be utilized when analyzing proteomic data where there is more than one observation per subject for a particular analyte. These individual differences can be modelled by assigning different random intercepts for each subject to represent the idiosyncratic variation(335).

In **Chapters 5, 6, 7 and 8**, a linear mixed model is used to model the random variation incurred by running multiple replicates on the flow cytometry and LCMS^E platforms described in **2.1**. Algebraically, this model is of the form:

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} + v_i + \varepsilon_i , \quad i = 1 \ldots . n$$

Where y_i represents the response variable for observation *i*, x_{i1} to x_{ik} represents the values of each of the *k* independent variables for observation *i*,β_1 to β_k are the regression coefficients for each of the independent variables,β_0 is the fixed intercept of the regression model, v_i is the random intercept for observation *i*,*n* represents the number of observations, and ε_i is an unobserved error term.

The model has several assumptions, including that the response variable is a linear function of the independent variables and the error term, the samples are drawn randomly from the population, and the random intercept and error term are independent of one another(335). The linear mixed model is implemented in R using the lme function in the nlme package with default settings(336).

2.2.12 Classification performance

In order to assess the classification performances of the statistical models in **Chapters 3 and 4**, ROC curves are plotted. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR). The TPR is also known as the sensitivity, and is the proportion of patients who are correctly identified by the model. The FPR is 1-specificity and is the proportion of controls who are wrongly classified as patients by the model. The line y=x on a ROC curve plot represents the performance which would be expected through random chance, and the greater the distance between a ROC curve and this line, the greater the classification performance(337).

Classification performance is assessed through three measures, measuring AUC, and computing the classifier sensitivity and specificity via the Youden's index(338).

The AUC represents the probability that a randomly chosen patient with schizophrenia will be ranked higher by the model than a randomly chosen healthy control(337). Previous studies using ROC curves to evaluate classification performance have defined the ranges of AUC values as follows; 0-9-1 (excellent), 0.8-0.9 (good), 0.7-0.8 (fair), 0.6-0.7 (poor), <0.6 (fail)(339–341).

Values for sensitivity and specificity are computed through identification of an optimal threshold point on the ROC curve. This was done using the commonly advocated Youden's index (*J*; calculated by $J = sensitivity + specificity - 1$)(338). The aim of the Youden's index is to maximise the difference between the TPR and FPR, thus maximizing the correct classification rate.

Classification performance is computed in R using the performance function in the ROCR package(342) with "auc", "sens" and "spec" for calculating AUC, sensitivity and specificity.

2.2.13 Goeman's global test

Goeman's global test is used to obtain further information on whether different microglial signalling pathways are dysregulated between patients and controls in **Chapters 5 and 6** by testing whether the combined expression profile of multiple individual epitopes in a pathway is associated with patient-control status. This is done through modelling individual epitope expressions as random effects in a logistic regression model(343). The global test was initially developed for microarray data to assess whether groups of genes were differentially expressed between controls and patients thus shifting analyses from the single gene level to the pathway level(343) and the method has subsequently been adopted for utility on proteomic datasets(344).

The assumption behind Goeman's global test is that all parameter coefficients in the regression are sampled from a common distribution with mean zero and variance τ^2. The null hypothesis is that the parameters are not differentially expressed, meaning that the parameter coefficients (β) are zero, and thus $\tau^2 = 0$. The global test model is defined algebraically for logistic regression as(343,345):

$$E(Y_i|\beta) = h^{-1}(\alpha + \sum_{j=1}^{m} x_{ij}\beta_j), \quad i = 1 \ldots n, \quad j = 1 \ldots . m$$

Where with a binary response variable *Y*, h^{-1} is the logit function, α is the intercept, and *x* is an *n x m* matrix of *n* subjects and *m* proteins, *β*are the regression coefficients. This simplifies to the following random effects model through the notation $r_i = \sum_j x_{ij}\beta_j$

$$E(Y_i|r_i) = h^{-1}(\alpha + r_i)$$

The null hypothesis for the global test can then be tested using a score test statistic which can be obtained from this model through procedures described at length in previous review papers(346,347). The score test statistic for each of the pathways examined in **Chapters 5 and 6** is a weighted average of the individual score test statistics for each epitope in the pathway. A p-value is then obtained for the pathway through permuting this test statistic. Goeman's global test is implemented in R using the gt function in the globaltest package with model="logistic" and 10,000 permutations(348).

2.2.14 Protein-protein interaction networks

In **Chapters 7 and 8**, a functional comparison is conducted between protein expression profiles in brain tissue from post-mortem samples and animal models of psychiatric disorders. This comparison is made possible through the construction of PPI networks, integrating individual protein abundances found to be significantly altered between disease and control groups with cellular network information on their first-degree interactors from multiple protein interaction data repositories. In each PPI network, proteins are represented as nodes, and interactions as edges. In **Chapters 7 and 8**, PPI networks were created using the software program Cytoscape(305) and the databases IntAct(349), MINT(350) and UniProt(351). Filtering was applied to ensure that only proteins expressed in the organism of interest were included in the network, and to exclude all interactions other than direct interactions or physical associations.

Functional annotation of different modules or clusters of interacting proteins within the network, is conducted through protein enrichment analysis using the Cytoscape plugin ClueGO(352). ClueGO conducts enrichment through the Gene Ontology (GO) database, identifying biological process GO terms annotated to each protein in the network, and using a hypergeometric test to assess which GO term annotations appear significantly more frequently than would be expected by chance(352). Subsequently, significant GO terms can be summarized in functional groups using the kappa score statistic which is a metric used to assess the similarity of GO terms based on shared underlying proteins.

More specific information on how PPI networks, GO term enrichment and kappa score grouping are used in this thesis and the other methodologies utilized for the functional comparison, can be found in **Chapters 7 and 8**.

Chapter 3 Comparison of statistical models for the classification and prediction of schizophrenia diagnosis through multiplex immunoassay profiling of serum

3.1 Introduction

In order to improve schizophrenia diagnosis, various studies in the past decade have described the potential benefits of a diagnostic test based on robust biomarkers in blood which can differentiate schizophrenia patients from healthy controls as a means of aiding psychiatrists reach an accurate conclusion(152,353). In addition, as discussed in **Chapter 1**, the clinical relevance of such a test would be enhanced if it also displayed the capacity to provide prognostic information regarding whether prodromal individuals at risk of developing schizophrenia would transition or not(152). Prodromal schizophrenia has become an area of increased research focus over the past two decades(12) as investigations have found that 20-30% of these individuals go on to develop schizophrenia over a two year period(14). If a biomarker test could be used to differentiate prodromal individuals who later experience a psychotic episode from those who do not, with a high classification performance as measured through the AUC, this could be greatly beneficial. Such a test would also need to be both highly sensitive and specific because, as mentioned in **Chapter 1**, previous research has suggested a minimum AUC, sensitivity and specificity threshold of 80% for a test to have clinical value(150). False positive predictions for prodromal individuals resulting in unnecessary treatment could have severe health consequences due to the side-effect profile of many conventional antipsychotics(165). An additional requirement is for such a test to require relatively few biomarkers as the cost of a test increases with the number of analytes measured. A previous attempt to develop a diagnostic test for schizophrenia did not reach clinical utility, as with 51 biomarkers, it was deemed too expensive(152).

The aim of this study is to identify a statistical model based on serum protein concentrations which has the potential to both reproducibly classify schizophrenia from healthy controls, and prodromal individuals who later develop psychosis (converters) from prodromal individuals who do not develop schizophrenia (non-converters). Proteomics has been used extensively as a means of identifying disease-associated biomarkers through altered expression levels.

The combination of a set of proteins as a biomarker panel is thought to potentially yield a specific signature for a particular disease(354), and an appropriate model based on such a panel could form the basis of a diagnostic or prognostic blood test(166,355,356). In this study protein concentrations are measured using the multiplex immunoassay platform described in **Chapter 2** which measures up to 225 proteins across various hormonal, inflammatory, and metabolic pathways, many of which have been previously implicated in schizophrenia(132,157). Immunoassays are thought to be more reliable than more traditional methods such as Western Blot which have lacked the ability to reliably detect different proteins across a broad dynamic range, while the development of ELISA assays are more costly and associated with a high failure rate(357). The platform used in this study has been applied in recent biomarker studies across a range of diseases including autoimmune disorders(358), coronary artery disease(359), epithelial ovarian cancer(281) and MDD(360).

This study aims to address some of the bioinformatics shortcomings highlighted by recent reviews of proteomic biomarker studies, as discussed in **Chapter 1**. In particular, many studies suffer from small sample sizes relative to the number of proteins measured, known as a "small n, large p" problem, leading to overfitting and identification of biomarkers and clinical demographic variables which do not reflect the wider disease population(186). Previous reviews have shown that psychiatry and clinical neuroscience as a whole suffer from small sample sizes, reducing the reliability and reproducibility of findings, yet obtaining larger cohorts is both costly and difficult(361). In the discovery stage of this study, a training dataset is created by combining five independent cohorts through meta analysis. The resulting meta-cohort consists of 204 controls and 127 first-onset antipsychotic naive schizophrenia patients. Previous research suggests that at least 5 EPV is required to avoid overfitting(189), and while the meta-cohort still has only 1.92 EPV, it is still greater than the cohorts typically used for identifying proteomic biomarkers in psychiatry. Recent discovery cohorts used for psychosis, MDD and bipolar disorder had 0.2 EPV(166), 0.79 EPV(191) and 1.64 EPV(362), respectively.

There are a preponderance of different statistical methodologies available for use in classification problems, but different studies have shown that the best performing algorithms are highly dependent on the individual dataset characteristics(363,364). In this study, models are fitted on the training data using a range of different statistical methods including Logistic Regression, Ridge Regression, the machine learning algorithms Random Forest and SVM, and the Bayesian methodologies Bayesian LASSO and BART. The rationale for applying a range of different methods to tackle this particular classification problem is to see which algorithm most accurately represents the underlying trends in the data, and thus can identify

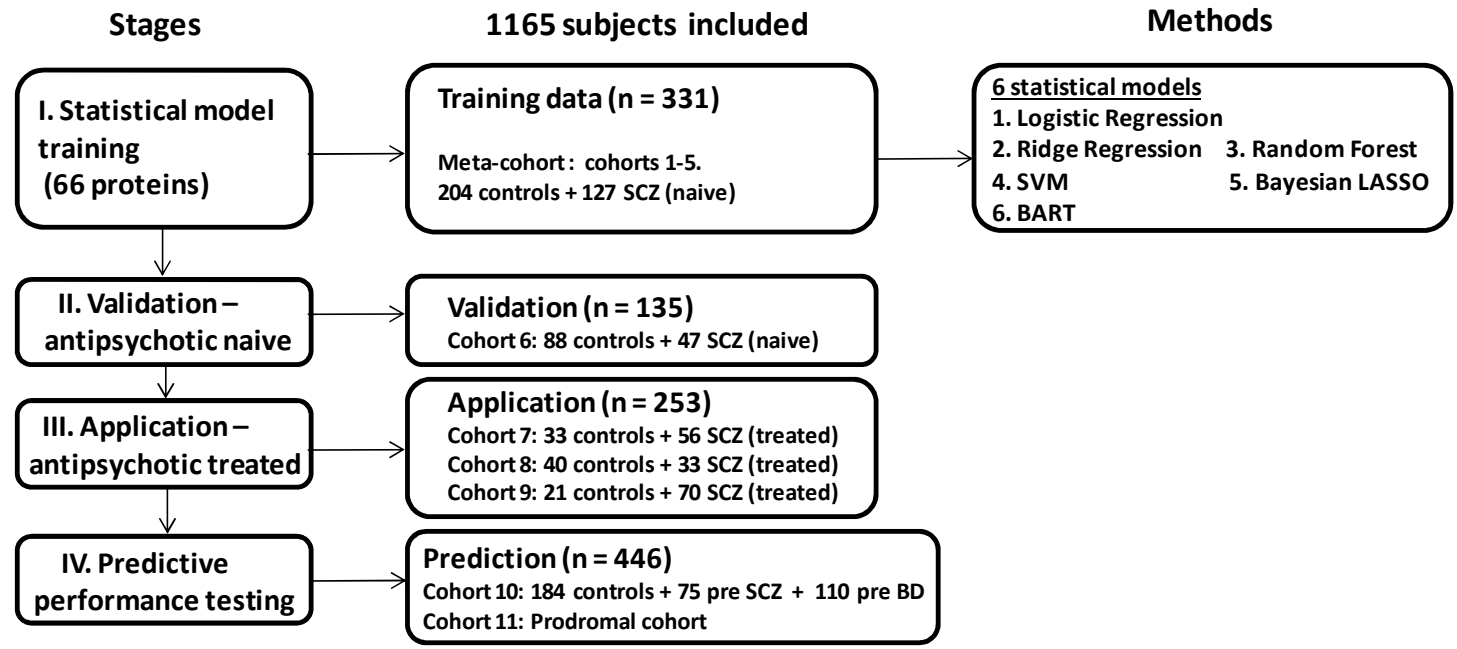
a model which produces robust classification performance on independent cohorts. Advantages and disadvantages of each of these methods are discussed in **3.2**.

As discussed in **Chapter 1**, relatively few studies which attempt to identify diagnostic protein biomarkers include independent validation cohorts, and some of those which do, apply statistical models incorrectly, resulting in inflated assessments of the real classification abilities of the underlying biomarkers(365). This study tests each model on an independent validation cohort consisting of 88 controls and 47 first-onset antipsychotic naive patients. In addition, while previous attempts to identify a diagnostic biomarker test for schizophrenia have focused mainly on first-onset antipsychotic naive schizophrenia patients(355), it is also relevant to examine the performance of a proposed test in classifying schizophrenia patients already on some form of antipsychotic treatment(152), a comparison which has not been assessed in past studies. This comparison is of interest because research has shown that many individuals thought to be at risk of the disorder are pre-treated with antipsychotic medication(162) so when a diagnosis is reached, they may have been taking antipsychotics for a period of time, and studies have found that antipsychotic medication can modulate peripheral biomarkers of the disorder(63,366). As such, this study design includes the testing of each model on three independent application cohorts of recent-onset antipsychotic treated schizophrenia patients and controls. To assess prognostic classification performance, each model is tested on two different prediction cohorts, one consisting of pre-symptomatic individuals and another consisting of prodromal individuals. In each case, serum was sampled in a timeframe of several months and several years prior to disease onset and diagnosis.

Lastly, the study considers practical application in terms of identifying a cost-effective test. Based on previous research, a test which measured all the proteins in the initial training data would likely be too expensive to be clinically viable(152), but if sufficient performance can be obtained on a smaller set of protein biomarkers, this may be more feasible. Thus the most informative disease biomarkers are obtained from each model by identifying the most important proteins contributing to the model fit through variable selection methods. Models are then trained on these smaller subsets of proteins and their performance is assessed on each of the independent cohorts.

3.2 Methods

The full workflow for this chapter is summarized in **Figure 3.1**.



To assess practical applicability, the following stages were included. The most important biomarkers were identified for each of the 6 models. Each model was then trained on these subsets of proteins, and performance assessed on the independent cohorts.

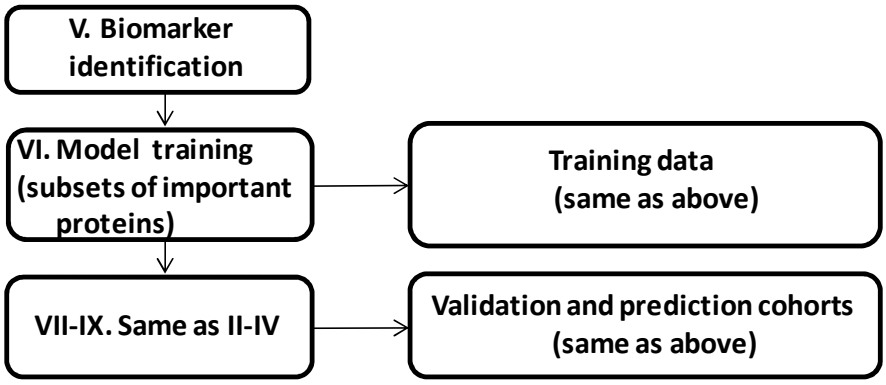


Figure 3.1. Workflow for Chapter 3.

3.2.1 Clinical samples

Serum samples were collected from 1165 individuals recruited from seven independent centres across Germany, Netherlands, Spain, United States and France as displayed in **Table 3.1**. The initial stage of the study is conducted on cohorts 1-5, consisting of healthy controls and either first or recent-onset antipsychotic naïve schizophrenia patients who were diagnosed as having the paranoid subtype of the disorder. Cohorts 1-5 came from three clinical centres in Germany and the Netherlands (cohort 1, Central Institute of Mental Health, Mannheim; cohorts 2-4, University of Magdeburg, Magdeburg; cohort 5, Erasmus University MC, Rotterdam).

The validation phase of the study is conducted on cohorts 6-9 from clinical centres in Spain, Germany and the Netherlands (cohort 6, University of Cantabria, Santander; cohort 7, Central Institute of Mental Health, Mannheim; cohort 8, Erasmus University MC, Rotterdam; cohort 9, University of Muenster, Muenster). Cohort 6 consists of healthy controls and first-onset antipsychotic naïve schizophrenia patients while cohorts 7-9 contains healthy controls and recent-onset schizophrenia patients treated with a mixture of first and second generation antipsychotic medications. For cohorts 1-9, DSM-4 diagnoses were performed by psychiatrists along with Positive and Negative Syndrome Scale (PANSS) testing. Patients and controls were excluded from this study if they met the following criteria: first degree relatives with a medical history of mental disease, diabetes, cardiovascular disease, immune and autoimmune disorders, infections, other neuropsychiatric/neurological disorders (multiple sclerosis, epilepsy, mental retardation), chronic (terminal) diseases affecting the brain (cancer, hepatic and renal insufficiency), alcohol or drug addiction, organic psychosis/organic affective syndromes, severe trauma, other psychiatric and non-psychiatric co-morbidity. The ethical committees at all involved centres approved the protocols of the study which were applied according to the Declaration of Helsinki. All participants provided written consent.

The prediction phase of the study is conducted on cohorts 10-11. Cohort 10 consists of samples selected from the US Department of Defence Serum Repository (DoDSR). Samples consist of pre-symptomatic individuals who subsequently presented with initial symptoms within 30 days of collection, and subsequently received a psychiatric diagnosis. They later received a diagnosis of either schizophrenia or bipolar disorder according to the DSM-4, several months or years later. In this study, they are referred to as ‘pre-schizophrenia’ or ‘pre-bipolar’ individuals. In addition, to complete cohort 10, healthy control samples were selected from the active duty military service population with no inpatient or outpatient psychiatric disorder diagnoses, as confirmed by current military records.

Cohort 11 consists of 77 prodromal individuals who were referred to the Adolescent and Young Adults Assessment Center (SHU, Paris, France) between 2009 and 2013 and enrolled in the ICAAR collaborative study. Individuals were included in this study if they displayed altered global functioning in the previous year, as defined by a score of less than 70 on the Social and Occupational Functioning Assessment Scale. Altered global functioning is associated with psychiatric symptoms and/or subjective cognitive complaints. The 77 individuals were evaluated by psychiatrists using the Comprehensive Assessment of At-Risk Mental State (CAARMS) threshold criteria(367). Of the 77 prodromal individuals assessed, 19 later developed psychosis after twelve months, and 58 did not. Psychosis conversion was defined through the CAARMS psychosis onset threshold, described as supra-threshold psychotic symptoms, meaning thought content and perceptual abnormalities, and/or disorganized speech.

As for cohorts 1–9, informed written consent was given by all the participants, and study protocols, collection and analysis of samples and all test methods were approved by the local Institutional Ethics Review Boards. As this cohort was specifically for prodromal individuals, individuals were excluded from the study if they already met DSM-4-defined criteria at the time of sample collection for psychosis, schizophrenia or schizo-affective disorders, pervasive developmental or bipolar disorders, as were individuals with other established diagnoses such as obsessive-compulsive disorders. Individuals were also excluded if they had currently been receiving antipsychotic treatment for more than 12 weeks, had psychoactive substance dependence or abuse during the previous year and/or more than 5 years, serious or evolutive somatic and neurological disorders or head injury and intelligence quotient <70.

Table 3.1 Patient and control demographics characteristics for the cohorts analyzed in Chapter 3. Values are presented as average ±standard deviation.

Stage	*Cohort*	*Number*	*Centre*	*Sex (M/F)*	*Age (years)*
Training	Cohort 1 (106)	52 CT 54 SCZ	Mannheim	27/25 32/22	30 ± 8 30 ± 10
	Cohort 2 (106)	73 CT 33 SCZ	Magdeburg	46/27 22/11	32 ± 9 31 ± 10
	Cohort 3 (39)	23 CT 16 SCZ	Magdeburg	10/13 8/8	33 ± 11 35 ± 11
	Cohort 4 (26)	16 CT 10 SCZ	Magdeburg	8/8 6/4	35 ± 11 37 ± 12
	Cohort 5 (54)	40 CT 14 SCZ	Rotterdam	33/7 11/3	26 ± 4 24 ± 6
Validation	Cohort 6 (135)	88 CT 47 SCZ	Santander	51/37 28/19	33 ± 8 30 ± 9
Application	Cohort 7 (91)	21 CT 70 SCZ	Mannheim	13/8 41/29	36 ± 9 38 ± 11
	Cohort 8 (73)	40 CT 33 SCZ	Rotterdam	33/7 25/8	27 ± 4 27 ± 8
	Cohort 9 (89)	33 CT 56 SCZ	Muenster	11/22 28/28	40 ± 13 38 ± 11
Prediction	Cohort 10 (369)	184 CT 75 pre-SCZ 110 pre-BD	US Military	136/48 67/8 70/40	22 ± 4 24 ± 5 21 ± 4
	Cohort 11 (77)	19 pre-SCZ 58 not pre-SCZ	Paris	11/7 33/25	20 ± 3 22 ± 4

3.2.2 Data pre-processing

The concentrations of 225 serum proteins are measured for all 1165 samples across cohorts 1-11 using the multiplex immunoassay platform described in **Chapter 2**. All serum proteins measured are run through a series of pre-processing steps. Proteins with more than 30% missed values are removed. Missing values arise either due to protein concentrations being below or above the limitation of detection, or due to low sample volume. To reflect the underlying biological reasoning behind missing values which were below/above the detection limit, they are replaced by the minimum or maximum protein concentration level in that particular cohort of samples. Missing values occurring due to low sample volume are replaced by the mean concentration for that protein in that cohort. All protein concentration values are log10-transformed to stabilize data variance and improve normality. In addition, quality control (QC) assessment is carried out for each cohort. Sample outliers are assessed using PCA through inspection of quantile-quantile plots. PCA is additionally used to check

for data artefacts. PCA is applied using the prcomp function in the R ggfortify package (302). Quantile-quantile plots are computed using the qqPlot function in the R car package.

Following the removal of proteins with more than 30% missing values from the initial immunoassay panel of 225, there are 66 proteins remaining which are measured across all 11 cohorts. A list of these 66 proteins and their abbreviations can be found in the **Appendix** (**Table A.3.1**). Batch effects arising due to technical variation are normalized using the ComBat algorithm(310) (**full details in Chapter 2**).

3.2.3 Statistical model training

Six statistical models, Logistic Regression, Ridge Regression, SVM, Random Forest, Bayesian LASSO and BART are trained on the metacohort formed from cohorts 1-5, consisting of 331 samples, 66 proteins and the demographic variables age and gender. The advantages and disadvantages of these methods are summarized in **Table 3.2**. Random Forest and SVM were chosen for this study as two of the most common machine learning techniques used in classification problems. Both have been previously applied to proteomic data(368). Studies have described one of the main advantages of Random Forest as its resilience to overfitting which makes it a good choice for a classifier on relatively small datasets such as those commonly studied in psychiatric research due to the difficulties of obtaining samples(369–372). SVM has the notable advantage of flexibility provided by the choice of kernel functions, enabling non-linear separations to be found between groups(373). Previous classification studies using proteomic data, found SVM to produce the most efficient classifier compared to other machine learning techniques such as artificial neural networks and partial least squares based methods(177). Bayesian LASSO and BART were chosen for this study due to the growing interest in the application of Bayesian methods to proteomic biomarker discovery and classification studies(187). The main advantage of Bayesian techniques compared to other approaches, is that they model the uncertainty inherent to the data which is thought to give them a better chance of validating on an independent cohort(374). Accounting for uncertainty is particularly applicable to schizophrenia datasets, given the high heterogeneity within the disease population. The Bayesian LASSO and BART models were chosen as two very different means of applying a Bayesian framework to proteomic data. Bayesian LASSO is a parametric test while BART is a non-parametric approach. The Bayesian LASSO has been applied to some sparse proteomic datasets in recent years(375–377) while BART has yet to be applied to proteomic data but has been used in gene expression studies(378).

Table 3.2 Advantages and disadvantages of statistical methods used in Chapter 3 as described in the literature(187,368,379,380)

Model	Advantages	Disadvantages
Logistic Regression	· Easily interpretable relationship between response and predictors through regression coefficients · Not computationally intensive	· Not so stable on data with a low EPV · Only assumes linear relationship between between response and predictors
Ridge Regression	· Easily interpretable relationship between response and predictors through regression coefficients · Penalizes unimportant predictors by shrinking coefficients, reducing overfitting · Not computationally intensive	· Only assumes linear relationship between between response and predictors
Random Forest	· Can handle large numbers of predictor variables · Can deal with missing data · Resilient to overfitting · Provides variable importance scores · Non-parametric test, so no assumptions about linearity	· More accurate models can require more trees which can increase run-time performance · Can't see relationship between response and predictors (e.g. through regression coefficients in logistic regression) which reduces interpretability
SVM	· Can find either linear/non-linear relationships between response and predictors through flexible choice of kernel functions · Regularisation parameter provides resilience to overfitting · Can handle large numbers of predictor variables	· Computationally intensive · Less interpretable than other models in terms of identifying important variables
Bayesian LASSO	· Works well with "small n, large p data" · Easily interpretable relationship between response and predictors through regression coefficients · Effectively eliminates non-predictive variables · Models data uncertainty through prior distribution	· Only assumes linear relationship between between response and predictors
BART	· Works well with "small n, large p data" · Automatically includes high order variable interactions · Effectively eliminates non-predictive variables · Models data uncertainty through prior distribution · Non-parametric test, so no assumptions about linearity · Provides variable importance scores	· Computationally intensive · Can't see relationship between response and predictors (e.g. through regression coefficients in logistic regression) which reduces interpretability

3.2.3.1 Model fitting

Each model is fitted using the R functions and packages defined in **Chapter 2**. The model parameters defined below are tuned through ten-fold cross-validation over a grid of values.

SVM: A non-linear SVM is used through the selection of a polynomial kernel function. The cost, scale and degree parameters of this kernel function are tuned.

Random Forest: The parameters mtry and number of trees.

BART: The hyper-parameters alpha, nu and k, and the parameter number of trees.

3.2.4 Evaluating classification performance on validation, application and prediction cohorts

The six statistical models trained in **3.2.3**, are first tested on the training meta-cohort to provide a comparison reference in terms of performance, for testing on the independent cohorts. They are subsequently tested on the validation cohort (Cohort 6 – antipsychotic naive schizophrenia patients and controls), and the three application cohorts (Cohorts 7-9 – antipsychotic treated schizophrenia patients and controls). Their performance is then tested on the two prediction cohorts, Cohort 10 (pre-schizophrenia individuals, pre-bipolar individuals, and controls) and Cohort 11 (prodromal individuals who later converted to psychosis and prodromal individuals who did not). In Cohort 10, the classification performance of each model was tested in three cases; pre-schizophrenia individuals against controls, pre-bipolar individuals against controls (to examine the disease specificity of the classification performance) and pre-schizophrenia individuals against pre-bipolar individuals. The latter comparison is thought to be of interest due to the high rates of misdiagnosis between the two disorders(139).

Classification performance is evaluated by plotting ROC curves (using the R functions and packages described in **Chapter 2**), computing AUC, sensitivity, and specificity. The AUC is thought to be an effective measure of the classification ability of a model, both in terms of evaluating its discriminatory ability between two groups, and to compare performance between models(381). Sensitivity and specificity values for the model's performance on a particular cohort were computed following identification of an optimal threshold point on the ROC curve via Youden's index (described in **Chapter 2**).

3.2.5 Biomarker identification

Following classification performance testing, the most informative biomarkers out of the 66 proteins are identified for each model using different variable selection methods. These biomarkers are the most important proteins which contributed to the model, maximising the discriminatory power between patients and controls. For the Bayesian LASSO model, the most informative biomarkers are defined as proteins whose 95% posterior credible interval does not contain zero(329). For BART, biomarker identification is done automatically by the model, using a variable selection algorithm explained in *Bleich et al* which utilizes the model 'variable inclusion proportions' that are computed as part of the tree-based structure(333). This is computed in R using the var_selection_by_permute_cv function in the bartMachine package. For Random Forest and SVM, biomarker identification is computed through recursive feature elimination (RFE) as in other studies, using Gini importance scores as the

ranking criterion for Random Forest and the square of the weights calculated by the model as the ranking criterion for SVM(382–384). RFE is applied in R using the rfeControl function in the caret package, with default settings(326). For Logistic Regression, biomarker identification is computed through stepwise selection which selects the most parsimonious set of proteins using the BIC (applied in R using the settings described in **Chapter 2**). To gain an indication of the most important biomarkers as determined through a penalized regression procedure, a LASSO regression model is used to conduct variable selection on the training data (applied in R using the settings described in **Chapter 2**).

3.2.6 Reduced model fitting

Following biomarker identification, the six methods are trained on each unique subset of biomarkers. Each method is trained on all biomarker sets to allow for comparability between methods on each set of proteins. Each model is trained on the biomarker set in question on the meta-cohort, and then tested on all of the validation, application and prediction cohorts. This obtains a spectrum of classification performances for each model, across all of the independent cohorts, for a series of different subsets of the 66 proteins. This makes it possible to determine whether similar or better performance could be achieved with a smaller, more optimal subset of protein predictors, which would be cheaper to incorporate into a clinical test.

3.3 Results

The study presented in this chapter includes a total of 1165 participants across 11 sample cohorts, comprising 331 in the training meta-cohort (cohorts 1-5), 135 in the validation cohort (cohort 6), 253 in the three application cohorts (cohorts 7-9), and 446 in the prediction cohorts (cohorts 10 and 11).

3.3.1 Data pre-processing

As **Figure 3.2 (a)** shows, there is initially some variation between the cohorts. These differences are likely to be a result of either technical variation in the initial handling and processing of samples, or variation from the cohorts being analysed on the multiplex immunoassay platform at different times. The ComBat algorithm is used to normalize this variation by borrowing information across the 66 proteins to calculate and then adjust for additive and multiplicative batch effects (as described in **Chapter 2**). The eleven cohorts with adjusted values for these 66 proteins following ComBat normalization are displayed in

Figure 3.2 (b). These adjusted protein values are then used for the subsequent analyses in this chapter.

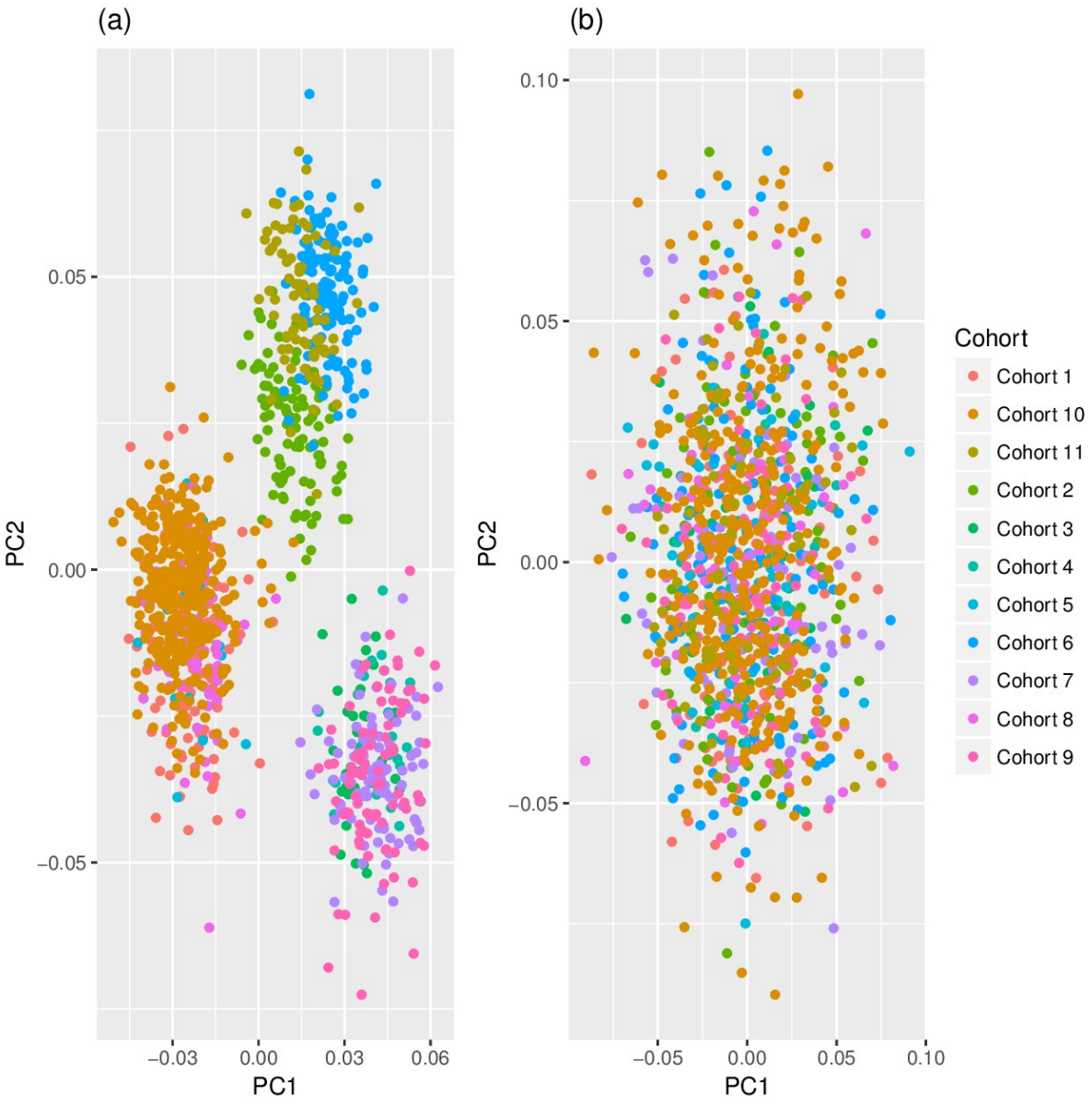


Figure 3.2 Principal component analysis (PCA) plots showing the 11 cohorts (a) before and (b) after ComBat normalisation to eliminate batch effects arising due to technical variation from the cohorts being run on different plates/days.

3.3.2 Evaluating classification performance on validation, application and prediction cohorts

3.3.2.1 Validation and application cohorts

The classification performance of the six models is tested on the training meta-cohort (Cohorts 1-5) to provide a reference in terms of AUC, sensitivity and specificity for the subsequent independent cohorts. The models are then tested on the independent validation cohort (Cohort 6 – antipsychotic naive schizophrenia patients and controls) and the three application cohorts (Cohorts 7-9 – antipsychotic treated schizophrenia patients and controls). Performance analysis across the training data, validation cohort and application cohorts is summarized in **Tables 3.3 and 3.4** while **Figure 3.3** shows ROC curves for all models on the validation and application cohorts. As described in **Chapter 2**, classification performance is typically categorized as "very good" (AUC > 0.9), "good" (AUC: 0.8-0.9), "fair" (AUC: 0.7-0.8), "poor" (0.6-0.7) and "fail" (AUC < 0.6).

SVM is the only model which produces at least a "good" performance as well as sensitivity and specificity of at least 0.8 on both the training meta-cohort (Cohorts 1-5) and the validation cohort (Cohort 6). SVM produces a "very good" performance on the meta-cohort (AUC: 0.97, sensitivity: 0.87, specificity: 0.96) and a "good" performance on cohort 6 (AUC: 0.88, sensitivity: 0.81, specificity: 0.81). In comparison, while the other five models all produce "very good" performances (AUCs: 0.96-1) and high sensitivities and specificities (sensitivities: 0.9-1, specificities: 0.9-1) on the meta-cohort, and "good" performances (AUCs: 0.81-0.88) on cohort 6, each model has either suboptimal sensitivity or specificity on cohort 6 (sensitivities: 0.74-0.96, specificities: 0.64-0.83). There is a decrease in classification performance between the meta-cohort and cohort 6 for the SVM model. This is expected on an independent test cohort, the model's performance still falls in the range of what could be deemed clinically useful.

However when the SVM model is tested on the three application cohorts, the performance is more variable. Classification performance is "poor" on Cohort 7 (AUC: 0.74, sensitivity: 0.76, specificity: 0.67) and the model fails on Cohort 9 (AUC: 0.56, sensitivity: 0.66, specificity: 0.52). While the model produces a "good" performance on Cohort 8 (AUC: 0.87, sensitivity: 0.76, specificity: 0.95), it is not sufficiently sensitive. These results indicate that the SVM model would not be clinically useful in diagnosing treated patients. The same trends across the three cohorts are seen for the other five models with "poor" performances on Cohort 7 (AUCs: 0.66-0.74, sensitivities: 0.54-0.76, specificities: 0.52-0.81), "good"-"very good" performances on Cohort 8 (AUCs: 0.8-0.96, sensitivities: 0.67-0.94, specificities: 0.75-0.98),

and failing on Cohort 9 (AUCs: 0.5-0.59, sensitivities: 0.34-0.84, specificities: 0.36-0.97). While models perform well on Cohort 8 in terms of AUC, they are typically not sensitive enough (<0.8) apart from the Random Forest model which is highly sensitive and specific on Cohort 8 (sensitivity: 0.94, specificity: 0.98). However this model does not perform as well on the other two application cohorts.

3.3.2.2 Prediction cohorts

The classification performances of the six models is subsequently tested on the prediction cohorts (Cohorts 10 and 11) to assess whether the models are able to predict schizophrenia before disease onset, based on the relative protein concentrations between first onset patients and controls. In Cohort 10, all models are used to classify between pre-schizophrenia individuals and healthy controls, pre-bipolar individuals and controls (to detect whether any classification ability seen between pre-schizophrenia individuals and controls was disease-specific), and between pre-schizophrenia individuals and pre-bipolar individuals (rates of misdiagnosis between these two psychiatric disorders is extremely high). In Cohort 11, all models are used to classify between prodromal individuals who developed psychosis within the following 12 months (converters) and prodromal individuals who did not develop psychosis within 12 months (non-converters). Classification performance across cohorts 10 and 11 is summarized in **Table 3.5**, while **Figure 3.4** shows ROC curves for all models for these cohorts.

As shown in **Table 3.5**, the performance of the SVM model (the only model which could classify patients from controls in the meta-cohort and cohort 6 with high AUC, sensitivity and specificity), on the prediction cohorts means that it would not be useful as a prognostic test.

Overall Random Forest is the only model which produces a “fair” classification performance (AUC: 0.7, sensitivity: 0.59, specificity: 0.73) on the comparison between pre-schizophrenia individuals and controls in Cohort 10. The other five models produce “poor” classification performances (AUCs: 0.61-0.68, sensitivities: 0.44-0.69, specificities: 0.57-0.79) with the sensitivity range indicating that models particularly struggle to correctly classify pre-schizophrenia individuals. All six models produce “poor” performances in the comparison between pre-schizophrenia individuals and pre-bipolar individuals (AUC: 0.6-0.65, sensitivities: 0.32-0.69, specificities: 0.52-0.91) with the sensitivity range again indicating that models struggle to accurately classify pre-schizophrenia individuals.

All six models fail on the classification comparisons in Cohort 10 between pre-bipolar individuals and controls (AUC: 0.5-0.55), and in Cohort 11 between psychosis converters

and non-converters (AUC: 0.38-0.43). While it may seem strange that the AUCs for Cohort 11 appear to be worse than random chance (AUC < 0.5), this is a mathematical quirk of ROC analysis which can occur if the relationships the models are using from the training data are not at all generalizable to that test cohort(337). Some research studies on ROC analysis have recommended inversing model classification decisions (i.e. AUC' = 1-AUC) for AUC < 0.5, but this would only be applicable in this case if there were no a priori information regarding whether a randomly chosen converter would be likely to ranked more highly by the models than a randomly-chosen non-converter(337).

Table 3.3 Classification performance of the six statistical models across the training meta-cohort (Cohorts 1-5) and the validation cohort (Cohort 6). Logistic; Logistic Regression. Ridge; Ridge Regression. RF; Random Forest. BL; Bayesian LASSO.

Performance Metric	Centre	Cohort	Logistic	Ridge	RF	SVM	BL	BART
AUC	Meta Cohort	Cohorts 1-5 (Training) - Antipsychotic naive SZ & Ctrls	0.98	0.96	1.00	0.97	0.97	0.99
	Santander	Cohort 6 (Validation) - Antipsychotic naive SZ & Ctrls	0.83	0.85	0.88	0.88	0.86	0.81
Sensitivity	Meta Cohort	Cohorts 1-5 (Training) - Antipsychotic naive SZ & Ctrls	0.94	0.90	1.00	0.87	0.91	0.99
	Santander	Cohort 6 (Validation) - Antipsychotic naive SZ & Ctrls	0.87	0.74	0.96	0.81	0.94	0.85
Specificity	Meta Cohort	Cohorts 1-5 (Training) - Antipsychotic naive SZ & Ctrls	0.93	0.90	1.00	0.96	0.93	0.99
	Santander	Cohort 6 (Validation) - Antipsychotic naive SZ & Ctrls	0.69	0.83	0.67	0.81	0.68	0.64

Table 3.4 Classification performance of the six statistical models across the three application cohorts (Cohorts7-9). Logistic; Logistic Regression. Ridge; Ridge Regression. RF; Random Forest. BL; Bayesian LASSO.

Performance Metric	Centre	Cohort	Logistic	Ridge	RF	SVM	BL	BART
AUC	Cologne	Cohort 7 (Application) - Antipsychotic treated SZ & Ctrls	0.69	0.67	0.69	0.74	0.72	0.66
	Rotterdam	Cohort 8 (Application) - Antipsychotic treated SZ & Ctrls	0.86	0.80	0.96	0.87	0.85	0.91
	Muenster	Cohort 9 (Application) - Antipsychotic treated SZ & Ctrls	0.57	0.50	0.58	0.56	0.55	0.59
Sensitivity	Cologne	Cohort 7 (Application) - Antipsychotic treated SZ & Ctrls	0.54	0.66	0.56	0.76	0.67	0.76
	Rotterdam	Cohort 8 (Application) - Antipsychotic treated SZ & Ctrls	0.67	0.76	0.94	0.76	0.70	0.79
	Muenster	Cohort 9 (Application) - Antipsychotic treated SZ & Ctrls	0.75	0.34	0.46	0.66	0.75	0.84
Specificity	Cologne	Cohort 7 (Application) - Antipsychotic treated SZ & Ctrls	0.81	0.76	0.76	0.67	0.81	0.52
	Rotterdam	Cohort 8 (Application) - Antipsychotic treated SZ & Ctrls	0.98	0.75	0.98	0.95	0.93	0.88
	Muenster	Cohort 9 (Application) - Antipsychotic treated SZ & Ctrls	0.48	0.97	0.79	0.52	0.45	0.36

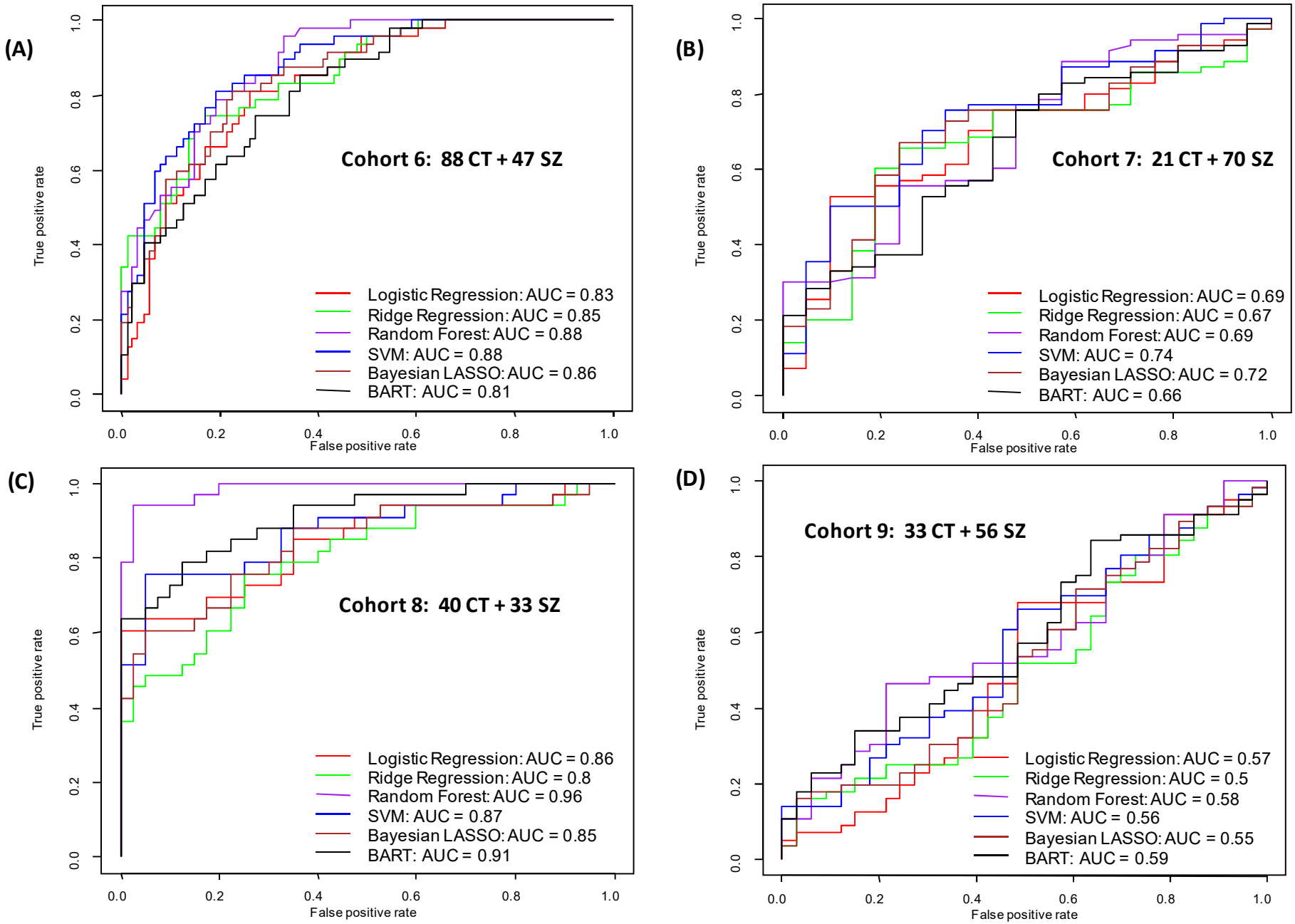


Figure 3.3. ROC curves showing classification performance for the six statistical models across the validation cohort; (A) Cohort 6 – antipsychotic naive schizophrenia (SZ) patients and controls (CT), and the 3 application cohorts; (B-D) Cohorts 7-9 - antipsychotic treated SZ patients & CT

Table 3.5 Classification performance of the six statistical models across the two prediction cohorts (Cohorts10-11). Logistic; Logistic Regression. Ridge; Ridge Regression. RF; Random Forest. BL; Bayesian LASSO.

Performance Metric	Centre	Cohort	Logistic	Ridge	RF	SVM	BL	BART
AUC	US Military	Cohort 10 (Prediction) - pre-SZ & Ctrls	0.66	0.65	0.70	0.61	0.67	0.68
	US Military	Cohort 10 (Prediction) - pre-BD & Ctrls	0.50	0.51	0.55	0.51	0.51	0.53
	US Military	Cohort 10 (Prediction) - pre-SZ & pre-BD	0.65	0.63	0.64	0.60	0.65	0.64
	Paris	Cohort 11 (Prediction) - psychosis converters & non-converters	0.41	0.41	0.42	0.41	0.43	0.38
Sensitivity	US Military	Cohort 10 (Prediction) - pre-SZ & Ctrls	0.69	0.44	0.59	0.51	0.64	0.53
	US Military	Cohort 10 (Prediction) - pre-BD & Ctrls	0.62	0.27	0.24	0.34	0.37	0.35
	US Military	Cohort 10 (Prediction) - pre-SZ & pre-BD	0.49	0.55	0.32	0.44	0.40	0.69
	Paris	Cohort 11 (Prediction) - psychosis converters & non-converters	0.95	1.00	1.00	1.00	0.68	0.53
Specificity	US Military	Cohort 10 (Prediction) - pre-SZ & Ctrls	0.57	0.79	0.73	0.70	0.63	0.78
	US Military	Cohort 10 (Prediction) - pre-BD & Ctrls	0.45	0.8	0.89	0.75	0.70	0.79
	US Military	Cohort 10 (Prediction) - pre-SZ & pre-BD	0.77	0.66	0.91	0.73	0.85	0.52
	Paris	Cohort 11 (Prediction) - psychosis converters & non-converters	0.21	0.03	0.03	0.02	0.36	0.48

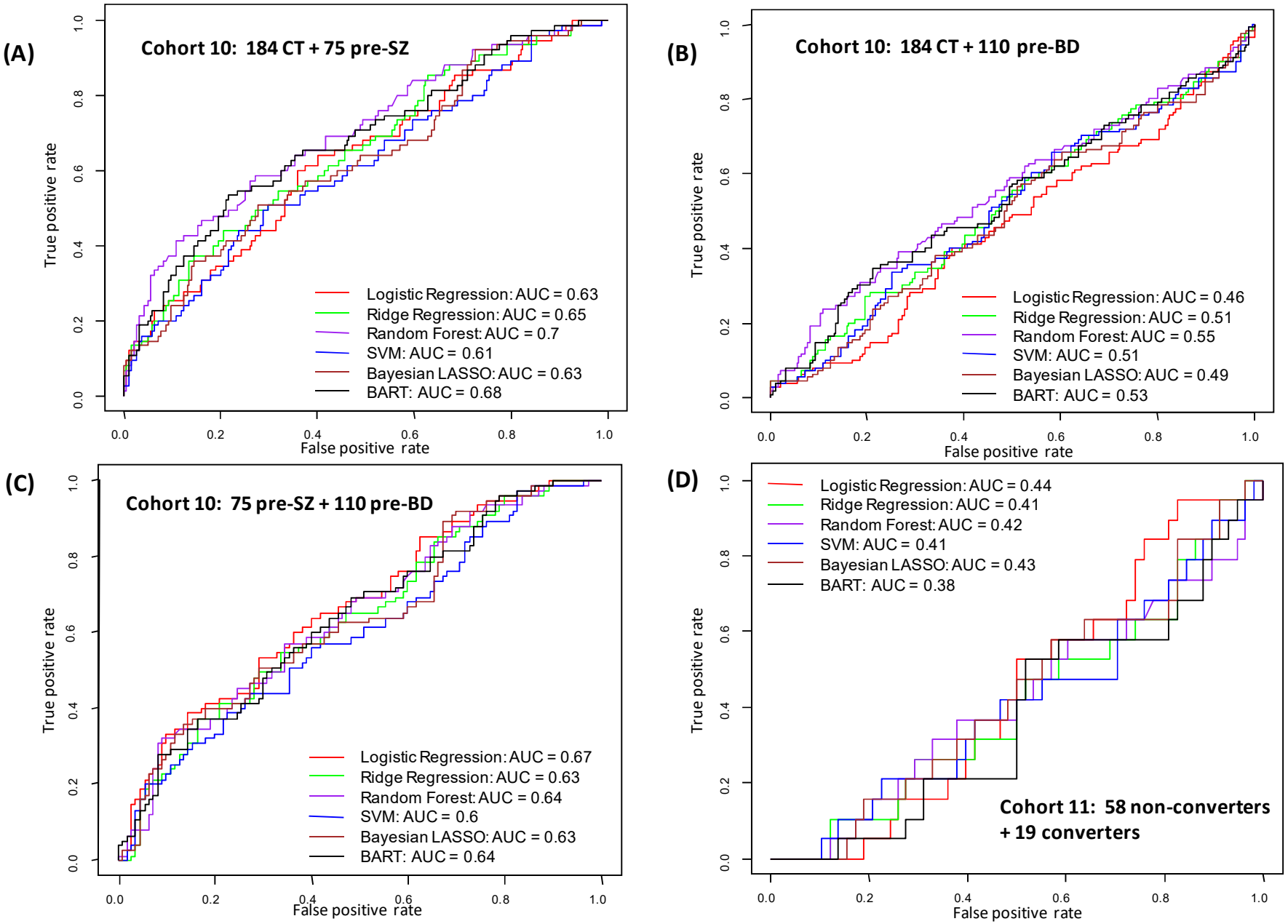


Figure 3.4. ROC curves showing predictive performance for the six statistical models across prediction cohorts 9 and 10; (A) Cohort 9 (US Military) – pre schizophrenia (SZ) individuals and controls (CT), (B) Cohort 9 (US Military) – pre bipolar disorder (BD) individuals and CT, (C) Cohort 9 (US Military) – pre SZ individuals and pre BD individuals, (D) Cohort 10 (Paris) – prodromal individuals who converted to psychosis and prodromal individuals who did not

3.3.3 Biomarker identification

The most informative biomarkers out of the 66 proteins in the training meta-cohort are identified using different variable selection methods for each model. Stepwise selection identifies 31 biomarkers and the variable gender for Logistic Regression. LASSO identifies 28 biomarkers, as an indication of the most important proteins as selected by a penalized regression procedure. RFE identifies 28 biomarkers for Random Forest and all 66 proteins are selected as biomarkers for SVM. The Bayesian LASSO model identifies 11 biomarkers, fewer than any other model. BART identifies 13 biomarkers. There are 7 biomarkers which are identified by all six methods. These proteins are Alpha-1 antitrypsin, Cortisol, Factor VII, Ferritin, Macrophage migration inhibitory factor, Pancreatic polypeptide, and Receptor for advanced glycosylation end products. **Table 3.6** lists the biomarkers identified by each method, highlighting the overlap.

3.3.4 Reduced model analysis

With 66 proteins, there are 2^{66} possible models for a given statistical method. As it would not be practical to examine the performance of all of these, this study uses different variable selection methods to identify sets of proteins which represent the most important biomarkers contributing to each model. In the case of Ridge Regression, LASSO is used for variable selection to obtain a set of biomarkers through a penalized regression method. As shown in **Table 3.6**, the six methods identified 5 unique sets of biomarkers (running RFE with SVM identified all 66 proteins as biomarkers) from the original 66 proteins. In addition, there are 7 common proteins to all these biomarker sets. Subsequently, models are trained using each statistical algorithm, on these 7 biomarkers, and then on each of the five unique sets of biomarkers identified by the models. These new models are then tested on each of the independent cohorts to examine whether similar or better classification performance could be achieved with a smaller, optimal set of proteins.

Figures 3.5-3.6 show the performances of each model in terms of AUC on the meta-cohort and the independent cohorts, for each set of biomarkers. The performance with all 66 proteins is included for comparison purposes. Plots for sensitivity and specificity are included in the **Appendices**. Cohorts 9 (application), 10 (prediction – pre bipolar vs control group), and 11 (prediction) are not included in these figures as all models “fail” when classifying these comparisons, for all sets of biomarkers.

Table 3.6 Biomarkers selected by the six variable selection methods. The 7 biomarkers selected by all models are highlighted in bold.

Biomarker Name	Stepwise Selection (Logistic Regression)	LASSO	RFE (Random Forest)	RFE (SVM)	Bayesian LASSO	BART
Age				✓		
Gender	✓			✓		
Adiponectin				✓		
Alpha-1 antitrypsin	✓	✓	✓	✓	✓	✓
Alpha-2 macroglobulin		✓	✓	✓		✓
Angiopoietin 2	✓		✓	✓		
Apolipoprotein A1	✓		✓	✓		
Apolipoprotein CIII		✓	✓	✓		✓
Apolipoprotein H	✓	✓	✓	✓		✓
AXL receptor tyrosine kinase		✓	✓	✓		✓
Beta-2 microglobulin	✓	✓	✓	✓		
Brain-derived neurotrophic factor				✓		
C reactive protein	✓	✓		✓	✓	
CD40 antigen				✓		
CD40 ligand				✓		
Chemokine CC4				✓		
Complement C3				✓		
Cortisol	✓	✓	✓	✓	✓	✓
Creatine Kinase MB			✓	✓		□
EN-RAGE	✓	✓	✓	✓		✓
Eotaxin				✓		
Epidermal growth factor		✓	✓	✓		
Epithelial derived neutrophil activating protein 78				✓		
Factor VII	✓	✓	✓	✓	✓	✓
FASLG receptor				✓		
Ferritin	✓	✓	✓	✓	✓	✓
Follicle stimulating hormone	✓	✓	✓	✓	✓	
Haptoglobin	✓	✓	✓	✓	✓	
Hepatocyte growth factor				✓		
Immunoglobulin A	✓			✓		
Immunoglobulin M				✓		
Insulin-like growth factor binding protein 2		✓	✓	✓		
Intercellular adhesion molecule 1				✓		
Interleukin-16	✓			✓		
Interleukin-18				✓		
Leptin	✓	✓	✓	✓		✓
Lipoprotein (a)				✓		

Macrophage derived chemokine				✓		
Macrophage inflammatory protein 1 beta		✓		✓		
Macrophage migration inhibitory factor	✓	✓	✓	✓	✓	✓
Matrix metalloproteinase 3	✓			✓		
Monocyte chemotactic protein 1				✓		
Myeloperoxidase	✓			✓		
Myoglobin			✓	✓		
Pancreatic polypeptide	✓	✓	✓	✓	✓	✓
Plasminogen activator inhibitor 1	✓	✓		✓		
Platelet derived growth factor				✓		
Progesterone	✓	✓	✓	✓		
Prolactin		✓		✓		
Pulmonary and activation regulated chemokine	✓			✓		
Receptor for advanced glycosylation end products	✓	✓	✓	✓	✓	✓
Resistin				✓		
Serum amyloid P component				✓		
Sex hormone binding globulin				✓		
Sortilin				✓		
Stem cell factor		✓		✓		
Superoxide dismutase			✓	✓		
RANTES	✓	✓		✓		
Tenascin C	✓	✓	✓	✓		
Testosterone			✓	✓		
Thrombospondin 1	✓			✓		
Thyroid stimulating hormone			✓	✓		
Thyroxine binding globulin				✓		
Tissue inhibitor of metalloproteinases 1	✓			✓		
Tumor necrosis factor receptor like 2	✓			✓		
Vascular cell adhesion molecule 1	✓			✓		
Vascular endothelial growth factor	✓	✓		✓		
von Willebrand factor	✓	✓	✓	✓	✓	

One of the main findings of this analysis is that the SVM model based on 66 proteins, identified in **3.3.2.1**, remains the only method which achieves at least a “good” classification performance and sensitivity and specificity of at least 0.8 on both the training meta-cohort and the validation cohort (cohort 6). Thus this is the only model which has sufficient performance for clinical use as a diagnostic test.

While all 6 statistical methods fit models on 13 proteins which produce “good”-“very good” performances on the meta-cohort (AUCs: 0.92-1, sensitivities: 0.8-1, specificities: 0.77-1) and cohort 6 (AUCs: 0.8-0.84, sensitivities: 0.79-0.91, specificities: 0.64-0.72), in terms of AUC, and 13 proteins would represent a far cheaper option for a diagnostic test than 66 proteins, in each case either sensitivity or in particular, specificity, is suboptimal. A similar trend is seen when fitting models on both sets of 28 proteins.

As observed with 66 proteins in **3.3.2.1**, classification performance varies substantially between the application cohorts. While all models fail on Cohort 9, performances range from “poor”-“fair” on Cohort 7, and “poor”-“very good” on Cohort 8. The most consistent models are the SVM model with 11 proteins which produces both a “fair” performance (AUC: 0.7, sensitivity: 0.7, specificity: 0.71) on Cohort 7 and a “good” performance on Cohort 8 (AUC: 0.81, sensitivity: 0.76, specificity: 0.8), and the Random Forest model with 31 proteins which produces a “fair” performance (AUC: 0.7, sensitivity: 0.84, specificity: 0.52) on Cohort 7 and a “very good” performance on Cohort 8 (AUC: 0.97, sensitivity: 0.97, specificity: 0.93) but while this latter performance is interesting, and shows it is possible to achieve high performance on an independent cohort of treated patients and controls, further research is needed to understand exactly why biomarker performance is so variable on other cohorts of treated patients.

The performance ranges on the prediction cohort comparisons from Cohort 10 shown in **Figures 3.5-3.6** are almost all in the “poor” category, as was observed in **3.3.2.2**. Along with the results from Cohort 11 (not included in these figures as all models fail) this appears to provide conclusive evidence that it is not possible to develop a prognostic test based on the protein concentrations examined in this study.

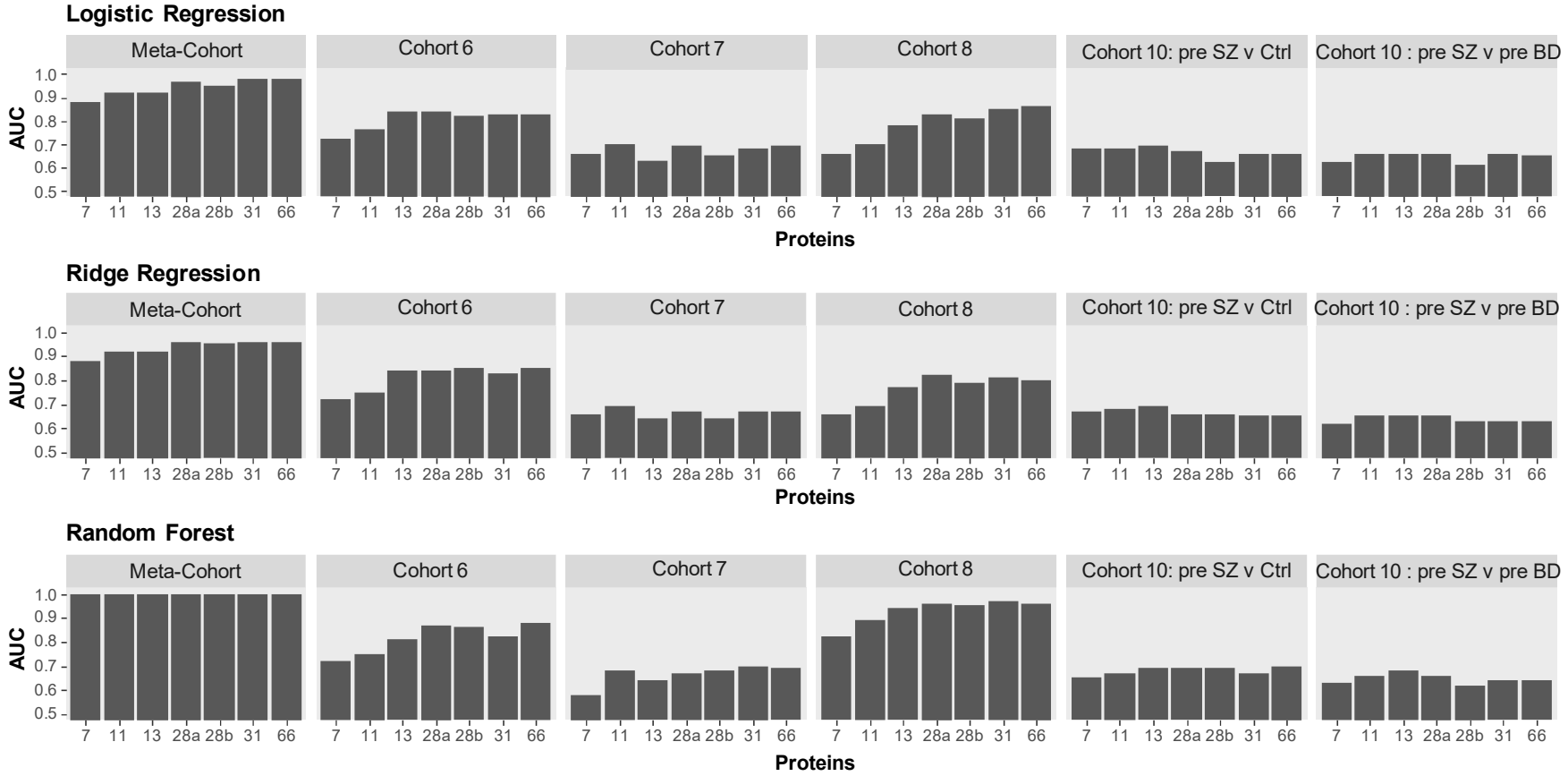


Figure 3.5 Classification performances for Logistic Regression, Ridge Regression and Random Forest across the training data (Meta-Cohort), and the independent validation (Cohort 6), application (Cohorts 7 & 8) and prediction (Cohort 10) cohorts. Performance was measured across models fitted on all unique protein biomarker sets identified in Table 3.8, the 7 biomarkers identified by all six models and the full set of 66 proteins for comparison. 28a and 28b refers to the set of 28 biomarkers identified by LASSO (28a) and the 28 biomarkers identified by RF-RFE (28b).

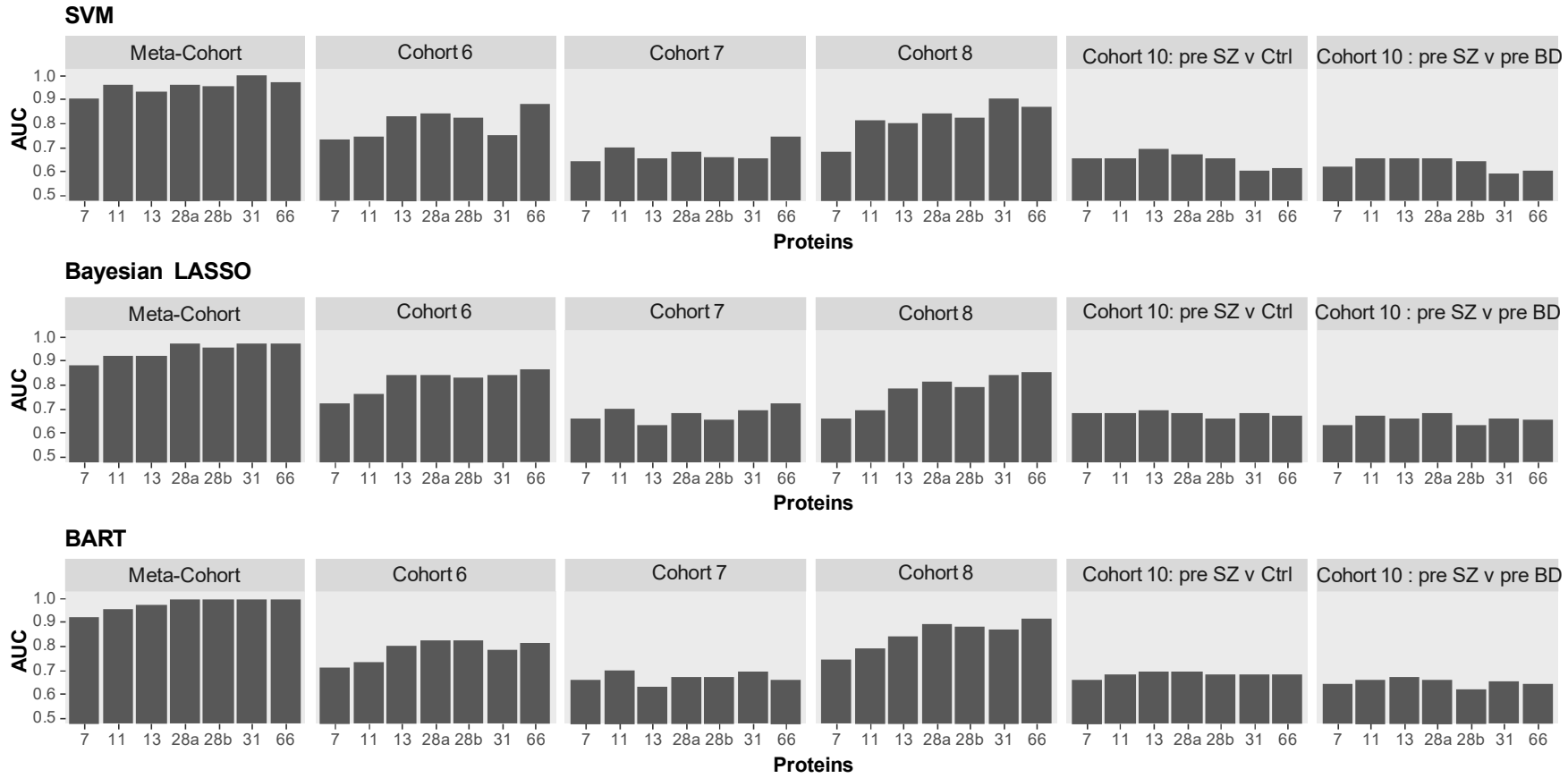


Figure 3.6 Classification performances for SVM, Bayesian LASSO, and BART across the training data (Meta-Cohort), and the independent validation (Cohort 6), application (Cohorts 7 & 8) and prediction (Cohort 10) cohorts. Performance was measured across models fitted on all unique protein biomarker sets identified in Table 3.8, the 7 biomarkers identified by all six models and the full set of 66 proteins for comparison. 28a and 28b refers to the set of 28 biomarkers identified by LASSO (28a) and the 28 biomarkers identified by RF-RFE (28b).

3.4 Discussion

In summary, this chapter presents the application of multiple statistical techniques towards the problem of identifying models which can classify schizophrenia through molecular profiling of serum. In this chapter, this problem is examined through profiling serum protein concentrations using multiplex immunoassay technology. Six statistical models were trained on a meta-cohort of antipsychotic naive schizophrenia patients and controls, tested on an independent cohort of antipsychotic naive patients and controls, and applied to three independent cohorts of antipsychotic treated patients and controls. The models were then tested as to whether they could predict early indications of the disease in cohorts of pre-symptomatic or prodromal individuals.

To our knowledge, the work presented in this chapter is one of the largest studies to date which has looked at the problem of classifying schizophrenia through serum proteomic profiling. The study design has a number of advantages, including the use of a large meta-cohort of antipsychotic naive schizophrenia patients and controls for training the models. Combining five independent cohorts from different clinical centres in this way helps to reduce the overfitting problem, and ensures that the model fits are based on a far larger disease population than customary in psychiatric studies. Recruiting first-onset antipsychotic naive patients for such studies is extremely difficult as even large psychiatric centres can only recruit 20-30 of these patients each year, and few centres follow strict standard operating procedures for sample collection(12).

3.4.1 Classification performance

The main finding of this study is that the SVM model based on the serum concentrations of 66 proteins can classify antipsychotic naive schizophrenia patients from healthy controls on both the training dataset and an independent cohort with an AUC of 0.88-0.97, and sensitivity and specificity of 0.81-0.87 and 0.81-0.96. This is an important result as previous proteomic studies looking at classifying schizophrenia through the levels of serum proteins have either not included independent validation cohorts for testing or applied inaccurate statistical methods for doing so(12,166,355,385). In addition, this finding matches up well against alternative approaches for classifying antipsychotic naive patients from controls which have been investigated in recent years. Studies attempting to build SVM classifiers using MRI data have reported a sizeable variation in classification performance ranging from “fair” to “excellent” (AUC: 0.7-0.92)(134). In particular, classification performance has been lower in larger MRI studies, for example *Nieuwenhuis et al*(386), where an SVM classifier

was trained on a cohort of 111 controls and 128 patients, and tested on an independent cohort of 122 controls and 155 patients, yielding an AUC of 0.7.

However, this model is unlikely to be suitable as a clinical diagnostic tool as it requires the measurement of 66 proteins. As stated in **3.1**, a previous diagnostic test for schizophrenia was withdrawn as it required the measurement of 51 proteins which was found to be too expensive(152). The subsequent analysis of models trained on smaller sets of proteins, consisting of the most informative biomarkers among the 66 proteins for differentiating patients from controls, found that these models were not sufficiently sensitive or specific enough.

In addition, the variability of the classification performances obtained through applying the different models examined in this study to multiple independent cohorts of recent onset, antipsychotic treated schizophrenia patients illustrate some of the practical challenges involved in attempting to develop a serum diagnostic test for schizophrenia which would be applicable for medicated patients as well as antipsychotic naive patients. Classification performances for the SVM model based on 66 proteins, for example, range from AUC: 0.56-0.87 across the three cohorts. This variability in performance has been observed in previous studies, with peripheral microRNA biomarkers achieving similarly variable classification performance (AUC: 0.69-0.85), when tested on a cohort of controls and recent onset patients treated with a mixture of first and second generation antipsychotics(154). There are several possible reasons for this variability, firstly the fact that antipsychotic medication is known to have a confounding effect on the levels of circulating proteins which can impact on classification accuracy(33). Hence some antipsychotics may be strongly modulating the levels of the proteins measured in this study, especially key predictors involved in driving the differences between antipsychotic naïve patients and controls. Unfortunately it is not possible to draw firmer conclusions as to the exact role of antipsychotics in this study's results as exact treatment information was not available for all patients in these cohorts. As such, we cannot discount the role of other factors in driving this variation in performance, including duration of illness, and the heterogeneity which exists between different populations of schizophrenia patients.

It is notable that none of the models examined in this study achieved sufficient classification performance on the prediction cohort comparisons for use as a prognostic test. The majority of the models produced "poor" performances in classifying pre-symptomatic individuals who later developed schizophrenia from healthy controls, and the sensitivities of these models were typically very low, thus demonstrating little ability to correctly classify the disease group. Classification performances were worse on the prodromal cohort, with all models

failing when attempting to classify psychosis converters from non-converters, even though these individuals were already displaying initial symptoms when samples were collected. These results are considerably different to those found by *Chan et al*(12), where the classification performances of a panel of 22 serum proteins measured through multiplex immunoassay were reported as ranging from AUC: 0.82-0.9 on different prediction cohorts of pre-onset individuals. However, that study is thought to be inaccurate as the biomarker coefficients were optimized for each test dataset and therefore did not reflect the true performance of the panel. The poor performances of the models in this study could be because they represent differences in the concentrations levels of these proteins between patients and controls at a timepoint when the patients were fully symptomatic, and these differences are more subtle at an earlier stage of the disease. In addition, it may be that different proteins are expressed strongly both early in the development of the disorder and after disease onset. One of the limitations of this study is that only 66 serum protein concentrations are measured in all eleven of the cohorts analyzed. Further limitations are discussed in **Chapter 9**.

3.4.2 Further work

While multiplex immunoassay platforms are one of the most conventional proteomic profiling technologies used in classification and biomarker discovery studies, the proteins measured, and thus the pathways examined, tend to be biased towards those already available in commercial assays been designed with a broad range of diseases in mind(283). The 66 proteins measured in this study are involved in pathways including inflammation, immune system function, lipid transport, hormonal and growth factor signalling, which have also been investigated in biomarker discovery studies in cancer(281), cardiovascular disease(359), and arthritis(358).

While this study has been necessary to explore the diagnostic and prognostic potential of these proteins in schizophrenia using robust statistical methods, it may be that a more targeted proteomic analysis (for example using MRM technology to measure proteins already known to be implicated in schizophrenia pathophysiology through existing studies) is required to detect proteins substantially expressed in both the prodromal phase and after disease onset. This approach is examined in **Chapter 4**.

Chapter 4 Comparison of statistical models for the classification and prediction of schizophrenia through targeted mass spectrometry profiling of serum

4.1 Introduction

Multiplex immunoassays are one of the most conventional proteomic profiling technologies currently used for biomarker identification. However, one of their disadvantages is that the only proteins measured tend to be those commonly available on commercial panels, which biases studies towards certain protein classes eg: proinflammatory cytokines(283,387). A multiplex immunoassay platform was used to profile the serum concentrations of 66 proteins across a training meta-cohort and independent validation, application and prediction cohorts in the study presented in **Chapter 3**. While this study identified an SVM model based on the concentrations of these 66 proteins, which produced a “good” (AUC > 0.8) classification performance as well as high sensitivity and specificity on an independent validation cohort, measuring this many proteins would not be viable for a clinical test. Subsequent testing of models based on 7-31 proteins found that they were either not sensitive or specific enough to be considered for clinical use. In addition, the SVM model based on 66 proteins “failed” (AUC < 0.6) at classifying psychosis converters from non-converters in one of the prediction cohorts. In the other prediction cohort, it only produced a “poor” (AUC: 0.6-0.7) performance when classifying pre-symptomatic individuals who later developed schizophrenia, from controls. This prediction ‘failure’ indicates that there is no relationship between the differential concentrations of those particular proteins between first-onset patients and controls, and psychosis converters and non converters. As such, different circulating proteins may be more important mechanistically, both in the early stages of disease progression, and following disease onset.

Following on from **Chapter 3**, the study in this chapter aims to identify a statistical model which can reproducibly classify schizophrenia from healthy controls, and predict disease conversion in pre-onset individuals with an AUC, sensitivity and specificity of at least 80%, through a different proteomic profiling approach. MRM is used to measure the abundances of a pre-selected panel of 147 peptides, corresponding to 77 proteins previously identified as

being associated with psychiatric disorders. This panel contains many peptides associated with proteins not typically measured in multiplex immunoassays, including members of the complement cascade and various neurotrophic factors(283). In addition, MRM is regarded as more reliable, yielding fewer false positive and false negative detections as it does not have the batch to batch antibody variation which is a drawback of multiplex immunoassays(284,292).

This is the first study to apply an MRM-based proteomic approach to the problem of identifying a diagnostic and prognostic biomarker test for schizophrenia. Through the targeted profiling of these serum analytes, this study aims to identify statistical models with an improved performance to those assessed in **Chapter 3**. Once again, an independent validation cohort of first-onset antipsychotic naive patients and controls is included in the study design, in addition to an application cohort of recent-onset antipsychotic treated patients and controls, and two prediction cohorts, one consisting of pre-symptomatic individuals and another consisting of prodromal individuals. As in **Chapter 3**, this study trains multiple statistical methods on a dataset of first-onset antipsychotic naive patients and controls to see which algorithm most accurately represents the underlying trends in the data. In this chapter, an additional statistical approach is considered which incorporates prior-based information. Bayesian approaches were used purely with non-informative priors in **Chapter 3**, but in this section BART is used with non-informative and informative priors. The benefits of informative priors in providing a more stable and accurate model by incorporating external information about the parameters have been previously discussed in the proteomic literature(187). Previous proteomic studies using Bayesian Belief Networks to classify patients with different subtypes of T-cell Leukemia Virus type 1 found that using a Bayesian model with informative priors outperformed other statistical approaches in terms of prediction accuracy(388). Lastly, as in **Chapter 3**, this study considers the trade-off between classification performance and cost which is a crucial practical consideration for all biomarker tests. The most informative disease biomarkers are identified from each model by identifying the peptides which contribute most to the classification performance. Models are then subsequently trained on these smaller subsets of peptides to see whether they produce similar performances on the independent cohorts to those trained on the original training data.

Identifying a model which can accurately classify prodromal converters from non-converters would have great clinical relevance as previous research studies have established a strong association between early therapeutic intervention and improved patient outcomes in the treatment of psychosis(163). While progress has been made in establishing clinical criteria for individuals with ultra-high-risk syndrome, there is currently no tool to aid psychiatrists

detect the 20-35% who will experience a psychotic episode over two to three years(14). As described in **Chapter 1**, the only two previous studies which have attempted to identify such a classification model have both had various limitations. *Perkins et al.* constructed a classifier to identify analytes in plasma which could differentiate converters from non-converters, but the study lacked an independent dataset in which to validate the performance(190). *Chan et al*. identified a panel of 22 serum proteins which could predict prodromal conversion in an independent dataset with an AUC of 0.9 when combined with clinical scores, but this result is thought to be unreliable as the coefficients of these proteins were optimized for the test dataset and therefore did not reflect the real predictive performance of this panel(12). This is one of the statistical shortcomings highlighted **Chapter 1** as contributing to the failure of many proteomic biomarkers to validate.

4.2 Methods

R packages, functions and settings were applied as in **Chapter 3**, unless otherwise stated.

4.2.1 Clinical samples

Serum samples were collected from 639 individuals recruited from five clinical centres across Germany, The Netherlands, United States and France as displayed in **Table 4.1**.

The training phase of the study is conducted on cohort 1, consisting of healthy controls and first-onset antipsychotic naïve schizophrenia patients from the Department of Psychiatry, University of Cologne. The validation phase of the study is conducted on cohort 2, consisting of healthy controls and first-onset antipsychotic naïve schizophrenia patients from Erasmus University Medical Centre, Rotterdam.

The application phase of the study is conducted on cohort 3, consisting of healthy controls and recent-onset schizophrenia patients with an average disease duration of approximately one year, treated with a mixture of first and second generation antipsychotic medications from the Department of Psychiatry, University Medical Centre Utrecht.

The prediction phase of the study is conducted on cohorts 4 and 5. Cohort 4 is the the same as Cohort 10, described in **3.2**. For Cohort 5, psychosis converters and non-converters were determined through the same criteria described for Cohort 11 in **3.2**.

Table 4.1 Patient and control demographics characteristics for the cohorts analyzed in Chapter 4. Values are presented as average ±standard deviation.

Cohort	*Number*	*Centre*	*Sex (M/F)*	*Age (years)*	*PANSS positive*	*PANSS negative*	*PANSS general*	*PANSS total*
Cohort 1 (139)	79 CT 60 SCZ	Cologne	43/36 31/29	31 ± 8 31 ± 10	NA 23 ± 6	NA 23 ± 7	NA 49 ± 10	NA 96 ± 20
Cohort 2 (21)	12 CT 9 SCZ	Rotterdam	12/0 9/0	27 ± 7 28 ± 6	NA 20 ± 5	NA 22 ± 7	NA 39 ± 6	NA 80 ± 15
Cohort 3 (32)	17 CT 15 SCZ	Utrecht	14/3 12/3	26 ± 4 25 ± 4	NA 11 ± 3	NA 14 ± 3	NA 27 ± 6	NA 53 ± 9
Cohort 4 (369)	184 CT 75 pre-SCZ 110 pre-BD	USA DoDSR	136/48 67/8 70/40	22 ± 4 24 ± 5 21 ± 4	NR NR NR	NR NR NR	NR NR NR	NR NR NR
Cohort 5 (78)	24 pre-SCZ 54 not pre-SCZ	Paris	15/9 28/26	20 ± 3 22 ± 4	16 ± 6 12 ± 4	18 ± 6 16 ± 7	42 ± 10 39 ± 10	75 ± 19 67 ± 18

4.2.2 Data pre-processing

The abundances of up to 147 peptides are measured for all 639 samples across cohorts 1-5 using the MRM platform described in **Chapter 2**. All peptide abundances measured are run through a series of pre-processing steps based on an optimal MRM analysis methodology outlined in *Ozcan et al*(283). The abundances of each peptide and its underlying transitions can be affected by various forms of non-biological variation, originating from the experimental settings. Therefore quality control steps were conducted to detect inaccurate transitions, so only the most reliable transitions were taken forward for analysis. Based on the recommended procedure as defined in *Ozcan et al*., to ensure reproducibility, the most abundant peptide-transitions, with a minimum of 80% consistency between endogenous and isotopically-labelled peptides, across mass spectrometry runs are selected(283). For the most abundant peptide-transitions with less than 80% consistency, peptides are visually inspected for interference from the biological matrix, and the most abundant transition was manually chosen based on quality control samples. Subsequently, the relative abundance ratio is computed for each peptide between endogenous and isotopically-labelled peptides. All peptide abundances are log transformed. In addition, sample outliers are assessed using PCA through inspection of quantile-quantile plots. PCA is additionally used to check for data artefacts. Two control samples are detected as outliers in cohort 1, and removed. Batch effects between each of the five sample cohorts, are normalized using the ComBat algorithm.

4.2.3 Exploratory analysis

To improve the EPV (defined in **Chapter 1**) and reduce overfitting, an exploratory univariate analysis is conducted on Cohort 1 using a linear regression model (applied in R as described in **Chapter 2**) to identify significantly ($p < 0.05$) changed peptides between first-onset antipsychotic naive schizophrenia patients and controls. The significantly ($p < 0.05$) altered peptides identified in this analysis are used to stratify the training data. Thus subsequently each statistical model is only trained on this reduced set of peptides, improving the EPV.

4.2.4 Statistical model training

Six statistical models, Logistic Regression, Ridge Regression, SVM, Random Forest, Bayesian LASSO and BART, are trained on 21 peptides measured across 137 samples (77 controls and 60 first-onset patients) and the demographic variable age, in cohort 1. In this instance, BART is applied with an uninformative prior. Model parameter tuning is carried out using ten-fold cross validation over a grid of values as described in **Chapter 3**.

4.2.4.1 Informative prior

To explore the potential of an informative prior for providing a more stable and accurate model, BART is subsequently trained on cohort 1 with an informative prior. Prior information is set as follows:

A BART model is trained on cohort 2 with uniform weights (the default setting) to obtain variable inclusion proportions for each of the 21 peptides. These proportions are used as weights determining how often each peptide should be proposed when training a BART model on cohort 1, so the splits proposed in tree growing and training for the model took into account which peptides are important in an independent dataset of antipsychotic naive patients and controls.

4.2.5 Evaluating classification performance on validation, application and prediction cohorts

The statistical models trained in **4.2.4**, are subsequently initially tested on the training data (Cohort 1) purely to provide a comparison reference in terms of performance, for testing on the independent cohorts. They are subsequently tested on the validation cohort (Cohort 2 – antipsychotic naive schizophrenia patients and controls), although for the BART model with an informative prior, it should be noted that this model's performance will be biased, because prior information was obtained from this cohort. The models are then tested on the

application cohort (Cohort 3 – antipsychotic treated schizophrenia patients and controls), and the prediction cohorts (Cohort 4 – pre schizophrenia individuals, pre bipolar individuals, and controls, and Cohort 5 - prodromal individuals who converted to psychosis after one year, and prodromal individuals who did not). In Cohort 4, the classification performance of each model is tested in three cases; pre-schizophrenia individuals against controls, pre-bipolar individuals against controls and pre-schizophrenia individuals against pre-bipolar individuals. Classification performance across the validation, application and prediction cohorts is again evaluated through plotting ROC curves and computing the associated metrics outlined in **Chapter 3**.

4.2.6 Biomarker identification

Following classification performance testing on the independent validation, application and prediction cohorts, the most important biomarkers are identified for each model. These biomarkers are the peptides which contribute most to the model performance, maximising the discriminatory power between schizophrenia patients and controls. Biomarker identification is conducted as described in **Chapter 3**.

4.2.7 Reduced model analysis

Following biomarker identification, the seven methods are trained on each unique subset of biomarkers. Each method is trained on all biomarker sets to allow for comparability between methods on each set of peptides. Each model is trained on the biomarker set in question on cohort 1, and then tested on all of the validation, application and prediction cohorts. This obtains a spectrum of classification performances for each model, across all of the independent cohorts, for a series of different subsets of the 21 peptides. This makes it possible to determine whether similar or better performance could be achieved with a smaller, more optimal subset of peptide predictors, which would be cheaper to incorporate into a clinical test.

4.3 Results

The study presented in this chapter includes a total of 639 participants across 5 sample cohorts, comprising 137 in the training data (cohort 1), 21 in the validation cohort (cohort 2), 32 in the application cohort (cohort 3), and 447 in the pre-onset prediction cohorts (cohorts 4 and 5). The ComBat algorithm is used to normalize batch effects between the five cohorts, and the adjusted peptide abundances are used for the subsequent analyses in this chapter.

4.3.1 Exploratory analysis

Exploratory analysis conducted on cohort 1 using a linear model, identifies 21 significantly ($p < 0.05$) changed peptides (corresponding to 16 unique proteins) between first-onset patients and controls, shown in **Table 4.2**. These peptides are used to stratify the training data to reduce the propensity for overfitting.

Table 4.2 Significantly ($p < 0.05$) altered peptide abundances between first-onset antipsychotic naive schizophrenia patients and healthy controls in cohort 1. These 21 peptides correspond to 16 unique proteins.

Protein	Abbreviation	Peptide Sequence	P Value
Vitronectin	VTNC	DWHGVPGQVDAAMAGR	0.0004
Haptoglobin	HPT	DYAEVGR	0.0042
Haptoglobin	HPT	VTSIQDWVQK	0.0047
Hemoglobin subunit alpha	HBA	MFLSFPTTK	0.0058
Antithrombin-III	ANT3	LPGIVAEGR	0.0070
Apolipoprotein H	APOH	VSFFCK	0.0073
Antithrombin-III	ANT3	FDTISEK	0.0089
α-1-antichymotrypsin	AACT	EQLSLLDR	0.0140
Apolipoprotein C-III	APOC3	DALSSVQESQVAQQAR	0.0160
Apolipoprotein C-I	APOC1	EFGNTLEDK	0.0190
Inter-α-trypsin inhibitor heavy chain H4	ITIH4	GPDVLTATVSGK	0.0216
Complement C4-A	CO4A	VLSLAQEQVGGSPEK	0.0220
α-2-antiplasmin	A2AP	FDPSLTQR	0.0245
Ficolin-3	FCN3	YGIDWASGR	0.0257
Apolipoprotein A-IV	APOA4	ALVQQMEQLR	0.0269
Complement component C9	CO9	VVEESELAR	0.0301
Apolipoprotein C-III	APOC3	GWVTDGFSSLK	0.0329
Apolipoprotein A-IV	APOA4	ISASAEELR	0.0387
Complement C4-A	CO4A	ITQVLHFTK	0.0410
Retinol-binding protein 4	RET4	YWGVASFLQK	0.0427
Complement C2	CO2	HAIILLTDGK	0.0457

4.3.2 Evaluating classification performance on validation, application and prediction cohorts

4.3.2.1 Validation and application cohorts

The classification performance of the seven models is tested on the training data (Cohort 1) to provide a reference in terms of AUC, sensitivity and specificity for the subsequent independent cohorts. The models are then tested on the independent validation cohort (Cohort 2 – antipsychotic naive schizophrenia patients and controls) and application cohort (Cohort 3 – antipsychotic treated schizophrenia patients and controls). Performance analysis across cohorts 1-3 is summarized in **Table 4.3** while **Figure 4.1** shows ROC curves for all models on cohorts 2 and 3. As described in **Chapter 3**, classification performance is typically categorized as "very good" (AUC > 0.9), "good" (AUC: 0.8-0.9), "fair" (AUC: 0.7-0.8), "poor" (0.6-0.7) and "fail" (AUC < 0.6).

SVM is the only model which produces at least a "good" performance plus sensitivity and specificity of at least 0.8, on both the training (cohort 1) and validation (cohort 2) cohorts. SVM produces a "very good" performance on cohort 1 (AUC – 0.92, sensitivity – 0.85, specificity – 0.86), and a "good" performance on cohort 2 (AUC – 0.87, sensitivity – 0.89, specificity – 0.92).

In comparison, while Random Forest and BART produce only "fair" performances on cohort 2 (AUC – 0.79), Bayesian LASSO, BART prior, Logistic Regression and Ridge Regression, produce "good" – "very good" (Cohort 1 AUCs: 0.89-0.93, Cohort 2 AUCs: 0.79-0.93) performances on cohorts 1 and 2. However each model has either suboptimal sensitivity or specificity compared to SVM (Cohort 1 sensitivities: 0.72-0.92, Cohort 1 specificities: 0.83-0.95, Cohort 2 sensitivities – 0.56-0.89, Cohort 2 specificities – 0.75-1) on one of the cohorts.

Table 4.3 Classification performance of the seven statistical models across cohorts 1-3 (training, validation and application). Logistic; Logistic Regression. Ridge; Ridge Regression. RF; Random Forest. BL; Bayesian LASSO. BART; BART – uninformative prior. BART prior; BART – informative prior.

Performance Metric	Centre	Cohort	Logistic	Ridge	RF	SVM	BL	BART	BART prior
AUC	Cologne	Cohort 1 (Training) - Antipsychotic naive SZ & Ctrls	0.93	0.90	1.00	0.92	0.92	0.89	0.91
	Rotterdam	Cohort 2 (Validation) - Antipsychotic naive SZ & Ctrls	0.84	0.93	0.79	0.87	0.87	0.79	0.8
	Utrecht	Cohort 3 (Application) - Antipsychotic treated SZ & Ctrls	0.67	0.62	0.73	0.63	0.63	0.66	0.66
Sensitivity	Cologne	Cohort 1 (Training) - Antipsychotic naive SZ & Ctrls	0.92	0.72	1.00	0.85	0.75	0.82	0.77
	Rotterdam	Cohort 2 (Validation) - Antipsychotic naive SZ & Ctrls	0.89	0.89	0.89	0.89	0.89	0.78	0.56
	Utrecht	Cohort 3 (Application) - Antipsychotic treated SZ & Ctrls	0.63	0.37	0.84	0.42	0.42	0.74	0.74
Specificity	Cologne	Cohort 1 (Training) - Antipsychotic naive SZ & Ctrls	0.81	0.95	1.00	0.86	0.94	0.96	0.92
	Rotterdam	Cohort 2 (Validation) - Antipsychotic naive SZ & Ctrls	0.75	0.92	0.58	0.92	0.92	0.75	1.00
	Utrecht	Cohort 3 (Application) - Antipsychotic treated SZ & Ctrls	0.76	1.00	0.59	0.94	1.00	0.71	0.71

Figure 4.1 ROC curves illustrating classification performance for each of the seven models on (a) cohort 2 (validation) – antipsychotic naive schizophrenia patients & controls; (b) cohort 3 (application) – antipsychotic treated patients & controls. Logistic; Logistic Regression. Ridge; Ridge Regression. RF; Random Forest. BL; Bayesian LASSO. BART; BART – uninformative prior. BART prior; BART – informative prior.

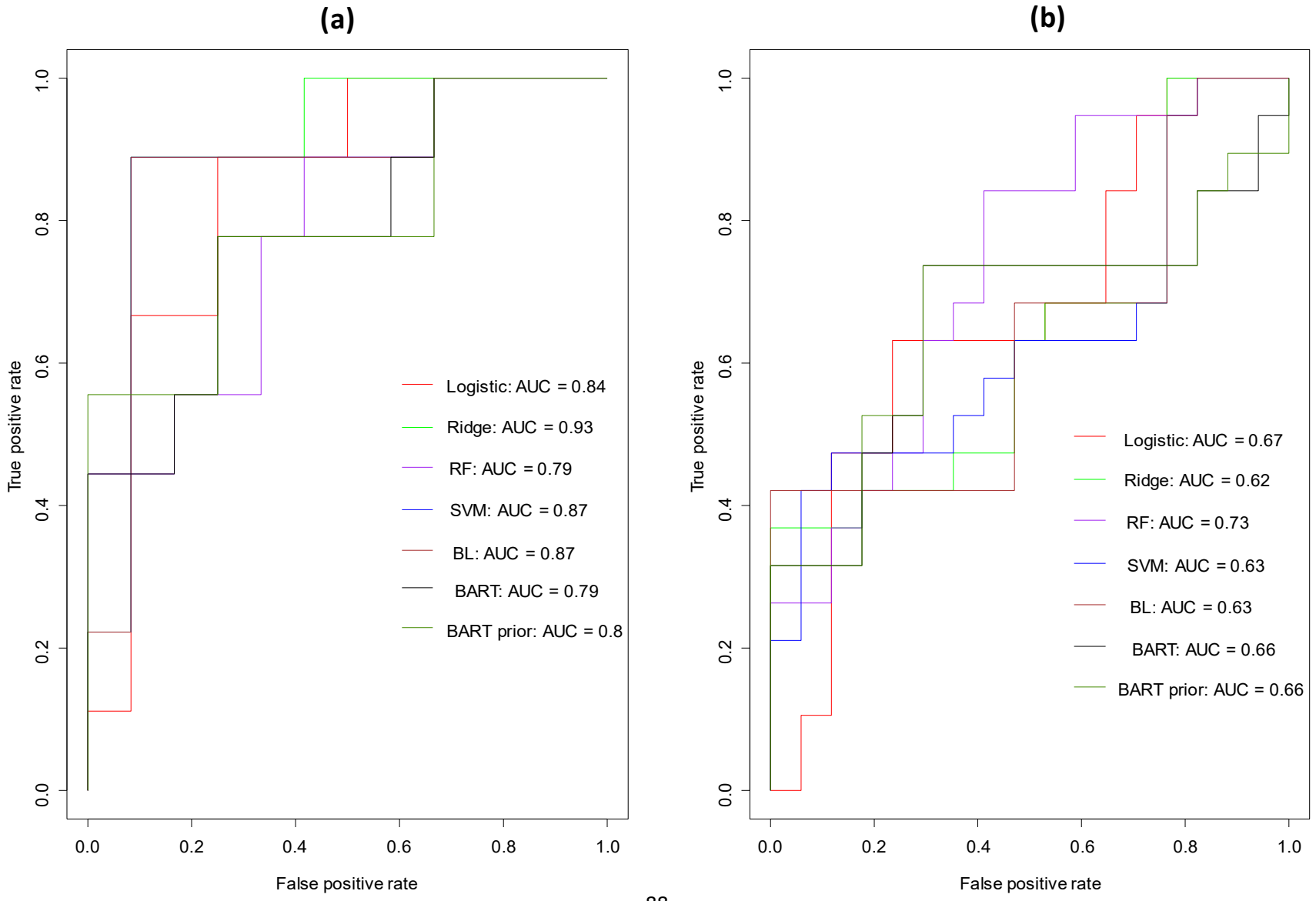


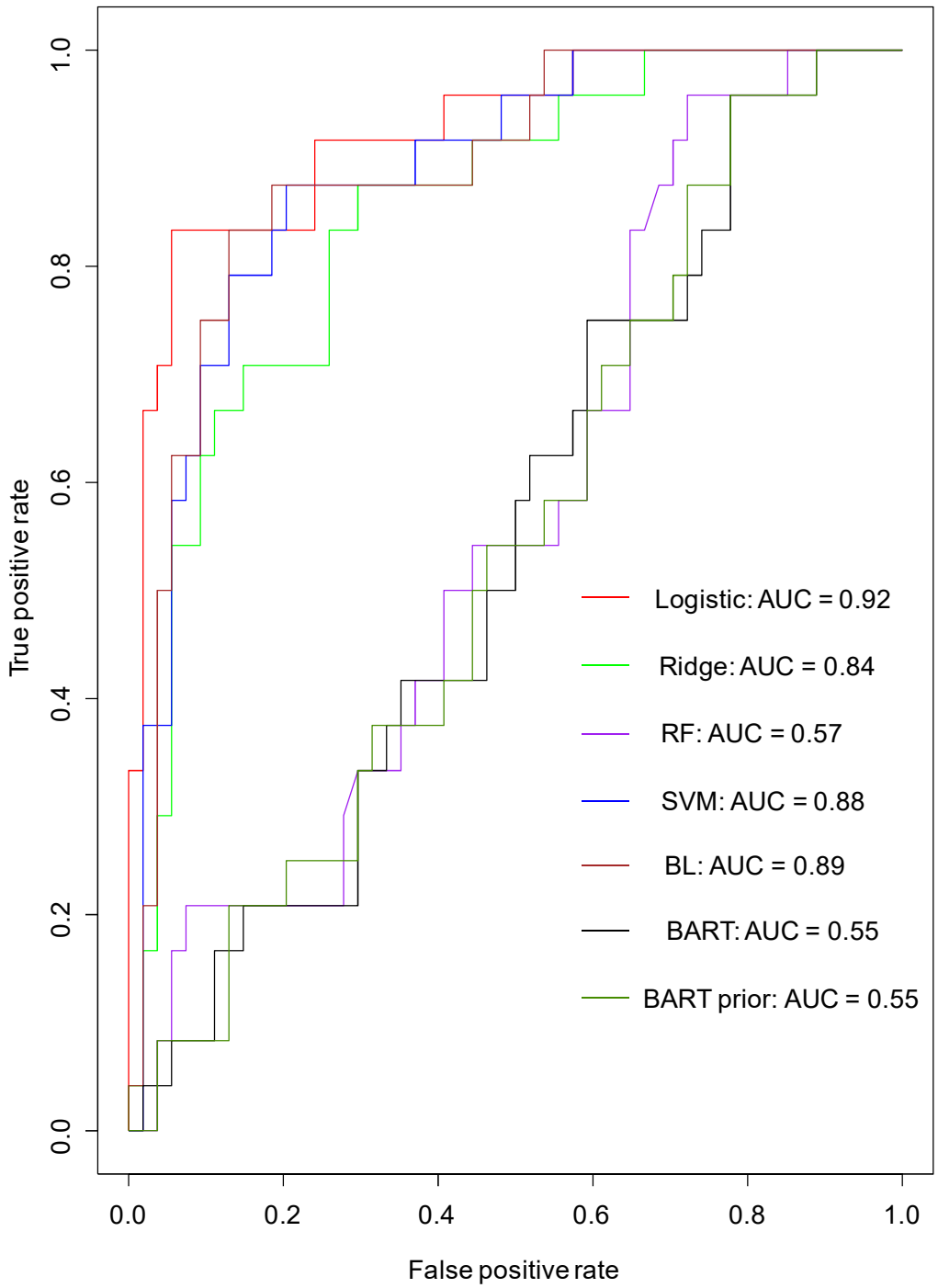
Table 4.4 Classification performance of the seven statistical models between pre-schizophrenia individuals and controls, pre-bipolar individuals and controls, and pre-schizophrenia individuals and pre-bipolar individuals. Samples came from cohort 4 (prediction). Logistic; Logistic Regression. Ridge; RidgeRegression. RF; Random Forest. BL; Bayesian LASSO. BART – uninformative prior. BART prior; BART – informative prior.

Performance Metric	Centre	Cohort	Logistic	Ridge	RF	SVM	BL	BART	BART prior
AUC	Military	Cohort 4 (Prediction) - pre-SZ & Ctrls	0.42	0.41	0.43	0.42	0.40	0.45	0.47
	Military	Cohort 4 (Prediction) - pre-BD & Ctrls	0.46	0.51	0.56	0.51	0.51	0.57	0.58
	Military	Cohort 4 (Prediction) - pre-SZ & pre-BD	0.46	0.40	0.38	0.41	0.39	0.38	0.39
Sensitivity	Military	Cohort 4 (Prediction) - pre-SZ & Ctrls	0.00	0.01	0.02	0.03	0.00	0.03	0.50
	Military	Cohort 4 (Prediction) - pre-BD & Ctrls	0.07	0.66	0.42	0.46	0.62	0.62	0.70
	Military	Cohort 4 (Prediction) - pre-SZ & pre-BD	0.81	1.00	0.99	1.00	0.96	0.99	0.99
Specificity	Military	Cohort 4 (Prediction) - pre-SZ & Ctrls	1.00	1.00	0.99	1.00	1.00	0.99	0.54
	Military	Cohort 4 (Prediction) - pre-BD & Ctrls	0.94	0.42	0.74	0.59	0.46	0.59	0.51
	Military	Cohort 4 (Prediction) - pre-SZ & pre-BD	0.22	0.02	0.06	0.02	0.06	0.06	0.06

Table 4.5 Classification performance of the seven statistical models between converters (prodromal individuals who later developed psychosis) and non-converters (prodromal individuals who did not). Samples came from cohort 5 (prediction). Logistic; Logistic Regression. Ridge; RidgeRegression. RF; Random Forest. BL; Bayesian LASSO. BART – uninformative prior. BART prior; BART – informative prior

Performance Metric	Centre	Cohort	Logistic	Ridge	RF	SVM	BL	BART	BART prior
AUC	Paris	Cohort 5 (Prediction) - converters & non-converters	0.92	0.84	0.57	0.88	0.89	0.55	0.55
Sensitivity			0.83	0.83	0.96	0.88	0.83	0.96	0.96
Specificity			0.94	0.72	0.28	0.80	0.87	0.22	0.22

Figure 4.2 ROC curves illustrating performance for each of the seven models on cohort 5 (prediction) – psychosis converters and non converters. Logistic; Logistic Regression. Ridge; Ridge Regression. RF; Random Forest. BL; Bayesian LASSO. BART; BART – uninformative prior. BART prior; BART – informative prior.



The SVM model's performance on the validation cohort (cohort 2) compares well to its performance on the training data (cohort 1), showing that classification ability is not substantially reduced when it is tested on an independent cohort of antipsychotic naive patients and controls. However this model produces a "poor" performance (AUC – 0.63) when tested on the application cohort (cohort 3) of recent-onset antipsychotic treated patients and controls. A drop in performance is expected on this cohort as the models were trained on a different category of patients, antipsychotic naive patients and controls, but this result suggests that the SVM model would not be clinically useful in diagnosing treated patients. All other models produce "poor" classification performances (AUCs: 0.62-0.67) on this cohort, apart from Random Forest which is the only model to produce a "fair" performance (AUC – 0.73). Apart from the three tree-based models, Random Forest, BART and BART prior, SVM and the other models have much lower sensitivities than specificities on cohort 3, showing that they struggle to correctly classify patients.

4.3.2.2 Prediction cohorts

The classification performances of the seven models are subsequently tested on the prediction cohorts (Cohorts 4 and 5) to assess the ability of the models to predict schizophrenia before disease onset. In Cohort 4, all models are used to classify between pre-schizophrenia individuals and controls, pre-bipolar individuals and controls (to detect whether any classification ability seen between pre-schizophrenia individuals and controls was disease-specific), and between pre-schizophrenia individuals and pre-bipolar individuals (rates of misdiagnosis between these two psychiatric disorders is extremely high). In Cohort 5, all models are used to classify between prodromal individuals who later developed psychosis (converters) and prodromal individuals who didn't develop the disorder (non-converters). Performance analysis across cohorts 4 and 5 is summarized in **Tables 4.4 and 4.5**, while **Figure 4.2** shows ROC curves for all models for the comparison in Cohort 5.

As shown in **Table 4.4**, all models fail on the classification comparisons investigated in Cohort 4. The performance ranges are AUC: 0.4-0.47 between pre schizophrenia individuals and controls, AUC: 0.46-0.58 between pre-bipolar individuals and controls, and AUC: 0.39-0.46 between pre schizophrenia individuals and pre-bipolar individuals. The issue of AUCs being less than 0.5 is discussed earlier in **Chapter 3**.

However, the most notable result of this study is found through the comparison investigated in Cohort 5, as shown in **Table 4.5** and **Figure 4.2**. The SVM model (the only model which could classify patients from controls in cohorts 1 and 2 with high AUC, sensitivity and specificity), is also able to classify psychosis converters from non converters with a "good"

performance and high sensitivity and specificity (AUC - 0.88, sensitivity – 0.88, specificity - 0.8).

The Logistic Regression and Bayesian LASSO models are also able to classify psychosis converters from non-converters with “good”-“very good” AUCs ranging from 0.89-0.92, and high sensitivity and specificities of 0.83 and 0.87-0.94. There is a notable difference in performance between the three tree-based models (Random Forest and the two BART models) and the other four models, which can be seen clearly in **Figure 4.2**. The three tree-based models “fail” when attempting to classify converters from non-converters with AUCs of less than 0.6.

4.3.3 Biomarker identification

All seven models are initially trained on 21 peptides and the demographic variable age in cohort 1. The same variable selection methods applied in **Chapter 3** are utilized to identify different subsets of biomarkers from these 21 peptides. The purpose of identifying biomarkers is for clinical applicability, to see whether each model can achieve a similar performance on the independent cohorts with a smaller subset of peptides, thus making a prospective clinical test cheaper to run. **Table 4.6** shows the biomarkers identified for each of the models. RFE identifies 20 biomarkers for Random Forest and 14 for SVM. Stepwise selection identifies 9 biomarkers for Logistic Regression, and LASSO selects 7 biomarkers. Bayesian LASSO selects 3 biomarkers, and BART with uninformative and informative priors selects the same 2 biomarkers. It may be notable in the context of schizophrenia pathogenesis that the peptide DWHGVPGQVDAAMAGR, corresponding to the protein Vitronectin (VTNC), is the only peptide to be selected as a biomarker by all seven models. VTNC has previously been shown in rodent studies to play a crucial role in the developing CNS in interacting with and determining the functions of proteins directly involved in synaptic plasticity, synaptogenesis, neural differentiation and survival, and neural regeneration following injury(389,390). Several of these proteins have been linked to schizophrenia and other psychiatric disorders in humans, suggesting that changes in the abundance of VTNC may play a role in the pathophysiological mechanisms underpinning these disorders(389).

4.3.4 Reduced model analysis

With 21 peptides, there are 2^{21} possible models for a given statistical method such as Logistic Regression. As it would not be practical to examine the performance of all of these, this study uses different variable selection methods to identify sets of peptides which represent the most important biomarkers contributing to each model. In the case of Ridge

Regression, LASSO is used for variable selection to obtain a set of biomarkers through a penalized regression method. As shown in **Table 4.6**, 7 unique sets of biomarkers were identified from the 21 peptides in the original training data. Cohort 1 is then used to train each statistical method on each of these sets of biomarkers. These models are then tested on each of the independent cohorts to examine whether similar or better classification performance could be achieved with a smaller, optimal set of peptides.

Table 4.6 The seven sets of biomarkers identified using stepwise selection with logistic regression, LASSO, recursive feature elimination with Random Forest and SVM, and biomarker selection strategies for Bayesian LASSO, and both BART models.

Biomarker Name	Protein Name	Stepwise Selection (Logistic Regression)	Ridge	RFE (Random Forest)	RFE (SVM)	BL	BART	BART prior
Age	NA							
DWHGVPGQVDAAMAGR	VTNC	✓	✓	✓	✓	✓	✓	✓
DYAEVGR	HPT			✓	✓			
VTSIQDWVQK	HPT	✓	✓	✓	✓			
MFLSFPTTK	HBA	✓	✓	✓	✓	✓		
LPGIVAEGR	ANT3			✓				
VSFFCK	APOH			✓	✓			
FDTISEK	ANT3			✓				
EQLSLLDR	AACT							
DALSSVQESQVAQQAR	APOC3	✓	✓	✓	✓			
EFGNTLEDK	APOC1			✓	✓			
GPDVLTATVSGK	ITIH4			✓	✓			
VLSLAQEQVGGSPEK	CO4A			✓	✓			
FDPSLTQR	A2AP			✓				
YGIDWASGR	FCN3			✓	✓			
ALVQQMEQLR	APOA4	✓	✓	✓	✓			
VVEESELAR	CO9	✓		✓				
GWVTDGFSSLK	APOC3	✓		✓				
ISASAEELR	APOA4	✓		✓	✓		✓	✓
ITQVLHFTK	CO4A		✓	✓	✓			
YWGVASFLQK	RET4	✓	✓	✓		✓		
HAIILLTDGK	CO2			✓	✓			

Figures 4.3-4.5 show the performances of each model in terms of AUC on cohort 1 and the independent cohorts, for each set of biomarkers. The performance with all 21 peptides is included for comparison purposes. The same plots for sensitivity and specificity are included

in the **Appendices**. Cohort 4 (prediction), is not included in these figures as all models "fail" on this cohort for all sets of biomarkers.

One of the main findings of this analysis is that, as in **4.3.2.1**, SVM is the only method which achieves at least a "good" classification performance and sensitivity and specificity of at least 0.8 on both the training cohort (cohort 1) and the validation cohort (cohort 2), with a reduced set of peptides. SVM achieves a comparable performance to the full 21 peptides with just 7 peptides (VTNC, HPT, HBA, APOC3, APOA4, CO4A, RET4). The SVM model with 7 peptides produces a "very good" performance on cohort 1 (AUC – 0.92, sensitivity – 0.8, specificity – 0.91), and a "good" performance on cohort 2 (AUC – 0.88, sensitivity – 0.87, specificity – 0.83).

In comparison, when fitting models on these 7 peptides using the other six statistical methods, the BART, BART prior and Ridge Regression models produce "very good" performances on cohorts 1 and 2 (Cohort 1 AUCs: 0.9-0.93, Cohort 2 AUCs: 0.91) but for each model, sensitivity is less than 0.8 for either cohort 1 or 2 ,indicating that the ability to correctly classify patients consistently may be suboptimal. The same trend is true for the Bayesian LASSO, Random Forest and Logistic Regression models with 7 peptides. Again the models produce classification performances in the "good"-"very good" range on cohorts 1 and 2 but sensitivities are less than 0.8 on one of the cohorts.

These results suggest that when it comes to diagnosing antipsychotic naive patients from controls, the SVM model with 7 peptides is the most optimal for fulfilling the desired criteria of consistently high classification performance, high sensitivity and specificity, and a relatively few number of biomarkers. While all 7 statistical methods still produce "good" performances on both cohorts 1 and 2 (Cohort 1 AUCs: 0.83-1, Cohort 2 AUCs: 0.83-0.85) with just 3 peptides, in each case the sensitivity (Cohort 1 sensitivities: 0.53-1, Cohort 2 sensitivities: 0.89-1) and/or specificity values (Cohort 1 specificities: 0.84-1, Cohort 2 specificities: 0.67-0.83) on the two cohorts are not sufficient for clinical use.

However the SVM model with 7 peptides produces a "poor" performance (AUC: 0.67, sensitivity: 0.53, specificity: 0.88) on the application cohort (cohort 3). None of the other models produce sufficient performances to be considered as clinically relevant tools, but the Random Forest model with 9 peptides (VTNC, HPT, HBA, APOC3, APOA4, CO9, APOC3, APOA4, RET4) is the only model out of all, including the models with 21 peptides, which achieves at least a "fair" classification performance as well as sensitivity and specificity of at least 0.7 (AUC – 0.76, sensitivity – 0.79, specificity – 0.71). Most models have high specificity but low sensitivity across the sets of biomarkers, showing the difficulty in classifying patients in this cohort based on the relative peptide abundances of antipsychotic

naive patients and controls. The tree-based models, BART, BART prior and Random Forest achieve the highest classification performances for cohort 3 with both BART models producing an AUC of 0.78 with 7 peptides, but Random Forest is the only model which achieves both “fair” performances and higher sensitivities for this cohort.

For the prediction cohort (cohort 5), the SVM model with 7 peptides has a low sensitivity (AUC: 0.82, sensitivity: 0.63, specificity: 0.89). Therefore, the SVM model with 21 peptides remains the only model which produces “good”-“very good” classification performances in discriminating both antipsychotic naive schizophrenia patients from controls across cohorts 1 and 2, and psychosis converters from non-converters in cohort 5, as well as having sensitivities and specificities for these comparisons of at least 0.8.

However from a biomarker cost perspective, an important finding could be that the Bayesian LASSO model with 3 peptides (VTNC, HBA and RET4) produces a “very good” performance and high sensitivity and specificity (AUC – 0.9, sensitivity – 0.92, specificity – 0.87) on this cohort. While this model is not sensitive or specific enough for utility as a diagnostic test between antipsychotic naive patients and controls (Cohort 1 – AUC: 0.83, sensitivity: 0.67, specificity: 0.84. Cohort 2 – AUC: 0.85, sensitivity: 1, specificity: 0.67), it may be worth further investigation as a purely prognostic test for detecting psychosis conversion, as this model is more sensitive and specific on cohort 5 than the SVM model with 21 peptides, and requires the measurement of far fewer biomarkers.

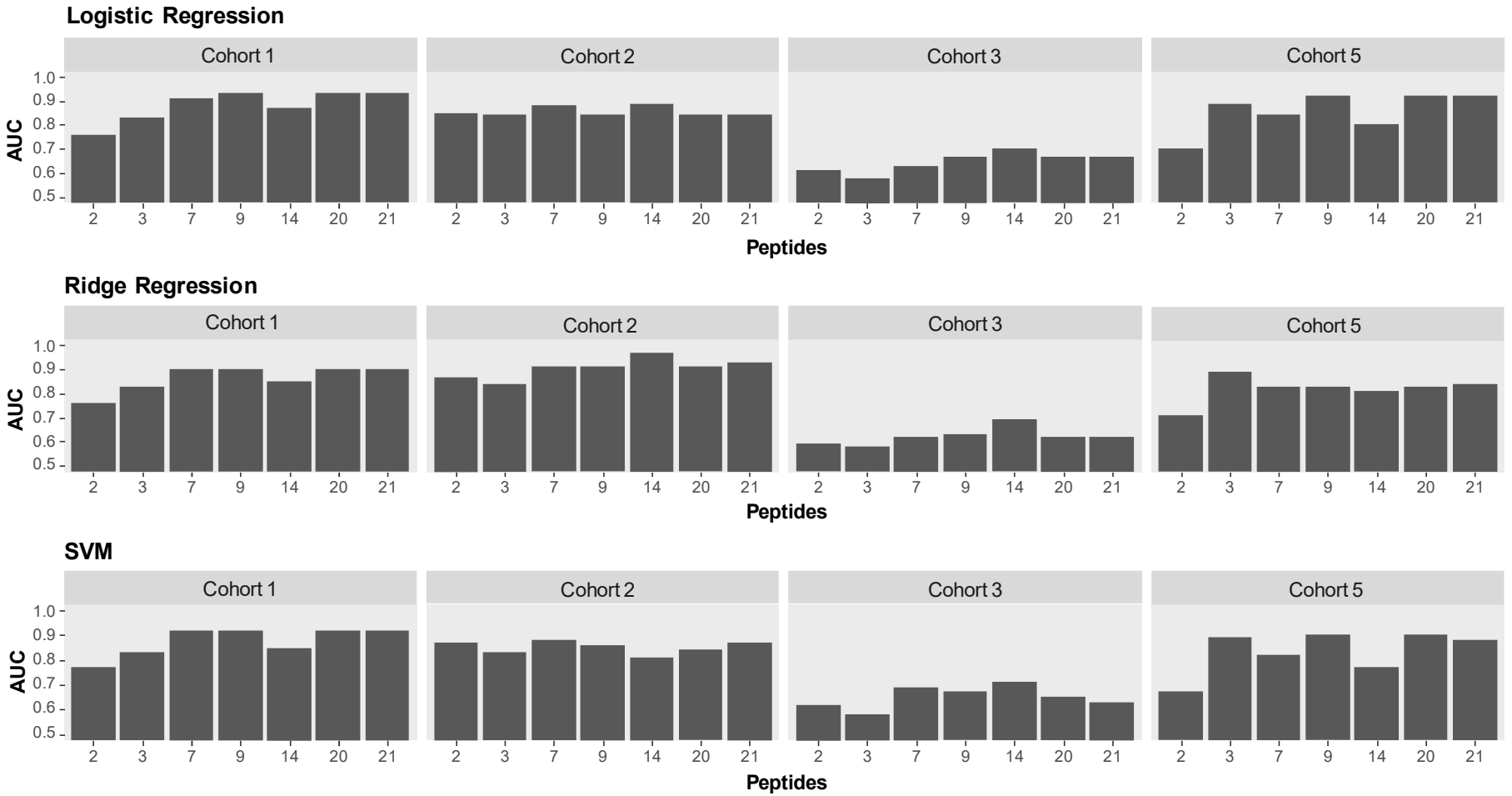


Figure 4.3 Classification performance on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts for Logistic Regression, Ridge Regression and SVM. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparative purposes.

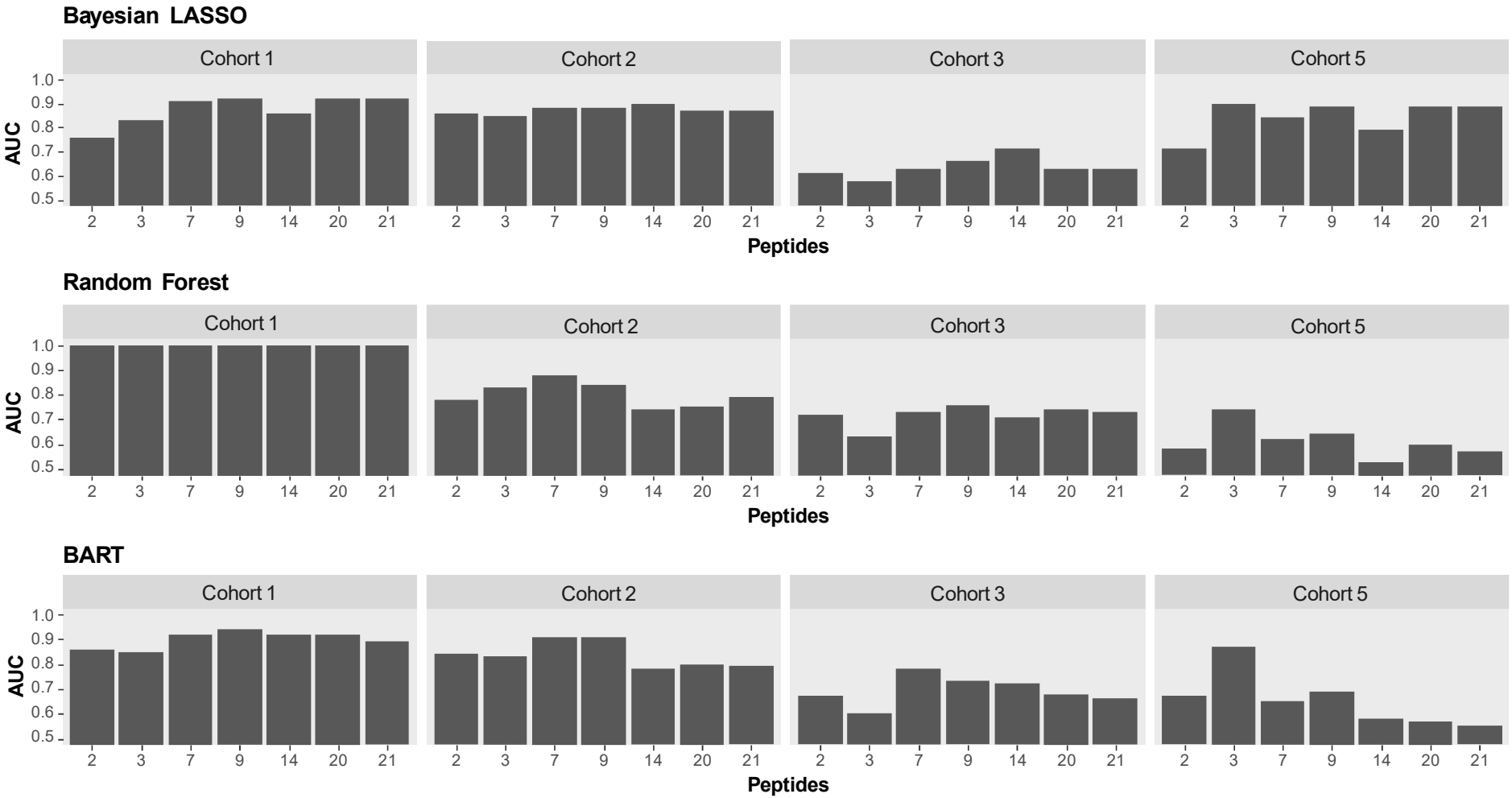


Figure 4.4 Classification performance on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts for Bayesian LASSO, Random Forest and BART. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparative purposes.

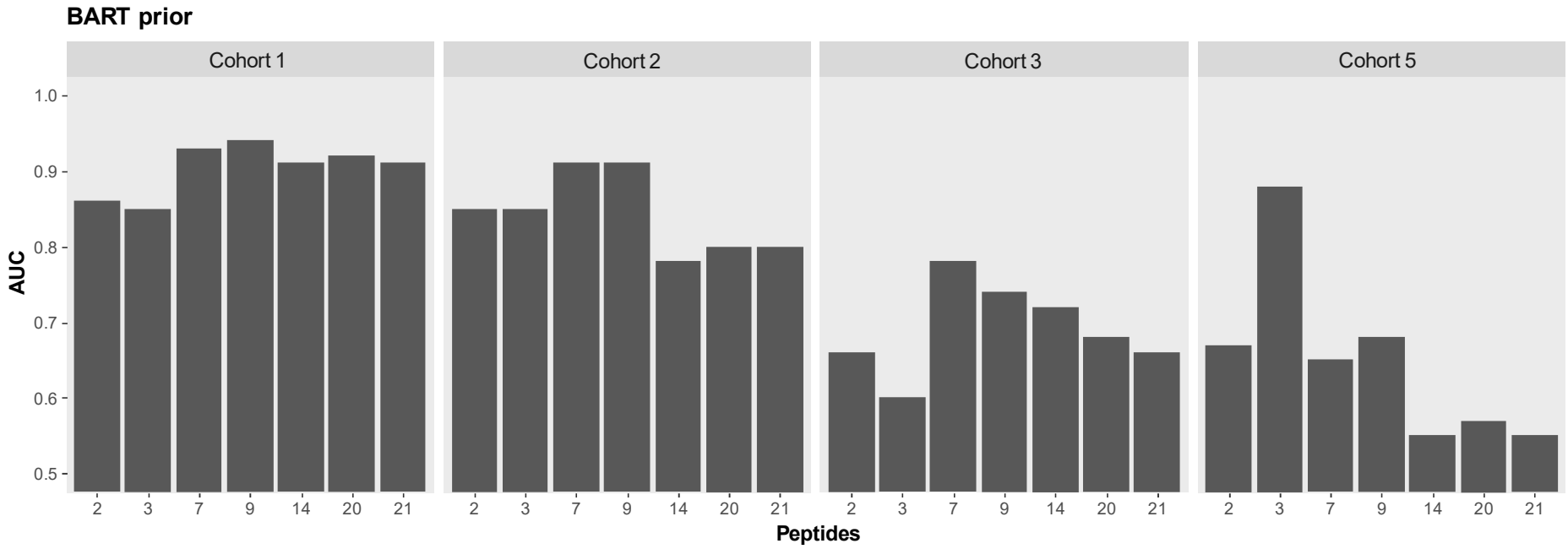


Figure 4.5 Classification performance on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparative purposes.

4.4 Discussion

In summary, this chapter presents the application of multiple statistical methods towards the problem of identifying models which can classify schizophrenia and predict the onset of the disease through molecular profiling of serum. This problem was initially examined in **Chapter 3** through the concentration levels of 66 proteins, measured using a multiplex immunoassay platform. The study presented in this chapter differs from **Chapter 3** as (a) it measures the abundances of peptides in serum using targeted mass spectrometry, and (b) its main focus is on identifying serum analytes which can predict the development of psychosis before disease onset. Hence a targeted approach is applied, measuring the abundances of a pre-selected panel of 147 peptides linked to psychiatric disorders, many of which are associated with different proteins to those studied in **Chapter 3**.

4.4.1 Classification performance with 21 peptides

There are two major findings of this study. The first is the identification of an SVM model based on the abundances of 21 peptides in serum which can classify antipsychotic naive schizophrenia patients from healthy controls with an AUC of 0.87-0.92 and sensitivity and specificity of 0.85-0.89 and 0.86-92. The second is that this model can also predict conversion to psychosis in prodromal individuals with an AUC of 0.88 (0.88 sensitivity and 0.8 specificity).

The latter finding is particularly notable, as thus far, no other study has been able to robustly identify a model based on a set of serum analytes which can predict psychosis conversion in an independent cohort. In **Chapter 3**, the statistical models based on the concentrations of 66 serum proteins, failed when it came to classifying converters from non-converters. In addition, independent studies which have claimed to identify protein biomarkers capable of predicting schizophrenia conversion have either lacked a validation cohort or demonstrated analytical flaws in the application of statistical models which have biased the results(12,190).

As such, the performance of this model in classifying psychosis converters from non-converters suggests that if this result survives further validation testing in similar cohorts, this model could form the basis of a blood biomarker assay for aiding clinicians in evaluating the likelihood of progression to schizophrenia in at-risk individuals, in conjunction with the currently used structured clinical interviews. Previous review papers have stated that such an assay would have the potential to be particularly useful, as there are currently no biological tools available to assist clinicians in this regard(152).

In light of these classification performances, the biological functions represented by these 21 peptides provides novel, useful information regarding disease pathophysiology. The results demonstrate that differences in their abundances are not only capable of discriminating the schizophrenia patient group from healthy controls once symptoms develop, thus representing a 'schizophrenia signal', but that this signal is present in the at-risk population even before disease onset, indicating that the relevant pathways may be involved at the earliest stages of schizophrenia pathogenesis. The 21 peptides in this panel are associated with the members of the apolipoprotein family, the complement cascade, and further proteins involved in inflammatory processes, metabolic function, coagulation, and cell adhesion. Of these associated proteins, apolipoprotein A-IV (APOA4)(391,392), apolipoprotein C-III (APOC3)(34), apolipoprotein H (APOH)(34), antithrombin-III (ANT3)(393,394), complement C2 (CO2)(25), complement C4-A (CO4A)(25,395), complement component C9 (CO9)(395), ficolin-3 (FCN3)(33), haptoglobin (HPT)(396,397), inter-α-trypsin inhibitor heavy chain H4 (ITIH4)(398), retinol-binding protein 4 (RET4)(399), and VTNC(355) have all been previously linked to schizophrenia in either genomic or proteomic studies.

It should be noted that the SVM model based on these 21 peptides "fails" to classify pre-schizophrenia individuals from controls in another independent cohort. This indicates that the model is only capable of predicting disease conversion in individuals who are already symptomatic. The pre-schizophrenia individuals in this cohort were sampled before they began displaying symptoms, and were only diagnosed with schizophrenia at a much later date. In contrast, the psychosis converters in the prodromal cohort were already displaying initial symptoms at the time of sample collection, although further symptom progression to a clinical diagnosis of psychosis lasted between several months and a year.

The SVM model's classification performance on the application cohort of recent-onset antipsychotic treated patients and controls is "poor" and the performance of other models on this cohort, based on these 21 peptides, ranges from "poor" to "fair". While patients in this cohort had an average disease duration of around one year, they had lower severity scores than the first-onset patients in the training data. Therefore we can suggest that the modulatory effects of antipsychotic treatment could be the reason why none of the models achieve better performance. The patients in this cohort are treated with a mixture of first and second generation antipsychotics. Antipsychotics have been previously shown to modulate the levels of haptoglobin, and complement C4, two of the proteins associated with the panel of 21 peptides, in plasma of schizophrenia patients(400). These findings suggest that in order to identify a statistical model which can achieve classification performances sufficient for clinical utility in this particular patient population, it may be necessary to train and test different statistical models specifically on cohorts of antipsychotic treated patients.

4.4.2 Classification performance with reduced peptide sets

As stated earlier in this chapter, the more serum analytes required, the greater the cost of a biomarker test. Thus, the reduced model analysis section of this study assessed whether there are more cost-effective options available which still produce sufficient classification performance, sensitivity and specificity for classifying antipsychotic naive patients from controls and psychosis converters from non-converters.

One of the key findings of this analysis is that an SVM model based on the abundances of just 7 peptides (VTNC, HPT, HBA, APOC3, APOA4, CO4A, RET4) can classify antipsychotic naive patients from controls with an AUC of 0.88-0.92 as well as sensitivity and specificity of 0.8-0.87 and 0.83-0.91. However, while this model does not have sufficient sensitivity for clinical use when classifying psychosis converters from non-converters, it is also found that a Bayesian LASSO model based on the abundances of 3 peptides (VTNC, HBA and RET4) can classify converters with an AUC of 0.9, sensitivity of 0.92 and specificity of 0.87. If the aforementioned SVM model with 21 peptides is found to be too expensive, an alternative possibility could be to do further validation work on two separate biomarker tests, one specifically for classifying antipsychotic naive patients using the SVM model with 7 peptides, and another specifically for classifying psychosis converters using the Bayesian LASSO model with 3 peptides.

Moreover, the fact that differences in the abundances of peptides relating to VTNC, HBA and RET4 between first-onset schizophrenia patients and controls are important in differentiating psychosis converters from non converters at an earlier stage on the disease suggests these analytes are worthy of further investigation as prognostic biomarkers. Previous research findings have linked both VTNC and RET4 to various pathological processes in the development of the disease, with VTNC being of particular interest as the only analyte which was selected as a biomarker by all seven methods. VTNC has been found to interact with, and mediate the function of other proteins which have key roles in the developing CNS relating to neural cell proliferation, neuritogenesis, synaptogenesis, and synaptic plasticity(389,390). Elevated plasma levels of RET4 have previously been implicated in the development of insulin resistance and other metabolic adversities in schizophrenia patients(399).

4.4.3 Comparison of statistical approaches

As discussed in **3.1**, it has been previously found for classification problems that different statistical methods perform better on different datasets, hence the application of multiple methodologies in this chapter to assess which was most suited to this data. As described in this discussion, SVM performs best out of all seven approaches in terms of a model which has sufficiently high AUC, sensitivity and specificity, both for classifying antipsychotic naive patients from controls but converters from non converters. However, purely in terms of AUC, the trends seen between the SVM, Bayesian LASSO, Logistic Regression and Ridge Regression models across all test cohorts are broadly similar. The tree-based methodologies Random Forest, and both BART approaches do not perform so well on this classification problem, most notably when classifying converters from non converters. In the case of Random Forest, this may be because the models overfit on the training data to a greater extent than any other approach. This is suggested that the fact that all Random Forest models return an AUC, sensitivity and specificity of 1 when tested on this data. There are no real differences in terms of performance, between BART models fitted with and without informative priors, but one of the limitations of this study is that due to data availability, the prior used is based on one of the test cohorts.

Further limitations and future work relating to this study are discussed in **Chapter 9**.

Chapter 5 Identification of novel drug discovery targets through analysis of data from a cellular model of schizophrenia

5.1 Introduction

The development of novel medications for schizophrenia has come to a standstill, due to insufficient understanding of the affected molecular pathways in patients, and how these translate to disease symptoms. As such, relevant cellular models are now thought to form a vital part of the preclinical drug discovery pipeline, both for identification of disease-specific signatures, and as potential novel screening platforms for drug profiling.

As described in **1.6**, the suitability of microglia as the basis for a putative cellular model of schizophrenia derives from multiple post-mortem immunohistochemistry and PET studies which have pointed towards abnormal activation of microglia as a mechanism for schizophrenia pathogenesis in approximately 40% of patients(227,401,402). This activation profile has been accompanied by evidence of increased neuroinflammatory processes through upregulated proinflammatory cytokine expression in the brain and periphery, along with regional brain volume reductions, and treatment-resistant negative symptomatology(402–404). In addition, GWAS findings in schizophrenia patients suggest that genetic polymorphisms in secreted proteins, including members of the complement cascade, may alter synaptic pruning by microglial cells during critical neurodevelopmental phases already linked to schizophrenia(25,200–202).

However, so far the majority of data linking microglial activation to schizophrenia has been acquired through *in vivo* studies, quantifying circulating proteins, for example specific cytokines, and imputing their potential net effect on microglial function based on their BBB permeabilities(63,402,403,405). Therefore, there is a need for *in vitro* studies to assess the functional role assumed by microglia at the proteome level across multiple cell signalling pathways, in response to complex body fluids such as serum and CSF(406–408). This study aims to investigate the signalling pathways underpinning microglial activation in schizophrenia through analyzing data from an *in vitro* cellular model in which human microglial cells are exposed to serum samples from schizophrenia patients and controls. Fluorescent intensities from targeted intracellular signalling epitopes were acquired using a

high-content flow cytometry screening platform. Subsequently, serum protein alterations are accrued through targeted mass spectrometry analysis, to explore which molecules may stimulate the observed microglial signalling patterns in the disease state relative to controls.

One of the aims of this study is to provide new results regarding the intracellular signalling alterations underpinning the complex microglial activation phenotype, relating to schizophrenia. Activated microglia are broadly categorised as classically activated M1 or alternatively activated M2 microglia(406,409), The balance between M1 and M2 phenotypes is believed to be particularly relevant to the pathogenesis of schizophrenia as it is thought to determine the long-term consequences of brain inflammation(406,410). M2 microglia, induce tissue restoration by releasing anti-inflammatory cytokines, trophic factors and the phagocytosis of cellular debris(406,409). M1 microglia instigate a protective response against an external insult such as pathogens or acute CNS injury, inducing an inflammatory state through the release of proinflammatory cytokines and reactive oxygen species(406,409). However dysfunctional M1 microglia have been linked to the deleterious consequences of microglial activation in the CNS, in stroke, traumatic brain injury, and neurodegenerative diseases(406,409).

In this study, intracellular signalling differences between patients and controls are quantified using different statistical methodologies. Each serum sample is run in triplicate, and as such, a linear mixed model is used to account for the individual variation between measurements. Once patient-control differences in individual epitopes have been assessed, the global expression pattern of individual changes at the pathway level was examined using the commonly used global test approach(343). This method was initially developed for genetics studies to investigate hypotheses that the combined expression levels of a set of genes were connected to a particular molecular function. More recently, it has been utilized in proteomic studies, with *Hollander et al.*(411),using the global test to examine whether the global expression pattern of a panel of biomarkers identified through high throughput mass spectrometry was associated with heart function status. The use of the global test in this study is akin to more targeted approaches which focus on sets of features involved in the same biological function or cellular pathway(344).

The final part of this study looks to quantify whether the observed dysfunctional microglial signalling phenotypes can be utilized for novel drug discovery. This is investigated by targeting the individual epitopes found to be over-activated by schizophrenia patient serum, with a range of compounds, some of which have been previously hypothesized to be microglial M1 inhibitors, such as rapamycin(412).

Overall the exploratory analyses presented in this chapter represent the first attempts to create a preclinical cellular model of schizophrenia to explore the signalling mechanisms involved in abnormal microglial activation in the disease state. The involvement of microglia in schizophrenia disease pathogenesis has been widely speculated, and due to their perivascular localization, it has been hypothesized that microglia can be polarized by factors present in circulating blood. This chapter also presents the first use of a preclinical cellular model to examine the potential efficacy of various compounds to normalize dysfunctional microglial activation, by testing a range of current antipsychotics and existing compounds which have been proposed as potential novel treatments for schizophrenia.

5.2 Methods

5.2.1 Functional characterization of cellular model through detection of M1 vs M2 microglial polarization signals

Initially it is assessed whether the proposed cellular model is capable of detecting characteristic signalling responses to particular ligands known in the literature to specifically induce either M1 or M2 microglial phenotypes(409,413). If this capability is observed, this would validate the model for use in a serum exposure study. 42 intracellular signalling epitopes are stimulated by 4 different M1 ligands (IFN-γ, IL-23, TNF-α and IL-1β), and 4 different M2 ligands (IL-4, IL-13, TGF-β and BMP7) titrated at different concentrations, in addition to two broad spectrum positive controls (calyculin which upregulates cell signalling expression, and staurosporine which downregulates cell signalling expression). Experiments are run in triplicate. Ligand responses are computed relative to the vehicle condition, using the Wilcoxon rank-sum test (applied in R as described in **Chapter 2**). Ligand responses are only considered if they are significant ($p<0.05$), and survive filtering for background fluorescence. Fold changes for each epitope are defined as ligand response/vehicle response.

5.2.2 Microglial exposure to serum from first-onset antipsychotic naive schizophrenia patients

5.2.2.1 Clinical samples

Serum samples were collected from 139 individuals, 60 antipsychotic naive, first onset schizophrenia patients, and 79 healthy controls, recruited from the Department of Psychiatry, University of Cologne, Germany. Diagnoses of schizophrenia were based according to the DSM-4. Exclusion criteria for patients and controls included additional neuropsychiatric diagnoses other than schizophrenia, other neurological disorders including epilepsy, mental retardation, multiple sclerosis, immune/autoimmune disorders, infectious disease, metabolic disorders including diabetes, obesity (body mass index above 30), cardiovascular disease, hepatic and renal insufficiency, gastrointestinal disorders, endocrine disorders including hypo-/hyperthyroidism and hypo-/hypercortisolism, respiratory diseases, cancer, severe trauma, substance abuse including psychotropic drugs and alcohol, somatic medication with brain side-effects, somatic medication affecting the immune system including glucocorticoids, anti-inflammatory/immunomodulating drugs and antibiotics. **Table 5.1** displays demographic information for the cohort.

Table 5.1 Demographic characteristics and determination of clinical samples used in the microglial serum exposure study with serum from first-onset antipsychotic naive schizophrenia patients and controls. PANSS - Positive and Negative Syndrome Scale. NA – not applicable. Values are presented as average ±standard deviation.

	SCZ	*Control*
N	60	79
Age (years)	30.72 ± 10.46	31.15 ± 8.32
Gender (male/female)	31/29	43/36
PANSS Positive	22.96 ± 5.91	NA
PANSS Negative	23.20 ± 7.34	NA
PANSS General	49.37 ± 9.95	NA
PANSS Total	95.54 ± 20.49	NA

5.2.2.2 Data pre-processing

The median fluorescent intensities of 62 cellular epitopes spanning key microglial cell signalling pathways, are measured across all 139 samples with 3 replicates, using the flow cytometry platform described in **Chapter 2**. A list of these 62 epitopes, their abbreviations, and their signalling pathways can be found in **Table 5.2.**

All epitope fluorescent intensities are run through a series of pre-processing steps. All epitope fluorescent intensity values are log10-transformed to stabilize data variance and improve normality. In addition, quality control assessment is carried out. Sample outliers are assessed using PCA through inspection of quantile-quantile plots (applied in R as described in **Chapter 3**). PCA is additionally used to check for data artefacts. Missing data points which occurred due to low cell count are excluded in all univariate analyses, and replaced with the mean value for that epitope for PCA and multivariate analyses.

5.2.2.3 Stimulant responses

Stimulant (schizophrenia patient or control serum) responses relative to the vehicle condition, for each of the 62 epitopes, are computed using the Wilcoxon rank-sum test. Only responses which were significant ($p < 0.05$), and survive adjustment for background autofluorescence are considered. For comparison, stimulant responses relative to the vehicle for positive control compounds staurosporine and calyculin across all 62 epitopes are computed. Calyculin induces widespread up-regulation while staurosporine induces widespread down-regulation. Fold changes for each epitope are defined as response to stimulant/response to vehicle.

Table 5.2 List of microglial cell signaling epitopes measured in this study and their corresponding genes and cell signalling pathway classes(296)**.**

Epitopes	Gene	Signalling Pathway
CREB (pS133) / ATF-1 (pS63)	CREB1	PKA
PKA RIIα (pS99)	PRKAR2A	PKA
PKA RIIβ (pS114)	PRKAR2B	PKA
p120 Catenin (pS268)	CTNND1	PKC
p120 Catenin (pS879)	CTNND1	PKC
p120 Catenin (pT310)	CTNND1	PKC
PKC-α	PRKCA	PKC
PKC-α (pT497)	PRKCA	PKC
PKC-θ	PRKCQ	PKC
PLC-γ1	PLCG1, PLCG2	PKC
PLC-γ1 (pY783)	PLCG1	PKC
PLC-γ2	PLCG2	PKC
PLC-γ2 (pY759)	PLCG2	PKC
4EBP1 (pT36/pT45)	EIF4EBP1	Akt/mTORC1
4EBP1 (pT69)	EIF4EBP1	Akt/mTORC1
Akt (pS473)	AKT1	Akt/mTORC1
Akt (pT308)	AKT1	Akt/mTORC1
Akt1	AKT1	Akt/mTORC1
β-Catenin (pS45)	CTNNB1	Akt/mTORC1
CD221 (pY1131)	IGF1R	Akt/mTORC1
eIF4E (pS209)	EIF4E	Akt/mTORC1
Ezrin (pY353)	EZR	Akt/mTORC1
GSK-3α/β	GSK3B	Akt/mTORC1
GSK-3β (pS9)	GSK3B	Akt/mTORC1
GSK-3β (pT390)	GSK3B	Akt/mTORC1
IRS-1 (pY896)	IRS1	Akt/mTORC1
PDPK1 (pS241)	PDPK1	Akt/mTORC1
S6 (PS235/PS236)	RP26	Akt/mTORC1
S6 (PS240)	RPS6	Akt/mTORC1
Bcl-2 (pS70)	BCL2	MAPK
ERK1/2 (pT202/pY204)	MAPK1, MAPK3	MAPK
MAPKAPK-2 (pT334)	MAPKAPK	MAPK
MEK1 (pS218)/MEK2 (pS222)	MAP2K1, MAP2K2	MAPK
MEK1 (pS298)	MAP2K1	MAPK
p38 MAPK (pT180/pY182)	MAPK14, MAPK13, MAPK12	MAPK
p53 (acK382)	TP53	MAPK
p53 (pS37)	TP53	MAPK
Ikβα	NFKBIA	IL1R/ TLR

IRF-7 (pS477/pS479)	IRF7	IL1R/ TLR
NF-kB p65 (pS529)	RELA	IL1R/ TLR
SHP2 (pY542)	PTPN11	JAK/Stat
Stat1 (N-Terminus)	Stat1	JAK/Stat
Stat1 (pS727)	Stat1	JAK/Stat
Stat1 (pY701)	Stat1	JAK/Stat
Stat3	Stat3	JAK/Stat
Stat3 (pS727)	Stat3	JAK/Stat
Stat3 (pY705)	Stat3	JAK/Stat
Stat4 (pY693)	Stat4	JAK/Stat
Stat5 (pY694)	Stat5A, Stat5B	JAK/Stat
Stat6 (pY641)	Stat6	JAK/Stat
c-Cbl (pY700)	CBL	IR/AR
c-Cbl (pY774)	CBL	IR/AR
Pyk2 (pY402)	PTK2B	IR/AR
SLP-76 (pY128)	LCP2	IR/AR
Src (pY418)	SRC	IR/AR
WIP (pS488)	WIPF1	IR/AR
Zap70 (pY319)/Syk (pY352)	SYK, ZAP70	IR/AR
CD140b (pY857)	PDGFRB	Other
CrkL (pY207)	CRKL	Other
Rb (pS780)	RB1	Other
Smad2 (pS465/pS467)/Smad3 (pS423/pS425)	SMAD2	Other

5.2.2.4 Univariate patient-control analysis

A linear mixed model (applied in R as described in **Chapter 2**) with a random intercept to account for the variation between replicates, and a stepwise selection procedure (applied in R as described in **Chapter 2)** to adjust for confounding variables age and gender, is used to compute differential responses between schizophrenia patients and controls. Fold changes for each epitope are defined as response to schizophrenia serum/response to control serum. Only epitopes which display significant serum responses relative to the vehicle in at least one clinical group are included in this analysis.

Permutation testing is used to evaluate p-values for each epitope by comparing a test statistic computed from the original dataset, with a distribution of permuted values obtained on a large number of simulated datasets constructed by randomly permuting the dataset patient-control labels 10,000 times. Each p-value is calculated as the fraction of permutation values which are at least as extreme as the original test statistic, thus representing the probability of obtaining a result at least as extreme as the test statistic given that the null hypothesis is true(414). As has been stated in previous research papers, permutation procedures are widely used in bioinformatics for statistical hypothesis testing in

underpowered studies such as this one, where the distribution of the test statistic is perceived to be unreliable due to insufficient sample size(415). The popularity of permutation procedures derives from the non-parametric nature of these tests, whereby a sampling distribution is estimated for the test statistic due to a lack of sufficient evidence to assume a particular model for the biological data under investigation(414).

Q-values are subsequently obtained by adjusting the p-values for multiple testing according to the Benjamini Hochberg procedure(416).

5.2.2.5 Multiplex immunoassay and MRM profiling of serum

To explore whether the identified microglial activation phenotype can be explained by differences between patients and controls in the relative concentrations or abundances of serum proteins known to be associated with psychiatric disorders, two distinct methodologies are used. Multiplex immunoassays (for low abundance proteins) are used to measure serum concentrations in all patient and control samples, of 17 immunomodulatory proteins previously linked to neuropsychiatric disorders, including complement factors and proinflammatory cytokines. These proteins are Complement C1q, Complement C3, Complement C3b, Complement C4, Granulocyte macrophage colony-stimulating factor, Interferon-γ, IL-10, IL-12(p70), IL-13, IL-1β, IL-2, IL-4, IL-6,Transforming growth factor-β1, Transforming growth factor-β2, Transforming growth factor-β3, and Tumor necrosis factor-α. MRM (for medium-high abundances proteins) is used to quantify abundances of a panel of 147 peptides (corresponding to 77 proteins previously linked to neuropsychiatric disorders)(283). These 147 peptides and corresponding proteins are detailed in the **Appendix**.

Data pre-processing for multiplex immunoassay data is carried out using the methodology already detailed in **Chapter 3**, while pre-processing for MRM data was carried out using the methodology detailed in **Chapter 4**. PCA plots are used to check for data artefacts. Two control samples are identified as outliers in the MRM data, and removed. Linear regression (applied in R as described in **Chapter 2**) is used to quantify protein concentration differences and peptide abundance differences respectively for multiplex immunoassay data and MRM data, between patients and controls while accounting for confounding variables age and gender using stepwise selection. Permutation testing to compute the distribution of p-values, and multiple testing correction to obtain q-values, are applied as described in **5.2.2.4**.

5.2.2.6 Global test analysis of signalling pathways

The epitopes analyzed in **5.2.2.4** are grouped into their associated signalling pathways and Goeman's global test (applied in R as described in **Chapter 2**) with a logistic regression model is applied to test whether the global expression pattern of all the epitopes in each of these pathways is associated with patient-control status. Permutation testing is used to compute p-values, and q-values are obtained through correcting for multiple testing as described in **5.2.2.4**. In this case, the null hypothesis is that patient-control status depends on none of the epitopes in the pathway. The relative contribution of individual epitopes to significantly altered pathways is examined visually through an influence plot which decomposes the overall model test statistic into the contributions made by individual epitopes(348).

5.2.2.7 Screening of microglial activation inhibitors on signalling epitopes

The final stage of this study explores whether the significant changes observed in epitope signalling in response to serum exposure from schizophrenia patients relative to controls (as accrued through the analysis outlined in **5.2.2.4**), are amenable to drug screening. As part of a targeted screening experiment, microglial cells are stimulated at these sites with two compounds known to inhibit microglial activation in previous *in vivo* studies, rapamycin and the tetracycline antibiotic minocycline. Median fluorescent intensities of these sites in response to compound stimulation are measured across six replicates, responses are computed relative to the vehicle condition using the Wilcoxon rank-sum test and adjusted for background fluorescence. Fold changes for each epitope are defined as response to compound/response to vehicle. Q-values are obtained through adjusting the Wilcoxon p-values for multiple testing, as described in **5.2.2.4**.

5.3 Results

5.3.1 Functional characterization of cellular model through detection of M1 vs M2 microglial polarization signals

The functional validity of the cellular model as a basis for detecting M1 versus M2 microglial polarization signals is examined through screening a variety of M1 specific and M2 specific ligands at different titrations across 42 microglial signalling epitopes. Results are displayed in **Figure 5.1**. Responses are observed which are known to be specific to the four M1 ligands (IFN-γ, IL-23, TNF-α and IL-1β) at epitope sites Stat1 (pY701), Stat4 (pY693), NF-κB p65

(pS529) and IκBα, and to the four M2 ligands (IL-4, IL-13, TGF-β and BMP7) at epitope sites Akt (pS473), Stat6 (pY641), Smad2 (pS465/pS467)/Smad3 (pS423/pS425) and PKA RIIα (pS99)(409). In addition, convergent cell signalling responses are seen within both M1 and M2 ligand classes. For example M1 ligands IL-23, TNF-α and IL-1β all induce downregulation of IκBα, and M2 ligands IL-4, IL-13, TGF-β and BMP7 all induce phosphorylation of Stat6 (pY641).

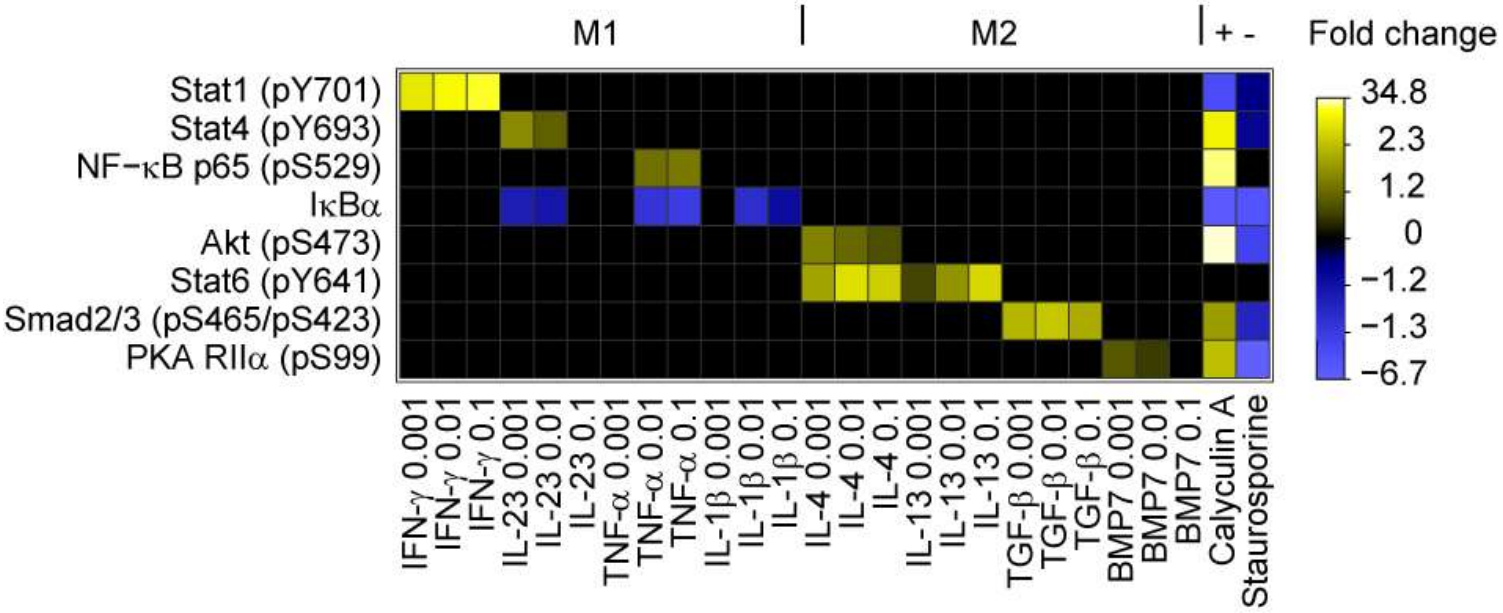


Figure 5.1 Characteristic responses at selected microglial signalling epitopes following stimulation with ligands which specifically induce either M1 or M2 phenotypes. Only significant responses (P<0.05) which survived filtering for background fluorescence are shown. Responses are represented as mean fold change in epitope expression, calculated as mean median fluorescence intensity (MFI) of the ligand/mean MFI of the vehicle across triplicate experiments. For down-regulated epitopes, the legend shows -1/fold change. Legend labels are distributed evenly across the quantile range for negative and positive fold changes separately. Ligands and epitopes are grouped by association to either M1 or M2 microglial signalling.

5.3.2 Microglial exposure to serum from first-onset antipsychotic naive schizophrenia patients

5.3.2.1 Stimulant responses

Median fluorescent intensities of 62 microglial signalling epitopes are measured across 60 "markedly ill" (417) (95.54 ± 20.49 PANSS total score) schizophrenia patients and 79 control samples. Stimulant responses relative to the vehicle condition were measured for patient and control serum, and positive controls calyculin and staurosporine for all epitopes as shown in **Figure 5.2**. There are 39 epitopes which both survived background fluorescence

filtering and showed significant responses relative to the vehicle in at least one clinical group.

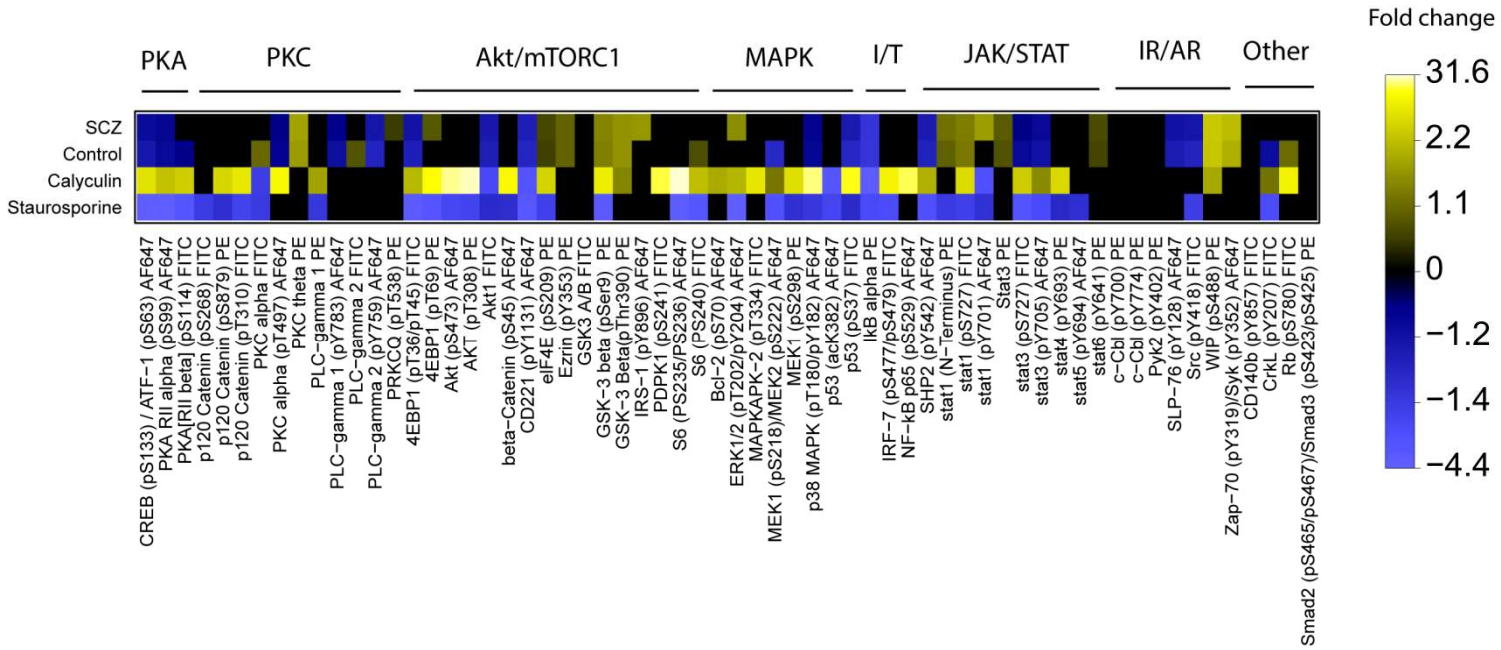


Figure 5.2 Stimulant responses across 62 microglial signalling epitopes for antipsychotive naive schizophrenia patient serum, control serum and positive controls calyculin and staurosporine. Only significant responses (P<0.05) which survived filtering for background fluorescence are shown. Responses are represented as mean fold change in epitope expression, calculated as mean median fluorescence intensity (MFI) of the ligand/mean MFI of the vehicle across triplicate experiments. For down-regulated epitopes, the legend shows -1/fold change. Legend labels are distributed evenly across the quantile range for negative and positive fold changes separately. The 62 epitopes are grouped by signalling pathway class (top).

5.3.2.2 Univariate patient-control analysis

A linear mixed model is applied to the 39 epitopes which show significant responses relative to the vehicle in at least one clinical group in **Figure 5.2**, to quantify differential epitope intensities in response to serum from schizophrenia patients and controls. Six epitopes (4EBP1 (pT36/pT45), 4EBP1 (pT69), stat3 (pY705), SHP2 (pY542), eIF4E (pS209), and Stat3) are found to show significant ($p < 0.05$) differential responses between patient and control serum as shown in **Table 5.3**, with 4EBP1 (pT36/pT45), 4EBP1 (pT69), stat3 (pY705), and SHP2 (pY542) remaining significant ($q < 0.05$) following correction for multiple testing.

These six epitopes are clustered on adjacent proteins in either the Akt/mTORC1 or JAK/STAT pathways, as shown in **Figure 5.3**. The direct mTORC1 substrates 4EBP1 (pT36/pT45) and 4EBP1 (pT69), along with 4EBP1 substrate eIF4E (pS209) comprise the

sites on the Akt/mTORC1 pathway. In addition, the JAK/STAT pathway sites consist of the activation site stat3 (pY705), the regulatory site SHP2 (pY542) and total Stat3 (independent of phosphorylation). The responses to serum from schizophrenia patients relative to controls at these sites suggests an increase in the activation status of both pathways.

Table 5.3 Alterations in microglial signalling epitopes in response to serum exposure from first-onset antipsychotic naive schizophrenia patients relative to controls.

Epitope	P Value	Q Value	Response Direction	Fold Change
4EBP1 (pT36/pT45)	0.0001	0.0001	↑	1.07
4EBP1 (pT69)	0.0001	0.0001	↑	1.07
stat3 (pY705)	0.0012	0.0152	↑	1.04
SHP2 (pY542)	0.0049	0.0466	↑	1.04
eIF4E (pS209)	0.0384	0.2432	↑	1.02
Stat3	0.0357	0.2432	↑	1.04

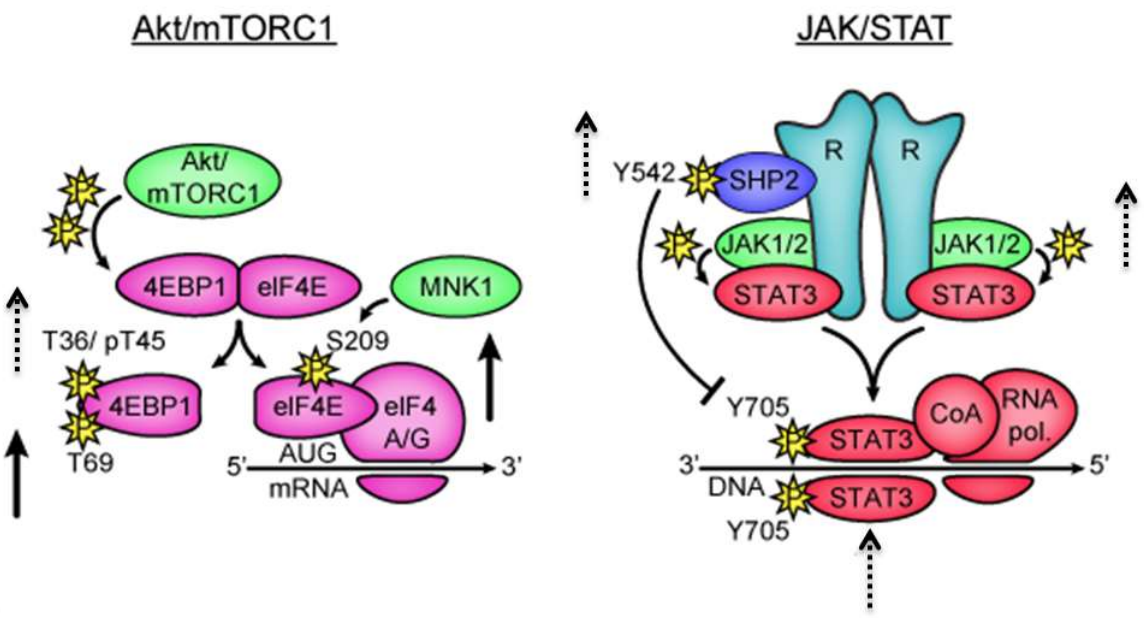


Figure 5.3 Epitopes which responded differently to antipsychotic naive schizophrenia patient serum exposure relative to healthy controls are shown in the mechanistic context of adjacent proteins in Akt/mTORC1 and JAK/STAT signalling pathways. Full vertical arrows represent the increased activation status of epitopes resulting from an increased potentiation of expression changes at those epitopes in response to SCZ serum relative to healthy controls. Dotted vertical arrows represent the increased activation status of epitopes resulting from a reduced attenuation of expression changes at those epitopes in response to SCZ serum relative to healthy controls. In

the Akt/mTORC1 pathway, mTORC1 phosphorylates 4EBP1 at residues T36 and T69, which triggers the dissociation of eIF4E. eIF4E is then phosphorylated by MNK1 at residue S209 and combines with eIF4A and G to initiate cap-dependent translation. In the JAK/STAT pathway, activation of cytokine or growth factor receptors induces the phosphorylation of STAT3 at residue Y705 by JAK1/2 kinases. Subsequent dimerization of STAT3 pY705 recruits transcriptional co-activators which enhance the activity of RNA polymerase at specific genomic loci. Phosphorylation of SHP2 at Y542 negatively regulates STAT3 Y705 phosphorylation. Proteins are coloured with respect to their cellular function: green (kinase), blue (phosphatase), pink (translation), red (transcription) and turquoise (receptor). P - phosphate group.(*Figure reproduced with permission of Dr Santiago Lago).*

The activation of the Akt/mTORC1 and JAK/STAT pathways has been implicated as an important molecular switch controlling the conversion of microglia from a resting phenotype to a deleterious proinflammatory M1-activated phenotype in animal models of neurodegeneration, stroke and traumatic brain injury(413,418–420). To obtain further understanding of the mechanisms within these pathways, implicated by this study, **Table 5.4** shows the serum response for each of the clinical groups at each of the 6 significant epitopes, relative to the vehicle condition, as displayed in **Figure 5.2**. These values show that there is a potentiation of expression changes in schizophrenia patients, compared to controls, for 4EBP1 (pT69), eIF4E (pS209) and Stat3. In addition, they show that the positive fold changes observed in **Table 5.3** for 4EBP1 (pT36/pT45), stat3 (pY705) and SHP2 (pY542), represent an attenuation of expression changes in patients relative to controls at these sites, ie: in controls there is a greater inhibitory effect taking place, which is reduced in patients. The consequential effect of this is a net functional increase in the activation status of the Akt/mTORC1 and JAK/STAT pathways. The over-activation of these signalling pathways in response to schizophrenia patient serum could represent a cellular phenotype driving M1 activation. For example, activation of Stat3 transcription factors downstream of proinflammatory cytokine receptors, is thought to trigger the further synthesis and secretion of characteristic M1 cytokines, for example IL-6, IL-23, IL-1β, TNF-α(413,418). At the same time, Akt/mTORC1 pathway activation through the epitopes 4EBP1 and eIF4E, may represent a metabolic shift towards the translation of proteins implicated in M1 microglial reactivity, for example IL-1β and TNF-α(419,420).

Table 5.4 Serum responses for antipsychotic naive schizophrenia (SCZ) patient serum and control (HC) serum relative to the vehicle condition at each of the six significant (p<0.05) epitopes listed in Table 5.3.

Epitope	SCZ response	HC response
4EBP1 (pT36/pT45)	0.963	0.934
4EBP1 (pT69)	1.061	1.038
stat3 (pY705)	0.988	0.97
SHP2 (pY542)	0.951	0.933
eIF4E (pS209)	1.027	1.01
Stat3	1.062	1.051

5.3.2.3 Multiplex immunoassay and MRM profiling of serum

Multiplex immunoassays and MRM are used to profile 17 serum proteins and 147 peptides (associated with 77 proteins) respectively to analyze whether the observed M1 phenotype could be explained by differences between the clinical groups in the concentrations or abundances of serum analytes.

This reveals significant ($p<0.05$) alterations in the concentrations of cytokines IFN-γ and TGF-β, and the abundances of peptides associated with several proteins from the apolipoprotein family (AII, AIV, CI, CIII and H subtypes), the complement cascade (C1 inhibitor, C4a, C9 and ficolin-3), and various coagulation factors (haptoglobin, antithrombin-III, inter-α-trypsin inhibitor heavy chain H4, α-1-antichymotrypsin and α-2-antiplasmin). However, it should be noted that only peptides associated with haptoglobin remain significant ($q<0.05$) following correction for multiple testing. The full table of results is included in the **Appendix. Figure 5.4** summarizes the differential peptide abundances between patients and controls, for the full panel of 147 peptides profiled in the MRM analysis.

These results are consistent with previously reported findings showing alterations in these serum analytes in schizophrenia patients(63,421). In addition, there is a direct link between these particular proteins and microglia, suggesting that the M1 phenotype of dysregulated microglial activation seen in this study may be induced by circulating proteins, a mechanism which may play a role in the pathogenesis of schizophrenia in the CNS. The cytokines IFN-γ and TGF-β are thought to be mediators of M1 and M2 microglial polarization, respectively(413). In the CNS microglia are the primary cell type expressing receptors for

complement factors (CR1, CR3 and CR4)(421), apolipoprotein (TREM2)(422) and haptoglobin (CD163)(423), and these protein classes have been linked to M1 microglial activation through increased synaptic pruning and neuroinflammation in response to cellular debris at crucial neurodevelopmental stages(421–424). **Table 5.5** summarizes the significant ($p<0.05$) serum analyte findings from both multiplex immunoassay and MRM profiling along with their functional pathways, microglial activation capacity and links to schizophrenia as reported in previous studies.

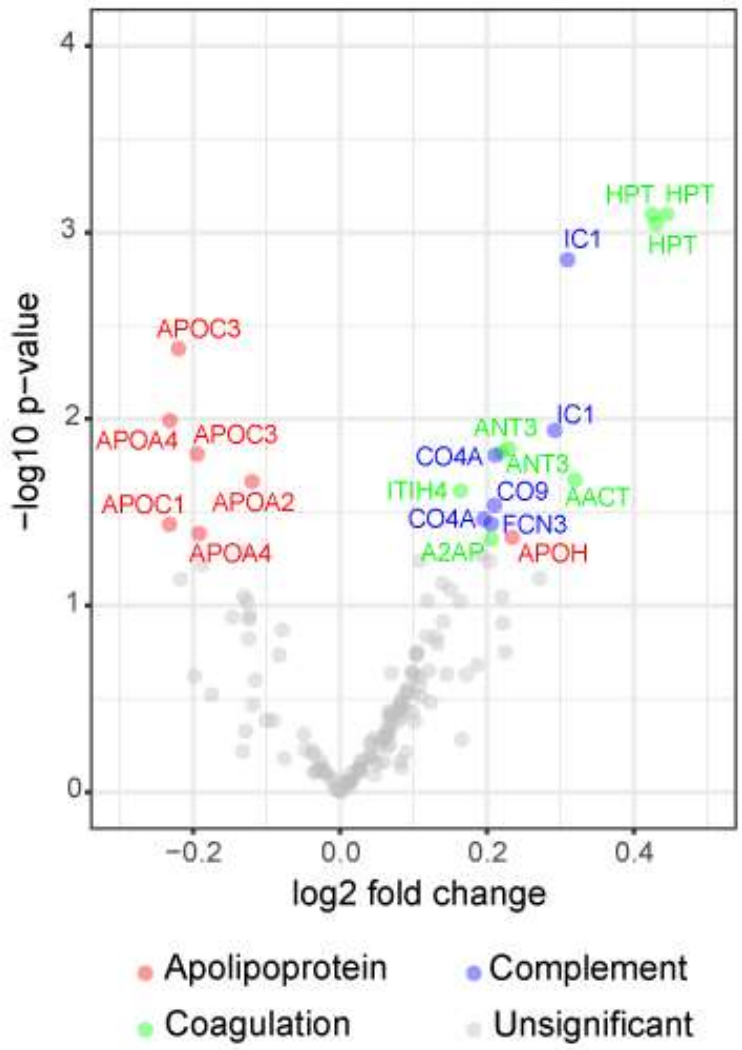


Figure 5.4 Volcano plot illustrating the relationship between log_2 fold change (x-axis) and statistical significance (y-axis) for differential peptide abundance changes between antipsychotic naive schizophrenia patients and controls as measured through multiplex reaction monitoring (MRM) mass spectrometry profiling. Peptides highlighted are significantly different ($p<0.05$) between patients and controls, and labelled with their associated proteins.

Table 5.5 Summary of significant (p<0.05) serum analyte findings from analysis presented in Section 5.3.2.3, along with an overview of associated functional pathways, links to schizophrenia as reported in previous research studies, and links to studies which report whether microglia respond to that particular protein, and how it may be associated with microglial polarization. NA – information not available.

			Functional pathways							Microglial polarization		
Protein	Abbreviation	Direction of change	Blood coagulation	Complement activation	Metabolic process	Acute phase/ inflammatory response	Metal/ ion binding	Lipid binding/ transport	References to SCZ	References to microglia	M1	M2
α-2-antiplasmin	A2AP	↑	✓			✓			(425)	NA		
α-1-antichymotrypsin	AACT	↑	✓		✓	✓			(227,426,427)	(428,429)	✓	
Antithrombin-III	ANT3	↑	✓			✓	✓		(394,425)	NA		
Apolipoprotein A-II	APOA2	↓			✓	✓		✓	(32,425)	(422)		✓
Apolipoprotein A-IV	APOA4	↓			✓	✓		✓	(32,391,425,430)	NA		
Apolipoprotein C-1	APOC1	↓			✓			✓	(32,425)	(431)		✓
Apolipoprotein C-III	APOC3	↓			✓			✓	(425)	NA		
Apolipoprotein H	APOH	↑			✓			✓	(156)	NA		
Complement C4-A	C4A	↑		✓		✓			(421)	(421)	✓	
Complement component C9	C9	↑		✓					(33)	NA		
Ficolin-3	FCN3	↑		✓			✓		(33)	NA		
Haptoglobin	HPT	↑			✓	✓			(391,432,433)	(434)		✓
Plasma protease C1 inhibitor	IC1	↑	✓	✓		✓			(426)	NA		
Interferon-γ	IFNG	↓				✓			(63,435,436)	(408,437)	✓	
Inter-α-trypsin Inhibitor heavy chain H4	ITIH4	↑			✓	✓			(75,438,439)	(440)		✓
Transforming growth factor-β1	TGFB1	↓				✓			(63,436)	(408,441,442)		✓

5.3.2.4 Global test analysis of signalling pathways

The 39 epitopes analyzed in **5.3.2.2** are grouped into six microglial signalling pathways based on the pathway classifications defined in **Table 5.2**; Akt/mTORC1, IR/AR, JAK/Stat, MAPK, PKA, and PKC. Goeman's global test is used to quantify whether the global patterns of the epitopes measured in these pathways was significantly changed in patients relative to controls.

As shown in **Table 5.6**, a significant ($p<0.05$) differential response is seen in the Akt/mTORC1 pathway between patients and controls, based on the combined patterns of responses from the ten epitopes, 4EBP1 (pT36/pT45), 4EBP1 (pT69), Akt1, CD221 (pY1131), eIF4E (pS209), Ezrin (pY353), GSK-3 beta (pSer9), GSK-3 Beta(pThr390), IRS-1 (pY896), and S6 (PS240). However, none of the pathways remain significant ($q<0.05$) following correction for multiple testing.

Table 5.6 Summary of microglial signalling pathway global test results. P-values are based on 10,000 permutations, significance level = 0.05

Pathway	Number of epitopes	P Value	Q Value
Akt/mTORC1	10	0.04	0.24
IR/AR	8	0.4	0.504
JAK/Stat	8	0.42	0.504
MAPK	3	0.12	0.36
PKA	3	0.51	0.51
PKC	7	0.37	0.504

Not all of the ten epitopes may contribute towards this significant ($p<0.05$) response, and thus the influence plot in **Figure 5.5** for the Akt/mTORC1 pathway provides more individual-level information about the contributions of each epitope. It illustrates the relative importance of each epitope to the overall test statistic through their individual p-values as covariates in the model.

As can be observed, while all epitopes are positively associated with schizophrenia, this overall dysregulation of the Akt/mTORC1 pathway appears to be mostly driven by the mTORC1 substrate 4EBP1 (pT69) which makes the greatest contribution (p-value: 0.01).

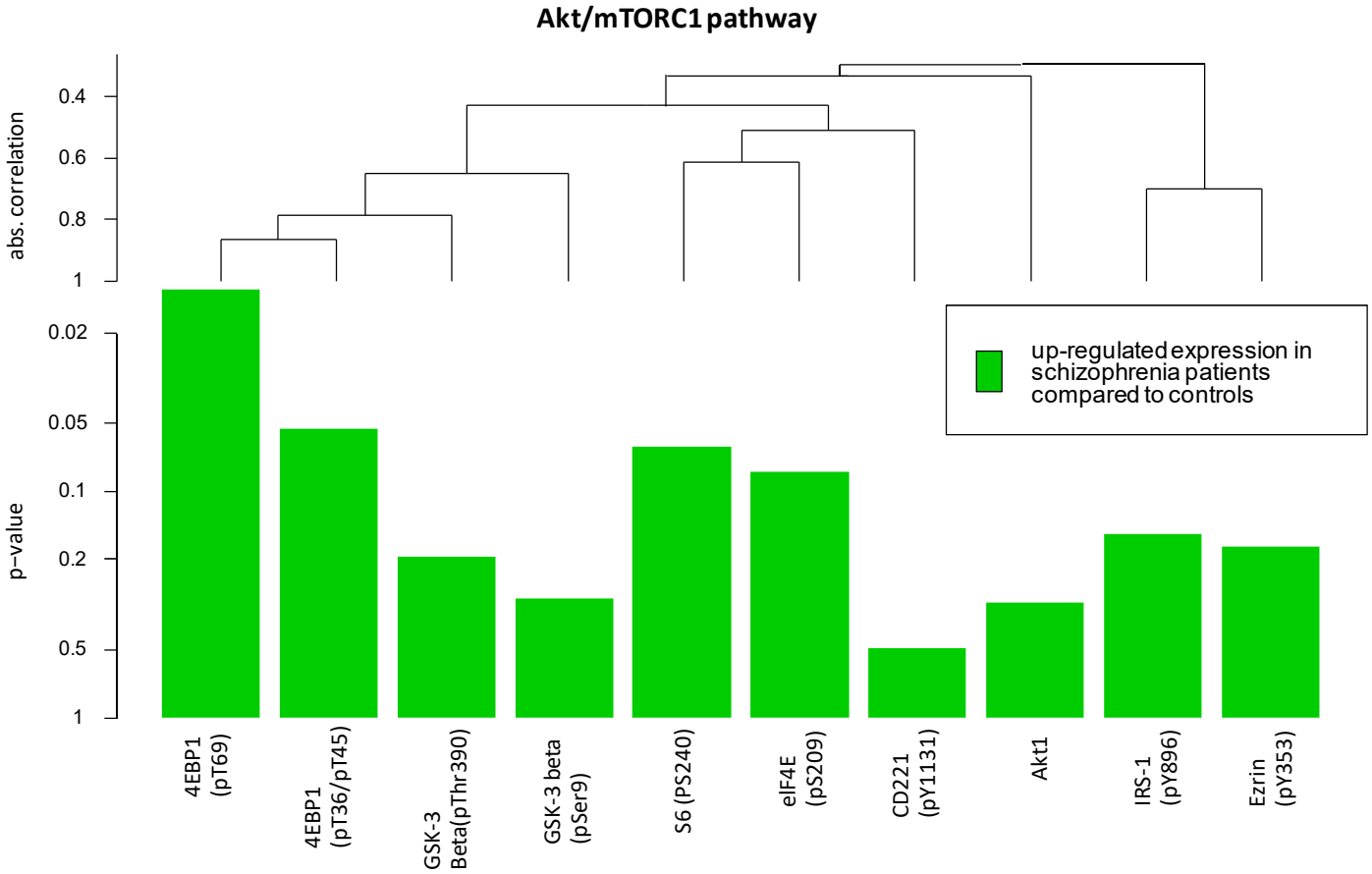


Figure 5.5 Influence plot for the ten epitopes measured in the Akt/mTORC1 pathway. Epitopes are plotted according to hierarchical clustering based on absolute linkage. Absolute correlation distance is the measure used to arrange the epitopes. The height of the bars represents the individual p-value for each epitope as a covariate in the model, thus illustrating its contribution to the overall test statistic.

5.3.2.5 Screening of microglial activation inhibitors on signalling epitopes

The compounds rapamycin and minocycline are screened across the six epitopes which showed significant differential signalling responses between patient and control serum in **5.3.2.2**. **Figure 5.6** shows the responses at each of these epitopes, and **Table 5.7** summarizes the compound responses relative to the vehicle condition.

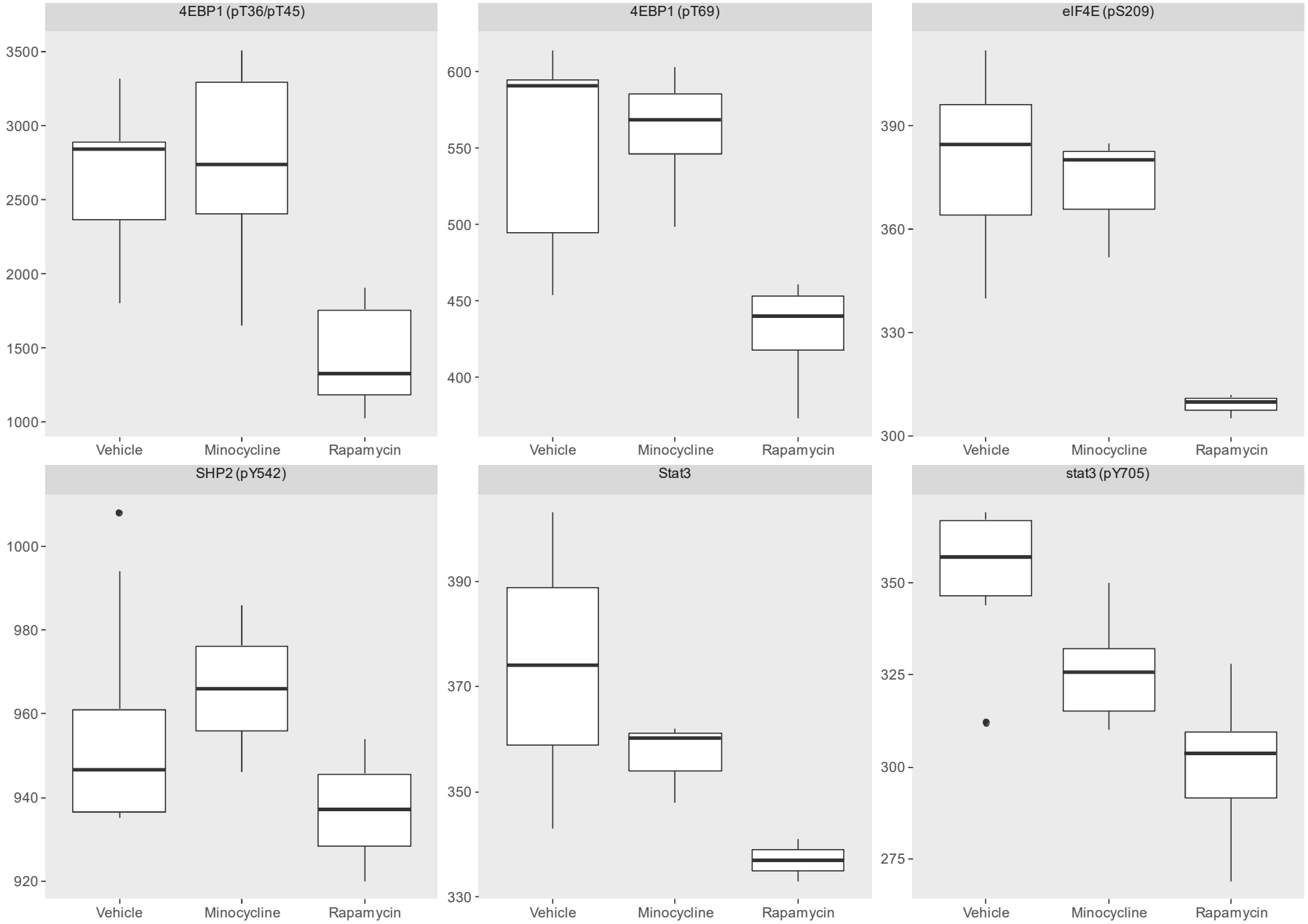


Figure 5.6 Responses to minocycline and rapamycin in comparison to the vehicle condition, for each of the six microglial signalling epitopes which showed differential responses between patient and control serum exposure.

Table 5.7 Compound responses relative to the vehicle condition across the six microglial signalling epitopes which show differential responses between patient and control serum exposure.

	Minocycline			Rapamycin		
Epitope	**Fold Change**	**P Value**	**Q Value**	**Fold Change**	**P Value**	**Q Value**
4EBP1 (pT36/pT45)	0.97	0.71	0.73	0.47	0.004	0.0048
4EBP1 (pT69)	0.96	0.36	0.54	0.75	0.004	0.0048
stat3 (pY705)	0.91	0.03	0.18	0.85	0.002	0.004
SHP2 (pY542)	1.02	0.73	0.73	0.99	0.33	0.33
eIF4E (pS209)	0.96	0.27	0.54	0.81	0.002	0.004
Stat3	0.95	0.11	0.33	0.9	0.002	0.004

Rapamycin and minocycline are selected as they have previously shown to inhibit microglial activation in vivo(233,419,443). Rapamycin displays broad inhibition of the M1 phenotype which is thought to be represented by these findings, significantly ($p<0.05$) inhibiting all epitopes apart from the JAK/Stat pathway regulatory site SHP2 (pY542), responses which remain significant following correction for multiple testing ($q<0.05$). Rapamycin displays particularly strong inhibition of the Akt/mTORC1 pathway epitopes 4EBP1 (pT36/pT45) and 4EBP1 (pT69) which are both phosphorylated mTORC1 substrate sites. This is expected as rapamycin is a known mTOR inhibitor(419). Minocycline displays more selective activity, only significantly inhibiting the expression of the JAK/Stat pathway activation site, phosphorylated stat3 (pY705). These findings suggest that phenotypes identified by this cellular model could form the basis for novel drug discovery approaches for schizophrenia.

5.4 Discussion

In summary, this chapter presents the application of a novel cellular model to investigate microglial signalling alterations in schizophrenia. Microglia are exposed to serum from schizophrenia patients and controls across 62 different epitopes spanning seven key signalling pathways. Epitope expression levels are measured using a flow cytometry platform, and univariate and multivariate statistical techniques are applied to quantify signalling changes between patients and controls. Finally, this cellular model is used to screen two known microglial inhibitors against epitopes found to be differentially expressed in the disease group.

This is the first study to suggest that circulating blood serum from schizophrenia patients can have a direct effect on the intracellular activation phenotypes of microglia in vitro. This is notable as it has been previously hypothesized that circulating serum factors can access the CNS and initiate local immune activation through a variety of mechanisms. These include leakage through the circumventricular organs which lack a BBB(204), disruption of the BBB due to high levels of peripheral inflammation resulting in abnormal trafficking of inflammatory molecules between the periphery and the brain(203), and peripheral interactions between cytokines and vagal nerve afferents which relay cytokines signals to relevant brain regions(444). In addition, activated immune cells such as monocytes and macrophages, can be recruited from the periphery to the brain parenchyma, and these cells can synthesize cytokines in the brain(444). Because of the perivascular localization of microglia to the BBB, it is thought that microglial activation can be initiated through such interactions between peripheral serum factors and the CNS.

As such, this study provides new information regarding potential causality between peripheral and CNS disease mechanisms. While alterations in circulating serum factors have often been considered a secondary effect of CNS abnormalities, the data presented in this study suggests that they may be sufficient, or at least reciprocal, for inducing cell signalling alterations in key CNS cell subpopulations which are involved in the pathogenesis of schizophrenia. Previous studies investigating the implications of dysfunctional JAK/STAT and mTORC1 signalling suggest that this particular interaction between the periphery and CNS may play a profound role in the disease aetiology in at least a subgroup of patients.

5.4.1 Akt/mTORC1 and JAK/STAT activation

One of the strongest findings of this study is the activation of the Akt/mTORC1 pathway as observed through the univariate and multivariate analysis. This appears to be driven by three

particular epitopes, most notably 4EBP1 (pT69), but also 4EBP1 (pT36/pT45) and eIF4E (pS209), although the latter does not remain significant ($q<0.05$) following multiple testing correction. Previous studies have repeatedly associated the activation of this pathway with the deleterious consequences of microglial M1 polarization in preclinical models of neurodegeneration(445–447). While Akt/mTORC1 signalling alone is not thought to be sufficient to initiate microglial activation, it is believed to act as a metabolic priming of key translational proteins, resulting in an exaggerated response when subsequently presented with particular proinflammatory stimuli(448). In addition, there is evidence of an association between increased activation of Akt/mTORC1 signalling and impairments in autophagy(449). It has been hypothesized that the latter play a role in the pathogenesis of psychiatric disorders following studies which identified a comparative reduction in autophagic regulator beclin-1 expression in post-mortem brain samples from schizophrenia patients, compared to controls(450). Induction of autophagy has also been suggested as a potential means of improving efficacy of various pharmacological treatments for MDD and bipolar disorder(451,452).

Activation of the JAK/STAT pathway is also observed through increased signalling at stat3 (pY705), SHP2 (pY542) and total Stat3 (independent of phosphorylation) in schizophrenia patients. The multivariate analysis results indicate that this may not be as strong a response as the Akt/mTORC1 pathway activation, but more epitopes would need to be measured along both of these pathways to reach a definite conclusion. The increased expression levels of these JAK/STAT epitopes is particular of note because Stat3 transcription factor activation in microglia is mainly associated with synthesis and secretion of proinflammatory cytokines which mediate M1 polarization such as IL-6, IL-23, IL-1β and TNF-α(406). These proinflammatory cytokines can additionally trigger further activation of Stat3 downstream of their receptors. As such, studies of stroke, neurodegeneration and traumatic brain injury have described Stat3 as a key molecular switch which mediates microglial transition from a resting state to an M1 activation state(419,420,453).

However increased Stat3 signalling is not purely restricted to proinflammatory cytokine secretion. Recent rodent studies have suggested that Stat3 is particularly highly expressed in the brain relative to the other Stat isotypes(454), and is strongly increased in CNS subpopulations at key stages of postnatal development associated with critical periods of synaptic pruning(455). In rodent models of MDD, Stat3 knockout specifically in microglia cells has been found to lead to increased synaptic plasticity and anti-depressive like behavioural endophenotypes(456). These findings, combined with the results of this study, suggest that the role of Stat3 activation, provoked by circulating serum factors, during critical

neurodevelopmental stages warrants further investigation, in the context of schizophrenia aetiology(457).

While it may be noted that signalling pathways such as Akt/mTOR and JAK/STAT are ubiquitous, their implication in this study in the context of schizophrenia is particularly informative as it adds cumulative evidence to previous hypotheses. In particular, recent publications have postulated a novel mTOR based hypothesis of the neuropathology of schizophrenia(458). As discussed in **Chapter 1**, there is evidence for a neurodevelopmental component to the pathophysiology of schizophrenia, and previous research has found that genetic and epigenetic factors can disrupt mTOR signalling, with adverse effects on neuronal growth and connectivity, a risk factor for the onset of schizophrenia in adulthood(458,459). As such it is particular notable that altered Akt/mTOR signalling has been observed in a model of the disease created using a CNS cell population. Likewise, the critical role of JAK/STAT signalling in the CNS for functions such as synaptic plasticity, means that many studies have hypothesised that dysregulation of the JAK/STAT pathway is at the heart of most brain disorders(460). This study is the first to imply that altered JAK/STAT signalling in microglial cells may play a role in the development of schizophrenia.

5.4.2 Alterations in circulating serum proteins

Immunoassay and mass spectrometry profiling of schizophrenia patient and control serum resulted in several findings which validate recently published studies. In particular, increased complement C4-A and complement component C9 in patient serum relative to controls matches the findings of several studies which have found associations between the complement cascade and schizophrenia. Variation at the C4 locus is one of the known genetic risk factors for schizophrenia(26), and this has been found to correlate with impaired C4-A expression in the brain and serum of patients(25,395,461). In addition, rodent studies have linked increased C4 with increased microglial synaptic pruning(462,463), which has been postulated as a potential mechanism in the development of schizophrenia during late adolescence and early adulthood(464,465).

In addition, alterations were found in several apolipoproteins, another finding that validates previous studies. Alterations in apolipoproteins in schizophrenia patient serum have been reported in various analyses, along with significant correlations between apolipoprotein levels and cognitive deficits and alterations in hippocampal brain volume(466). Furthermore, a rodent model of neurodegeneration found that overexpression of microglial apolipoprotein receptor TREM2, resulted in a neuroprotective shift in microglia towards M2 polarization, and a reduction in proinflammatory cytokine secretion(422). As such, we can hypothesize that

the reductions in the levels of Apolipoprotein A-2, A-4, C-1 and C-III observed in schizophrenia patient serum in this study, may possibly reflect a compromised neuroprotective profile in the CNS.

It should be noted that caution must be drawn regarding the interpretation of these serum alterations, as few of the significant ($p<0.05$) findings survive multiple testing correction, save for peptides associated with haptoglobin. However, because many of these differentially expressed analytes have already been identified in other biomarker studies of schizophrenia, further work in understanding the potential functional relationships between apolipoproteins, the complement cascade, and Akt/mTORC1 and JAK/STAT signalling in microglia, may prove worthwhile.

5.4.3 Inhibitory profiles of rapamycin and minocycline

Rapamycin and minocycline are both compounds which have previously been associated with inhibitory effects on microglial activation, and rapamycin has been proposed as a potential novel treatment for schizophrenia following the results of clinical trials in disorders with high schizophrenia comorbidity such as autism spectrum disorder(412). The potential efficacy of minocycline as a treatment for schizophrenia is currently being investigated in various clinical trials(467). Minocycline has already been used as an add-on treatment for schizophrenia and meta-analyses have suggested that it displays efficacy at treating the negative symptoms of the disease, an aspect which is known to be under-addressed by existing antipsychotics(468).

The findings of this study suggest that rapamycin in particular can be used to target epitopes which are overactivated by schizophrenia serum exposure and thus potentially inhibit microglial M1 polarization. While minocycline has a more selective profile, it is of particular interest that minocycline significantly ($p<0.05$) inhibits stat3 (pY705) signalling, although this finding does not survive adjustment for multiple testing. Many studies have suggested that Stat3 in microglia may represent a crucial target for addressing treatment resistant negative symptomatology common to schizophrenia and other psychiatric disorders. Stat3 may be one of the mechanisms by which minocycline achieves its reported therapeutic benefits in this regard, and thus this may be worthy of further investigation.

Overall, the findings of this study, and the fact that microglia have been increasingly implicated in the pathogenesis of psychiatric disorders and treatment response, suggests that they have the potential to provide an underexploited target for novel treatments with the potential for modification of the disease course, as is discussed further in **Chapter 9**.

Chapter 6 Analysis of data from functional imaging and a cellular model of schizophrenia to investigate microglial signalling in antipsychotic treated schizophrenia patients

6.1 Introduction

To date, the only available methodology for investigating microglia cells *in vivo* is the application of PET imaging which can be used to quantify TSPO expression(469). Both altered microglia density and the microglial activation phenotype are associated with elevated TSPO levels(302), which can be assessed by measuring TSPO binding with various TSPO PET ligands across various brain regions(470). However so far, PET studies investigating TSPO binding in schizophrenia have yielded inconsistent results(302). Among the reasons suggested for this are methodological differences, and variation in the stage of the disease for each cohort of patients. Some studies have analyzed individuals in the prodromal phase(401), while others have examined patients in early stages of the disease(470) or chronic antipsychotic treated patients(471). In addition, the ability to interpret these studies and draw conclusions about microglial function has been limited due to the absence of any orthogonal experimental methodologies. As such it has not been possible to make inferences regarding alterations in microgial signalling pathways or patient inflammatory profiles potentially associated with these brain changes(302).

This study aims to investigate microglial activation in a cohort of recent onset antipsychotic treated schizophrenia patients and healthy controls by using PET to measure BPND in total gray matter and multiple gray matter regions of interest. In addition, alterations in the inflammatory status of these patients are examined through the use of targeted mass spectrometry to profile serum samples obtained from the study participants at the time of PET imaging, by measuring a panel of analytes known to be associated with psychiatric disorders. Concurrently, a human microglial cell line was exposed to serum from these patients and controls, through the same methodology outlined in **Chapter 5**. This cellular model is utilized to dissect the cell signalling pathways which may predispose microglia to phenotypic switching in these patients.

An additional aim of this study is to utilize this cellular model to further investigate microglial signalling differences in antipsychotic treated schizophrenia patients relative to controls, by exposing microglia to serum from two separate cohorts of treated patients and controls. Univariate and global test analyses are conducted on the data from the individual cohorts before both datasets are combined in a meta analysis to increase the statistical power.

As indicated by the results of **Chapters 3 and 4**, it is likely that the levels of serum analytes in recent-onset antipsychotic treated patients may be different to those of first onset antipsychotic naive patients. Hence it is necessary to characterize CNS processes in this patient group. In **Chapter 5**, signalling alterations were identified which pointed to microglial activation in antipsychotic naive patients, and findings from both rodent model studies(472), and human PET studies(473) have suggested that chronic exposure to antipsychotic medication may stimulate microglial activation in some recent-onset treated schizophrenia patients. This suggests that antipsychotics could in fact perpetuate neuroinflammation in this patient group, possibly through different intracellular mechanisms to those observed in antipsychotic naive patients. As such, the final part of this study looks to provide more information on whether antipsychotics themselves appear to be directly associated with the observed signalling alterations. The individual epitopes found to be differentially expressed in treated patients compared to controls, are targeted with a range of first and second generation antipsychotics to explore whether they induce the same direction of expression changes relative to the vehicle condition.

6.2 Methods

R packages, functions and settings were applied as in **Chapter 5**, unless otherwise stated.

6.2.1 Clinical samples

Serum samples were collected from 53 individuals recruited from two clinical centres in the Netherlands as shown in **Table 6.1.**

Cohort 1 consists of 15 recent onset schizophrenia patients treated with a mixture of first and second generation antipsychotics, and 17 healthy controls recruited from the Department of Psychiatry, University Medical Centre Utrecht, The Netherlands. Patients displayed moderate symptoms at the time of serum collection, and the average duration of antipsychotic treatment was one year. Cohort 2 consists of 9 recent onset schizophrenia patients and 12 controls recruited from Erasmus University Medical Centre, Rotterdam, The Netherlands. Patients displayed moderate symptoms at the time of serum collection, and had been treated with olanzapine for six weeks.Diagnoses of schizophrenia were based

according to the DSM-4. The same exclusion criteria described in **Chapter 5** was applied to patient and control groups.

Table 6.1 Demographic characteristics and determination of clinical samples for the cohorts analyzed in Chapter 6. Values are presented as average ± standard deviation.

	Cohort 1	**Cohort 2**
Centre	Utrecht	Rotterdam
Number	17 CT 15 SCZ	12 CT 9 SCZ
Age (years)	26 ± 4 25 ± 4	28 ±8 27 ± 7
Sex (M/F)	14/3 12/3	12/0 9/0
PANSS Positive	NA 11 ± 3	NA 15 ± 5
PANSS Negative	NA 14 ± 3	NA 19 ± 5
PANSS General	NA 27 ± 6	NA 31 ± 8
PANSS Total	NA 53 ± 9	NA 64 ± 16
Antipsychotic treatment (SCZ only)		
Clozapine	3	0
Risperidone	2	0
Olanzapine	5	9
Other	5	0

6.2.2 PET imaging

PET imaging is conducted for all patients and controls from cohort 1, using the method described in **Chapter 2**. PET is used to measure the BPND of TSPO ligand (R)-[^{11}C]PK11195 across total gray matter and five gray matter regions of interest (frontal cortex, temporal cortex, parietal cortex, striatum and thalamus). These regions of interest are selected based on widespread TSPO availability in the brain.

A linear regression model is used to explore whether total gray matter or any of the five gray matter regions showed significantly ($p < 0.05$) changed TSPO binding between patients and controls. The analysis accounts for age and gender, covariates which are selected using

stepwise selection (described in **Chapter 2**). Q-values are obtained by adjusting the p-values for multiple testing, as described in **Chapter 5**.

6.2.3 Cellular model data pre-processing

The cellular model described in **Chapter 5 i**s applied to cohorts 1 and 2 in separate experiments. For each cohort, a human microglial cell line is exposed to serum from patients and controls and the median fluorescent intensities of 62 cellular epitopes (listed in **Chapter 5**), are measured across all samples with 3 replicates, using flow cytometry. Pre-processing steps are carried out as described in **Chapter 5**.

6.2.4 Stimulant responses and univariate analysis

For both cohorts 1 and 2, stimulant (patient or control serum) responses relative to the vehicle condition, for each of the 62 epitopes, are computed using the Wilcoxon rank-sum test. Only responses which were significant ($p < 0.05$), and survive adjustment for background autofluorescence are considered. Positive controls calyculin and staurosporine are included for comparison as in **Chapter 5**. Univariate analysis is subsequently conducted on epitopes which displayed significant ($p < 0.05$) serum responses relative to the vehicle in at least one clinical group, using a linear mixed model as described in **Chapter 5**. Fold change definitions are the same as in **Chapter 5**.

6.2.5 Global test analysis of signalling pathways

For cohorts 1 and 2, the epitopes analyzed in **6.2.4** are grouped into their associated signalling pathways and Goeman's global test (described in **Chapter 2**) with a logistic regression model is used to test whether the global expression pattern of all the epitopes in each of these pathways is associated with patient-control status as described in **Chapter 5**. The relative contribution of individual epitopes to significantly ($p < 0.05$) altered pathways is examined visually through an influence plot which decomposes the overall model test statistic into the contributions made by individual epitopes. Q-values are obtained by adjusting the p-values for multiple testing, as described in **Chapter 5**.

6.2.6 Correlation analysis

For patients in cohort 1, correlation analysis is conducted between PANSS scores representing clinical severity, and PET regions in which TSPO BPND was found to be significantly ($p < 0.05$) different between patients and controls. Correlations are computed

using the Pearson correlation coefficient. Similarly, correlation analysis is conducted between PANSS scores, and epitope expression for epitopes either found to be significantly ($p < 0.05$) different between patients and controls (in cohort 1) through univariate analysis or found to be significantly ($p < 0.05$) contributing to differences between patients and controls in a significantly altered pathway in the global test analysis. Finally correlations are conducted between these epitopes and PET regions. Correlations are computed separately for control and patient groups. Q-values are obtained by adjusting the correlation p-values for multiple testing, as described in **Chapter 5**.

6.2.7 MRM profiling of serum

To explore whether the cell signalling phenotypes observed in cohorts 1 or 2 can be explained by differences between patients and controls in the relative abundances of serum proteins associated with psychiatric disorders, MRM is once again used to quantify abundances of the panel of 147 peptides utilized in **Chapter 5**. Both data pre-processing and linear regression analysis are carried out as described in **Chapter 5**.

6.2.8 Meta analysis

Due to the relatively small sample sizes of cohorts 1 and 2, the two cohorts are analyzed together as a meta-cohort to investigate whether the gain in statistical power yielded further insight regarding microglial signalling alterations and serum alterations between antipsychotic treated patients and controls. Univariate and global test analysis was conducted on the meta-cohort using the same methodologies applied in **6.2.4**, **6.2.5** and **6.2.7**. The ComBat algorithm (described in **Chapter 2**) was utilized to normalize batch effects due to technical variation resulting from the cellular model experiments for cohorts 1 and 2 being run at different times.

6.2.9 Microglial exposure to antipsychotics at key signalling epitopes

The final stage of this study looks to provide further evidence towards whether antipsychotics are involved in inducing the observed cell signalling alterations. As part of a targeted screening experiment, microglial cells are stimulated at epitopes found to be significantly ($p < 0.05$) altered between patients and controls in univariate and multivariate analyses across cohorts 1, 2 and the meta-cohort, with several common first and second generation antipsychotics. Median fluorescent intensities of these sites in response to

compound stimulation are measured across six replicates, responses were computed relative to the vehicle condition using the Wilcoxon rank-sum test and adjusted for background fluorescence. Fold changes for each epitope are defined as response to antipsychotic/response to vehicle. Q-values are obtained by adjusting the Wilcoxon p-values for multiple testing, as described in **Chapter 5**.

6.3 Results

6.3.1 PET imaging

Following PET imaging scans of patients and controls in cohort 1, (R)-[^{11}C]PK11195 BPND was quantified across total gray matter and five gray matter regions – frontal cortex, temporal cortex, parietal cortex, striatum and thalamus. BPND, and thus TSPO expression, is found to be significantly ($p < 0.05$) increased in the temporal cortex only, as shown in **Table 6.2**, although this result does not survive multiple testing correction.

Table 6.2 Differences in (R)-[^{11}C]PK11195 binding potential (BPND) between patients and controls in cohort 1 across total gray matter and five gray matter regions.

	Binding potential			
Region of interest	**Control**	**SZ**	**P Value**	**Q Value**
Frontal	0.12±0.07	0.13±0.07	0.37	0.39
Temporal	0.08±0.09	0.13±0.06	**0.04**	0.24
Parietal	0.15±0.07	0.16±0.07	0.35	0.39
Striatum	0.09±0.11	0.10±0.10	0.39	0.39
Thalamus	0.20±0.12	0.13±0.13	0.19	0.38
Total grey matter	0.14±0.09	0.15±0.06	0.14	0.38

6.3.2 Stimulant responses

Median fluorescent intensities of 62 microglial signalling epitopes are measured across patient and control samples in cohorts 1 and 2. Stimulant responses relative to the vehicle condition are measured for patient and control serum, and positive controls calyculin and staurosporine for all epitopes as shown in **Figure 6.1.** There are 25 epitopes which both survive background fluorescence filtering and show significant responses relative to the vehicle in at least one clinical group for cohort 1, and 34 epitopes for cohort 2.

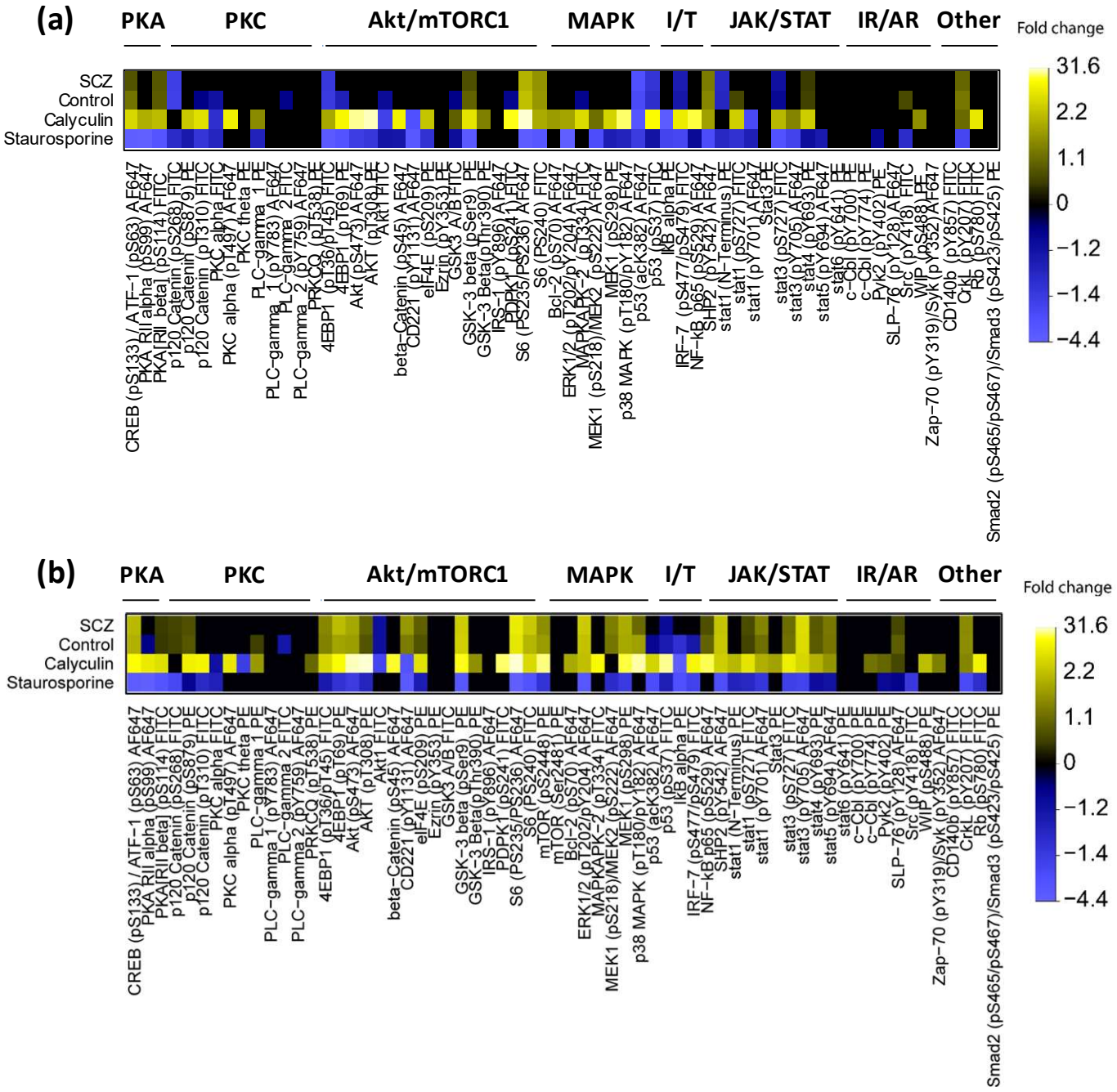


Figure 6.1 Stimulant responses for cohort 1 (a) and cohort 2 (b) across 62 microglial signalling epitopes for microglial exposure to antipsychotic treated schizophrenia patient serum, control serum and positive controls calyculin and staurosporine. Only significant responses (P<0.05) which survived filtering for background fluorescence are shown. Responses are represented as mean fold change in epitope expression, calculated as mean median fluorescent intensity (MFI) of the ligand/mean MFI of the vehicle across triplicate measurements. For down-regulated epitopes, the legend shows -1/fold change. Legend labels are distributed evenly across the quantile range for negative and positive fold changes separately. The 62 epitopes are grouped by signalling pathway class.

6.3.3 Univariate patient-control analysis

For cohorts 1 and 2, a linear mixed model is applied to the epitopes which show significant responses relative to the vehicle in at least one clinical group in **Figure 6.1**, to quantify differential epitope intensities in response to serum exposure from schizophrenia patients and controls. For cohort 1, the epitopes (stat3 (pS727) and CrkL (pY207)) are found to show significant ($p<0.05$) differential responses between patient and control serum, and for cohort 2, the epitope p38 MAPK (pT180/pY182) was found to be significant ($p<0.05$) between patients and controls as summarized in **Table 6.3**, although none of these findings remain significant ($q<0.05$) following adjustment for multiple testing. This table also displays the serum response for each of the clinical groups at both significant epitopes, relative to the vehicle condition.

Table 6.3 Alterations in microglial signalling epitopes for cohorts 1 and 2 in response to serum from recent-onset antipsychotic treated schizophrenia (SCZ) patients relative to healthy controls (HC). SCZ response and HC response refers to the serum response for each of the clinical groups relative to the vehicle condition.

	Epitope	P Value	Q Value	Response Direction	Fold Change	SCZ response	HC response
Cohort 1	**stat3 (pS727)**	0.025	0.451	↓	0.95	0.88	0.92
	CrkL (pY207)	0.036	0.451	↓	0.96	1.1	1.15
Cohort 2	**p38 MAPK (pT180/pY182)**	0.03	0.935	↓	0.87	1.12	1.28

The fold changes in **Table 6.3** show a downregulated response in patients compared to controls at the indicated epitope sites. Mechanistically, the serum response values at these sites suggest the changes may be due to a potentiation of expression changes in schizophrenia patients at stat3 (pS727), and an attenuation of expression changes in patients at CrkL (pY207) and p38 MAPK (pT180/pY182).

6.3.4 Global test analysis of signalling pathways

For cohort 1, 25 epitopes are analysed in **6.3.3**, and for cohort 2, 34 epitopes are analysed. These epitopes are grouped into microglial signalling pathways based on the pathway classifications defined in **Chapter 5**. Goeman's global test is used to quantify whether the global patterns of the epitopes measured in these pathways were significantly changed in patients relative to controls.

As shown in **Table 6.4**, a significant (q<0.05) differential response is seen in the JAK/STAT pathway in cohort 1, based on the combined patterns of responses from the five epitopes stat3 pS727, stat4 pY693, stat1 pS727, SHP2 pY542, and stat1 N Terminus. No significant responses are found in the pathways analysed in cohort 2.

Table 6.4 Summary of microglial signalling pathway global test results for cohorts 1 and 2. P-values are based on 10,000 permutations, significance level = 0.05

Cohort 1				Cohort 2			
Pathway	**Number of epitopes**	**P Value**	**Q Value**	**Pathway**	**Number of epitopes**	**P Value**	**Q Value**
Akt/mTORC1	8	0.19	0.32	**Akt/mTORC1**	12	0.31	0.78
JAK/Stat	5	0.0017	0.0085	**IL1R/TLR**	3	0.09	0.63
MAPK	3	0.49	0.49	**IR/AR**	2	0.56	0.78
PKA	2	0.16	0.32	**JAK/Stat**	8	0.92	0.92
PKC	4	0.3	0.38	**MAPK**	6	0.43	0.78
				PKA	3	0.68	0.79
				PKC	5	0.55	0.78

The relative contributions of the five JAK/STAT pathway epitopes towards this significant response is summarized in the influence plot in **Figure 6.2**. The overall dysregulation of this pathway appears to be strongly associated with the changes observed in three epitopes, stat3 pS727 (p-value: 0.001), SHP2 pY542 (p-value: 0.001), and stat4 pY693 (p-value: 0.003). As in the univariate analysis, the responses of each epitope are downregulated in treated patients compared to controls. The stimulant responses relative to the vehicle from **Figure 6.1** point to a potentiation of expression in patients at stat3 pS727, and an attenuation of expression at SHP2 pY542, and stat4 pY693.

6.3.5 Correlation analysis

Correlations are computed between PET regions, epitopes found to be significant in either the univariate or global test analyses on cohort 1, and PANSS scores. The significant ($p<0.05$) findings are summarized in **Table 6.5**. A positive correlation is found between TSPO expression in the temporal cortex and PANSS general, while negative correlations are found between TSPO expression in the temporal cortex and stat3 (pS727) expression for both clinical groups. The latter correlations remain significant ($q<0.05$) following adjustment for multiple testing.

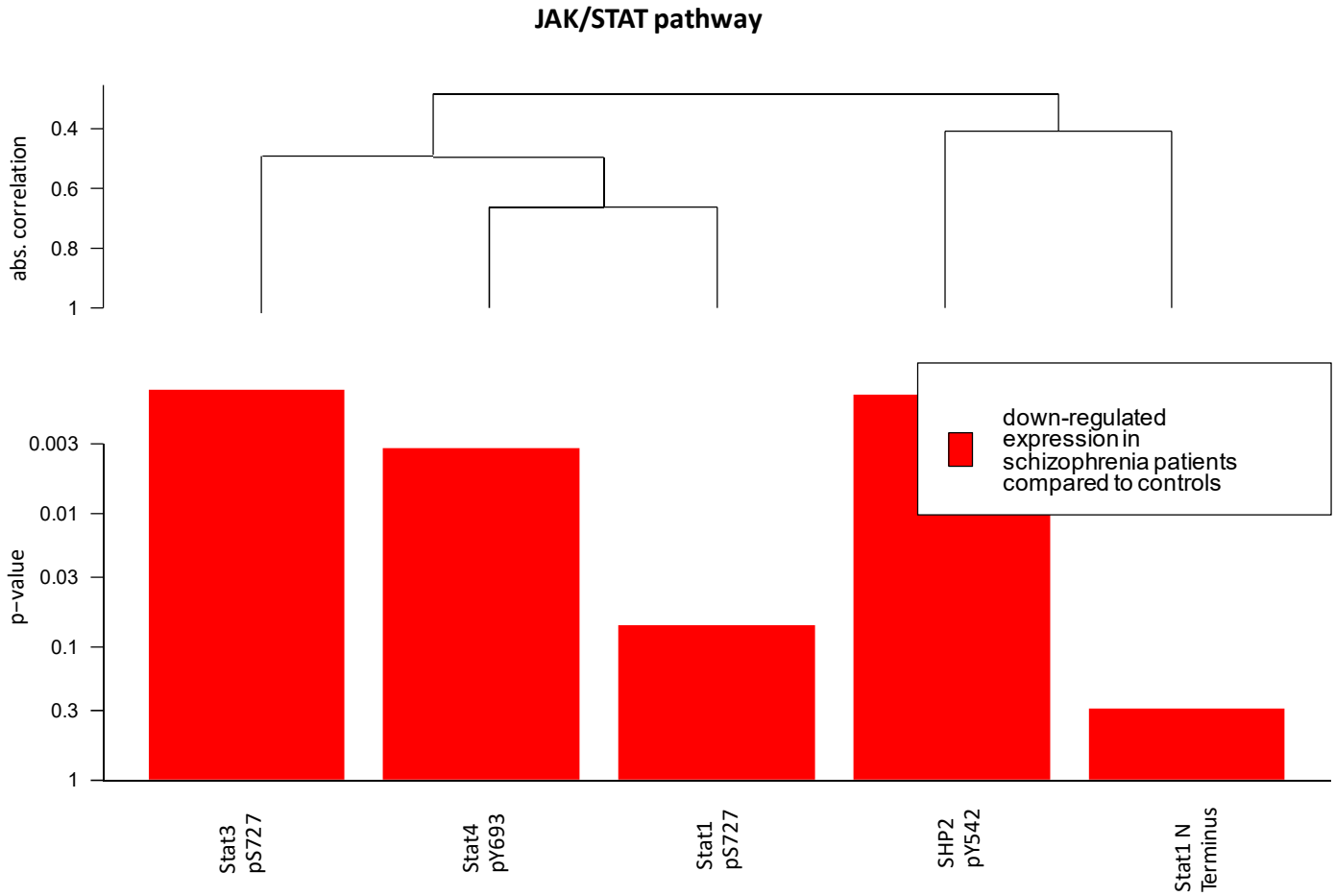


Figure 6.2 Influence plot for the five epitopes measured in the JAK/STAT pathway in cohort 1. Epitopes are plotted according to hierarchical clustering based on absolute linkage. Absolute correlation distance is the measure used to arrange the epitopes. The height of the bars represents the individual p-value for each epitope as a covariate in the model, thus illustrating its contribution to the overall test statistic.

Table 6.5 Summary of correlation analysis between PET regions found to have significant (p<0.05) TSPO expression between patients and controls, epitopes identified as significant through univariate or multivariate analysis, and PANSS scores.

		Group	Pearson Coefficient	P Value	Q Value
Temporal	**PANSS general**	SZ	0.54	0.046	0.18
Temporal	**stat3 (pS727)**	SZ	-0.64	0.01	0.03
Temporal	**stat3 (pS727)**	Ctrl	-0.72	0.001	0.003

6.3.6 MRM profiling of serum

MRM is used to profile 147 peptides in the serum samples from cohort 1 and cohort 2 to analyze whether the phenotypes observed in each cohort could be explained by differences in the abundances of serum analytes between patients and controls. The results are detailed in **Table 6.6**. For cohort 1, significant (p<0.05) changes are observed in the abundances of 15 peptides including those associated with proteins from the apolipoprotein family (L1, C1, E and H subtypes), and various coagulation factors (A1AT, HEP2, HPT and HRG). For cohort 2, only the antibody IGHG2 shows significant differential abundance between patients and controls. None of these findings remain significant following adjustment for multiple testing.

Table 6.6 Significant (p<0.05) alterations in peptide abundances between antipsychotic treated schizophrenia patients and healthy controls as measured through multiplex reaction monitoring (MRM) mass spectrometry in cohorts 1 and 2. The associated proteins are included alongside each peptide sequence.

	Cohort 1					Cohort 2			
Protein	**Peptide sequence**	**P Value**	**Q Value**	**Fold change**	**Protein**	**Peptide sequence**	**P Value**	**Q Value**	**Fold change**
Apolipoprotein L-1	VNEPSILEMSR	0.0011	0.077	1.41	**Immunoglobulin heavy constant gamma 2**	GLPAPIEK	0.004	0.616	0.44
Apolipoprotein L-1	LNILNNNYK	0.0011	0.077	1.58					
Apolipoprotein L-1	VTEPISAESGEQVER	0.0061	0.266	1.43					
Histidine-rich glycoprotein	DSPVLIDFFEDTER	0.0076	0.266	1.33					
Heparin cofactor 2	IAIDLFK	0.0158	0.329	1.46					
α-1 antitrypsin	SPLFMGK	0.0200	0.329	1.52					
α-1 antitrypsin	SVLGQLGITK	0.0208	0.329	1.71					
Apolipoprotein F	SLPTEDCENEK	0.0220	0.329	0.62					
Apolipoprotein C-I	EWFSETFQK	0.0229	0.329	0.72					
Histidine-rich glycoprotein	ADLFYDVEALDLESPK	0.0246	0.329	1.27					
Alpha-2-HS-glycoprotein	HTLNQIDEVK	0.0259	0.329	1.21					
Haptoglobin	VGYVSGWGR	0.0426	0.398	1.55					
Pigment epithelium-derived factor	TVQAVLTVPK	0.0442	0.398	1.29					
Immunoglobulin heavy constant gamma 2	GLPAPIEK	0.0443	0.398	0.58					
Apolipoprotein E	AATVGSLAGQPLQER	0.0483	0.398	1.50					

6.3.7 Meta analysis

To obtain further information on microglial signalling differences between antipsychotic treated patients and controls, cohorts 1 and 2 are combined to increase statistical power and a meta analysis was conducted. The ComBat algorithm is used to remove batch differences resulting from technical variation between the two cohorts. 19 epitopes are measured across both cohorts. Univariate and global test analysis is computed as in **6.3.3** and **6.3.4**. Results are summarized in **Table 6.7**.

Table 6.7 Summary of univariate and global test results for the meta analysis on cohorts 1 and 2.

	Epitope	P Value	Q Value	Response Direction	Fold Change
Univariate analysis	**stat4 (pY693)**	0.03	0.31	↓	0.96
Global test analysis	**Pathway**	**P Value**	**Q Value**	**Number of Epitopes**	
	Akt/mTORC1	0.61	0.76	7	
	JAK/STAT	0.02	0.1	4	
	MAPK	0.96	0.96	2	
	PKA	0.38	0.76	2	
	PKC	0.58	0.76	2	

As in the univariate and global test analysis of cohort 1, the results show alterations in the JAK/STAT pathway, although in the univariate meta-cohort analysis, stat4 (pY693) is found to be significantly ($p<0.05$) different between patients and controls, rather than stat3 (pS727).

The influence plot for the global test analysis on the meta-cohort in **Figure 6.3** again shows that, as in cohort 1, the key individual epitopes driving the significant ($p<0.05$) dysregulation of the JAK/STAT pathway are stat4 (pY693), stat3 (pS727), and SHP2 (pY542), except that this time stat4 (pY693) is the most important driver, while in cohort 1 alone, the three epitopes appear to have relatively equal contributions to the observed effect.

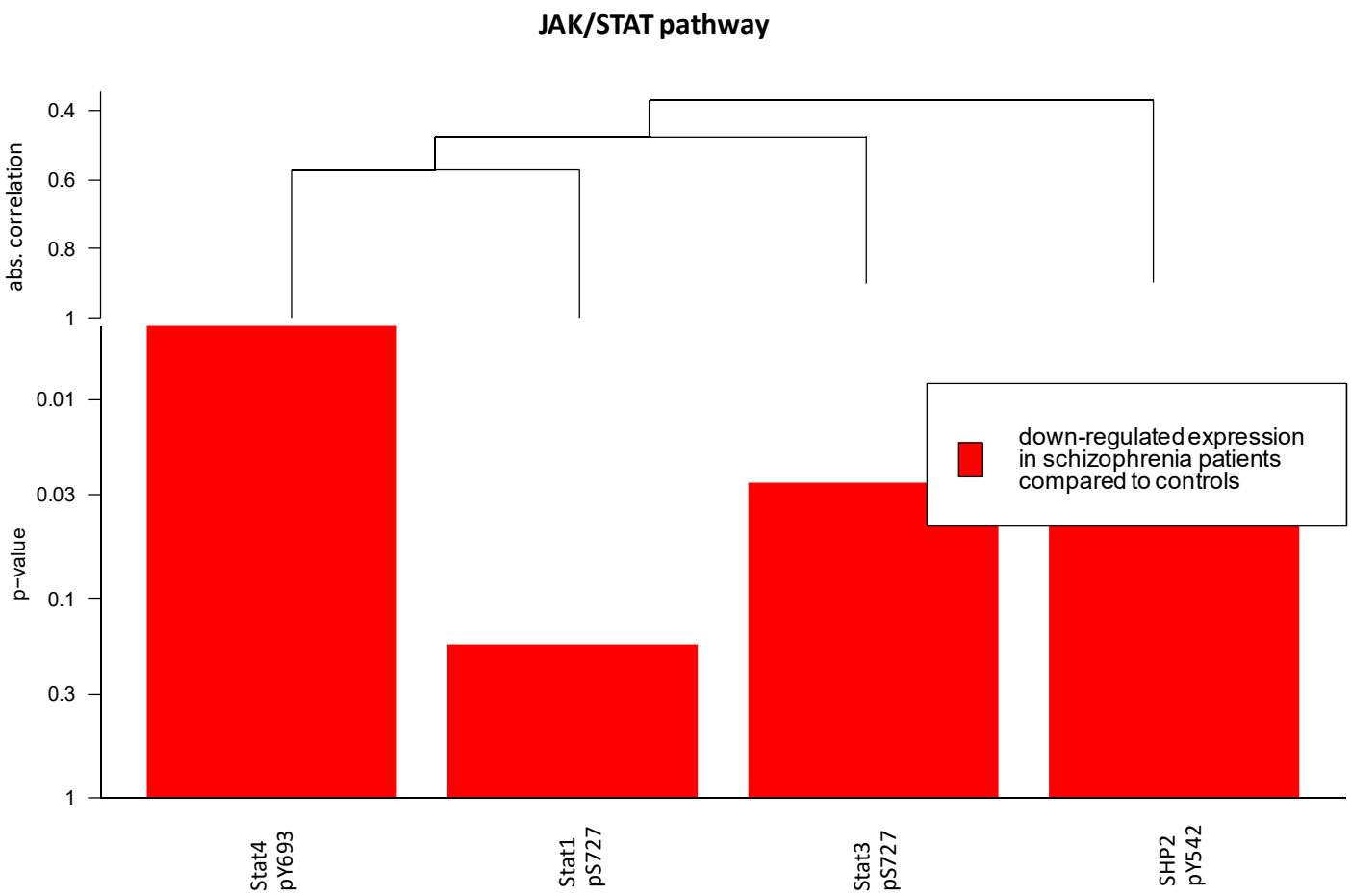


Figure 6.3 Influence plot for the five epitopes measured in the JAK/STAT pathway in the meta-cohort. Epitopes are plotted according to hierarchical clustering based on absolute linkage. Absolute correlation distance is the measure used to arrange the epitopes. The height of the bars represents the individual p-value for each epitope as a covariate in the model, thus illustrating its contribution to the overall test statistic.

The MRM datasets for cohorts 1 and 2 are then combined and analyzed as a meta-cohort to see whether the increased statistical power could shed further insights on the serum differences between patients and controls driving the changes in epitope expression. ComBat is again used to normalize technical variation between the cohorts. Results are summarized in **Table 6.8**.

As in cohort 1, significant ($p<0.05$) changes are observed in the abundances of proteins from the apolipoprotein family (L1 and C1 subtypes), and the coagulation factor HRG. As in cohort 2, the antibody IGHG2 is significant. However, again, none of these findings are significant ($q<0.05$) following adjustment for multiple testing.

Table 6.8 Significant alterations in peptide abundances between antipsychotic treated schizophrenia patients and healthy controls as measured through multiplex reaction monitoring (MRM) mass spectrometry in the meta-cohort. The associated proteins are included alongside each peptide sequence.

Protein	Abbreviation	Peptide sequence	P Value	Q Value	Fold change
Immunoglobulin heavy constant gamma 2	IGHG2	GLPAPIEK	0.001	0.169	0.52
Apolipoprotein L-1	APOL1	VNEPSILEMSR	0.004	0.266	1.30
Apolipoprotein L-1	APOL1	LNILNNNYK	0.014	0.604	1.35
Apolipoprotein C-I	APOC1	EWFSETFQK	0.018	0.604	0.75
Pigment epithelium-derived factor	PEDF	TVQAVLTVPK	0.025	0.604	1.24
Apolipoprotein C-I	APOC1	EFGNTLEDK	0.026	0.604	0.78
Apolipoprotein L-1	APOL1	VTEPISAESGEQVER	0.034	0.687	1.22
Histidine-rich glycoprotein	HRG	DSPVLIDFFEDTER	0.046	0.808	1.20

6.3.8 Microglial exposure to antipsychotics at key signalling epitopes

The final stage of this study explores whether there is evidence to suggest that antipsychotics themselves can potentiate the changes in epitope expression observed through univariate and global test analyses in this study. To investigate this, the first and second generation antipsychotics haloperidol, aripiprazole, clozapine, olanzapine and risperidone are screened against epitopes stat3 (pS727), CrkL (pY207), p38 MAPK (pT180/pY182), stat4 (pY693), and SHP2 (pY542). **Table 6.9** summarizes the antipsychotic responses relative to the vehicle condition.

The most notable findings are that olanzapine and haloperidol significantly ($p<0.05$) downregulate the expression of all five of these epitopes, especially as these results remain significant ($q<0.05$) following adjustment for multiple testing. They thus appear to be exacerbating the attenuation or potentiation phenotypes observed for the disease group at each of these epitopes in this study. This is illustrated in **Figure 6.4**. The other three antipsychotics do not have such a broad effect, but risperidone significantly ($q<0.05$) downregulates p38 MAPK (pT180/pY182), the epitope found to be downregulated between patients and controls in cohort 2, while clozapine significantly ($p<0.05$) downregulates p38 MAPK (pT180/pY182) and stat4 (pY693), the main epitope which was altered between patients and controls in the meta analysis.

Table 6.9 Microglial responses to antipsychotic exposure relative to the vehicle condition across the five microglial signalling epitopes which were implicated in inducing differential responses between treated patients and controls in this study.

	Aripiprazole			Clozapine			Haloperidol			Olanzapine			Rispiridone		
Epitope	FC	P Value	Q Value	FC	P Value	Q Value	FC	P Value	Q Value	FC	P Value	Q Value	FC	P Value	Q Value
CrkL (pY207)	0.95	0.048	0.12	0.96	0.2	0.22	0.95	0.047	0.047	0.92	0.007	0.012	0.94	0.17	0.17
p38 MAPK (pT180/pY182)	0.96	0.16	0.27	0.92	0.002	0.01	0.88	0.002	0.005	0.83	0.002	0.005	0.94	0.009	0.045
stat3 (pS727)	1	0.46	0.58	0.99	0.22	0.22	0.92	0.002	0.005	0.95	0.03	0.03	0.97	0.048	0.12
stat4 (pY693)	1.04	0.93	0.93	0.96	0.048	0.12	0.95	0.041	0.047	0.88	0.002	0.005	0.96	0.097	0.1375
SHP2 (pY542)	0.98	0.022	0.11	0.98	0.11	0.18	0.93	0.022	0.037	0.9	0.02	0.025	0.98	0.11	0.1375

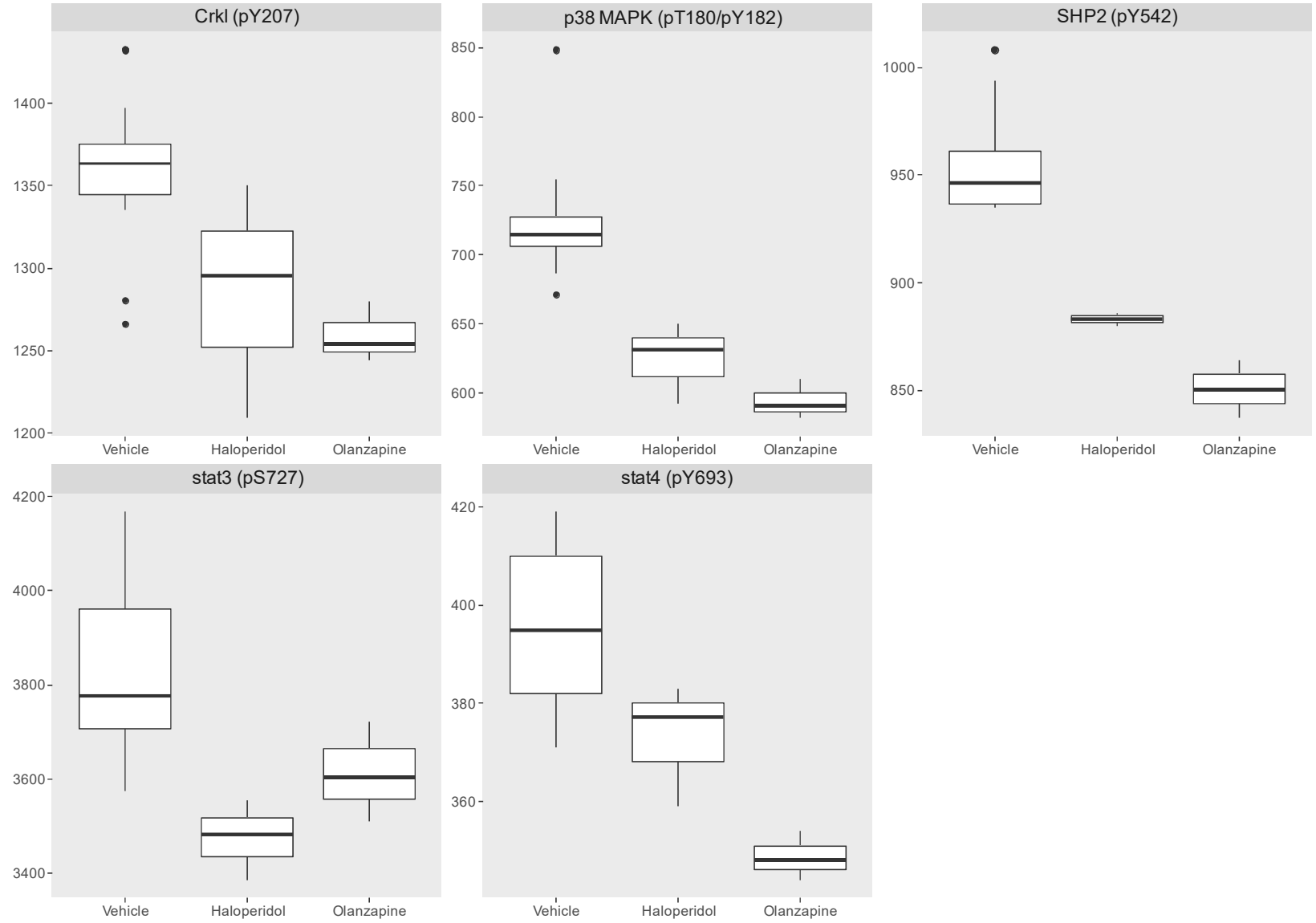


Figure 6.4 Microglial responses to haloperidol and olanzapine in comparison to the vehicle condition, for each of the five microglial signalling epitopes which showed significant ($p < 0.05$) differential responses from serum exposure to treated patients and controls in this study

6.4 Discussion

In summary, this chapter investigates microglial disturbances in antipsychotic treated schizophrenia patients through the analysis of PET imaging of TSPO expression levels across several gray matter regions. This is the first study to provide evidence regarding dysfunctional CNS mechanisms in addition to PET data from different brain regions. The lack of information regarding the cellular processes which could be contributing to the observed PET signal has been widely described as a limitation of previous PET studies(302,401,471). In this study, this is achieved through the use of a cellular model to investigate intracellular signalling alterations between patients and controls by quantifying differential expression patterns across a number of epitopes in various microglial signalling pathways.

This study conducts a further investigation of the mechanisms behind signalling alterations in antipsychotic treated patients through a meta analysis, combining microglial signalling data from two separate cohorts of recent onset antipsychotic treated patients to gain statistical power. Finally the study examines whether antipsychotics themselves have the potential to induce these signalling changes in microglia.

6.4.1 Increased microglial activation in gray matter

One of the most notable findings of this study is evidence for significantly ($p<0.05$) increased TSPO expression in the temporal region in recent onset antipsychotic treated schizophrenia patients. TSPO expression levels are low in the healthy brain, but increased in activated microglia, and thus they have been used as a marker of microglial activation and neuroinflammation in schizophrenia(474). While this result does not remain significant ($q<0.05$) following adjustment for multiple testing, this finding could still be of interest as it corresponds to the results of previously published PET studies which have identified significantly ($p<0.05$) increased microglial activity in frontal and temporal cortex gray matter in patients with schizophrenia and individuals in the prodromal phase(401,471). A contributing factor to the significant ($p<0.05$) finding in this study may be the inclusion of the hippocampus in the temporal region studied, as previous studies using the TSPO ligand (R)-[^{11}C]PK11195 to measure TSPO expression have reported significant increases in the hippocampal region in patients(471). The hippocampal region was included in the temporal region rather than being assessed on its own, in order to minimize the partial volume effect. In PET studies, this effect can be problematic as it causes the signal of the voxels at the edge of the region being measured, to be underestimated. Because the relative influence of

the partial volume effect is smaller for larger brain regions, one method of reducing the effect is to pool brain regions together in this manner(302).

As in previous studies(401,475), a positive association was found between increased temporal TSPO expression and symptom severity, as assessed through PANSS scores. This result provides additional evidence that disturbances in microglial activation may be associated with increasing symptom severity, and possibly with disease pathogenesis. Previous PET studies have suggested that the deleterious effects of aberrant microglial activation are particularly involved in the early phases of schizophrenia, with one study suggesting that microglia activation is present in patients with an illness duration of less than five years(302,470). This corresponds to the cellular model findings in **Chapter 5** which point towards microglial activation in first-onset antipsychotic naive patients, and the PET findings in this study, conducted on patients with an average disease duration of approximately one year.

Mechanistically, this study suggests that the intracellular signalling alterations in the CNS underpinning microglia activation are different in recent-onset treated patients to first-onset antipsychotic naive patients. While in **Chapter 5**, a net functional activation of the Akt/mTORC1 pathway is strongly implicated in the observed phenotype, alongside activation of several epitopes within the JAK/STAT pathway; in this study impaired JAK/STAT signalling appears to be the primary affected mechanism, albeit through a different pattern of expression changes to those observed in **Chapter 5**. As noted in **5.4**, the role of this particular pathway is of interest as it has been described as a molecular switch mediating microglial polarization(419,420,476). However, this study provides evidence that this switch may be mediated by different mechanisms in recent-onset treated patients, compared to those of antipsychotic naive patients in **Chapter 5**. Both univariate and global test analysis of cohort 1 implicate changes in the expression of stat3 (pS727) as a key component of JAK/STAT dysregulation. The comparative responses indicates that while stat3 (pS727) is downregulated in patients and controls, there is a greater reduction in stat3 (pS727) expression in the disease group. In addition, correlation analysis identifies a significant negative correlation between stat3 (pS727) and TSPO expression in the temporal region, suggesting that increased downregulation of stat3 (pS727) leads to greater TSPO expression and thus microglial activation. This observed link between stat3 and TSPO is particularly interesting due to previous studies which have suggested that the PKC – ERK1/2 – Stat3 signal transduction pathway is the main regulator of TSPO expression in both schizophrenia and control tissues(477).

Global test analysis implicates reductions in the expressions of the epitopes stat4 (pY693) and SHP2 (pY542) as contributing to this cellular phenotype. As explained further in **6.4.3**, there are suggestions that these expression changes may be a consequence of antipsychotic treatment. Similarly, this may be the case for CrkL (pY207), which was identified as being significantly downregulated in patients in the univariate analysis. A previous study using PBMCs as a cellular model of schizophrenia found a reduction in expression levels of CrkL (pY207) in patients after six weeks of olanzapine treatment(296). However it is not possible to conclude definitively whether these changes are a marker of disease state or a consequence of treatment exacerbating a pathological phenotype. For example, Crkl has also previously been linked to schizophrenia through genetic studies which have linked alterations in the *CRKL* gene locus to risk of developing the disease(478,479).

As such, the mechanistic differences between this study, and those presented in **Chapter 5**, may be due to a variety of factors. Due to the heterogeneity of schizophrenia, there may be different mechanisms of microglia disturbances across different patient populations. In addition, the involvement of different epitopes along the JAK/STAT pathway may be a consequence of the difference in disease duration between the patients studied in **Chapter 5** and those in this study. There is evidence that microglia shift in both function and morphology during the course of schizophrenia(480).

Mass spectrometry profiling of patient and control serum in cohort 1 identified a number of overlaps with findings from **Chapter 5** including alterations in Apolipoprotein C-I and Haptoglobin. The increased abundances of Haptoglobin, α-1 antitrypsin, and Apolipoprotein E points towards a pro-inflammatory M1 activation phenotype. Haptoglobin is known to modulate many aspects of the acute-phase response(481), while Apolipoprotein E is the major apolipoprotein in the brain and is synthesized by activated microglia(482). Studies have found evidence that overproduction of Apolipoprotein E may exacerbate inflammation(483,484), while as noted in **Chapter 5**, the reduced levels of Apolipoprotein C-1 may represent compromised neuroprotective mechanisms.

As in **Chapter 5**, it should be noted that a major caveat to this study is that a number of the results - including the observed significant alterations in epitopes stat3 (pS727), CrkL (pY207) and p38 (pT180/pY182) in the univariate analyses of cohorts 1 and 2, and the significant alterations in serum analytes measured through mass spectrometry profiling of cohort 1 – do not remain significant ($q<0.05$) following multiple testing correction. However, this study is very underpowered in terms of sample size, and thus should be viewed largely as an exploratory analysis, gathering evidence towards future studies with larger cohorts. In

addition, global test analysis of cohort 1 implicated impaired JAK/STAT signalling in patients, a finding which did remain significant ($q<0.05$) following multiple testing, indicating that this particular signalling pathway does play a relevant role in the disease group.

One experimental limitation of this study which should be noted in the context of these findings is the fact that TSPO ligands such as (R)-[^{11}C]PK11195 do not exclusively bind to microglia, but may also be indicative of TSPO expression in other CNS cell types such as astrocytes(485,486). Future developments are needed in terms of ligands which are more specific exclusively to microglia.

6.4.2 Dysregulation of the JAK/STAT pathway in a meta analysis of antipsychotic treated patients

While JAK/STAT may appear to be a fairly generic signalling pathway, as stated in **Chapter 5**, identifying alterations in JAK/STAT signalling in a microglial cell line with relation to schizophrenia is notable as dysregulation of this pathway in the CNS has been increasingly postulated in being involved in a range of neurological disorders(460,487). This is because JAK/STAT is one of the most important signalling pathways in the regulation of neural function, which highlights the importance of understanding how it can influence the fate and function of brain cells in the context of schizophrenia.

As such, to further investigate the microglial signalling phenotype in recent onset antipsychotic treated patients, cellular model and mass spectrometry serum profiling data from cohorts 1 and 2 were analyzed as a meta analysis to gain statistical power. Few conclusions were possible from analysis of cohort 2 alone, likely due to lack of power, but univariate analysis of cellular model data from this cohort identified a significant attenuation in expression levels of the mitogen-activated protein kinase (MAPK) epitope p38 (pT180/pY182) in patients compared to controls. This again points towards a compromised neuroprotective state as preclinical rodent studies have found that in a healthy state, increased levels of p38 are crucial to regulating microglial response to brain injury(488) through attenuating and balancing the increased levels of pro-inflammatory cytokines before the acute-phase response causes harmful and persisting consequences(489).

One of the limitations of this study is that p38 (pT180/pY182) did not survive data pre processing in cohort 1, and thus is not included in the meta analysis. As with cohort 1, univariate and global test analysis of the meta-cohort points towards dysregulation of the JAK/STAT pathway driven by stat3 (pS727), stat4 (pY693) and SHP2 (pY542). The two phenotypes identified in cohorts 1 and 2 may be linked as rodent studies suggest that the

MAPK and JAK/STAT pathways tend to converge during periods of cellular stress(490). In particular p38 is thought to directly modify epitopes on the JAK/STAT pathway(491,492).

6.4.3 Stimulation of dysfunctional microglial signalling by first and second generation antipsychotics

There is growing evidence that antipsychotics may cause or perpetuate neuroinflammation(493). The screening of a range of first and second generation antipsychotics on the microglial epitopes which showed altered signalling in this study, may point towards this, suggesting that antipsychotics have the potential to exacerbate dysfunctional microglial signalling through these mechanisms.

In the results of the meta analysis of cohorts 1 and 2, it is possible that olanzapine has a substantial effect as it was used to treat all patients in cohort 2. Evidence of this is provided by the screening analysis in which olanzapine is found to significantly ($q<0.05$) downregulate the expression of all five epitopes, an effect which remains following adjustment for multiple testing. The biggest reductions in expression are seen in p38 (pT180/pY182) which was found to be significantly ($p<0.05$) downregulated in patients in cohort 2, and stat4 (pY693) which is found to be significantly ($p<0.05$) downregulated in univariate analysis of the meta-cohort. Rodent studies have found that olanzapine has a causal effect in terms of JAK/STAT pathway activation(494), thus its profile across these five epitopes may provide both new information on its mechanism of action in microglia.

In addition, haloperidol significantly ($q<0.05$) downregulates all five epitopes. This is interesting in light of previous rodent findings which suggest that typical antipsychotics can induce microglia activation(493). These rodent studies have also found that the volume of brain regions where microglial cells are activated shrink following antipsychotic treatment(495). Previous studies analyzing TSPO expression in atypical antipsychotics have suggested that clozapine has the potential to induce microglia activation(496). As such, it is interesting that clozapine significantly downregulates p38 (pT180/pY182) and stat4 (pY693), thus possibly accentuating neuroinflammation through those mechanisms.

In future, these findings may prove to be notable in the context of the side effect profiles of first and second generation antipsychotics. While there have yet to be studies investigating a causal relationship between microglial activation and side effects relating to antipsychotic treatment, it may be plausible that processes resulting from dysfunctional microglial signalling could contribute to these adverse symptoms. Extrapyramidal symptoms are a common side effect of haloperidol, and are thought to be a consequence of dysregulation of the dopaminergic system, and damage to GABA-ergic neurons(497). Abnormal microglial

signalling is thought to result in extensive neurotransmitter dysregulation, exacerbating the disinhibition of dopamine neurons, as well as altering GABA-ergic signalling pathways(241). As such it is possible that extrapyramidal symptoms are a downstream consequence of dysfunctional microglial activation, stimulated by haloperidol. Concurrently, more recent theories have suggested that microglial activation can induce the hypoglutamatergic states associated with the pathophysiology of schizophrenia(498). While atypical antipsychotics such as olanzapine and clozapine act on the glutamate system, a common side effect of this is metabolic disturbances such as impaired glucose tolerance and weight gain(499). Further studies are needed to investigate whether potentiation of microglial signalling mechanisms as a consequence of atypical antipsychotic treatment could be associated with these side effects.

Further limitations and future work relating to this study are discussed in **Chapter 9**.

Chapter 7 Development of a proteomic systems methodology for the evaluation of molecular brain changes in rodent models compared to psychiatric disorders

7.1 Introduction

Despite the considerable disease burden of psychiatric disorders, current treatment options remain limited. Schizophrenia patient response rates to first and second generation antipsychotics are typically low(17,18) while response rates to treatment are equally poor in other psychiatric disorders with only one third of MDD patients reaching remission criteria following initial therapy(500). In order to develop novel and improved treatments, animal models of psychiatric disorders are investigated to evaluate the potential benefits of promising new drug candidates.

The aetiology of schizophrenia, MDD and other psychiatric disorders is not well understood, the onset of these disorders is most likely precipitated by a combination of genetic and environmental factors, leading to a diverse range of clinical phenotypes(501–503). Exposure of animals to various forms of genetic modifications and selective trait breeding, environmental stressors, or pharmacological challenges can induce behavioural disturbances which may represent certain neurocognitive or neurobehavioural endophenotypes(255,504–506). Although no rodent model can reflect a given psychiatric disorder in its complexity, they may provide insights into certain aspects of the underlying molecular pathology associated with different symptom dimensions. Of all animal models, rodent models are most commonly used for preclinical drug evaluations due to the advantages of a homogenous breeding background and the potential for controlled experiments(255).

However such models of psychiatric disorders have yet to yield the expected breakthroughs in terms of novel pharmacological treatments. Among the main reasons postulated for this is that it is difficult to reliably or convincingly assess human behavioural symptoms of these disorders in animals, and examining the translational relevance of animal models with the

human pathology has proven difficult(507). As described in **1.7.3**, animal models of psychiatric disorders have varying degrees of construct, face and predictive validity, but a novel step has been to attempt to identify molecular disease hallmarks of distinct endophenotypes in these models as an essential part of their integration into the drug development pipeline(272,507,508).

The work presented in this chapter is the first study to develop and apply a novel methodology for directly evaluating and comparing the molecular changes associated with rodent models of psychiatric disorders to human post-mortem brains on the functional level. Direct proteomic comparisons between animal models and human tissue have yielded relatively little information for the study of psychiatric disorders due to the limited overlap between the mouse and human proteome(272). Alternatively, an analysis methodology based on GO annotations to protein abundance changes enables a cross-species comparison between functional patterns associated with the molecular changes between species and models. The application of this methodology is based on proteomics data obtained from post-mortem brain tissue and equivalent data from commonly used preclinical models. This data is subsequently integrated and evaluated using a network biology approach, resulting in the identification of pathophysiological features associated with the psychiatric disorder in question.

In this chapter, this methodology is developed on rodent models of MDD (based on different common environmental stressors, i.e. social defeat (SD), chronic mild stress (CMS), prenatal stress (PNS)) and human MDD post-mortem brains. Environmental stress is thought to be one of the most significant risk factors for the development of MDD, especially in an interaction with genetic risk factors, early life stress and ongoing stress may determine an individual's vulnerability to develop depression(501). The methodology is subsequently applied to rodent models of schizophrenia and human post-mortem tissue in **Chapter 8**.

7.2 Methods

Figure 7.1 provides a step-by-step illustration of the general methodology outlined in this chapter for conducting a systems approach for comparing proteomic data from animal models of a psychiatric disorder to human post-mortem brain tissue.

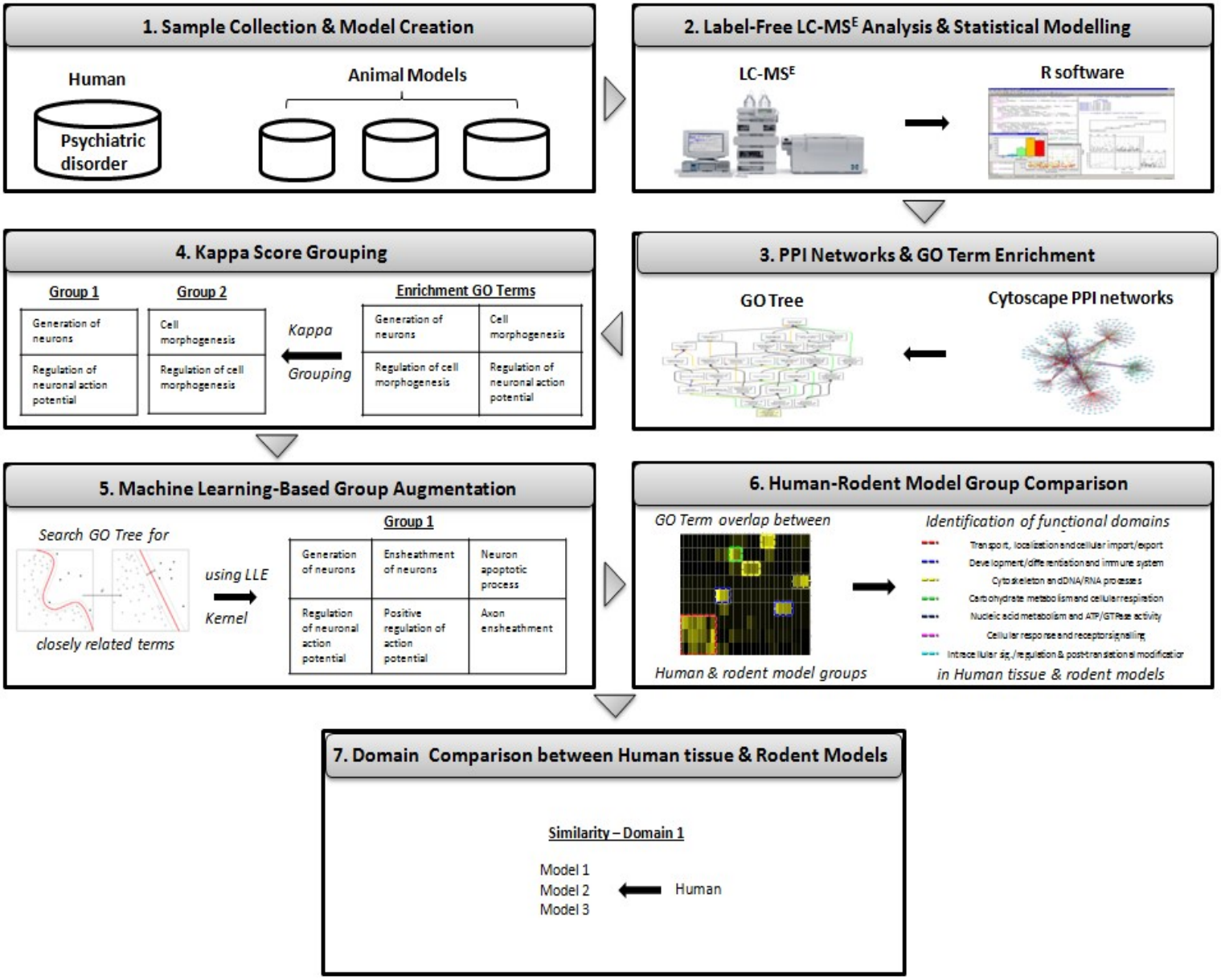


Figure 7.1 Flowchart representing an outline of the methodology developed in this chapter

7.2.1 Clinical samples

35 post-mortem anterior prefrontal cortex (Brodmann area 10 - BA10) brain samples were obtained from the Stanley Medical Research Institute(509). Samples are derived from 12 MDD patients (with purely affective diagnoses and no psychotic features in their disease course) and 23 healthy control subjects.

Tissue collection took place with the full informed consent of a first-degree relative to comply with the Declaration of Helsinki. As outlined previously(510), there were no differences in grey and white matter volumes between samples, and no significant differences in brain side, gender or secondary axis diagnosis of alcohol abuse/dependency and drug abuse/dependency between patients and controls.

Three stress-based rodent models of MDD commonly used in preclinical research are generated applying the standard protocols outlined below. Rats (Rattus norvegicus) are used in the CMS and PNS models and mice (Mus musculus) are used in the SD model. At the end of each protocol, rodents are sacrificed by decapitation, and frontal cortex tissue samples are collected from both stressed samples and a matching number of unstressed control samples. As in previous studies(511), the frontal cortex is defined as the anterior portion of the cortex up to 2.15 mm rostral from bregma.

7.2.1.1 Social defeat

7-9 week old animals are exposed individually to 9-13 month old socially dominant aggressor animals for ten minutes each day, for ten successive days, in a 1:1 social dominance constellation as described previously(512,513). The two mice are kept in the same cage separated by a Plexiglas screen for the remaining 24 hours for constant visual and olfactory cues, prolonging the defeat procedure. A different aggressor mouse is used each day to minimize inter-aggressor variability. On day eleven, all mice exposed to SD stress are subjected to the established social interaction test(513) which employs a video-tracking system to score interaction/avoidance behaviours towards an aggressor mouse. Following this test, all mice exposed to SD stress are sorted into either susceptible or resilient groups, by comparing the interaction test score ratio for each animal to established behavioural scores associated with defeat and resilience(512,513). An interaction test score ratio of 1, in which SD animals have displayed an identical amount of interaction with a social and a non-social target in the presence of a formerly unknown mouse, has been previously defined as the threshold for dividing defeated mice into susceptible and resilient groups. Twelve animals deemed susceptible in accordance to the SD procedure (social

interaction score <1) described above, and 12 control animals which did not undergo defeat were sacrificed 24 hours after the final social interaction testing and 48 hours after the last defeat (SD mice only).

7.2.1.2 Chronic mild stress

24 adult male rats are housed individually with free access to food and water. 12 animals are maintained as controls and 12 animals are exposed to a mild stress regimen for three consecutive weeks, five days a week, based on a previously described CMS procedure(514,515). Stressors consist of repeated periods of confinement to small cages (24 × 10 x 9 cm), two periods of continuous overnight light exposure, one 18 hour period of food and water deprivation followed by 2 hours of restricted food access, one 18 hour period of water deprivation immediately followed by 1 hour's exposure to an empty bottle, and one 18 hour period of group housing in a soiled cage. The CMS protocol is evaluated through three commonly used behavioural tests. The sucrose preference test (SPT)(516) sees each animal receive free access to 150ml of sucrose solution and 150ml of water. The SPT tests for reduced responsiveness to positive stimuli, which is symptomatic of depression in humans. The forced swim test (FST)(517)is then used to analyse whether animals show inclinations towards despair by placing them in a situation where they may perceive that escape from the water is impossible(518). Finally the tail suspension test(519)is used as an additional quantifier of behavioural despair. Animals from the stress group display similar depression-related phenotypes in comparison to the control group as expected from previous work(520). All animals are sacrificed via decapitation following the end of the protocol.

7.2.1.3 Prenatal stress

Pregnant dams are obtained at gestation day 6-7 and randomly assigned to control or prenatal stress groups after one week of acclimatisation. The stress groups are exposed to varying stressors during their third trimester, from gestational day 14 until day 20, following an established protocol(521). A different stressor is applied for each of these days. Following birth, all dams and pups were housed in a single cage and left undisturbed for 20 days. On postnatal day 21, pups are separated from their mothers and kept single housed. In an evaluation of the PNS paradigm, young adult animals are subjected to stress evaluation behavioural tests between postnatal days 50-70(521). Locomotor and rearing/climbing activity was assessed over a 24h period as changes in circadian activity patterns can be symptomatic of depression-like behaviour. The FST is used as in the CMS procedure.

Similar to previous work(521), animals from the stress groups display a phenotype related to increased behavioural despair and decreased exploratory behaviour. Adult male offspring (10 from the control group and 10 from the stress group) are sacrificed at postnatal day 150.

7.2.2 Label-free LC-MS^E analysis of brain tissue

Using previously defined storage, preparation and measurement procedures(510), brain samples are analyzed individually in technical duplicates using label-free LC-MS^E. Approximately 12-16mg of mouse tissue per sample is used for the SD model and 22-28mg of rat tissue per sample is used for the CMS and PNS models. For the subsequent analysis, the Swiss-Prot human reference proteome (Uniprot release, Sep 2014; 20 209 entries) is used for peptide/protein identification. The protein sequences of the Mus musculus and Rattus norvegicus UniProt reference proteome files are merged to create a joint database (retrieved Sep 2014; total joint number of entries 24 577).

LC-MS^E raw data is processed using the ProteinLynx Global Server v.2.5. (Waters Corporation), and Rosetta Elucidator v.3.3 (Rosetta Biosoftware), applying settings and procedures as reported previously(510). Peptide signal intensities for each sample are exported for pre-processing and statistical analysis in the software R. Only peptides with an amino acid sequence ending in R or K are considered to avoid unspecific trypsin cleavage. Non-unique peptides are excluded. The first two principal components are used to identify sample outliers(522) resulting in the removal of two control samples and one MDD sample, one stress sample from the PNS model and one stress sample from the SD model. Log_2 transformation is applied to stabilize data variance.

Protein abundance changes for the human and rodent model comparisons (MDD compared to control, or stress versus non-stressed controls in the rodents, respectively) are determined using a linear mixed model (applied in R as described in **Chapter 2**). This model adjusted for covariates in the post-mortem brain samples regarding age, gender, diagnoses of alcohol or substance abuse, brain pH, brain side and post-mortem interval (PMI). For the human tissue, the false discovery rate (FDR) is controlled by adjusting p-values according to the Benjamini Hochberg procedure(416) with a cut-off of 0.05, using the p.adjust function in the R stats package(302). For each rodent model, the p-values for each protein are evaluated through the permutation testing procedure described in **Chapter 5**.

7.2.3 Protein-protein interaction networks

PPI networks are created for MDD and the SD, CMS, and PNS rodent models using the software package Cytoscape v3.2.1(523), enabling the comparison of MDD and rodent tissue on a functional level, based on the annotation of GO terms to the detected significant protein abundance changes described above. The databases MINT(350), IntAct(349) and UniProt(351)are used to retrieve all available known PPI between the significant (corrected p-value ≤ 0.05) protein abundance changes and their respective first-order protein interactors. Network nodes are filtered by taxonomy identifiers (9606 for Homo sapiens in the MDD network along with 10116 and 10090 for Rattus norvegicus and Mus musculus respectively in the rodent networks) while edges are filtered so that all connections other than direct interactions or physical associations between proteins are excluded. The structures of each network are characterized using three properties of complex networks. These are, average degree – the degree of each protein (the number of other proteins to which it is connected) averaged across the entire network(524); characteristic path length - the average of the shortest path lengths between all pairs of proteins (which is small for less informative networks)(525); and density – the number of edges in the network divided by the possible number of edges(525).

7.2.3.1 GO term enrichment

GO term enrichment is computed on each PPI network using the ClueGO(352) Cytoscape package, with default settings unless described below. The ontology category used is “Biological Process”. The Homo sapiens GO database is used for the terms of the MDD network. The Mus musculus database is used to evaluate all rodent networks to reduce a species-specific annotation bias. A two-sided hypergeometric distribution is used to compute the statistical significance of each GO term, describing the probabilities associated with sampling randomly without replacement from a finite network of proteins where all proteins have an equal chance of being drawn. This determines whether any GO terms occur at a frequency greater than would be expected by chance. For each term, p-values are corrected for multiple testing (q-values) by applying a Benjamini-Hochberg correction. Terms with a significant q-value ($q < 0.05$) are taken forward and terms with no significant enrichment ($q > 0.05$) or less than two proteins are removed automatically.

7.2.3.2 Kappa score grouping

Functional grouping is applied to the list of terms for each network using the kappa score(37), a metric which reflects the degree of the relationship between GO terms, based on shared underlying proteins. A kappa score of 0.7 is used, as this requires abundant shared proteins in order for terms to be grouped(352,526), ensuring that each group has a distinct biological functionality. SD, CMS or PNS functional groups which did not contain at least two GO terms are excluded from the analysis. A more stringent threshold is applied to the MDD functional groups, as the MDD PPI network is a factor of ten larger than the rodent networks. This size discrepancy derives from the information bias in GO annotations and protein databases between human and rodent proteins. This bias exists because far more studies have been conducted investigating human proteins, and experimentally annotating their functions compared to rodent proteins. This means that more information is known about human proteins and their interactors, and there are more GO annotations associated with each protein. The net result of this is that human PPI networks are going to be far larger, and thus this analysis yields far more functional groups for MDD, compared to the three models.

As such, to account for this, MDD functional groups which did not contain at least ten GO terms are excluded from the analysis. Groups are named according to the most significant GO term following the Benjamini Hochberg FDR correction mentioned in **7.4.1**.

7.2.3.3 Local linear embedding kernel group augmentation

The biological interpretation of each functional group is enhanced by augmenting the groups with closely related GO terms using a diffusion-type manifold embedding technique called a Local Linear Embedding (LLE) kernel, commonly used to group related proteins or genes(527,528). Kernels are manifold embedding techniques which are commonly used in bioinformatics to classify data points into particular categories(529). In this instance, kernel methods are used to compute similarity metrics through a geometric interpretation of manifold embedding where each GO term is treated as though positioned in a virtual two-dimensional space based on its place in the GO tree. The LLE kernel is chosen over other kernel techniques, as it emphasizes short-range interactions between terms(530).

The LLE kernel is used to embed the entire H.sapiens GO graph into Euclidean space, and then for every MDD term in a given functional group, it obtains distance similarity values between that term and any other term in the graph. These similarity values utilise the geometry of the GO graph to estimate the divergence in function between any two terms,

based on underlying proteins. A positive similarity value between each term in a functional group and any other term in the GO database meant this new term was related and could be added to the group. A negative value meant that there are no shared proteins, and so the new term was unrelated.

This was computed in R using the calc.diffusion.kernel function with method="diffKernelLLE", in the R GOSim package(531). Likewise with each SD, CMS, and PNS term, similarity metrics are computed to every term in the M. musculus GO tree. The concept behind the use of the LLE kernel to augment the existing functional groups is that the particular proteins underlying these additional terms will be associated with a variety of related biological functions to those already conveyed by the group. Hence, this method yields more informative functional groups of highly related terms.

7.2.4 Functional comparison between MDD and rodent models

Following augmentation, the percentage overlap of GO terms is computed between MDD functional groups and those of the three rodent models, using a Z score transformation. Related clusters of functional groups between MDD and each model were identified using hierarchical clustering with the Ward's criterion metric. This enables different functional domains of the disease which are represented across all three models, to be determined, where each domain consists of a vector of GO terms. The subsets of GO terms behind each domain which are completely unique to that domain were then identified, thus defining the biological functionality more precisely.

7.2.5 Domain comparison using GO term similarity

Having identified a series of unique functional domains of human MDD represented by the four models, it is necessary to quantify which model represents each domain most closely. This is done by modifying an approach from genetic research(532,533) to obtain a numerical quantification for the closeness of the models to the disease by computing a similarity score between rodent and human domains. These scores are obtained through evaluating the average of the best matching GO term similarity between the domain vectors, where the pairwise similarity scores between GO terms were obtained using the LLE kernel described in **7.2.3.3**.The rodent and human domains are more similar, the closer the similarity scores are to 1.

7.3 Results

7.3.1 Protein abundance changes for brain tissue comparisons

1280 unique quantifiable proteins are identified across all post-mortem brain samples. 875 were identified across all SD samples, 749 across all CMS samples and 887 across all PNS samples. 109 proteins were found to be differentially expressed in MDD patients compared to control individuals. For the rodent models, 68 proteins were found to be differentially expressed in stressed compared to control mice in the SD model, 43 in the CMS model and 30 in the PNS model. Appendix tables **A.7.1-A.7.4** display these proteins and their fold changes.

7.3.2 PPI networks and GO term enrichment analysis

By retrieving all available interactions from the protein databases UniProt, MINT and IntAct, PPI networks are created for MDD and all three rodent models from the significant proteins identified above and their respective first-order interactors. In order to characterize each of these networks, several commonly used structural properties of complex networks are computed, average degree, characteristic path length and density as summarized in **Table 7.1**.

Table 7.1 Structural properties of protein-protein interaction networks. Abbreviations: MDD, major depressive disorder; SD, social defeat; CMS, chronic mild stress; PNS, pre-natal stress; N = number of significant proteins, n = number of significant proteins included in the network

Condition/Model	N	n (% of N)	Number of nodes	Number of edges	Average degree	Characteristic path length	Density
MDD	109	92 (84%)	2218	4892	2.953	3.826	0.001
SD	69	36 (52%)	151	282	2.583	4.377	0.017
CMS	43	24 (56%)	70	280	4.229	3.126	0.061
PNS	30	16 (53%)	77	170	3.377	3.133	0.044

The CNS and PNS networks display a slightly greater connectivity and density, potentially because they are more compact, as indicated by their smaller characteristic path length. However, the biological conclusions which can be drawn from direct comparisons between structural properties of networks are inherently limited, due to the possibility that the observed effects are artefacts of information bias in the protein databases used to create the

networks. For example previous studies have shown that certain nodes may appear to have a higher degree simply because they have been investigated further and more is known about them than other proteins, rather than possessing any biological significance(534). As such, a functional comparison based on gene ontology annotations may prove more informative.

Following functional enrichment analysis of the networks, and Kappa score grouping of the resulting terms, 77 MDD functional groups, 52 CMS groups, 41 SD groups and 9 PNS groups are obtained. Each group corresponds to a specific biological process. **Table 7.2** displays the top 5 groups for MDD and each rodent model, in order of significance.

Table 7.2 Top 5 functional groups for each protein-protein interaction network. Abbreviations: MDD, major depressive disorder; SD, social defeat; CMS, chronic mild stress; PNS, pre-natal stress; GO, Gene Ontology; N = number of nodes in the network.

Condition/Model	N	Top five group names	No of GO terms	Group p-value	No of significant proteins per group
MDD	2218	Phosphorylation	22	8.6 E-104	43
		Phosphate-containing compound metabolic process	16	9.3 E-87	42
		Intracellular signal transduction	17	7.5 E-86	55
		Regulation of protein modification process	27	5.2 E-67	20
		Regulation of cellular process	13	6.0 E-55	59
SD	151	Organelle organization	11	5.7 E-22	17
		Chromosome organization	12	6.5 E-16	16
		Multicellular organismal development	10	1.3 E-14	21
		Cellular component assembly	12	2.2 E-13	13
		Transport	44	4.2 E-12	23
CMS	70	Protein localization	22	2.0 E-13	10
		Generation of neurons	35	3.2 E-13	15
		Establishment of localization in cell	5	2.8 E-12	10
		Organelle organization	32	3.1 E-12	14
		Cell morphogenesis involved in neuron differentiation	7	4.0 E-11	4
PNS	77	Cell morphogenesis	7	1.4 E-5	6
		Purine ribonucleoside metabolic process	12	1.9 E-5	8
		Synaptic vesicle endocytosis	7	2.6 E-5	5
		Cellular protein localization	5	5.4 E-5	3
		Regulation of receptor internalization	3	1.4 E-4	3

7.3.3 Identification of corresponding functional domains between MDD and rodent models

Kernel techniques are used to enhance the identified functional groups, as described in the methods section. Subsequently the percentage overlap is computed between MDD and rodent model groups. It is found that groups which cluster together and overlap are involved in closely related biological processes, resulting in the identification of seven functional domains of MDD in the post-mortem brains – “transport, localization and cellular import/export”, “development/differentiation and immune system”, “cytoskeleton and DNA/RNA processes”, “carbohydrate metabolism and cellular respiration”, “nucleic acid metabolism and ATP/GTPase activity”, “intracellular signalling/regulation & post-translational modification” and “cellular response and receptor signalling”, which are also represented by the GO terms based on protein changes detected in the rodent models. The CMS model represents all seven of these domains and the SD and PNS models represent five, respectively. The domains are shown in **Figure 7.2**.

Prior to quantifying which model was most representative of MDD for each functional domain, the individual groups of GO terms behind the seven domains for MDD and all three models are displayed by projecting them into Cytoscape. The resulting four networks of GO terms in **Figure 7.3**, in which nodes represent terms and edges connect the terms found to be related by the LLE kernel, display the subsets of GO terms behind each domain, which are completely unique to that domain.

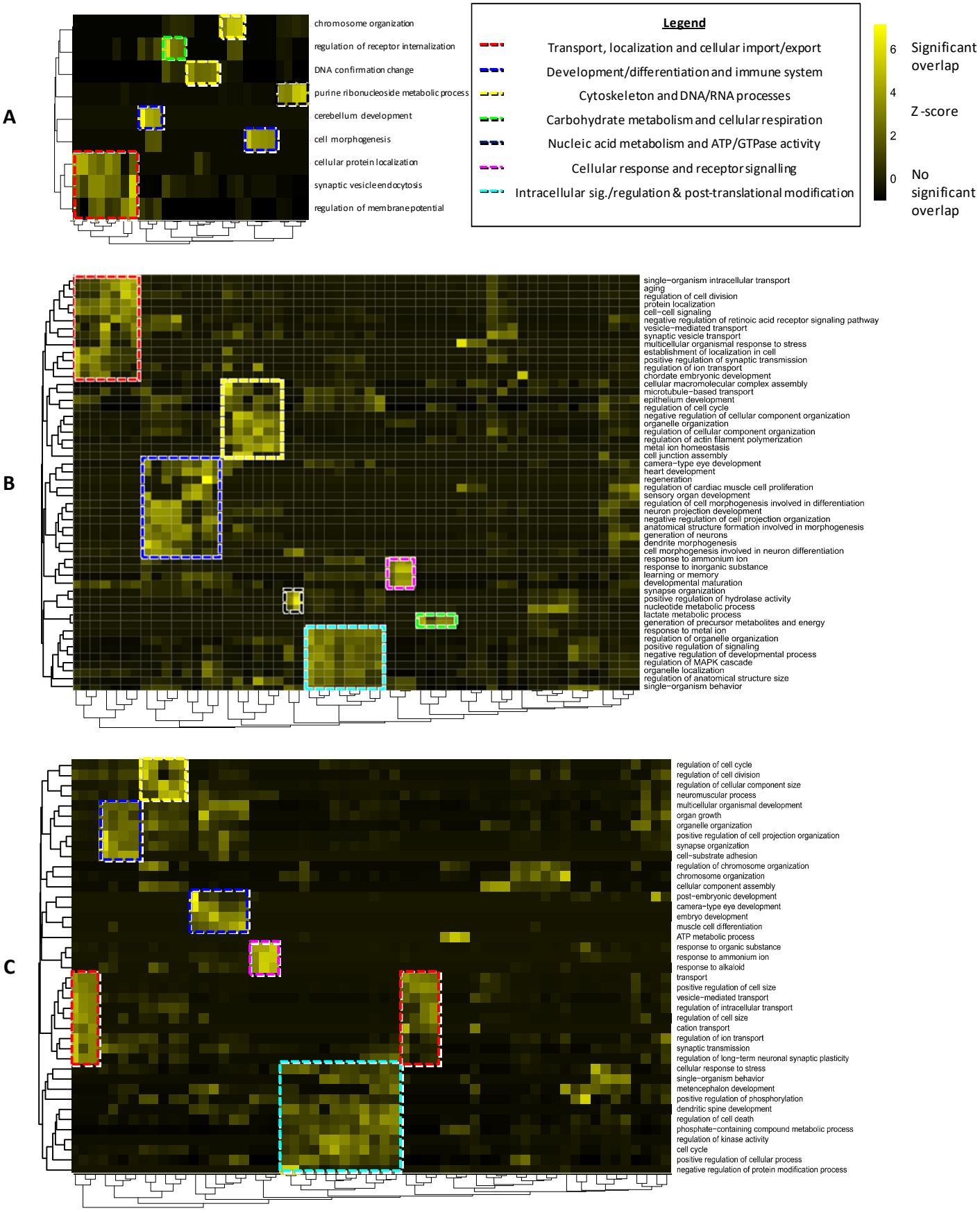


Figure 7.2 Identification of functional domains of human MDD represented across two or more rodent models. Vertical axes on the right side represent rodent functional groups. Hierarchical clustering using Ward's criterion was used on both vertical and horizontal axes to identify related clusters of groups for both MDD and each model. (A) Prenatal Stress (B) Chronic Mild Stress (C) Social Defeat

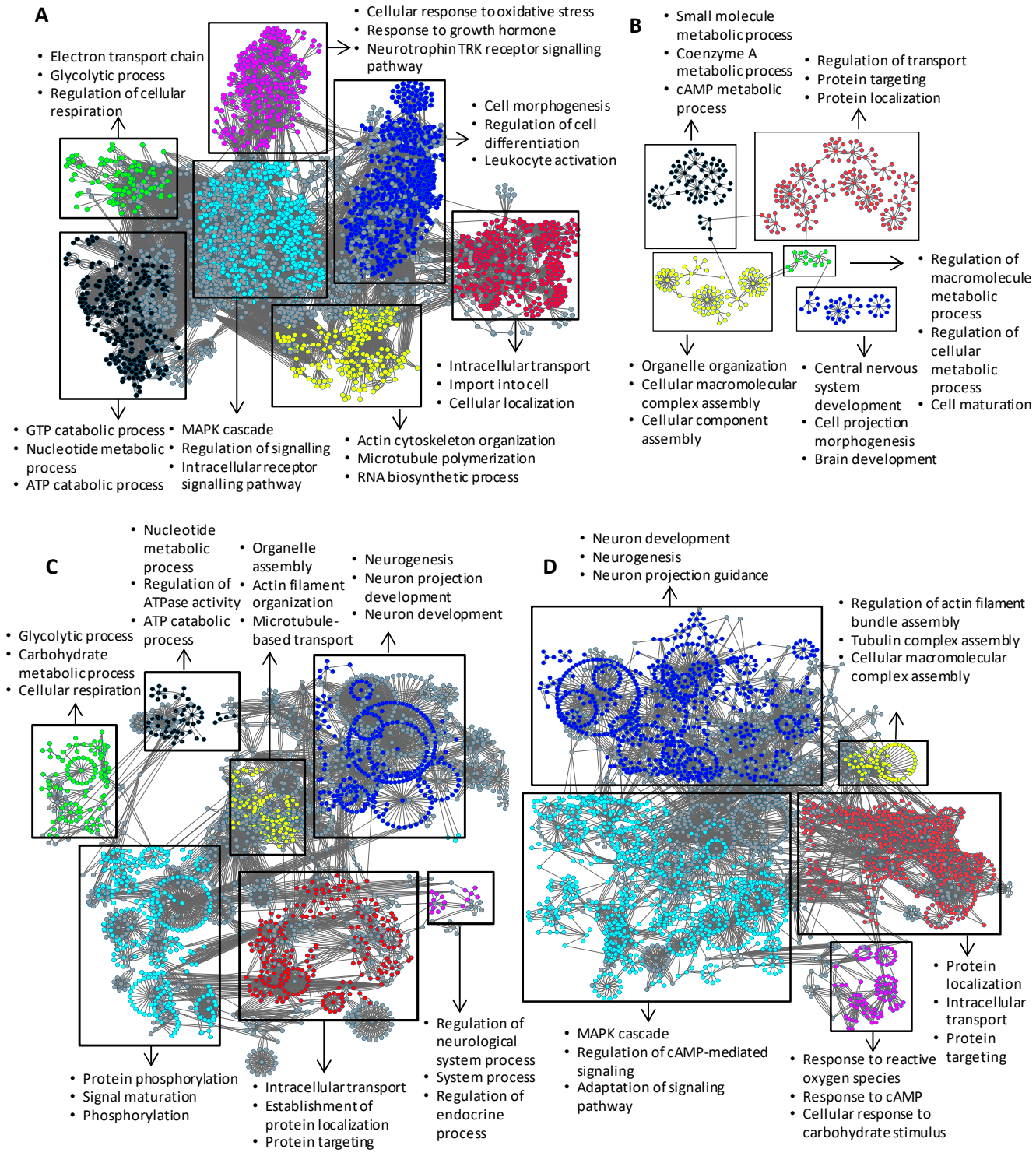


Figure 7.3 Networks of GO terms representing the seven functional domains of MDD which were identified across at least two rodent models. Projection of significantly overlapping GO terms into functional networks enabled the identification of a vector of unique terms underlying a particular domain for both MDD and each rodent model. Representative GO terms for each domain are highlighted. (A) MDD (B) Prenatal Stress (C) Chronic Mild Stress (D) Social Defeat

7.3.4 Quantification of most representative rodent model via GO term similarity methods

The vectors of GO terms for each domain are compared by computing scores based on GO term similarity, enabling the quantification of the model that represents MDD most closely. These scores are shown in **Table 7.3**.

The scores indicates that overall the SD model represents MDD most closely (to a greater extent than the other two models) for the functional domains “transport, localization & cellular import/export”, “cytoskeleton & DNA/RNA processes”, “intracellular signalling/regulation & post-translational modification” and “cellular response & receptor signalling”. The SD and CMS models are particularly close to MDD for three functional domains where the similarity scores between models and the disease are greater than 0.5 - “transport, localization & cellular import/export”, “development/differentiation & immune system” and “cellular response & receptor signalling.”

Table 7.3 Domain comparison to Major Depressive Disorder based on scores computed using Gene Ontology Term similarity. Abbreviations: SD, social defeat; CMS, chronic mild stress; PNS, pre-natal stress.

Functional Domain	SD	CMS	PNS
Transport, localization & cellular import/export	0.62	0.45	0.22
Development/differentiation & immune system	0.62	0.65	0.33
Cytoskeleton & DNA/RNA processes	0.35	0.3	0.1
Nucleic acid metabolism & ATP/GTPase activity	N/A	0.4	0.53
Carbohydrate metabolism & cellular respiration	N/A	0.37	0.0004
Intracellular sig./regulation & post-translational modification	0.4	0.16	N/A
Cellular response & receptor signalling	0.62	7.9E-05	N/A

7.4 Discussion

Despite the high prevalence and considerable clinical impact of psychiatric disorders, progress in understanding their pathophysiology has proved difficult, thought to be due to the heterogeneous and complex nature of these disorders(535). Animal models of psychiatric disorders have been utilised for the identification of novel drug targets and have provided new insights into the genetic and molecular alterations which are thought to underpin these disorders(536). However, challenges have arisen in comparing alterations in animal models with the underlying molecular disease factors in psychiatric disorders measured in serum(508) or brain tissue(510). Therefore, most comparisons are based on behavioural phenotypes, sometimes additionally supplemented with histological and electrophysiological data. This has led to difficulties in bias and irreproducibility, as many of the behavioural characteristics for psychiatric disorders like MDD or schizophrenia cannot be translationally evaluated in animals(537).

Here, a novel methodology is presented which allows the comparison of brain changes from rodent models with those observed in human post-mortem tissue from individuals with a psychiatric disorder. In the application presented in this chapter, changes in three rodent models are compared with those observed in human MDD brains. Thus, it is possible to assess which model reflects changes in MDD most closely for various pathophysiological features of the disease. The approach seeks to compare human and animal brain tissue on a functional level using GO terms annotated to proteins. Proteomic analysis has several advantages over genomic/transcriptomic analysis, as protein changes represent the transcribed and translated genetic information resulting from epigenetic and mRNA modifications more closely reflect the disease pathophysiology(538). Twin studies, for example, have approximated the heritability of MDD at 37% suggesting that genetic investigations may not be the optimum basis for a comparison(539). Similar to the case of schizophrenia and other psychiatric disorders, evidence suggests that MDD is a polygenic disorder with the disease phenotypes arising as a consequence of many small risk genes together with environmental factors(503).

Through this approach, it is possible to quantify seven functional domains identified as altered in MDD post-mortem brain tissue, which are all represented across two or more of the investigated animal models. It is shown that the four functional domains “transport, localization & cellular import/export”, “cytoskeleton & DNA/RNA processes”, “intracellular sig./regulation & post-translational modification”, and “cellular response & receptor signalling”, are represented most closely by the SD model. While these domains may seem relatively generic, this confirms existing findings in the literature as previous rodent studies of

SD stress have found brain changes implying impaired cell proliferation mechanisms(540) which is consistent with the finding that the SD model is characterised especially through intracellular signalling and post-translational modification. In addition, it is demonstrated that the functional domains “development/differentiation & immune system” and “carbohydrate metabolism & cellular respiration” are represented most closely by the CMS model. It is unsurprising that immune dysfunction was represented strongly in this model as disturbed HPA function, dampened neurogenesis and increased oxidative stress have been reported in rat CMS models(541,542). The “development/differentiation & immune system” domain is also strongly represented in the SD model. This is in line with previous studies demonstrating that SD stress induces pro-inflammatory signalling including increased monocyte trafficking via a variety of immune-regulatory pathways found to be dysfunctional in the CNS of MDD patients(543,544).

This methodology is the first proteomic approach for a cross-species comparison of human and animal tissue which applies a novel functional analysis procedure using GO terms annotated to proteins rather than relying on a comparison based solely on the significantly altered proteins as measured through LC-MS^{E}. There are multiple strengths of the methodology including the use of carefully matched post-mortem brain samples while adjusting for covariates in a regression analysis via a stepwise selection procedure which accounted for the most commonly addressed confounders while avoiding overparameterization. The present approach involves the use of a joint rodent protein database at the LC-MS^{E} processing stage and the *Mus musculus* GO database at the GO term enrichment stage to avoid introducing any species-specific bias which could bias the results in favour of one particular model. The chosen systemic methodology involving enrichment of PPI networks is superior to the standard way of pathway analysis based on individual proteins which is most commonly used in preclinical studies. A systems approach leverages the signature proteins as a representation of changes in signalling pathways, characterizing the biological processes more precisely through the inclusion of first-degree interactors in the networks, instead of interpreting the relevance between each protein and phenotype, thus expanding the functional scope of the study. In addition, once functional groups have been determined within the network, an approach commonly used to compare gene functionality is adopted(545,546). This approach utilizes kernel methods, one of the most advanced techniques in machine learning, to compare vectors of GO terms through their closeness to each other in the GO hyperspace, thus making individual comparisons possible between models and the disease(529,547). By tackling a pressing question in translational research through the adaptation and application of machine learning based methods already established as a means of conducting functional comparisons in other

disciplines of medicine, decision processes in preclinical neuropsychiatry could benefit of synergies between different fields of molecular research.

In summary, it is shown that different animal models of a psychiatric disorder can represent individual functional aspects of the disease more closely than others. The development of a methodology which can quantify molecular similarity between preclinical models and the disorder in this way is needed, as it is generally agreed that individual models are unlikely to mirror the full extent of the human disease(548). Therefore, in future, the way forward in preclinical research may be separate pharmacological studies, each focusing on different models targeting different clusters of disease symptoms, and thus different underlying molecular changes. This approach is likely to be of increasing importance in the search for novel pharmacological compounds, based on theories that individual psychiatric disorders could be a constellation of diseases, manifesting in behavioural symptoms which correlate with different neurobiological adaptations(549). There is a pressing need for pharmacological interventions that differ from current approaches which have tended to focus on a specific signalling pathways such as glutamatergic neurotransmission for schizophrenia and monoaminergic neurotransmission for MDD(550). Future research should aim to re-evaluate these disease-model comparisons incorporating proteomic analysis from genetic and pharmacological models. The ability to apply this novel method to conduct a direct functional comparison between multiple preclinical models and a psychiatric disorder will help gain greater insights into the underlying molecular and cellular mechanisms behind behavioural abnormalities and their response to pharmacological interventions, as attempts are made to obtain a greater understanding of the consequences of differing stressors in the context of these disorders.

Chapter 8 Application of a proteomic systems methodology for the evaluation of the molecular validity of preclinical rodent models compared to schizophrenia brain pathology

8.1 Introduction

Antipsychotics represent the first line of pharmacotherapy for schizophrenia and predominantly target dopamine, noradrenaline and serotonin pathways(1). However these medications typically fail to treat the cognitive and negative symptoms, which contribute substantially to the morbidity of schizophrenia(551–553). Due to increased evidence that dysfunction of glutamatergic transmission is implicated in psychotic states(125,554,555), glutamatergic animal models of the disease have increasingly been used to test the potential efficacy of novel compounds(256). Such models enable the study of central biological processes associated with a particular group of symptoms and the testing and evaluation of novel treatments.

Glutamatergic mechanisms were initially implicated in schizophrenia pathogenesis after reduced cerebrospinal fluid levels were reported in patients(556), with later studies pointing to more complex mechanisms behind dysfunction of glutamate neurotransmission(261). Glutamatergic models of schizophrenia in animals include genetic manipulation of the NMDAR(557) and acute or chronic exposure to NMDAR antagonists such as PCP(558) and ketamine(559). Systemic treatment with these antagonists in preclinical studies was found to mimic negative symptoms of the disease alongside an increase in glutamate efflux in the prefrontal cortex(560,561). Administration of PCP and ketamine in clinical studies was found to induce psychotomimetic effects ranging from positive symptoms such as hallucinations and paranoia(261) to negative and cognitive symptoms in healthy volunteers(562,563), in addition to precipitating psychotic relapses in chronic stable schizophrenia patients(564).

Preclinical evaluation is an important step in the drug discovery pipeline for schizophrenia, and allows the prioritization of compounds for clinical trials. However, regardless of ongoing advances in other areas of medicine, the development of preclinical models for schizophrenia and other neuropsychiatric disorders is at a near standstill(270), and there

have been failures in finding new pharmacological treatments because the observed outcomes in behavioural screenings were not predictive of clinical outcomes. As mentioned in **Chapter 7**, the main reason for this is the fact that uniquely human behavioural symptoms cannot be reliably or convincingly mirrored or assessed in animals, explaining difficulties in face validity. Additionally, the face validity of genetic and pharmacological models has been criticized as more than a single gene is considered to be important in the etiology of schizophrenia and injections of compounds only induce transient phenotypes. As such, there is a need to find a way to quantitatively assess which animal models represent different pathophysiological aspects of schizophrenia most closely.

In this chapter, the methodology developed and outlined in **Chapter 7** is applied to post-mortem brain tissue from schizophrenia patients and four established rodent models of the disease, integrating this proteomic information in a novel systems based approach. The rodent models were based on either psychosis-inducing NMDAR-antagonists (Ketamine and PCP) or genetic modifications targeting the glutamate system (NR1-knockdown) were investigated. Rats were chosen over mice in the pharmacological models to increase brain tissue yield for proteomic extractions. Mice were chosen for the NR1 knockdown due to the availability of superior methods of genetic manipulation compared to rats. The anterior frontal cortex was chosen as the tissue of interest in both humans and rodents as a brain region which is strongly linked to psychotic disorders. The anterior prefrontal cortex plays a crucial role in the processing and evaluation of internally generated information across multiple cognitive operations(565), and the negative and cognitive symptoms of schizophrenia are characterized by impairments in executive functioning and socio-emotional cognition. As described in **Chapter 7**, through this methodology the biological processes affected in the anterior frontal cortices in each condition are characterized and compared across species and models. Based on protein-protein-interaction networks, key functional patterns are identified that allow the molecular similarity of the models with the human condition to be quantified, a novel way of interrogating translational preclinical validity.

8.2 Methods

R and Cytoscape packages, functions and settings were applied as in **Chapter 7**, unless otherwise stated.

Figure 8.1 provides a step by step illustration of the workflow for this chapter.

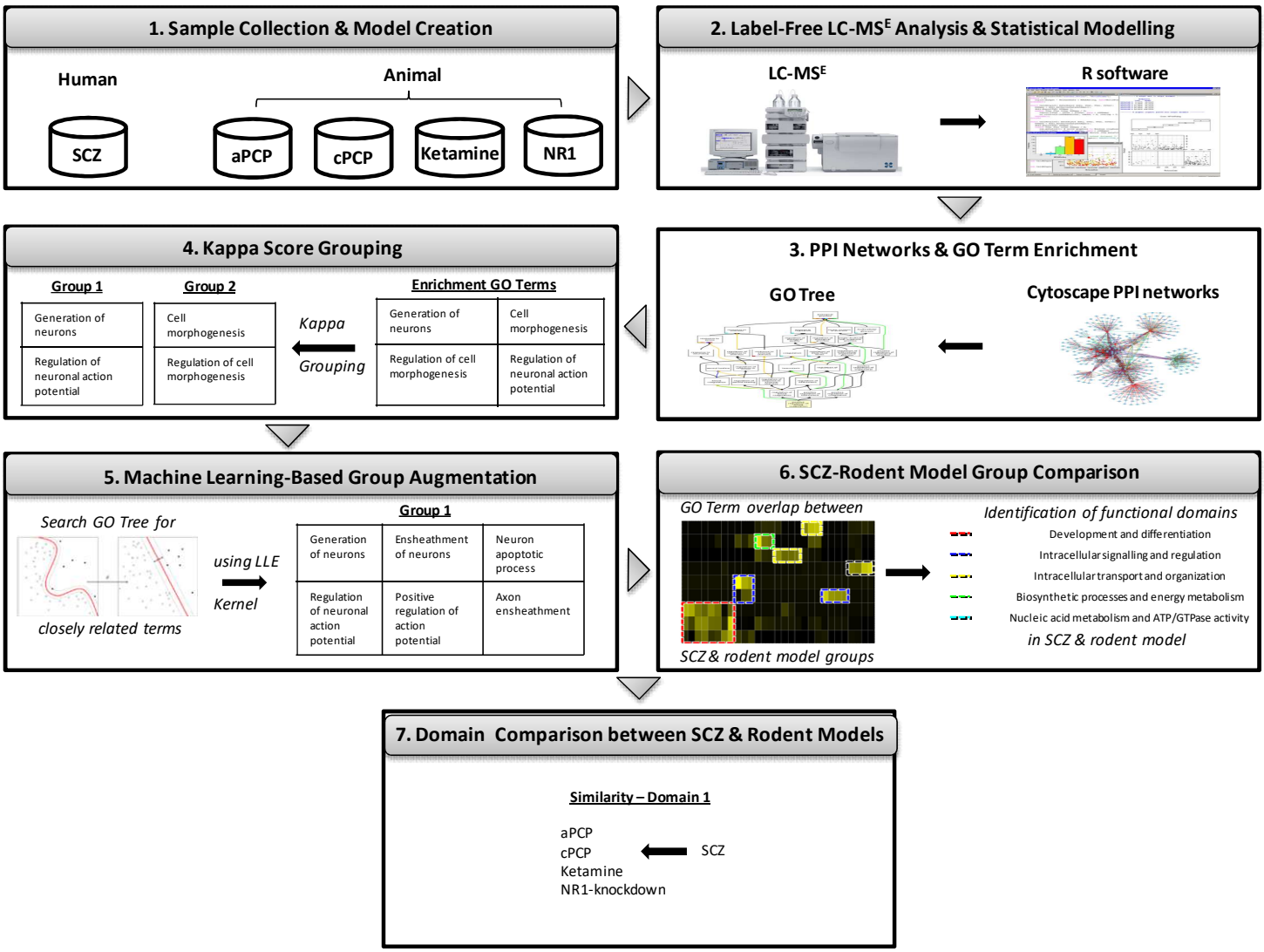


Figure 8.1 Flowchart representing methodology and experimental setup used in this chapter.

8.2.1 Clinical samples

46 post-mortem anterior prefrontal cortex (BA10) brain samples are obtained from the Stanley Medical Research Institute(509). Samples are derived from 23 schizophrenia patients and 23 control subjects.

Tissue collection took place with full informed consent of a first-degree relative to comply with the Declaration of Helsinki. As outlined previously(510), there are no differences in grey and white matter volumes between samples, and no significant differences in brain side, gender or secondary axis diagnosis of alcohol abuse/dependency and drug abuse/dependency between patients and controls.

Four glutamatergic animal models of schizophrenia commonly used in preclinical drug discovery and development are generated applying the protocols outlined below. Rats (Rattus norvegicus) are used for the ketamine, acute PCP (aPCP) and chronic PCP (cPCP) models and mice (Mus musculus) are used for the NR1 knockdown model. All animals are housed 4-5 per cage on a 12-h light/dark cycle in a temperature-controlled facility with food and water available ad libitum. At the end of each protocol, rodents are sacrificed by decapitation, and brain tissue samples are collected. A matching number of control samples are collected at the same time.

Ketamine: 10 control animals received a subcutaneous dose of saline and 10 animals received 10mg/kg ketamine subcutaneously as described previously(566). Dosage levels are based on previous research which examines dose/response levels based on locomotor activity ataxia, brain dialysis/neurotransmitter release and pharmacological magnetic resonance data(567–570). Doses are chosen which yield robust readouts while avoiding inducing anaesthesia. Sacrifice and tissue collection takes place place two hours after the last injection.

aPCP: 10 control animals received a subcutaneous dose of saline and 10 animals received 5mg/kg PCP hydrochloride subcutaneously as previously described(571,572). Sacrifice and tissue collection takes place 30min after the last injection.

cPCP: 10 control animals received subcutaneous doses of saline and 9 animals received 5mg/kg PCP hydrochloride subcutaneously for 15 consecutive days as previously described(573). Sacrifice and tissue collection takes place 30min after the last injection on day 15.

NR1 Knockdown: NR1 transgenic mice are bred and genotyped as previously described(269,557,574). 12 adult male homozygous mice and 12 wild-type littermates are used for this study.

8.2.2 Label-free LC-MS^E analysis of brain tissue

Using previously defined storage, preparation and measurement procedures(510), brain samples are analyzed individually in technical duplicates using label-free LC-MS^E. For the subsequent analysis, a human proteome database is obtained from UniProt (retrieved 2015-14-10, number of entries 20,196). A joint *Mus musculus* and *Rattus norvegicus* database is created by merging the protein sequences of the respective UniProt reference proteome files (retrieved 2015-14-10, number of entries 24,664). MS raw data and ProteinLynx Global Server v2.5 search results are imported into the Rosetta Elucidator software (build 3.3, Rosetta Biosoftware). Settings and procedures are applied as stated previously(510), and peptide signal intensities for each sample are exported for pre-processing and statistical analysis in R. Only peptides with an amino acid sequence ending in R or K are considered to avoid unspecific trypsin cleavage. Non-unique peptides are excluded. Sample outliers are identified using the first two principal components(522) resulting in the removal of one control sample from the aPCP, cPCP, Ketamine and NR1-knockdown models.

Following $\log_2$ transformation to stabilize data variance, protein abundance changes for the human and rodent model comparisons are determined using a linear mixed model, with a stepwise selection adjustment for covariates age, gender, diagnoses of alcohol or substance abuse, PMI, brain pH and brain side. For the schizophrenia tissue, the FDR is controlled at a cut-off of 0.05 by adjusting the *p*-values according to the Benjamini Hochberg procedure(416). For each rodent model, p-values are evaluated through permutation testing as described in **Chapter 5**.

8.2.3 Protein-protein interaction networks

As described in **Chapter 7**, the analysis framework is designed to compare schizophrenia and rodent tissue on a functional level, based on the annotation of GO terms to the protein abundance changes described earlier. In order for these terms to represent the biophysical interactions which occur between sets of proteins(575), PPI networks for schizophrenia and the aPCP, cPCP, ketamine and NR1 knockdown rodent models are created using the software package Cytoscape v3.2.1(523). Each network is represented as a graph where the nodes are proteins and the edges are interactions between proteins.

Networks are constructed by retrieving all available known PPI information for proteins with significantly changed abundances following multiple testing or permutations as described in **8.2.2**, and their first-order protein interactors, from the databases MINT(350), IntAct(349) and UniProt(351). Filtering is applied to both the node and edge lists for all four networks. Nodes are filtered by taxonomy identifiers (9606 for *Homo sapiens* in the schizophrenia network, in addition to 10116 and 10090 for *Rattus norvegicus* and *Mus musculus* respectively in the rodent networks), while edges are filtered to exclude all connections other than direct interactions or physical associations between proteins. All unconnected subsets of nodes are removed from the network. The structures of each network are assessed using three common properties of complex networks, average degree, characteristic path length and density, as defined in **Chapter 7**.

8.2.3.1. GO term enrichment

The ClueGO(352) Cytoscape package is used to compute GO term enrichment on each PPI network. Settings and filtering methods are applied as described in **Chapter 7**. The statistical significance of each GO term is computed using a two-sided hypergeometric distribution, determining whether GO terms occur at a frequency greater than would be expected by each term. The significance of each term is adjusted for FDR using the Benjamini Hochberg correction with a cut-off of 0.05.

8.2.3.2. Kappa score grouping

The list of terms for each network are functionally grouped based on shared underlying proteins using a kappa score of 0.7(576). The kappa score metric reflects the degree of the relationship between two GO terms. A score of 0.7 or higher requires abundant shared proteins(352,526), ensuring that the groups are likely to be biologically similar. Functional groups which did not contain at least two GO terms are excluded from the analysis. Groups are named according to the most significant (q-value <= 0.05) GO term.

8.2.3.3. Local linear embedding kernel group augmentation

A LLE kernel is applied to enhance the biological interpretation and comparability of each functional group by augmenting them with closely related GO terms as previously outlined in **Chapter 7**. Manifold embedding techniques classify data points in particular categories and are commonly used in bioinformatics. The kernel applied here computes similarity scores based on a geometric interpretation of manifold embedding interpreting every GO term as a

point on a virtual two-dimensional GO tree. The LLE kernel is selected because it emphasizes short-range interactions between GO terms(530) and thus is typically used to group related proteins or genes(527–529). Every given GO term in a functional group is analysed for positive similarity values to other GO terms due to relative closeness on the GO tree. The LLE kernel group augmentation is based on the concept that the functional annotations added by this approach are highly likely to be related to biological functions already conveyed by that group (and closely related to the underlying PPIs). Therefore this procedure results in more informative function groups of highly interconnected GO terms.

8.2.4 Functional comparison between schizophrenia and rodent models

The enhanced groups of GO terms are used as the basis for a functional comparison between schizophrenia and the four rodent models. The percentage overlap of terms in terms of Z score is computed between each of the schizophrenia groups and the rodent model groups. Hierarchical clustering using Ward's criterion is employed to identify related clusters of groups for both schizophrenia and each model, enabling the identification of different functional domains in the disease represented across all four models, where each domain is a vector of GO terms. The subsets of GO terms behind each domain which are completely unique to that domain are then identified, thus defining the biological functionality more precisely.

8.2.5 Domain comparison using GO term similarity

For each unique functional domain of human schizophrenia represented across the four rodent models, a numerical quantification is obtained for which model represented the human domain most closely, by adapting an existing approach used in genetic research(532,533). A similarity score is computed between rodent and human domains by first evaluating the pairwise similarity scores between individual GO terms in the domain vectors using the LLE kernel, and then obtaining the average of the best matching GO term similarity between the domains. The rodent and human domains are more similar, the closer the similarity scores are to 1.

8.3 Results

8.3.1 Behavioural characteristics of rodent models

For both acute and chronic models, PCP doses are found to induce the expected abnormalities of hyperlocomotion, increased stereotypic behaviour and impaired attention and social interaction as reported previously(577,578). Ketamine doses induce hyperlocomotion, stereotypy, impaired information processing with abnormalities in cognitive function, and impaired social interaction, behavioural characterizations which has been described in previous acute ketamine models(567,579). NR1-knockdown mice display both hyperlocomotion and increased stereotypic behaviour in addition to impairments in cognition and escape behaviours as found previously(580).

8.3.2 Protein abundance changes for brain tissue comparisons

A total of 1280 quantifiable proteins are measured across all schizophrenia brain samples, 643 across all aPCP samples, 873 across all cPCP samples, 772 across all Ketamine samples and 409 across all NR1 knockdown samples. **Figure 8.2** summarizes the overlap in proteins measured between these groups. It is found that 159 proteins were differentially expressed in schizophrenia tissue compared to controls. For the rodent models, 47 proteins were found to be differently expressed in the aPCP-control comparison, 84 in the cPCP-control comparison, 93 in the Ketamine-control comparison and 80 in the NR1 knockdown-control comparison. Appendix tables **A.8.1-A.8.5** display these differentially expressed proteins and their fold changes. **Appendix A.8.6** displays the 162 proteins which were quantified across all schizophrenia and rodent model tissue samples, as indicated in **Figure 8.2**.

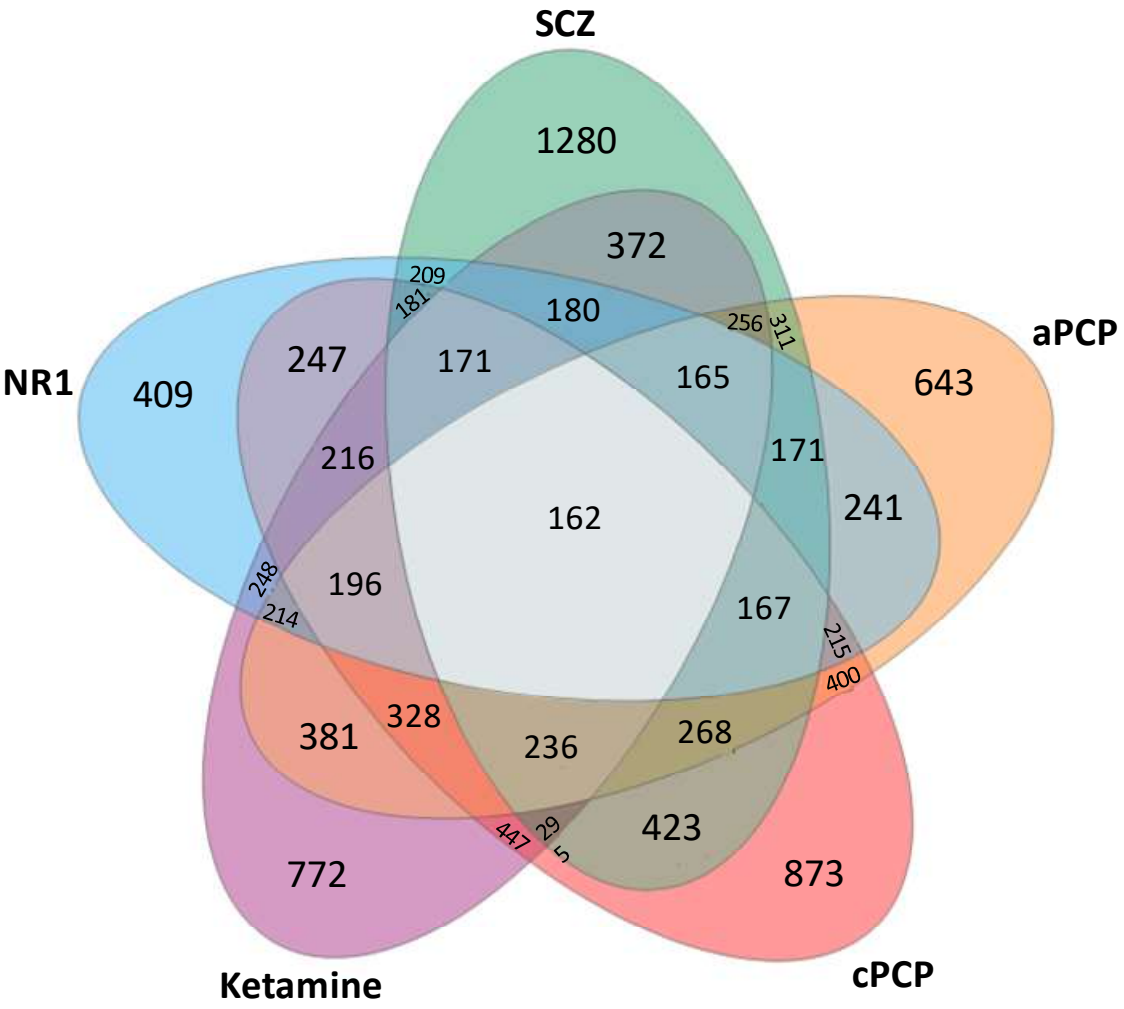


Figure 8.2 Overlaps in quantifiable proteins detected in schizophrenia, aPCP, cPCP, Ketamine and NR1-Knockdown brain tissue.

8.3.3 PPI networks and GO term enrichment analysis

PPI networks are created for schizophrenia and all four rodent models using PPI information from the UniProt, MINT and IntAct protein databases between the significant proteins identified in **8.3.2**, and their respective first-order interactors. The significantly changed abundances of these proteins is determined through multiple hypotheses testing (human samples) or permutation testing (rodent samples). All rodent model PPI networks, as well as the schizophrenia PPI network are characterised using several structural properties of complex networks - average degree, characteristic path length and density as summarized in **Table 8.1**. The structural properties of average degree and characteristic path length are most closely related between the schizophrenia and Ketamine networks. However, as

addressed previously in **Chapter 7** the biological conclusions which can be drawn from direct comparisons between structural properties of networks are inherently limited, and the subsequent functional comparison based on gene ontology annotations is most informative.

Table 8.1 Structural properties of protein-protein interactionnetworks. Abbreviations: SCZ, schizophrenia; aPCP, acute phencycline; cPCP, chronic phencycline; N = number of significant proteins, n = number of significant proteins included in the network.

Condition/Model	N	n (% of N)	Number of nodes	Number of edges	Average degree	Characteristic path length	Density
SCZ	159	115 (72%)	2163	3297	2.698	4.162	0.001
aPCP	48	29 (62%)	150	203	2.493	3.772	0.017
cPCP	85	59 (69%)	351	507	2.741	3.3	0.008
Ketamine	94	54 (57%)	189	271	2.698	4.278	0.014
NR1-knockdown	81	52 (64%)	349	537	2.951	3.368	0.008

Functional enrichment analysis of the networks, and grouping of the resulting terms according to a Kappa score of 0.7, yields 222 schizophrenia functional groups, 63 aPCP groups, 128 cPCP groups, 117 Ketamine groups and 119 NR1-knockdown groups, with each group corresponding to a specific biological process. **Table 8.2** displays the top 5 functional groups for schizophrenia and each rodent model, in order of significance.

Table 8.2 Top 5 functional groups for each protein-protein interaction network. Abbreviations: SCZ, schizophrenia; aPCP, acute phencycline; cPCP, chronic phencycline; N = number of nodes in the network.

Condition/Model	N	Top five group names	No of GO terms	Group p-value	No of significant proteins per group
SCZ	2163	Regulation of phosphorus metabolic process	14	4.7 E-115	24
		Phosphorylation	12	6.4 E-92	23
		Protein phosphorylation	4	1.1 E-70	38
		Cellular protein metabolic process	4	6.2 E-68	40
		Intracellular signal transduction	3	2.1 E-67	30
aPCP	150	Cellular component assembly	3	1.2 E-16	6
		Transport	3	4.0 E-15	14
		Regulation of cellular metabolic process	23	4.2 E-15	14
		Negative regulation of biological process	2	5.3 E-14	7
		Regulation of signalling	4	6.4 E-14	4
cPCP	351	Phosphorus metabolic process	2	1.6 E-40	31
		Organic substance catabolic process	3	2.4 E-40	30
		Single-organism catabolic process	23	3.2 E-40	31
		Transport	2	2.2 E-32	33
		Cell projection organization	15	1.7 E-25	21
Ketamine	189	Transport	2	1.4 E-26	20
		Regulation of localization	2	2.3 E-20	10
		Cell communication	2	2.7 E-19	22
		Cell-cell signalling	2	5.1 E-19	6
		Regulation of cell communication	4	5.2 E-19	13
NR1-knockdown	349	Phosphorus metabolic process	2	3.5 E-34	28
		Transport	2	1.3 E-31	21
		Establishment of localization in cell	9	3.4 E-23	14
		Cellular component assembly	3	1.7 E-22	15
		Regulation of transport	2	6.0 E-22	13

8.3.4 Identification of corresponding functional domains between schizophrenia and rodent models

Following the enhancement of these functional groups by kernel techniques, the percentage overlap is computed between schizophrenia and rodent model groups. It is found that groups which cluster together and overlap are involved in closely related biological processes, resulting in the identification of five functional domains of the disease – “development and differentiation”, “intracellular signalling and regulation”, “intracellular transport and organization”, “biosynthetic processes and energy metabolism”, and “nucleic acid metabolism and ATP/GTPase activity” - which are represented across all four models. These domains are shown in **Figure 8.3.**

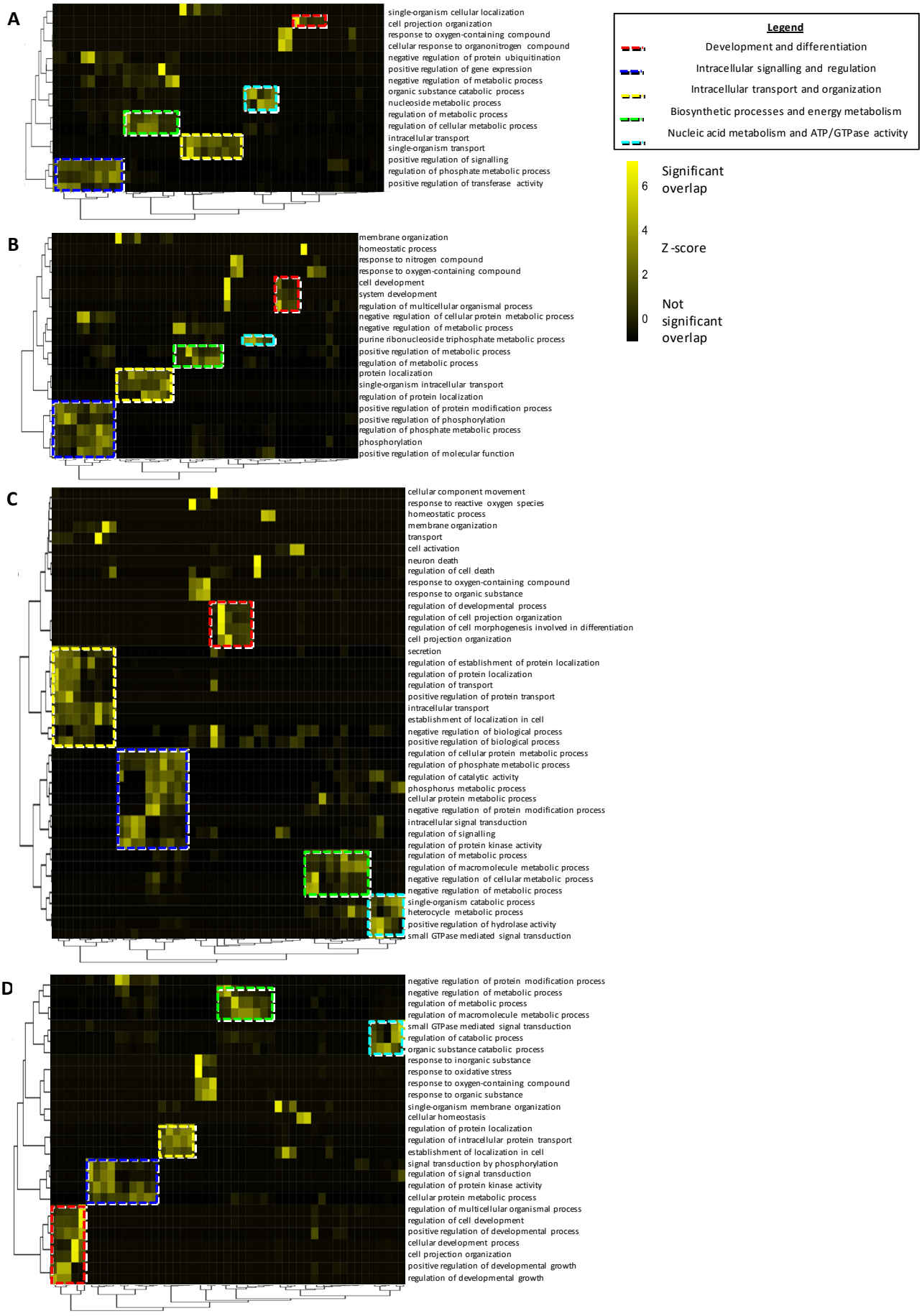


Figure 8.3 Identification of five functional domains of human schizophrenia represented across all four rodent models. Vertical axes on the right side represent rodent functional groups. Hierarchical clustering using Ward's criterion was used on both vertical and horizontal axes to identify related clusters of groups for both schizophrenia and each model. (A) Acute PCP (B) Ketamine (C) NR1-Knockdown (D) Chronic PCP

8.3.5 Identification of most representative rodent model via similarity based methods

Similarity-based methods are used to compare the vectors of GO terms for each domain, making it possible to quantify the model that is most representative of schizophrenia through similarity scores as shown in **Table 8.3**.

Table 8.3 Domain comparison to schizophrenia based on scores computed using gene ontology term similarity. Abbreviations: SCZ, schizophrenia; aPCP, acute phencycline; cPCP, chronic phencycline; N = number of nodes in the network.

Functional Domain	aPCP	cPCP	Ketamine	NR1-knockdown
Development and differentiation	0.398	0.561	0.405	0.445
Intracellular signalling and regulation	0.365	0.385	0.454	0.396
Intracellular transport and organization	0.639	0.688	0.644	0.496
Biosynthetic processes and energy metabolism	0.582	0.61	0.573	0.554
Nucleic acid metabolism and ATP/GTPase activity	0.68	0.681	0.648	0.658

These scores indicate that overall the cPCP model represents schizophrenia the most closely for the four functional domains, “development and differentiation”, “intracellular transport and organization”, “biosynthetic processes and energy metabolism”, and “nucleic acid metabolism and ATP/GTPase activity,” more than the other three models. The Ketamine model represents “intracellular signalling and regulation” most closely, although the similarity scores between models and the disease were generally not so close for this domain (less than 0.5).

8.4 Discussion

Our current understanding of the underlying molecular pathology associated with schizophrenia is limited. Post-mortem studies and animal models of the disease can provide new insights into the patterns of alterations at the genetic and protein level which play a role in the neuropathology of schizophrenia(510,581,582). However, so far it has proven hard to quantify the molecular similarity of the models with the human disease pathology(507). These characterization issues have meant that progress has been limited in developing new pharmacotherapies from animal studies(583). As a consequence, few new chemical entities have reached the clinic over the last decades.

In this chapter, the methodology developed in **Chapter 7** is applied to conduct a non-hypothesis driven integration of proteomic data on the systems biology level enabling the direct comparison of brain changes from four rodent models of schizophrenia with those observed in schizophrenia post-mortem brains. Previous approaches have tried to establish various phenotypic similarities between animal models and schizophrenia through construct, face and predictive validity, assessments predominantly based on behavioural paradigms often supplemented with histological or electrophysiological investigations(507). However, these approaches have struggled to establish a particular model as the 'best for use' for a particular aspect of the disease because assessing similarities between animal behaviour and schizophrenia patient characteristics is typically vague, leading to difficulties regarding bias and irreproducibility(584). In addition, demonstrating that molecular changes in schizophrenia blood serum or plasma alone are equivalent to that of a relevant animal model has proven to be challenging(272).

Collectively, the application of this systems-based methodology in this study enables the identification of functional aspects identified in schizophrenia post-mortem brain tissue which are represented across all four animal models – intracellular signalling and regulation, development and differentiation, intracellular transport and localization, biosynthetic processes and energy metabolism, nucleic acid metabolism and ATP/GTPase activity. The described approach revealed that the latter four functional domains are represented most closely by the cPCP model. Blocking the NMDAR through PCP treatment has been previously associated with neurodegenerative pathologies in both humans and animal models(585), but while acute NMDA antagonist treatment has mainly been linked to disinhibition of the cortical transmitters glutamate, dopamine and serotonin (5-HT)(586,587), chronic NMDA antagonist administration is associated with more complex molecular and behavioural adaptations, leading to a more defined cognitive deficit profile(588,589). This could explain why the cPCP model outperforms the ketamine and aPCP models across most

categories. Reduced rates of oxygen uptake into mitochondria isolated from brain tissue have been observed in previous PCP rat studies, hence it is not unexpected that chronic PCP treatment represented impairments in energy metabolism, a common trait of psychiatric diseases(590–592). It is interesting to note that despite precautions of introducing a species-specific bias (e.g. joint rodent protein database, one shared functional annotation pool), the NR1 knockdown mice show the lowest similarity scores across the domains. Although all rodent models investigated in this study support the theory of a dysfunctional hypoglutamatergic frontal cortex state in schizophrenia post mortem brains(593), it is possible that compensatory developmental mechanisms following a single gene knockdown are more likely to influence and potentially dilute a functional cross-species similarity, than the molecular reaction of the frontal cortex to NMDAR antagonist exposure in adult animals. As an example of these compensatory mechanisms, recent research has indicated that while NR1 is a key mediator of cortical ErbB4 signalling, it is not the sole transmitter(594). The ErbB4 pathway integrates input from multiple ligands, and evidence suggests that when NR1 signalling is impaired, NR2 can partially compensate for this state in the brain. Concurrently, it has recently been shown in preclinical studies that a cortical loss of NR1 results in an increase in the expression and sensitivity of postsynaptic GABA receptors on pyramidal neurons(594). In addition, the results of this study could reflect the fact that the NR1 knockdown model may not necessarily model a schizophrenia-like state, as while human risk haplotypes for schizophrenia map to non-coding regions of the NR1 gene, in practise the NR1 knockdown may therefore not reflect a comparable expression/proteome level effect to the identified human risk SNPs(595). Finally, while it is thought that an optimal level of NR1 is required for the development and homeostasis of synaptic neurotransmission, and an imbalance in the inhibitory/excitatory ratio could result in the deficits in cortical synchronization observed in schizophrenia(596), evidence remains inconclusive as to whether diminished or enhanced NR1-mediated ErbB4 signalling is most likely to trigger psychosis-like phenotypes, with both increased and decreased expression levels being found in schizophrenia human post-mortem tissue(597,598).

While the five functional domains identified in this study are not specific to schizophrenia, previous research papers have identified impairments in each of these processes in the context of the disorder(554,591). One of the advantages of this methodology is that being able to identify which model most closely recapitulates different functional characteristics of the disease, could result in more informed future preclinical studies, targeting those particular models as a means of yielding new insights regarding the molecular basis of these abnormalities.

The current consensus regarding animal models of psychiatric disorders is that no single model can completely recapitulate the full complexity of human conditions(599). As an example, auditory electrophysiological abnormalities recorded in studies involving the NR1 knockdown model are thought to more closely resemble those seen in autism than schizophrenia(600). Conversely, this model is a good proxy for behavioural phenotypes which invoke some of the negative symptoms of schizophrenia such as impairments in spatial cognitive performance(580) and reduced social interaction. Hence, it is felt that future models of schizophrenia should be focused on behavioural endophenotypes and more importantly molecular alterations, as more understanding is gained of the genetic and neurodevelopmental causes(271,601). In particular given the failures of current medications in treatment of negative and cognitive symptoms of schizophrenia, preclinical models for different symptom clusters are likely to play an increasingly important role in new pharmacological approaches(427,602–604). It is felt that proteomic analysis has advantages over a genomic/transcriptomic based approach, as this provides a greater indication of the functional alterations within tissue(538). Additionally, while the heritability of schizophrenia has been approximated as being between 50% and 90% beased on twin studies, the genetic landscape may not be the best framework for comparison as schizophrenia appears to be a polygenic disorder with disease phenotypes arising through the accumulation of multiple small risk genes(502,605,606).

Our results provide evidence that different models can represent functional aspects of schizophrenia more closely than others. This analysis is the first to directly compare multiple animal models to schizophrenia on a functional level, with the findings supporting the notion that a variety of available models, each reflecting different pathological molecular hallmarks of schizophrenia could be important for insights into the molecular and cellular basis of behavioural abnormalities relevant to schizophrenia, as well as testing their responsiveness to existing and upcoming medication. This need will become increasingly important to reach a broader understanding of the ramifications of a given genetic, environmental or pharmacological manipulation in the context of psychotic spectrum disorders.

Chapter 9 Final discussion

In summary, this thesis uses proteomic profiling of serum and brain samples to: (a) investigate new diagnostic strategies for schizophrenia (**Chapters 3 and 4**), (b) examine data from a novel cellular model of schizophrenia, thus identifying new potential targets in the CNS and quantifying the potential of various compounds for future clinical trials (**Chapters 5 and 6**), and (c) develop a novel methodology for conducting a functional comparison between post-mortem tissue and various animal models of schizophrenia, offering the potential for more clinically informative preclinical rodent studies (**Chapters 7 and 8**). Central to this work has been the application and integration of relevant bioinformatic strategies for tackling these diverse problems, and adapting to the various forms of proteomic data generated by the experimental methodologies summarized in **Chapter 2**.

To conclude this thesis, the main findings for each of the three above mentioned sub-sections are summarized, and conclusions are drawn about their significance to the field of schizophrenia research. Subsequently limitations are discussed, along with potential future studies which may yield further insights.

9.1 Classifying and predicting schizophrenia through serum profiling

9.1.1 Summary of findings

As described in **Chapter 1**, there is a need for alternative diagnostic strategies to improve and complement the existing clinical interview based diagnostic approaches for schizophrenia. Previous attempts to develop serum based diagnostic and prognostic tests for the disorder, through identifying statistical models based on panels of protein biomarkers, have proven too costly for clinical utility and been hindered by various limitations in study design eg: lack of validation cohorts, and flaws in statistical methodology, as summarized in **Chapter 1**.

Chapters 3 and 4 present two different proteomic approaches to try and identify a panel of serum analytes which can diagnose schizophrenia and predict the onset of the disease. In **Chapter 3** a multiplex immunoassay platform is used to measure the concentration levels of a range of proteins across hormonal, inflammatory and metabolic pathways. In **Chapter 4**, targeted mass spectrometry is used to measure the abundances of a panel of peptides which have previously been linked to psychiatric disorders, many of which are not typically measured in standard commercially available assays. In both studies, multiple statistical

models are trained using a variety of methods, including machine learning and Bayesian methodologies, to assess which method is most applicable to the data, and thus best suited to solving this problem. Validation cohorts of antipsychotic naive schizophrenia patients and controls are included in the study designs, in addition to independent application cohorts to assess model performance at classifying antipsychotic treated patients from controls. Both studies have two cohorts of pre-onset individuals to assess whether the models can predict disease conversion in pre-symptomatic individuals, and whether they can predict psychosis conversion in a group of prodromal subjects. In each study, classification performance is computed on the full training dataset before different variable selection methods are used to identify important subsets of analytes in the data. Each statistical method is used to train new models on these subsets of analytes to answer the practical question of whether sufficient classification performance for clinical utility can be achieved with a smaller, and cheaper set of analytes.

The study presented in **Chapter 3** identifies an SVM model based on the concentrations of 66 proteins which is the only model to have high sensitivity, specificity and at least a “good” performance (AUC: 0.88) at classifying antipsychotic naive patients from controls in an independent validation cohort. However this model only serves diagnostic, rather than prognostic value, as it produces a “poor” performance when classifying pre-symptomatic individuals who are later diagnosed with schizophrenia from healthy controls, and fails at classifying psychosis converters from non converters. In addition, it would likely not be suitable for clinical use as it requires too many proteins to be financially viable. While different models based on smaller sets of proteins identified through variable selection, achieve a “good” AUC on the validation cohort, they lack sufficient sensitivity and specificity to be clinically viable.

The work presented in **Chapter 4** finds an SVM model based on the abundances of 21 peptides which is the only model in this study to classify antipsychotic naive patients from controls in both the training data and a validation cohort, with at least a “good” performance (AUC: 0.87-0.92) and high sensitivity and specificity. In addition, most notably, this same model can classify converters from non converters with an AUC of 0.88, and high sensitivity and specificity, thus illustrating its potential value as a diagnostic and prognostic test in individuals who are symptomatic. The model, however fails at classifying individuals who were pre-symptomatic at the time of sample collection, and later (over a period of months or years) reach a diagnosis of schizophrenia. None of the models based on smaller sets of peptides, as identified through variable selection, achieve the same combination of high diagnostic and prognostic performance. However an SVM model based on the abundances of 7 peptides, achieves a comparable performance on the training data and validation

cohort, in terms of AUC, sensitivity and specificity, to the model with 21 peptides. Similarly, a Bayesian LASSO model based on the abundances of 3 peptides achieves a slightly better performance on the cohort of converters and non converters. As such, these models could represent a more cost-effective alternative worthy of further validation, as purely diagnostic and prognostic tests respectively.

Both **Chapter 3** and **Chapter 4** indicate the confounding effect that antipsychotic treatment appears to have on how well a test classifies patients from controls. In **Chapter 3**, the SVM model based on the concentrations of 66 proteins achieves AUCs ranging from 0.56-0.87 across three of these application cohorts, and in **Chapter 4**, the SVM model based on 21 peptide abundances performs only poorly (AUC: 0.63) on an application cohort of treated patients and controls.

9.1.2 Significance of findings

The studies in **Chapters 3 and 4** represent the most rigorous attempts to date, both in terms of methodology and design, to identify a statistical model based on serum protein biomarkers which can act as a diagnostic and prognostic tool for schizophrenia. This is important as reviews which have attempted to understand why so few biomarkers have been converted into clinical tests for schizophrenia, have criticised a lack of rigour in statistical model application, and lack of validation in the findings of previous studies(607). The studies presented in these chapters attempt to overcome some of the limitations in previous research by robustly applying a range of statistical methods to quantify which is the most applicable to this particular problem, and conducting the largest studies of this nature in terms of sample size. **Chapter 3** measures serum protein concentrations across 1165 samples, while **Chapter 4** measures peptide abundances across 676 samples.

The ability of the SVM model based on 21 peptide abundances (**Chapter 4**)to classify both antipsychotic naive patients from controls and psychosis converters from non converters, with high AUC, sensitivity and specificity is the most notable finding from these studies. From a prognostic perspective, this result is particularly relevant because previous attempts to identify such a panel have proven inadequate, as described in **Chapter 1** and **Chapter 4**. Previous reviews have stipulated that if a disease trace could be detected in the serum at an early stage, it could result in better patient outcomes. The current debates surrounding the prodromal phase of schizophrenia arise from the lack of diagnostic tools which can accurately identify the 20-30% of individuals who go on to develop the disorder over a two to three year period(608). Studies have stipulated that this area of unknown results in issues regarding stigmatization, and the potential for inappropriate treatment(608,609). Market

analysis has suggested that a blood test which possesses prognostic abilities to predict conversion in pre-onset individuals, and could thus be used in conjunction with the current structured clinical interviews, would be highly valued by psychiatrists(610).

From a diagnostic perspective, the SVM model based on these 21 peptides is more advantageous than the SVM model based on 66 proteins identified in **Chapter 3**, firstly because it is more sensitive and specific on an independent cohort, and secondly because it requires the measurement of fewer analytes. The SVM model based on 66 proteins would likely not be clinically viable as one of the reasons that a previous biomarker test for schizophrenia (measuring 51 proteins), was withdrawn, was because it was too expensive(152). In addition, the SVM model based on 21 peptides would be more practical because targeted mass spectrometry is more cost effective than immunoassays and requires less sample volume(283). As such, the comparative results between **Chapters 3 and 4** suggest that targeted mass spectrometry is a preferable method for future studies seeking to identify prognostic and diagnostic biomarker panels for psychiatric disorders. While immunoassay platforms are limited to commercially available assays, mass spectrometry provides the option of targeting analytes linked to these disorders such as VTNC and RET4 for which assays are not readily available.

In both chapters, SVM is the statistical method which comes closest to providing an optimal solution. SVM has been widely utilized to identify diagnostic and prognostic biomarker panels in other areas of medicine such as cancer research, and SVM classifiers are felt to have good generalizability, aiding the reliability of performance on independent cohorts(611). The other methods explored in these chapters identified models which produced sufficient performance in terms of the AUC, but did not have sufficient sensitivity and/or specificity. Previous reviews have stipulated that it is important for biomarker tests for schizophrenia to be both highly sensitive and specific, as described in **Chapter 1**, hence models which did not achieve at least 80% sensitivity and specificity were not considered. The performance of the SVM model with 21 peptides (**Chapter 4**) compares well with previously published proteomic biomarker studies for schizophrenia. It has higher sensitivity and specificity in classifying antipsychotic naive patients from controls, than the 51-biomarker assay which was briefly implemented as a clinical diagnostic for the disorder(132). In addition, the model's performance seems to compare favourably with the results of other diagnostic solutions being explored for schizophrenia. It appears to be more sensitive than a recently discovered panel of metabolomic biomarkers which reported a sensitivity and specificity of 82% and 89% in classifying patients from controls, while imaging techniques such as MRI, PET and single-photon emission computerized tomography have yet to achieve sufficient sensitivity and specificity for clinical use(612). In terms of understanding the pathophysiology of

schizophrenia, the 21 peptides provide novel information as the performance of the SVM model in classifying converters from non converters suggests they may be implicated in the development of psychosis during the prodromal phase. Most have been previously linked to schizophrenia through either GWAS or proteomic studies(25,34,355,391,392,395,398,613). The finding that this model fails at classification on the prediction cohort of pre-symptomatic individuals is notable in terms of connecting symptomatology to disease pathophysiology. It indicates that the abundances of these 21 peptides can only be used to classify individuals before disease onset, who are already displaying initial symptoms and signs of altered global functioning.

The studies presented in **Chapters 3 and 4** are also unique as they use different variable selection methods, based on the different statistical techniques, to prune the model space from the original set of analytes in the training data, and obtain several sets of what are deemed to be the most important proteins/peptides. While an alternative approach to this problem could be to utilize the variable importance scores or inclusion proportions computed in tree-based models such as Random Forest and BART, and rank the proteins/peptides before iteratively fitting models to the top 2 analytes, top 3 analytes etc and computing performance, such an approach would not be applicable to methods such as SVM or Bayesian LASSO. The findings of the model analysis on reduced sets of analytes in **Chapter 4** yields a couple of potential alternative solutions worthy of further investigation. In particular, the Bayesian LASSO model based on just 3 peptides could represent a financially viable prognostic test if it is found to validate on further prodromal cohorts of psychosis converters and non converters. In addition, Bayesian models offer useful practical advantages from a clinician's perspective as through the posterior distribution they can provide a credible interval for classification performance which may be more informative than just a point prediction(187).

9.1.3 Limitations

While the performance of the SVM model of 21 peptides, identified in **Chapter 4**, is a notable finding, there are not sufficient independent cohorts in this study to truly assess whether it may have value as a diagnostic and prognostic test in the wider disease population. However, as has been already stated, such cohorts, especially those of prodromal individuals, are difficult to obtain and the studies presented in **Chapter 3** and **Chapter 4** are larger than any other diagnostic study thus far for schizophrenia.

An inherent limitation of this work in the field of psychiatric disorders is the fact that it is difficult to get a training dataset which is large enough to gain an accurate representation of

the relative levels of these analytes between “disease” and “control” groups on a global level. As such, the size of the training datasets in **Chapter 3** and **Chapter 4** with 127 patients and 60 patients respectively, are unlikely to be sufficient. Likewise, it may be that a much wider range of analytes needs to be measured in order to identify a set of biomarkers truly fundamental to the pathophysiology of schizophrenia both throughout the prodromal phase and after disease onset. However, due to the relative paucity of research studies which have explored the potential of serum biomarkers as a diagnostic and prognostic tool for schizophrenia, these results can be seen as exploratory, pointing towards further investment in larger studies.

One of the technical limitations of the study presented in **Chapter 3** is that while up to 225 protein concentrations are measured by the multiplex immunoassay platform, only 66 proteins are measured across all 11 cohorts, and are able to be incorporated in the study. This is because in every cohort, a certain proportion of the proteins are excluded as part of pre-processing due to missing values.

Another limitation is the lack of information regarding details of antipsychotic medication for the application cohorts in **Chapter 3**. This limits the conclusions which can be drawn on why classification performance varies between these cohorts. Finally, while the demographic variables age and gender were incorporated as predictor variables where possible in **Chapters 3 and 4**, it was not possible to include other demographic variables such as body mass index(614) and cannabis consumption(615) which have been previously linked to schizophrenia and could be important predictors in conjunction with protein biomarkers, as they were either not recorded or only partially recorded during sample collection. Similarly, other demographic variables which were not considered included smoking, the use of contraceptives and menstrual cycle phase. These factors are known to potentially influence the levels of proteins involved in hormonal and metabolic pathways known to be dysregulated in at least a subset of schizophrenia patients(12). Future study designs should aim to account for these factors.

9.1.4 Future work

To further investigate the true predictive performance of the SVM model with 21 peptides (as a diagnostic and prognostic test), the SVM model with 7 peptides (as a diagnostic test), and the Bayesian LASSO model with 3 peptides (as prognostic test), it would be necessary to validate these models on further independent cohorts from different clinical centres. In particular, it would be necessary to investigate whether the prognostic tests still perform well when classifying converters at different stages of the prodromal phase.

Moreover, it may also be of interest to further investigate the role of the 21 peptides, and their associated proteins, in the development of schizophrenia. Elucidating the nature of how these alterations in the periphery represent pathophysiological changes occurring in the CNS during the prodromal phase of the disease may yield new insights regarding schizophrenia pathogenesis.

In addition, it would be of interest to determine whether each of these models can act as a disease-specific differential diagnostic. Most notably, schizophrenia is commonly misdiagnosed as bipolar disorder (139), and there have been cases of misdiagnosis with MDD (616). In addition, there is a distinct overlap in clinical characteristics between schizophrenia and other disorders such as obsessive compulsive disorder (OCD) (617), and attention deficit hyperactivity disorder (ADHD) (618).

The ability of the SVM model with 21 peptides, the SVM model with 7 peptides, or the Bayesian LASSO model with 3 peptides to act as differential diagnostic tests could be assessed by fitting them to independent cohorts of schizophrenia patients and bipolar disorder patients, as well as schizophrenia-MDD, schizophrenia-OCD, and schizophrenia-ADHD cohorts, and assessing performance. If the model performance is not sufficiently good in these comparisons, then this indicates that these peptides are not disease-specific to schizophrenia. In this instance, it may be necessary to obtain separate individual tests for schizophrenia-bipolar, schizophrenia-MDD, schizophrenia-OCD and schizophrenia-ADHD by training and testing new models on a range of cohorts for each of these comparisons, using a different panel of peptides that are thought to be uniquely implicated in schizophrenia pathophysiology based on existing biomarker studies. This work would be worth undertaking as tests which could elucidate schizophrenia from these other disorders would be invaluable to clinicians.

The results of **Chapter 3 and Chapter 4** indicate that a separate proteomic test would likely be required to accurately classify recent-onset antipsychotic treated patients from controls. Such a test may be clinically useful in instances where individuals have been pre-treated with antipsychotics during the prodromal phase. This study would involve training models on a cohort of recent-onset antipsychotic treated patients and controls, and then testing each model on independent cohorts of patients treated with either first generation or second generation antipsychotics or a mixture of both, to assess robustness across different forms of treatment. It could be useful if such a study design included independent cohorts of bipolar disorder and MDD patients treated with medication, to assess whether a proposed test could classify treated schizophrenia patients from treated patients with other psychiatric disorders.

Finally, it is hoped that proteomic biomarker tests may eventually be able to elucidate different subgroups from within the schizophrenia population, based on different pathophysiological mechanisms, such as those outlined in **Chapter 1**. If training and test datasets of sufficient sample size could be accrued, measuring a wide range of molecular pathways, then it would potentially be possible to conduct an exploratory study separating patients into different subgroups using clustering analysis, based on the levels of different proteins.

9.2 Investigating disturbances in microglial signalling through analyzing data from a novel cellular model of schizophrenia

9.2.1 Summary of findings

Despite the many hypotheses regarding the role of dysfunctional microglial activation in the etiology and pathophysiology of schizophrenia, the empirical evidence to suggest that microglial abnormalities may influence the course of the disease largely emanates from PET imaging data, which has yielded inconsistent findings, and post mortem studies. **Chapters 5and6** present analyses of data obtained from the first *in vitro* studies which attempt to characterize the intracellular mechanisms underpinning microglial activation in response to serum exposure from first-onset antipsychotic naive schizophrenia patients (**Chapter 5**) and recent-onset antipsychotic treated schizophrenia patients (**Chapter 6**). The cellular data in the latter study is analyzed in conjunction with PET readouts.

The two studies identify different cellular mechanisms of aberrant microglial activation, both pointing towards a pro-inflammatory state, classically regarded as M1 polarization. The findings in **Chapter 5** suggest that in antipsychotic naive schizophrenia patients this phenotype is primarily associated with increased activation of the Akt/mTORC1 pathway. Differential expression levels of the epitopes 4EBP1 (pT36/pT45), 4EBP1 (pT69) and eIF4E (pS209) appear to be a key factor in these signalling alterations. Concurrently, activation of the JAK/STAT pathway, as observed through increased expression levels of the stat3 (pY705), Stat3 and SHP2 (pY542) epitopes, may interact with impaired Akt/mTORC1 signalling, to switch microglia from a resting state to an activated pro-inflammatory phenotype. Serum profiling changes indicate that these signalling alterations may be induced by circulating serum factors including increased levels of members of the complement cascade, as well as reduced levels of apolipoproteins which have previously been associated with neuroprotective mechanisms in microglia. Screening of the known

microglial activation inhibitor, rapamycin, significantly reduces the expression of all epitopes apart from SHP2 (pY542), suggesting it may have the potential to impart a normalizing effect across the Akt/mTORC1 and JAK/STAT pathways, down-regulating epitope expression to the levels seen in the healthy state. The antibiotic minocycline displays a more selective profile, only significantly reducing the expression of stat3 (pY705).

In **Chapter 6**, the results of PET imaging data implies increased microglial activation associated with neuroinflammation in the temporal regions of recent-onset antipsychotic treated patients. This increased activation is found to be associated with increased symptom severity. Concurrently, cell signalling data from microglial exposure to serum from these patients, identifies impairment in JAK/STAT pathway signalling through reductions in the expression of epitopes stat3 (pS727), stat4 (pY693), and SHP2 (pY542). Impaired JAK/STAT signalling appears to be linked to dysfunctional microglial activation in these patients, as reduction in stat3 (pS727) expression is significantly correlated with the increased microglial activation seen in the PET data. Serum profiling changes identified by mass spectrometry suggest a pro-inflammatory state through increased abundances of haptoglobin, α-1 antitrypsin and Apolipoprotein E. There is evidence that increased expression of the latter in the brain exacerbates neuroinflammation(484). Meta analysis of two cohorts of recent-onset antipsychotic treated schizophrenia patients and controls provides further evidence as to impairment in JAK/STAT signalling in this patient population, and univariate analysis of the latter cohort suggests that reductions in the expression of p38 (pT180/pY182) may also be involved. In a healthy state, increased expression of this epitope is involved in regulating microglial response to insult and injury and attenuating neuroinflammation before it has deleterious consequences in the CNS(488). 14 of the 24 patients in these two cohorts are treated with olanzapine, a drug which has been previously shown to modulate JAK/STAT signalling and this study provides evidence that it, and other antipsychotics, may contribute to the signalling impairments observed in this particular study. When olanzapine and haloperidol are screened against the five epitopes identified as having altered expression in patients in either of the cohorts, they are found to significantly change expression levels across all of them in the same direction as in the disease state.

9.2.2 Significance of findings

The findings presented in **Chapter 5** are the first *in vitro* evidence that circulating serum analytes can induce signalling changes in microglia which may represent a dysfunctional activation phenotype contributing to the pathogenesis of schizophrenia. They are particularly notable in the context of the fact that several of the serum proteins (e.g. apolipoproteins)

identified as being altered in patients through immunoassay and mass spectrometry profiling, have been associated with increased BBB permeability in subgroups of schizophrenia patients(619).

In addition, this study provides the first direct experimental evidence of impaired Akt/mTORC1 signalling in a model of schizophrenia. While pathways such as Akt/mTOR and JAK/STAT may seem relatively generic, their implication in **Chapters 5 and 6** in a cellular model of schizophrenia, created using a CNS cell population, is particularly notable. This is because disrupted Akt/mTORC1 signalling has been shown to result in abnormal neurodevelopment and deficient synaptic plasticity, and as such, it has been hypothesized as a potential causal factor in the neuropathology of schizophrenia(620). Prior to this study, the only evidence associating Akt/mTORC1 with schizophrenia came from studies reporting over-activation or inhibition of this pathway in response to environmental stressors or upstream activators linked to schizophrenia(620). Similarly, this is the first direct evidence linking Stat3 signalling to schizophrenia. Microglial Stat3 signalling has previously been identified in other neuropsychiatric disorders such as MDD(403), and given the link between impaired microglial activation and cognitive deficits in schizophrenia(403), it is plausible that Stat3 may be involved in the negative symptomatology of the disease. The identification of both Akt/mTORC1 activation and JAK/STAT activation through Stat3 is notable as rodent studies have previously suggested that Akt/mTORC1 can stimulate a proinflammatory response via Stat3(621). Taken together, the findings of this study suggest that the impact of peripheral alterations on the CNS may have more profound implications in the pathophysiology of schizophrenia than previously anticipated. As such, the epitopes identified in this study as being key to impaired Akt/mTORC1 and JAK/STAT pathway activation in first-onset schizophrenia patients may represent novel drug targets, providing potential for therapeutic intervention. The capability of rapamycin of normalizing the expression of all but one of these epitopes, suggests that the modification of microglial function, either by direct pharmacological inhibition or the normalization of circulating ligands with microglial activation propensity, represents a potentially novel strategy of reversing dysfunctional signalling patterns in the CNS at a relatively early stage. The more specific profile of minocycline, targeting stat3 (pY705), is also notable as studies have suggested it displays efficacy at improving negative symptomatology which is notoriously treatment-resistant(468).

The findings presented in **Chapter 6** represent the first study to combine functional PET imaging information regarding microglia in schizophrenia, with cell signalling and serum proteomic data. It has been previously stipulated that PET studies should look to combine information regarding changes in microglia intracellular cascades to aid interpretation of the

results. The finding of increased microglial activation in the temporal region in recent-onset antipsychotic treated schizophrenia patients matches with the results of a previous PET study examining both treated patients and individuals in the prodromal phase of schizophrenia(401). It is notable that JAK/STAT signalling is once again implicated in dysfunctional microglial activation, and the changes in epitope expression suggests that this occurs through a different mechanism compared to first-onset antipsychotic naive patients. It appears possible that this mechanism is either induced or perpetuated by antipsychotic treatment, which corresponds to findings from previous preclinical and post-mortem studies implicating both typical antipsychotics and atypical antipsychotics such as olanzapine and clozapine in driving neuroinflammation. However to date the exact mechanisms by which this happens in the CNS remain poorly understood. The results of this study provide novel *in vitro* evidence for their mechanisms of action in microglia.

9.2.3 Limitations

To provide a reference point to previous work regarding microglial activation in schizophrenia, the conclusions of both **Chapters 5 and 6** point towards the classical microglial polarization definitions of M1 and M2(407). However, as has been discussed in the recent scientific literature, these broad definitions are likely to be an oversimplification, and in reality microglia are likely to assume multiple intermediate phenotypes, for example M2a M2b, M2c and Mox12(406,409,622,623). However as of yet, the full spectrum of phenotypes and their individual definitions in terms of function, remains little understood. One of the limitations of the study in **Chapter 6** it that it is not possible to conclude definitively whether the differences in the phenotype observed, compared to that of **Chapter 5**, are a result of antipsychotic treatment or disease duration. As will be discussed in **9.2.4**, the construction of future study designs could provide more information regarding these questions.

From a statistical perspective, a limitation of both studies in **Chapters 5 and 6** is the fact that a number of the significant findings do not survive correction for multiple testing. However, as the first *in vitro* studies to examine microglial activation in the context of schizophrenia, these should be viewed as exploratory analyses and the findings are intended as a starting point for future work (as discussed in **9.2.4**) rather than definitive proof. In addition, both studies are statistically underpowered and hence it is not possible to draw firm conclusions, especially in light of the heterogeneity of schizophrenia. However the results demonstrate that these cellular models have the potential to shed new light on microglial disturbances in schizophrenia, particularly in conjunction with other experimental techniques such as PET

imaging and thus are worthy of future investigation with larger cohorts. The small cohort sizes in both studies reflects the difficulty and high cost of obtaining serum or imaging data from schizophrenia patients, a common limitation of psychiatric research. The size of the cohort used for the PET analysis in **Chapter 6** is representative of typical cohorts in previously published PET research in schizophrenia(302,401,471,475). In addition, permutation testing was conducted for all univariate and multivariate analyses to ensure an accurate distribution of p-values. In future, once more studies have been conducted using these cellular models, meta analyses could be employed, for example across the results obtained on different populations of antipsychotic naive patients and controls. Meta analyses across studies may also yield more information from the existing PET studies of microglial activation in schizophrenia. One of the limitations of these studies is that they often fail to clearly define the patient population, which is crucial for comparison as microglia shift in morphology and function through the course of the disease, and with the effects of treatment(480). There may be substantial differences between individuals with prodromal schizophrenia to those with chronic schizophrenia who have been treated for many years. As observed in this thesis, the intracellular mechanisms in a cohort of first onset antipsychotic naive patients differ to a cohort of recent onset antipsychotic treated patients, hence meta analyses would need to be conducted between studies with similarly defined patient groups, for example first-onset antipsychotic naive patients.

9.2.4 Future work

As has been mentioned previously, microglial phenotypes are not static and continuously shift during the course of disease progression(480). Therefore, it could be of particular interest to use this cellular model to gather further data on microglial signalling alterations in response to serum collected at different points in the prodromal phase, in the five years following disease onset, and in chronic patients. Data collected early in the course of the disease may be of particular interest as recent studies suggest that many of the pathological changes attributed to microglial dysfunction, such as altered synaptic refinement, occur early in the course of schizophrenia, and only in a subset of patients(227,403,624). Therefore, from a perspective of developing new therapeutic strategies, it is of interest to identify the optimal time point for early intervention. If sufficient sample sizes could be accrued, it may be possible to use microglial pathology to characterize patient subgroups at different stages of the disease by using clustering techniques to identify groups of intracellular biomarkers.

In particular, the results in **Chapter 5** imply that more studies should be conducted to further characterize the process of Akt/mTORC1 pathway activation in microglia in relation to

schizophrenia, and the interaction with the JAK/STAT pathway. Given the identification of microglial Stat3 signalling as being impaired in multiple neuropsychiatric disorders, measurement of key epitopes in this pathway in response to circulating serum factors at critical neurodevelopmental phases could reveal new information regarding schizophrenia pathogenesis and in particular, the little understood negative and cognitive symptomatology. In addition, although many of the proteins reported in **Chapter 5**, such as complements and apolipoproteins, are consistent with previous reports of changes in circulating mediators of microglial function, further characterization is required to determine their relationship to the observed signalling changes in the Stat3 and mTORC1 pathways. Further to the results acquired in **Chapter 6**, study designs which examine changes in microglial phenotypes in the same patients before and after treatment, over the course of a longitudinal study, could reveal further information on the role of antipsychotics in altering microglial function and perhaps contributing to disease pathology.

Furthermore, one of the limitations of the current experimental methodology is that only three epitopes can be measured at a time. Alternative experimental techniques such as mass cytometry could be used to enable the simultaneous measurement of as many as 50 epitopes. This would allow the use of more complex statistical techniques such as Bayesian networks to make inferences regarding signalling links in key pathways such as Akt/mTORC1, JAK/STAT and MAPK. A similar methodology has been previously successfully applied to infer protein signalling networks in a breast cancer cell line(301). Finally, given the demonstrated ability of this model in providing a functional basis for understanding the cellular mechanisms of microglial specialization, it could in future be utilized to analyze serum responses to microglia obtained through reprogramming induced pluripotent stem cells (iPSCs) derived from schizophrenia patients and controls, thus offering the potential for patient-specific modelling of microglial dysfunction(199,625,626).

9.3 Developmental and application of a methodology for the functional comparison of proteomic changes in rodent models to those in psychiatric disorders

9.3.1 Summary of findings

The current bottleneck in drug development for schizophrenia and other psychiatric disorders is thought to come down to a variety of factors including the cost and length of time required to discover novel compounds, a lack of suitable targets due to reasons such as the failure of many proteomic biomarkers to validate, and an inability to relate changes observed

in rodent models of these disorders to the human condition(504). While the development of cellular models in recent years, such as the model of microglial activation in schizophrenia analyzed in **Chapters 5 and 6**, have provided an alternative for some preclinical work, rodent models remain a vital part of the drug discovery pipeline as they allow for the study of genetic or environmental manipulations associated with a disorder in fully developed neuronal systems. However, there remains a considerable need for new strategies to improve their utility by quantifying how a particular model represents different pathophysiological characteristics of a psychiatric disorder(270).

Chapters 7 and 8 present the development and application of a novel systems methodology to quantify the molecular similarity of protein alterations observed in these models with the human disease pathology. As outlined in **Chapter 7**, this methodology creates PPI networks from protein alterations between control and disease states in both human and animal tissue, thus leveraging the mass spectrometry-identified proteins as being representative of changes in signalling pathways by introducing validated molecular interaction partners. Subsequent GO enrichment enables a comparison on the functional level, separating identified GO terms into functional groups using the kappa score metric, based on the underlying proteins shared by overlapping terms. Hierarchical clustering was used to identify related constellations of GO term groups representing neuropathological domains which shared functional similarities between each model and the disease. Finally, kernel based machine learning methods were used to obtain a numerical quantification of how representative each model is of the disease for particular domains, by computing the pairwise similarity scores for the underlying vectors of GO terms based on their relative distance in the GO hyperspace.

Subsequently this methodology was applied to compare proteomic changes in brain samples from three commonly used environmental stress models to those observed in MDD post-mortem tissue (**Chapter 7**), and changes in four common glutamatergic models of schizophrenia to those observed in schizophrenia post-mortem tissue (**Chapter 8**). In **Chapter 7**, seven neuropathological domains associated with MDD and represented across at least two models were identified. Through statistical evaluation using kernel-based machine learning techniques, the social defeat model was found to represent MDD brain changes most closely for four of the seven domains. In **Chapter 8**, five neuropathological domains were identified in schizophrenia post-mortem brain tissue and represented across all four animal models. Of these domains, four were represented most closely by the cPCP model.

9.3.2 Significance of findings

The methodology described is the first approach to enable a cross-species comparison of proteomic data from human and animal tissue on the functional level. This has the potential to aid scientific decision-making regarding which models of affective and psychotic spectrum conditions warrant further prioritization. This novel approach could also enable more targeted studies by virtue of estimating which models most closely recapitulate particular neuropathological facets of interest for a psychiatric disorder.

Due to the limited overlap in proteome coverage comparing rodent and human tissue samples, previous neuropsychiatric evaluations have been almost entirely restricted to comparing observed behavioural characteristics between animals which were either thought to mirror disease-specific endophenotypes, or have been demonstrated to possess predictive probabilities tailored towards already known mechanisms of action of psychotropic medication(270). However it has proven challenging to quantify these comparisons, leading to issues of reproducibility. In contrast, the methodology introduced in this thesis enables the direct quantification of which models represent specific salient features of MDD or schizophrenia most closely, based on underlying pathological molecular hallmarks, thus representing a new advance in connecting brain proteomic fingerprints with behavioural patterns in animal models of major psychiatric diseases. As such, this may aid future drug development, as researchers will be able to determine which rodent model will be most suitable for preclinical testing of a novel compound aimed at a particular aspect of human pathophysiology (eg glutamatergic disturbances or axonal transport deficiencies).

The application of this methodology in **Chapters 7 and 8** also revealed that comparing and contrasting findings between even aetiologically diverse rodent models and a psychiatric disorder may allow for the identification of functional domains which are represented strongly by multiple different models. These common patterns of change may represent mechanisms which underlie fundamental neurobiological changes present in MDD or schizophrenia. As such, it has been suggested that the functional overlap of the underlying protein abundance changes in multiple models of different aetiology with protein abundance changes found in MDD or schizophrenia brain samples enhances the construct validity of each of these models(627).

While the five neuropathological domains of schizophrenia represented across the four rodent models in **Chapter 8**, may appear relatively generic, they correspond to previous preclinical findings relating to schizophrenia. In particular, the abnormalities identified relating to glycolysis and mitochondrial energy metabolism have appeared in multiple

studies, and have been postulated to potentially affect a variety of processes in the CNS which are dependent on adenosine triphosphate (ATP) such as presynaptic transmitter release and intracellular transport. This can lead to impaired information processing across brain regions, and thus these mechanisms have been implicated in psychiatric disorders, so it is not surprising to see them reflected to various degrees across all models(43). It is an advantage of the methodology that in both **Chapters 7 and 8**, a certain model (e.g. the chronic PCP model in **Chapter 8**) can be found to recapitulate the majority of the identified domains most strongly. This allows for more informed future studies which target that particular model, with the aim of understanding the molecular underpinning of these functional impairments in greater detail, in the context of the disease.

9.3.3 Limitations

One of the limitations of this methodology is the fact that in the present form, it cannot account for directionality in terms of protein signalling. While the effects on a particular function/process can differ depending on whether the proteins associated with a particular GO term are up/downregulated, there is as yet no standard annotation to account for this in PPI networks. In addition, one has to keep in mind that ultimately the identified domains were limited by the detected fraction of the proteome and therefore are less likely to contain GO terms based on proteins that are difficult to detect in whole-tissue approaches, for example membrane-integral proteins or proteins with very specific expression time windows (e.g. proteins involved in apoptosis signalling).

As with the studies discussed in **9.2**, the studies described in **Chapters 7 and 8** in which this method is applied, are statistically underpowered. This relates both to the limited availability of post-mortem brain tissue samples relating to psychiatric disorders, and the exploratory nature of the study meaning that only a limited amount of animal tissue was available. As such, while the results of each study should be interpreted with that in mind, the method outlined is intended as a template which could be used for future studies comparing further animal models of a psychiatric disorder to the human disease in question and would remain applicable with larger numbers of samples.

In addition, in the presented studies, despite taking steps to reduce a rodent species-specific bias through a joint protein database, it is likely that full comparability of preclinical models might only be achievable on a species specific level. As both mice and rats come with possible advantages to potential future experimental designs (tissue amount and sample size, available behavioural paradigms, ease of genetic manipulation) the results regarding the comparative validity have to be interpreted with caution. Future studies applying this

methodology should aim to use rodent models created in a single species. In addition, it should be noted that the applied proteome extraction protocol and the subsequent in silico analyses in both human post-mortem brain and rodent tissue can only deliver an approximate view of the underlying proteomes as multiple cell types are used simultaneously and are therefore represented jointly in the significant abundance changes. Subsequent analyses should consider the fractionization of samples, rendering the analyses of targeted sub-proteomes possible.

9.3.4 Future work

The flexible structure of the framework proposed in **Chapters 7 and 8** allows for the re-evaluation of cross-species network comparisons once new information is available. Future research should aim to introduce proteomic information of different putative MDD and schizophrenia animal models for comparison with the human disease pathology. While the analysis presented in **Chapter 8** focused on the glutamatergic system through examining three pharmacological models of schizophrenia and one gene-knockout model, future studies may look to include dopaminergic manipulations using direct and indirect dopamine agonists which have previously shown to induce behavioural phenotypes associated with positive and negative symptoms of schizophrenia such as supersenstivity to psychostimulants or persisting prepulse inhibition abnormalities and various other sensory gating deficits(255). In addition a future study could focus on comparing further models based on different genetic manipulations, for example the 22q11.2 deletion syndrome mouse model which is thought to model some of the negative and cognitive symptoms of schizophrenia. Moreover, while the analyses presented in **Chapters 7 and 8** are based on the profiling of tissue from the anterior prefrontal cortices in humans and rodents due to previous studies linking this region to altered top-down control, future studies may wish to investigate subcortical brain regions eg: the limbic system, including the hippocampus.

Finally, the significantly changed protein sets which underlie each of the identified functional domains could form the basis of future assays for drug development using techniques such as MRM. Once such key pathway candidate biomarkers have been identified, MRM can be applied to analyze their abundance levels in a selective, quantitative manner. This has been previously reported on in a study of anterior prefrontal cortex tissue from patients with schizophrenia, bipolar disorder and MDD, in which expression levels were quantified for a panel of 56 proteins suggested to be associated with various functional aspects of these

disorders , including for example alterations in cellular energy metabolism and dysfunction of neuronal differentiation(291).

Appendix

Chapter 3 - Appendices

Table A.3.1 66 serum proteins and their abbreviations which were measured across the 11 cohorts analyzed

Analyte	Abbreviation
Adiponectin	ADPN
Alpha-1 antitrypsin	A1AT
Alpha-2 macroglobulin	A2M
Angiopoietin-2	ANGPT2
Apolipoprotein A1	ApoA1
Apolipoprotein CIII	ApoCIII
Apolipoprotein H	ApoH
AXL receptor tyrosine kinase	AXL
Beta-2 microglobulin	B2M
Brain-derived neurotrophic factor	BDNF
C reactive protein	CRP
CD40 antigen	CD40
CD40 ligand	CD40L
Chemokine CC4	CCR4
Complement C3	C3
Cortisol	Cortisol
Creatine Kinase MB	CK-MB
Extracellular newly identified receptor for advanced glycation end-products binding protein	EN-RAGE
Eotaxin	Eotaxin
Epidermal growth factor	EGF

Epithelial derived neutrophil activating protein 78	ENA-78
Factor VII	FVII
FASLG receptor	FasR
Ferritin	Ferritin
Follicle stimulating hormone	FSH
Haptoglobin	HAPT
Hepatocyte growth factor	HGF
Immunoglobulin A	IgA
Immunoglobulin M	IgM
Insulin-like growth factor binding protein 2	IGFBP2
Intercellular adhesion molecule 1	ICAM
Interleukin 16	IL16
Interleukin 18	IL18
Leptin	Leptin
Lipoprotein (a)	Lp(a)
Macrophage derived chemokine	MDC
Macrophage inflammatory protein 1 beta	MIP1 beta
Macrophage migration inhibitory factor	MIF
Matrix metalloproteinase 3	MMP3
Monocyte chemotactic protein 1	MCP1
Myeloperoxidase	MPO
Myoglobin	MB
Pancreatic polypeptide	PPP
Plasminogen activator inhibitor 1	PAI1
Platelet derived growth factor	PDGF
Progesterone	P4
Prolactin	PRL
Pulmonary and activation regulated chemokine	PARC
Receptor for advanced glycosylation end products	RAGE

Resistin	Resistin
Serum amyloid p component	SAP
Sex hormone binding globulin	SHBG
Sortilin	Sortilin
Stem cell factor	SCF
Superoxide dismutase	SOD
T cell specific protein RANTES	RANTES
Tenascin C	TNC
Testosterone	TEST
Thrombospondin 1	THBS1
Thyroid stimulating hormone	TSH
Thyroxine binding globulin	TBG
Tissue inhibitor of metalloproteinases 1	TIMP1
Tumor necrosis factor receptor like 2	TNFR2
Vascular cell adhesion molecule 1	VCAM1
Vascular endothelial growth factor	VEGF
von Willebrand factor	VWF

Table A.3.2 Significant (q<0.05) serum protein concentration differences between antipsychotic naive schizophrenia patients and healthy controls as analyzed in Chapter 3

Molecular Function	Analyte	Abbreviation	Fold Change	Covariates	P-value	Q-value
Inflammatory response	Macrophage migration inhibitory factor	MIF	1.73	None	2.90E-10	1.91E-08
Hormonal signalling	Pancreatic polypeptide	PPP	1.92	None	8.25E-10	2.72E-08
Inflammatory response	Ferritin	Ferritin	1.82	None	5.96E-08	1.31E-06
Hormonal signalling	Leptin	Leptin	0.50	None	5.50E-07	9.07E-06
Inflammatory response	Extracellular newly identified receptor for advanced glycation end-products binding protein	EN-RAGE	1.71	None	8.08E-07	1.07E-05
Inflammatory response	Tenascin C	TNC	1.28	None	2.50E-06	2.75E-05
Growth factor signalling	Insulin-like growth factor binding protein 2	IGFBP2	1.31	None	5.0604E-06	4.77E-05
Clotting cascade	Factor VII	FVII	0.83	None	6.81E-06	5.6169E-05
Hormonal signalling	Cortisol	Cortisol	1.24	None	8.49E-06	6.2271E-05
Lipid transport	Apolipoprotein CIII	ApoCIII	0.83	None	0.000012	8.0341E-05
Inflammatory response	Receptor for advanced glycosylation end products	RAGE	0.79	None	0.00015	0.0009
Inflammatory response	Alpha-1 antitrypsin	A1AT	1.12	None	0.00027	0.0013
Inflammatory response	Haptoglobin	HAPT	1.45	None	0.00026	0.0013
Hormonal signalling	Progesterone	P4	1.23	None	0.00026	0.0013
Hormonal signalling	Follicle-stimulating hormone	FSH	1.41	None	0.00043	0.0019
Immune system	Superoxide dismutase	SOD	1.25	None	0.0011	0.0047
Inflammatory response	von Willebrand factor	VWF	1.19	None	0.003	0.01
Inflammatory response	Alpha-2 macroglobulin	A2M	1.10	None	0.004	0.01
Growth factor signalling	AXL receptor tyrosine kinase	AXL	0.90	None	0.004	0.01
Growth factor signalling	Angiopoietin-2	ANGPT2	1.18	None	0.006	0.02
Inflammatory response	Beta-2 microglobulin	B2M	0.95	None	0.009	0.03
Lipid transport	Apolipoprotein H	ApoH	1.08	None	0.01	0.03
Inflammatory response	Interleukin 16	IL16	1.12	None	0.01	0.03
Immune system	Myeloperoxidase	MPO	1.22	Gender	0.01	0.03
Hormonal signalling	Testosterone	TEST	1.25	None	0.01	0.03
Hormonal signalling	Apolipoprotein A1	ApoA1	0.91	None	0.01	0.04
Lipid transport	Immunoglobulin A	IgA	0.89	None	0.02	0.04
Immune system	Thyroid stimulating hormone	TSH	0.86	Gender	0.02	0.04
Growth factor signalling	Stem cell factor	SCF	0.91	None	0.02	0.05

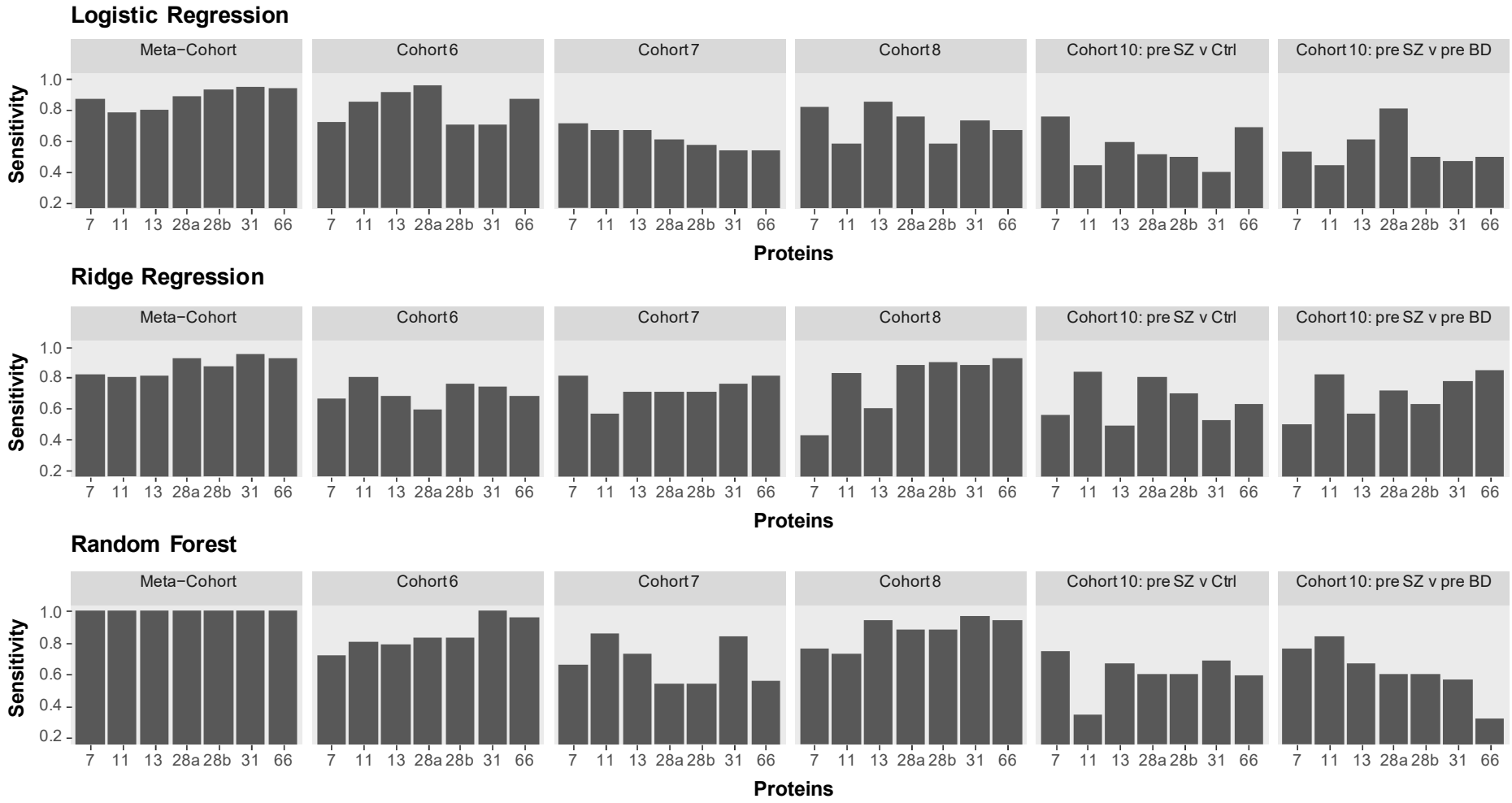


Figure A.3.1. Classification performances (sensitivity) for Logistic Regression, Ridge Regression and Random Forest across the training data (Meta-Cohort), and the independent validation (Cohort 6), application (Cohorts 7 & 8) and prediction (Cohort 10) cohorts. Performance was measured across models fitted on all unique protein biomarker sets identified in Table 3.8, the 7 biomarkers identified by all six models and the full set of 66 proteins for comparison. 28a and 28b refers to the set of 28 biomarkers identified by LASSO (28a) and the 28 biomarkers identified by RF-RFE (28b).

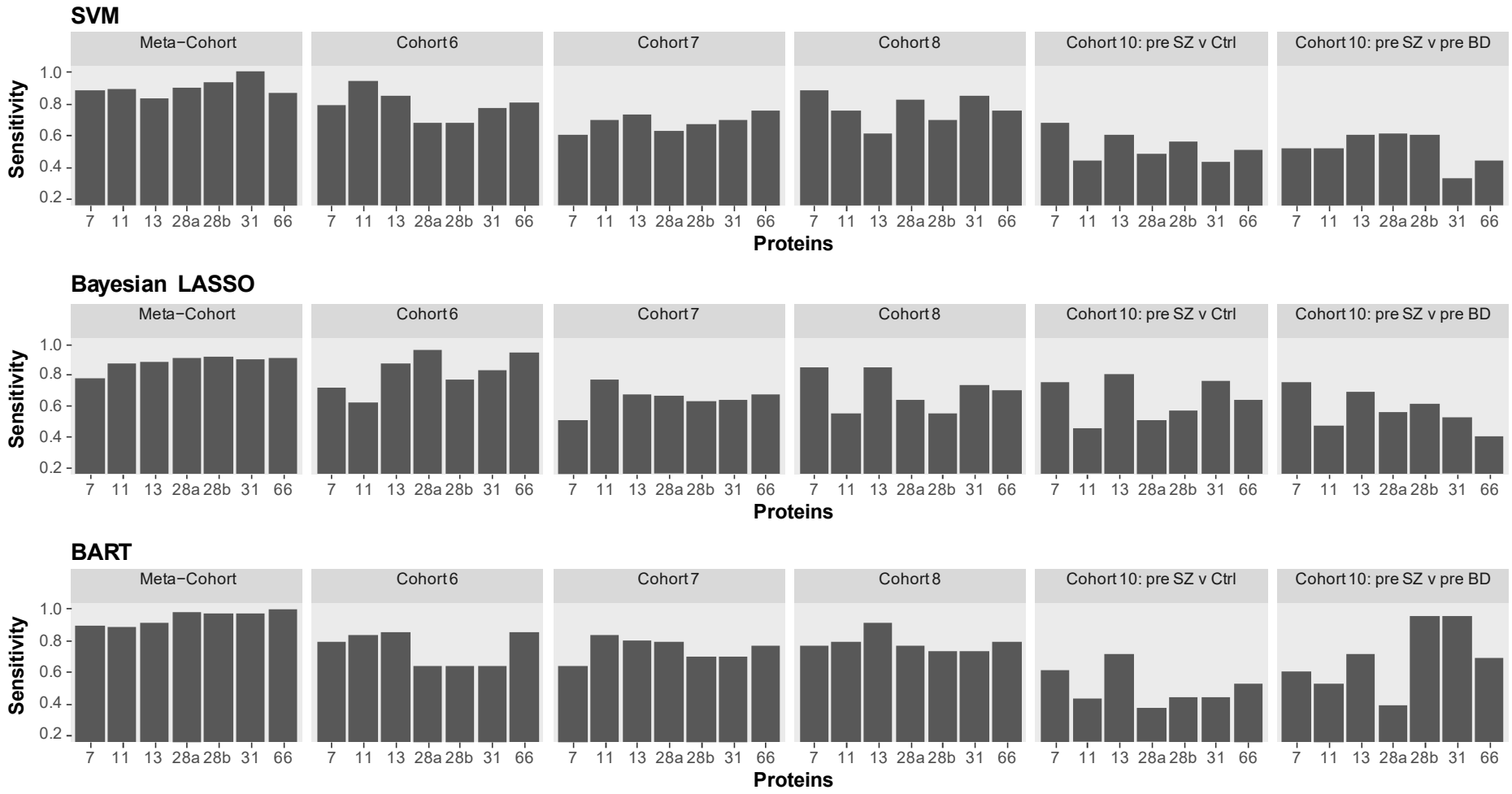


Figure A.3.2 Classification performances (sensitivity) for SVM, Bayesian LASSO, and BART across the training data (Meta-Cohort), and the independent validation (Cohort 6), application (Cohorts 7 & 8) and prediction (Cohort 10) cohorts. Performance was measured across models fitted on all unique protein biomarker sets identified in Table 3.8, the 7 biomarkers identified by all six models and the full set of 66 proteins for comparison. 28a and 28b refers to the set of 28 biomarkers identified by LASSO (28a) and the 28 biomarkers identified by RF-RFE (28b).

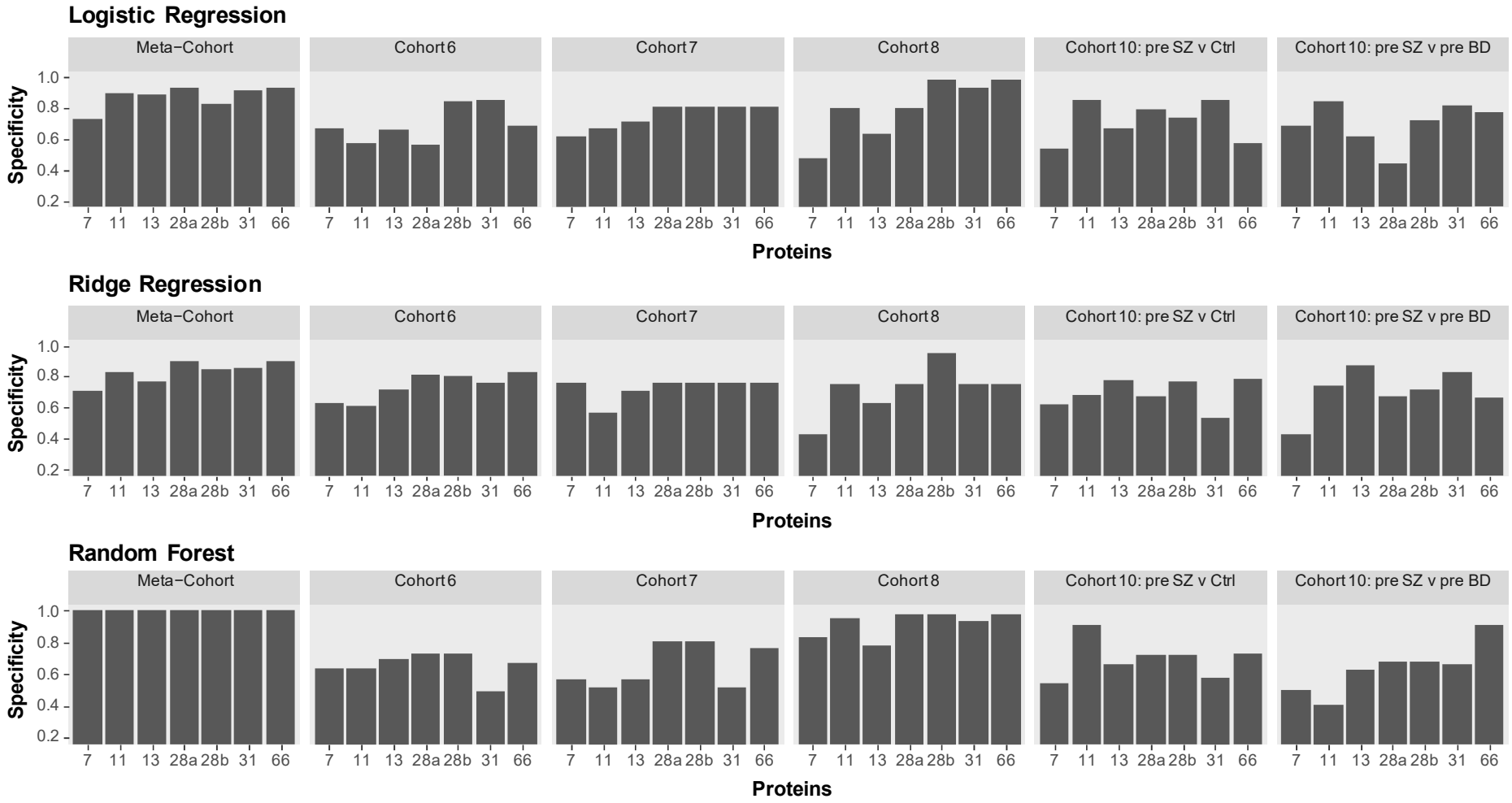


Figure A.3.3. Classification performances (specificity) for Logistic Regression, Ridge Regression and Random Forest across the training data (Meta-Cohort), and the independent validation (Cohort 6), application (Cohorts 7 & 8) and prediction (Cohort 10) cohorts. Performance was measured across models fitted on all unique protein biomarker sets identified in Table 3.8, the 7 biomarkers identified by all six models and the full set of 66 proteins for comparison. 28a and 28b refers to the set of 28 biomarkers identified by LASSO (28a) and the 28 biomarkers identified by RF-RFE (28b).

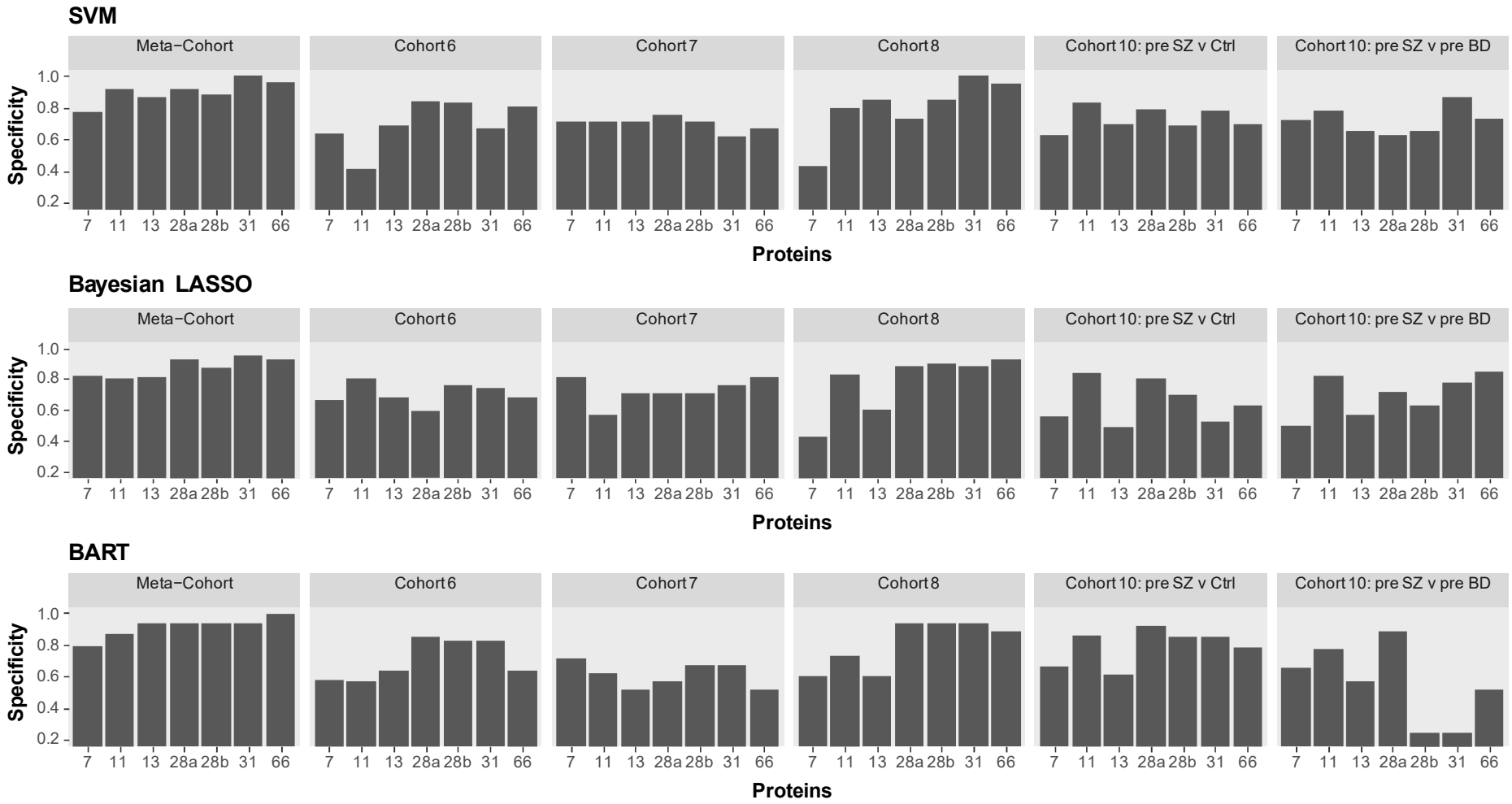


Figure A.3.4 Classification performances (specificity) for SVM, Bayesian LASSO, and BART across the training data (Meta-Cohort), and the independent validation (Cohort 6), application (Cohorts 7 & 8) and prediction (Cohort 10) cohorts. Performance was measured across models fitted on all unique protein biomarker sets identified in Table 3.8, the 7 biomarkers identified by all six models and the full set of 66 proteins for comparison. 28a and 28b refers to the set of 28 biomarkers identified by LASSO (28a) and the 28 biomarkers identified by RF-RFE (28b).

Chapter 4 - Appendices

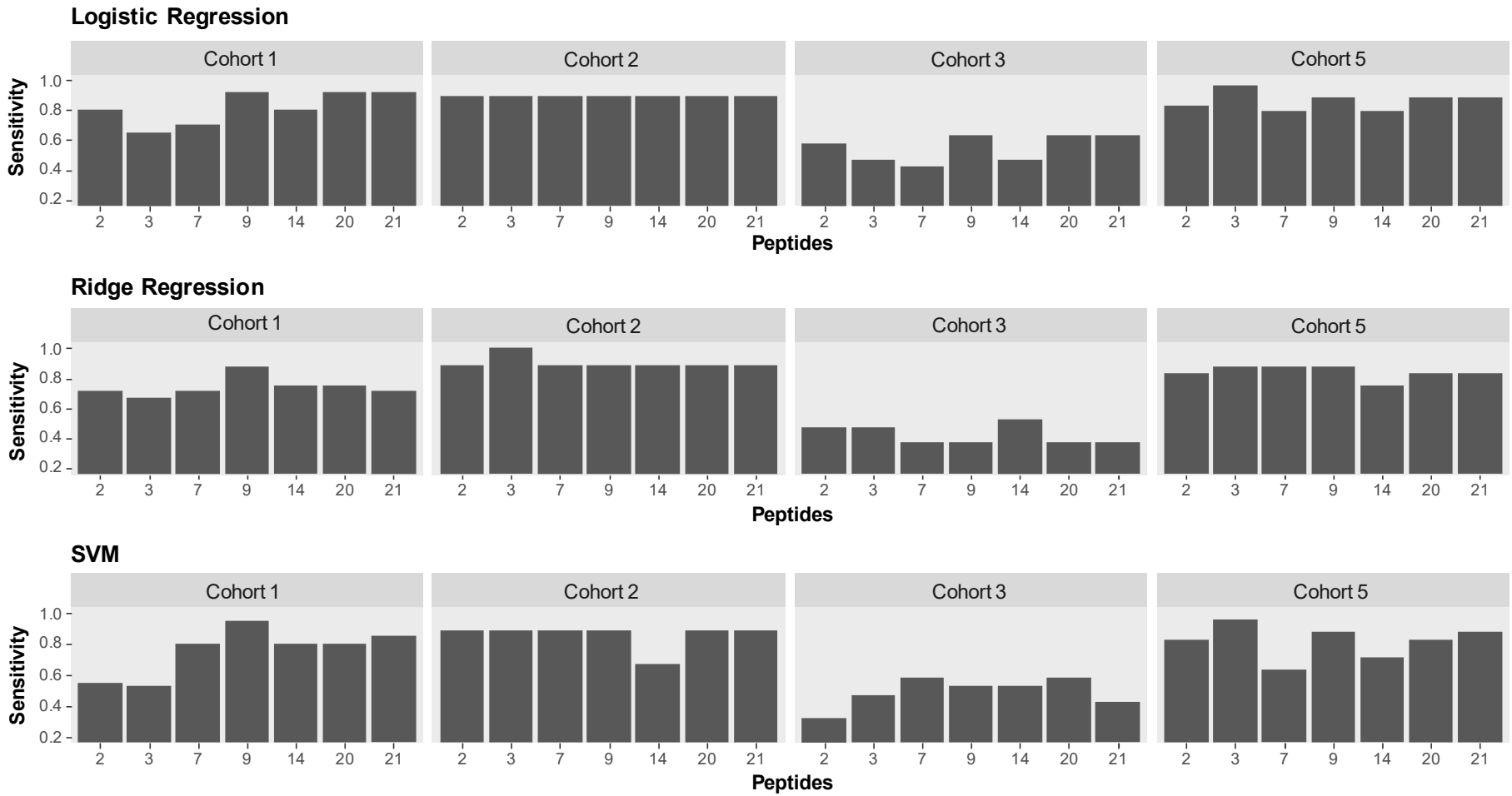


Figure A.4.1. Classification performance (sensitivity) on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts for Logistic Regression, Ridge Regression and SVM. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparison

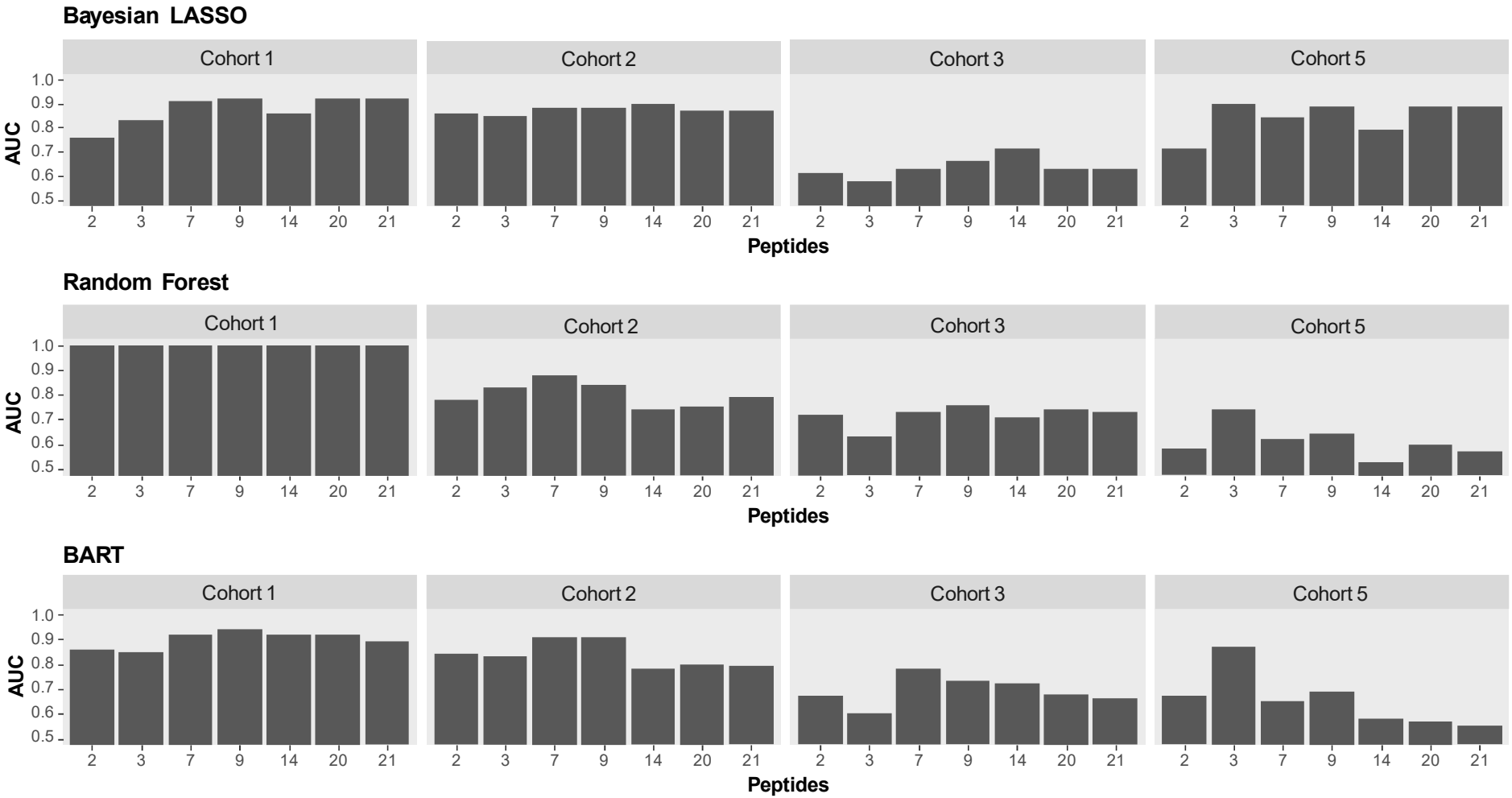


Figure A.4.2 Classification performance (sensitivity) on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts for Bayesian LASSO, Random Forest and BART. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparison

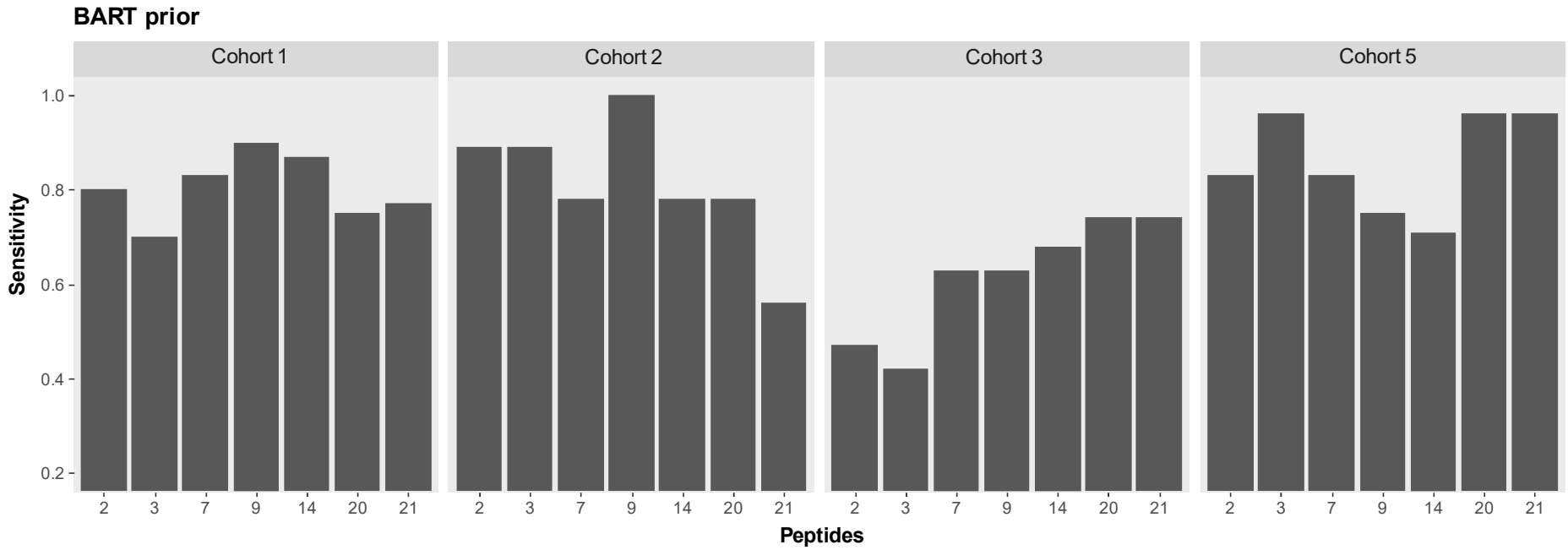


Figure A.4.3 Classification performance (sensitivity) on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts for BART prior. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparison

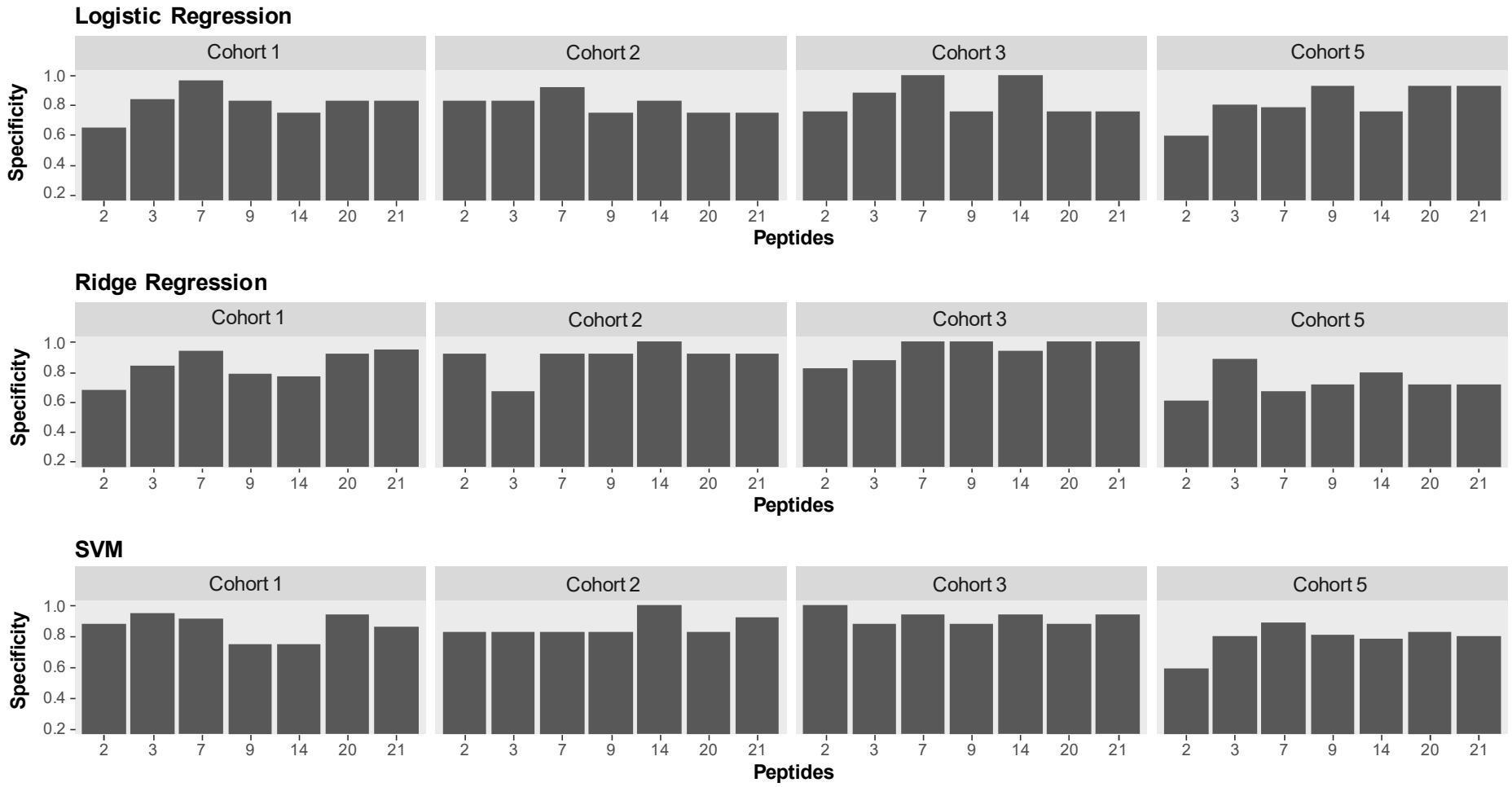


Figure A.4.4. Classification performance (specificity) on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts for Logistic Regression, Ridge Regression and SVM. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparison

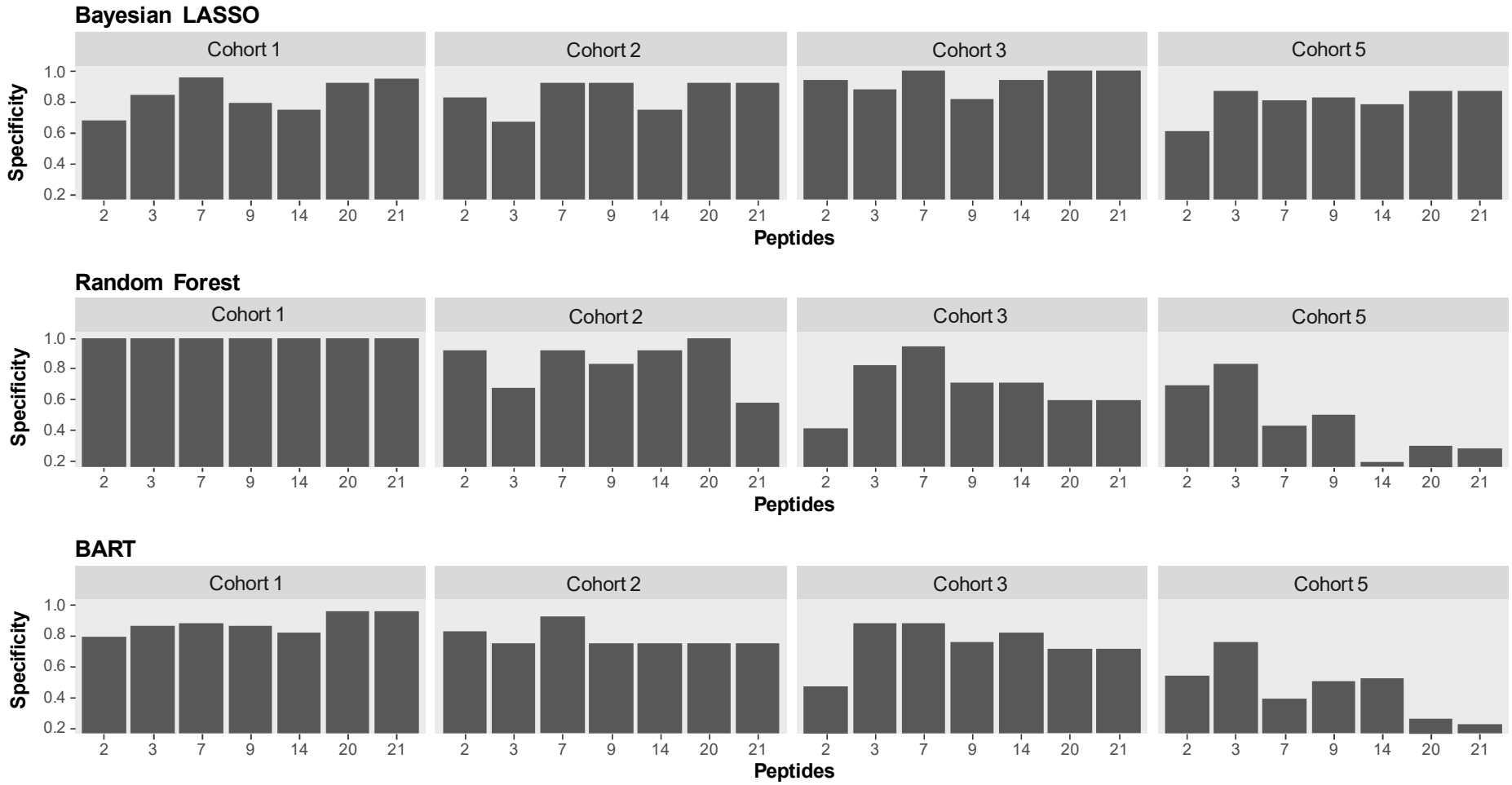


Figure A.4.5 Classification performance (specificity) on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts for Bayesian LASSO, Random Forest and BART. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparison

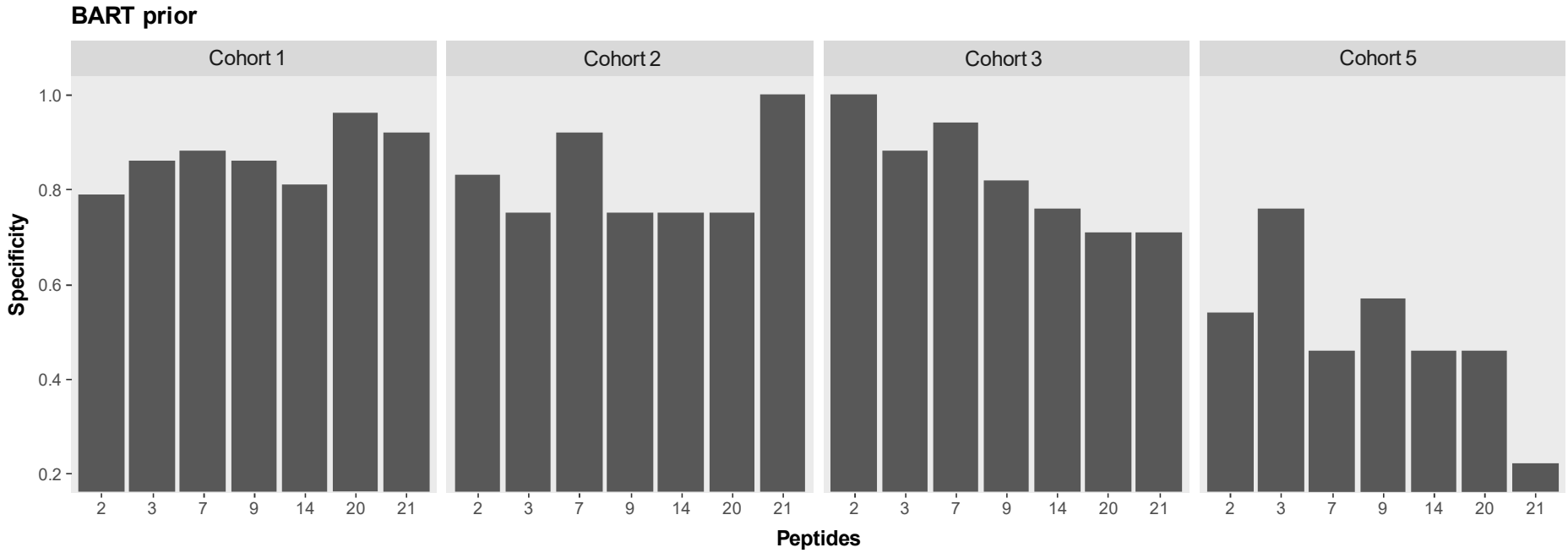


Figure A.4.6 Classification performance (specificity) on the training data (cohort 1), and the independent validation (cohort 2), application (cohort 3) and prediction (cohort 5) cohorts for BART prior. Performance was measured across models fitted on all unique peptide biomarker sets identified in Table 4.6 and the full set of 21 peptides for comparison

Chapter 5 - Appendices

Table A.5.1 The panels of 17 immunomodulatory proteins, and 147 peptides measured in the serum of 79 controls and 60 schizophrenia patients from Cologne using multiplex immunoassay and multiplex reaction monitoring (MRM) mass spectrometry, in Chapter 5.

Protein	Uniprot Entry	Uniprot ID	Peptide sequence	Assay
Complement C1q	NA	NA	NA	MI
Complement C3	CO3_HUMAN	P01024	NA	MI
Complement C3b	NA	NA	NA	MI
Complement C4	NA	NA	NA	MI
Granulocyte macrophage colony-stimulating factor	CSF2_HUMAN	P04141	NA	MI
Interferon-γ	IFNG_HUMAN	P01579	NA	MI
Interleukin-10	IL10_HUMAN	P22301	NA	MI
Interleukin-12(p70)	NA	NA	NA	MI
Interleukin-13	IL13_HUMAN	P35225	NA	MI
Interleukin-1β	IL1B_HUMAN	P01584	NA	MI
Interleukin-2	IL2_HUMAN	P60568	NA	MI
Interleukin-4	IL4_HUMAN	P05112	NA	MI
Interleukin-6	IL6_HUMAN	P05231	NA	MI
Transforming growth factor-β1	TGFB1_HUMAN	P01137	NA	MI
Transforming growth factor-β2	TGFB2_HUMAN	P61812	NA	MI
Transforming growth factor-β3	TGFB3_HUMAN	P10600	NA	MI
Tumor necrosis factor-α	TNFA_HUMAN	P01375	NA	MI
Angiotensinogen	ANGT_HUMAN	P01019	SLDFTELDVAAEK	MRM
Angiotensinogen	ANGT_HUMAN	P01019	FMQAVTGWK	MRM
Angiotensinogen	ANGT_HUMAN	P01019	ALQDQLVLVAAK	MRM
Antithrombin-III	ANT3_HUMAN	P01008	FDTISEK	MRM
Antithrombin-III	ANT3_HUMAN	P01008	LPGIVAEGR	MRM
Apolipoprotein A-I	APOA1_HUMAN	P02647	EQLGPVTQEFWDNLEK	MRM
Apolipoprotein A-I	APOA1_HUMAN	P02647	ATEHLSTLSEK	MRM
Apolipoprotein A-II	APOA2_HUMAN	P02652	SPELQAEAK	MRM
Apolipoprotein A-IV	APOA4_HUMAN	P06727	IDQNVEELK	MRM
Apolipoprotein A-IV	APOA4_HUMAN	P06727	ISASAEELR	MRM
Apolipoprotein A-IV	APOA4_HUMAN	P06727	ALVQQMEQLR	MRM
Apolipoprotein C-I	APOC1_HUMAN	P02654	EFGNTLEDK	MRM
Apolipoprotein C-I	APOC1_HUMAN	P02654	EWFSETFQK	MRM
Apolipoprotein C-II	APOC2_HUMAN	P02655	ESLSSYWESAK	MRM
Apolipoprotein C-II	APOC2_HUMAN	P02655	TAAQNLYEK	MRM
Apolipoprotein C-III	APOC3_HUMAN	P02656	GWVTDGFSSLK	MRM
Apolipoprotein C-III	APOC3_HUMAN	P02656	DALSSVQESQVAQQAR	MRM
Apolipoprotein C-IV	APOC4_HUMAN	P55056	AWFLESK	MRM
Apolipoprotein D	APOD_HUMAN	P05090	VLNQELR	MRM
Apolipoprotein E	APOE_HUMAN	P02649	LEEQAQQIR	MRM
Apolipoprotein E	APOE_HUMAN	P02649	ALMDETMK	MRM

Apolipoprotein E	APOE_HUMAN	P02649	AATVGSLAGQPLQER	MRM
Apolipoprotein E	APOE_HUMAN	P02649	LGPLVEQGR	MRM
Apolipoprotein E	APOE_HUMAN	P02649	SELEEQLTPVAEETR	MRM
Apolipoprotein F	APOF_HUMAN	Q13790	SLPTEDCENEK	MRM
Apolipoprotein L1	APOL1_HUMAN	O14791	VNEPSILEMSR	MRM
Apolipoprotein L1	APOL1_HUMAN	O14791	LNILNNNYK	MRM
Apolipoprotein L1	APOL1_HUMAN	O14791	VTEPISAESGEQVER	MRM
Apolipoprotein M	APOM_HUMAN	O95445	SLTSCLDSK	MRM
Apolipoprotein M	APOM_HUMAN	O95445	AFLLTPR	MRM
C4b-binding protein α chain	C4BPA_HUMAN	P04003	EDVYVVGTVLR	MRM
C4b-binding protein α chain	C4BPA_HUMAN	P04003	YTCLPGYVR	MRM
C4b-binding protein α chain	C4BPA_HUMAN	P04003	FSAICQGDGTWSPR	MRM
Carbonic anhydrase 1	CAH1_HUMAN	P00915	ADGLAVIGVLMK	MRM
Carboxypeptidase B2	CBPB2_HUMAN	Q96IY4	YPLYVLK	MRM
Carboxypeptidase B2	CBPB2_HUMAN	Q96IY4	DTGTYGFLLPER	MRM
CD5 antigen-like	CD5L_HUMAN	O43866	EATLQDCPSGPWGK	MRM
Ceruloplasmin	CERU_HUMAN	P00450	NNEGTYYSPNYNPQSR	MRM
Ceruloplasmin	CERU_HUMAN	P00450	EVGPTNADPVCLAK	MRM
Clusterin	CLUS_HUMAN	P10909	FMETVAEK	MRM
Clusterin	CLUS_HUMAN	P10909	IDSLLENDR	MRM
Coagulation factor XII	FA12_HUMAN	P00748	CFEPQLLR	MRM
Coagulation factor XII	FA12_HUMAN	P00748	VVGGLVALR	MRM
Complement C1q subcomponent subunit C	C1QC_HUMAN	P02747	TNQVNSGGVLLR	MRM
Complement C1r	C1R_HUMAN	P00736	YTTEIIK	MRM
Complement C1r subcomponent-like protein	C1RL_HUMAN	Q9NZP8	GSEAINAPGDNPAK	MRM
Complement C1s subcomponent	C1S_HUMAN	P09871	TNFDNDIALVR	MRM
Complement C1s subcomponent	C1S_HUMAN	P09871	LLEVPEGR	MRM
Complement C2	CO2_HUMAN	P06681	HAIILLTDGK	MRM
Complement C3	CO3_HUMAN	P01024	VYAYYNLEESCTR	MRM
Complement C3	CO3_HUMAN	P01024	AGDFLEANYMNLQR	MRM
Complement C4-A	CO4A_HUMAN	P0C0L4	VLSLAQEQVGGSPEK	MRM
Complement C4-A	CO4A_HUMAN	P0C0L4	ITQVLHFTK	MRM
Complement C4-A	CO4A_HUMAN	P0C0L4	DFALLSLQVPLK	MRM
Complement component C6	CO6_HUMAN	P13671	TLNICEVGTIR	MRM
Complement component C6	CO6_HUMAN	P13671	SEYGAALAWEK	MRM
Complement component C8 α chain	CO8A_HUMAN	P07357	MESLGITSR	MRM
Complement component C9	CO9_HUMAN	P02748	VVEESELAR	MRM
Complement component C9	CO9_HUMAN	P02748	LSPIYNLVPVK	MRM
Complement factor B	CFAB_HUMAN	P00751	EELLPAQDIK	MRM
Complement factor B	CFAB_HUMAN	P00751	DISEVVTPR	MRM
Complement factor B	CFAB_HUMAN	P00751	YGLVTYATYPK	MRM
Complement factor B	CFAB_HUMAN	P00751	DLLYIGK	MRM
Complement factor H	CFAH_HUMAN	P00751	CFEGFGIDGPAIAK	MRM
Corticosteroid-binding globulin	CBG_HUMAN	P08185	ITQDAQLK	MRM
Corticosteroid-binding globulin	CBG_HUMAN	P08185	GTWTQPFDLASTR	MRM
Fibronectin	FINC_HUMAN	P02751	YSFCTDHTVLVQTR	MRM

Ficolin-3	FCN3_HUMAN	O75636	YGIDWASGR	MRM
Gelsolin	GELS_HUMAN	P06396	AGALNSNDAFVLK	MRM
Gelsolin	GELS_HUMAN	P06396	SEDCFILDHGK	MRM
Haptoglobin	HPT_HUMAN	P00738	DYAEVGR	MRM
Haptoglobin	HPT_HUMAN	P00738	VTSIQDWVQK	MRM
Haptoglobin	HPT_HUMAN	P00738	VGYVSGWGR	MRM
Hemoglobulin subunit α	HBA_HUMAN	P69905	MFLSFPTTK	MRM
Hemoglobulin subunit α	HBA_HUMAN	P69905	FLASVSTVLTSK	MRM
Hemoglobulin subunit γ-1	HBG1_HUMAN	P69891	MVTAVASALSSR	MRM
Hemopexin	HEMO_HUMAN	P02790	VDGALCMEK	MRM
Hemopexin	HEMO_HUMAN	P02790	NFPSPVDAAFR	MRM
Heparin cofactor 2	HEP2_HUMAN	P05546	IAIDLFK	MRM
Heparin cofactor 2	HEP2_HUMAN	P05546	FAFNLYR	MRM
Histidine-rich glycoprotein	HRG_HUMAN	P04196	ADLFYDVEALDLESPK	MRM
Histidine-rich glycoprotein	HRG_HUMAN	P04196	DSPVLIDFFEDTER	MRM
Igα-1 chain C region	IGHA1_HUMAN	P01876	DASGVTFTWTPSSGK	MRM
Igα-1 chain C region	IGHA1_HUMAN	P01876	TPLTATLSK	MRM
Igα-2 chain C region	IGHA2_HUMAN	P01877	DASGATFTWTPSSGK	MRM
Igγ-1 chain C region	IGHG1_HUMAN	P01857	FNWYVDGVEVHNAK	MRM
Igγ-2 chain C region	IGHG2_HUMAN	P01859	GLPAPIEK	MRM
Igγ-2 chain C region	IGHG2_HUMAN	P01859	TTPPMLDSDGSFFLYSK	MRM
Igγ-3 chain C region	IGHG3_HUMAN	P01860	DTLMISR	MRM
Igγ-3 chain C region	IGHG3_HUMAN	P01860	NQVSLTCLVK	MRM
Igμ chain C region	IGHM_HUMAN	P01871	YAATSQVLLPSK	MRM
Igμ chain C region	IGHM_HUMAN	P01871	QIQVSWLR	MRM
Inter-α-trypsin inhibitor heavy chain H1	ITIH1_HUMAN	P19827	LDAQASFLPK	MRM
Inter-α-trypsin inhibitor heavy chain H1	ITIH1_HUMAN	P19827	GSLVQASEANLQAAQDFVR	MRM
Inter-α-trypsin inhibitor heavy chain H2	ITIH2_HUMAN	P19823	FYNQVSTPLLR	MRM
Inter-α-trypsin inhibitor heavy chain H2	ITIH2_HUMAN	P19823	IQPSGGTNINEALLR	MRM
Inter-α-trypsin inhibitor heavy chain H4	ITIH4_HUMAN	Q14624	GPDVLTATVSGK	MRM
Inter-α-trypsin inhibitor heavy chain H4	ITIH4_HUMAN	Q14624	ETLFSVMPGLK	MRM
Kininogen-1	KNG1_HUMAN	P01042	DFVQPPTK	MRM
Kininogen-1	KNG1_HUMAN	P01042	DIPTNSPELEETLTHTITK	MRM
Lumican	LUM_HUMAN	P51884	SLEDLQLTHNK	MRM
N-acetylmuramoyl-L-alanine amidase	PGRP2_HUMAN	Q96PD5	GCPDVQASLPDAK	MRM
N-acetylmuramoyl-L-alanine amidase	PGRP2_HUMAN	Q96PD5	TFTLLDPK	MRM
Phosphatidylinositol-glycan-specific phospholipase D	PHLD_HUMAN	P80108	NQVVIAAGR	MRM
Pigment epithelium-derived factor	PEDF_HUMAN	P36955	LQSLFDSPDFSK	MRM
Pigment epithelium-derived factor	PEDF_HUMAN	P36955	TVQAVLTVPK	MRM
Pigment epithelium-derived factor	PEDF_HUMAN	P36955	ELLDTVTAPQK	MRM
Pigment epithelium-derived factor	PEDF_HUMAN	P36955	DTDTGALLFIGK	MRM
Plasma kallikrein	KLKB1_HUMAN	P03952	LSMDGSPTR	MRM
Plasma protease C1 inhibitor	IC1_HUMAN	P05155	TNLESILSYPK	MRM
Plasma protease C1 inhibitor	IC1_HUMAN	P05155	FQPTLLTLPR	MRM
Plasminogen	PLMN_HUMAN	P00747	FVTWIEGVMR	MRM
Protein AMBP	AMBP_HUMAN	P02760	TVAACNLPIVR	MRM

Protein AMBP	AMBP_HUMAN	P02760	ETLLQDFR	MRM
Prothrombin	THRB_HUMAN	P00734	SGIECQLWR	MRM
Prothrombin	THRB_HUMAN	P00734	ELLESYIDGR	MRM
Retinol-binding protein 4	RET4_HUMAN	P02753	YWGVASFLQK	MRM
Retinol-binding protein 4	RET4_HUMAN	P02753	QEELCLAR	MRM
Serotransferrin	TRFE_HUMAN	P02787	EGYYGYTGAFR	MRM
Serum albumin	ALBU_HUMAN	P02768	AAFTECCQAADK	MRM
Serum albumin	ALBU_HUMAN	P02768	ETYGEMADCCAK	MRM
Serum albumin	ALBU_HUMAN	P02768	QNCELFEQLGEYK	MRM
Serum amyloid P-component	SAMP_HUMAN	P02743	IVLGQEQDSYGGK	MRM
Sex Hormone-binding globulin	SHBG_HUMAN	P04278	IALGGLLFPASNLR	MRM
Tetranectin	TETN_HUMAN	P05452	EQQALQTVCLK	MRM
Transthyretin	TTHY_HUMAN	P02766	AADDTWEPFASGK	MRM
Transthyretin	TTHY_HUMAN	P02766	VLDAVR	MRM
Vitronectin	VTNC_HUMAN	P04004	DVWGIEGPIDAAFTR	MRM
Vitronectin	VTNC_HUMAN	P04004	DWHGVPGQVDAAMAGR	MRM
α-1-antichymotrypsin	AACT_HUMAN	P01011	EQLSLLDR	MRM
α-1-antichymotrypsin	AACT_HUMAN	P01011	EIGELYLPK	MRM
α-1-antichymotrypsin	AACT_HUMAN	P01011	ADLSGITGAR	MRM
α-1-antitrypsin	A1AT_HUMAN	P01009	LSITGTYDLK	MRM
α-1-antitrypsin	A1AT_HUMAN	P01009	SVLGQLGITK	MRM
α-1-antitrypsin	A1AT_HUMAN	P01009	SPLFMGK	MRM
α-1B-glycoprotein	A1BG_HUMAN	P04217	CLAPLEGAR	MRM
α-1B-glycoprotein	A1BG_HUMAN	P04217	ATWSGAVLAGR	MRM
α-1B-glycoprotein	A1BG_HUMAN	P04217	SGLSTGWTQLSK	MRM
α-2-antiplasmin	A2AP_HUMAN	P08697	FDPSLTQR	MRM
α-2-antiplasmin	A2AP_HUMAN	P08697	DFLQSLK	MRM
α-2-antiplasmin	A2AP_HUMAN	P08697	DSFHLDEQFTVPVEMMQAR	MRM
α-2-HS-glycoprotein	FETUA_HUMAN	P02765	HTLNQIDEVK	MRM
α-2-HS-glycoprotein	FETUA_HUMAN	P02765	FSVVYAK	MRM
α-2-macroglobulin	A2MG_HUMAN	P01023	NEDSLVFVQTDK	MRM
α-2-macroglobulin	A2MG_HUMAN	P01023	AIGYLNTGYQR	MRM
β-2-glycoprotein 1	APOH_HUMAN	P02749	EHSSLAFWK	MRM
β-2-glycoprotein 1	APOH_HUMAN	P02749	VSFFCK	MRM

Table A.5.2 Alterations in serum analytes between schizophrenia patients and controls as described in Section 5.3.2.3.

Protein	Peptide sequence	Assay	Covariates	Permuted P	Q Value	Fold change
Interferon-γ (IFN-γ)	NA	MI	-	0.018	0.159	-1.07
Transforming growth factor-β1 (TGF-β1)	NA	MI	-	0.045	0.091	-1.05
Antithrombin-III	FDTISEK	MRM	gender + age	0.016	0.208	1.26
Antithrombin-III	LPGIVAEGR	MRM	gender + age	0.015	0.208	1.25
Apolipoprotein A-II	SPELQAEAK	MRM	-	0.024	0.244	-1.13
Apolipoprotein A-IV	IDQNVEELK	MRM	gender	0.008	0.208	-1.26
Apolipoprotein A-IV	ISASAEELR	MRM	-	0.041	0.308	-1.21
Apolipoprotein C-I	EFGNTLEDK	MRM	age	0.033	0.299	-1.26
Apolipoprotein C-III	GWVTDGFSSLK	MRM	gender	0.004	0.123	-1.25
Apolipoprotein C-III	DALSSVQESQVAQQAR	MRM	-	0.011	0.208	-1.21
Apolipoprotein H	EHSSLAFWK	MRM	-	0.047	0.308	1.26
Complement C4-A	VLSLAQEQVGGSPEK	MRM	age	0.016	0.208	1.24
Complement C4-A	ITQVLHFTK	MRM	age	0.044	0.299	1.22
Complement component C9	VVEESELAR	MRM	-	0.027	0.284	1.23
Ficolin-3	YGIDWASGR	MRM	-	0.038	0.299	1.23
Haptoglobin	DYAEVGR	MRM	age	0.000	0.044	1.56
Haptoglobin	VTSIQDWVQK	MRM	-	0.001	0.044	1.54
Haptoglobin	VGYVSGWGR	MRM	age	0.002	0.044	1.53
Inter-α-trypsin inhibitor heavy chain H4	GPDVLTATVSGK	MRM	gender	0.024	0.254	1.18
Plasma protease C1 inhibitor	TNLESILSYPK	MRM	gender + age	0.001	0.051	1.36
Plasma protease C1 inhibitor	FQPTLLTLPR	MRM	gender + age	0.013	0.208	1.34
α-1-antichymotrypsin	EQLSLLDR	MRM	gender + age	0.018	0.244	1.38
α-2-antiplasmin	FDPSLTQR	MRM	gender + age	0.043	0.308	1.23

Chapter 7 - Appendices

Table A.7.1 Proteins significantly (q<0.05) altered in MDD post-mortem brains compared to healthy controls.

Protein Name	Fold Change	Q Value
CATD	1.010	0.001
CH60	1.007	0.001
ANR16	1.031	0.001
AT1A4	1.009	0.002
PRDX3	1.013	0.002
GBB1	1.009	0.002
FBX2	1.018	0.002
FUT4	1.063	0.002
KCC2A	1.008	0.002
HPRT	1.013	0.002
NEGR1	1.034	0.002
PP1A	1.026	0.002
MAP2	1.009	0.003
SODC	1.008	0.003
DHE3	1.006	0.003
HS905	1.018	0.003
DHPR	1.005	0.003
CO1A1	0.996	0.003
TTC25	1.031	0.003
WDR20	1.047	0.004
AT2B1	1.014	0.004
DPYL2	1.004	0.004
VATF	1.050	0.004
TBA1B	1.004	0.005
PGDH	1.061	0.005
MYL6	1.040	0.005
ALBU	1.005	0.005
HXK1	1.007	0.005
ENOG	1.004	0.006
E41L3	1.009	0.006
HS90B	1.007	0.007
ACTB	1.010	0.007
VDAC2	1.017	0.007
GLNA	1.013	0.007
ARC1A	1.010	0.007
TPPP	1.006	0.007
THIO	1.013	0.009
GNAO	1.007	0.010
HSP7C	1.006	0.010

GBB4	1.010	0.011
LSAMP	1.015	0.011
SYN2	1.007	0.011
ODPB	1.009	0.011
NDRG4	1.024	0.011
CR1L	0.984	0.012
DPOLN	1.031	0.012
CADM2	1.011	0.013
OTUB1	1.004	0.013
CTDP1	1.025	0.013
TRI13	1.014	0.014
APOE	1.040	0.015
PGK1	1.006	0.016
TBA8	1.006	0.016
IGHG1	1.015	0.016
MTMR6	1.018	0.016
PP2AB	1.006	0.016
MDHM	1.005	0.017
HXC9	1.140	0.017
YI007	0.984	0.020
CDC42	1.011	0.020
DDAH1	1.003	0.020
VPP1	1.010	0.020
GT251	0.964	0.020
LIPE	1.018	0.020
SNP25	1.005	0.020
CA138	0.948	0.020
DPPA3	1.026	0.021
SNAA	1.016	0.021
AP2A2	1.009	0.022
1433S	1.012	0.022
TRI42	1.026	0.022
MIDN	1.039	0.022
BCAS1	1.010	0.023
TCPA	1.015	0.023
TBB8	1.009	0.024
ALDOA	1.006	0.024
POTEE	1.007	0.024
K1C10	0.987	0.024
CNTN1	1.006	0.024
BIN1	1.004	0.024
HCD2	1.028	0.024
APT	1.027	0.024
ATIF1	1.067	0.024
SCND1	1.081	0.024
POTEF	1.012	0.024

MK01	1.010	0.024
ANKS3	1.013	0.025
NRIP1	0.984	0.026
MCMBP	0.999	0.028
KCRB	1.005	0.030
CALL3	1.011	0.030
1433G	1.008	0.031
QCR1	1.008	0.031
KCC2G	1.007	0.032
STXB1	1.004	0.036
BASI	0.969	0.038
MPIP3	1.018	0.039
MDHC	1.006	0.039
HNRPQ	1.040	0.044
TY3H	1.016	0.045
RAC1	1.010	0.046
DYN2	1.005	0.046
ATPK	1.015	0.046
EF1A2	1.008	0.046
RASK	1.012	0.047
BAG3	0.940	0.047
K2C4	0.996	0.047
PP1G	1.018	0.047
HAUS8	1.103	0.047

Table A.7.2 Proteins significantly (p<0.05) altered in social defeat model brain tissue compared to control animals.

Protein Name	Fold Change	P Value
MTERF	0.960	0.002
NASP	0.983	0.002
ADT4	0.987	0.003
TRAP1	0.970	0.004
NOP58	0.978	0.006
GNAI1	0.991	0.006
CEP72	0.984	0.007
CX7A2	0.962	0.008
PCNP	1.035	0.010
SFXN1	0.946	0.011
RL8	0.956	0.012
TEKT1	1.015	0.012
EFC4A	1.040	0.012
FIGL1	0.979	0.013
GNAT3	0.988	0.013

TBB3	0.991	0.015
CALM	0.990	0.016
MAP6	0.984	0.016
H2B1A	0.961	0.017
ZC12A	0.977	0.018
CX6B1	0.993	0.018
STX1B	0.989	0.018
EIF3I	0.954	0.019
HS90A	0.991	0.019
HS90B	0.988	0.019
VAMP3	0.978	0.022
TM11F	0.966	0.023
MTPN	0.985	0.023
GPM6A	0.991	0.024
PP1G	1.006	0.024
NCAM1	0.991	0.025
ACTN3	0.985	0.025
RAB13	0.947	0.026
RAB14	0.983	0.026
KCC2B	0.983	0.027
MARCS	0.973	0.027
PACN1	0.994	0.028
1433E	0.991	0.029
LMNB1	0.928	0.030
NFM	0.990	0.031
TPPP	0.984	0.031
GNAS1	0.986	0.032
GATA4	0.973	0.033
HBB1	0.985	0.034
CENPQ	0.946	0.034
VP9D1	0.975	0.035
ARF5	0.978	0.036
PPIA	0.994	0.037
GRP78	0.990	0.037
MUS81	0.988	0.038
AT12A	0.989	0.039
BIN1	0.977	0.039
GNAS2	0.979	0.040
LRC8E	1.051	0.041
MYO1A	0.942	0.043
NIBL1	1.039	0.043
PPCT	1.057	0.044
H33	0.965	0.045
VDAC1	0.984	0.046
K1C20	0.957	0.048
EFTU	0.985	0.049

SYDC	0.955	0.049
HS12A	0.983	0.049
CENPO	0.925	0.049
FIL1L	0.977	0.049
RAP1B	0.976	0.050
H2B2E	0.980	0.050
RAB5C	0.979	0.050

Table A.7.3 Proteins significantly (p<0.05) altered in chronic mild stress model brain tissue compared to control animals.

Protein Name	Fold Change	P Value
IP3KA	1.054	0.001
HEM1	1.094	0.001
MCM7	0.974	0.006
ADPRH	1.036	0.006
RAB3A	1.005	0.007
TBA3	0.936	0.010
AIF1L	1.010	0.011
AT2A2	1.017	0.012
COX41	0.995	0.013
DHSB	1.024	0.013
OFUT1	1.029	0.013
PEBP1	1.004	0.015
PPM1E	0.987	0.017
TRAP1	0.990	0.018
AMPD1	0.991	0.021
KPYM	0.996	0.022
NP1L2	1.044	0.025
ACTN1	1.010	0.028
GFAP	0.996	0.028
TPIS	0.996	0.029
CAZA2	1.010	0.030
CEP57	0.978	0.032
M3K8	0.975	0.033
CALR	0.993	0.033
NCAM1	0.995	0.033
FKB1A	1.006	0.034
NFM	0.996	0.038
DREB	0.993	0.039
AT5F1	0.996	0.039
ACSL6	0.979	0.039
OCSTP	1.054	0.043
LDHB	1.007	0.043
PROF1	1.003	0.044

WWC2	0.992	0.046
ODP2	0.994	0.046
K6PF	0.993	0.046
IL18	0.988	0.047
S10A8	0.982	0.048
GNAT2	1.077	0.048
PCCA	0.908	0.048
SYUG	1.011	0.048
RAB1A	1.008	0.049
GALT	0.983	0.050

Table A.7.4 Proteins significantly (p<0.05) altered in prenatal stress model brain tissue compared to control animals.

Protein Name	Fold Change	P Value
PP2BB	1.034	0.003
HBB1	0.981	0.004
CC114	0.938	0.008
CNTRB	1.030	0.010
GNAL	1.030	0.011
K2C73	1.090	0.011
FRMD6	0.952	0.011
CAP2	1.018	0.013
FA92B	1.094	0.013
HKDC1	1.022	0.014
CIP4	1.084	0.015
AT2A1	1.057	0.016
CPLX2	1.019	0.017
H33	0.979	0.017
RAB14	1.013	0.019
ARC1A	1.082	0.025
RM37	1.050	0.026
GDIR1	1.021	0.027
NFASC	1.018	0.031
TTC23	1.085	0.032
MATR3	0.876	0.037
CYB5B	1.025	0.040
FA11	0.968	0.040
GCM2	1.033	0.043
CNTN1	1.008	0.044
SH3G2	1.014	0.044
SMAG1	1.065	0.045
NDRG2	1.027	0.046
M3K8	1.039	0.048
KIZ	1.057	0.048

Chapter 8 - Appendices

Table A.8.1 Proteins significantly (q<0.05) altered in schizophrenia post-mortem brains compared to healthy controls.

Protein Name	Fold Change	Q Value
HBB	1.003	0.000
AT1A2	1.005	0.000
HBA	1.000	0.000
AP2A2	1.004	0.000
HS90B	1.003	0.000
ALDOA	1.007	0.000
BIN1	1.007	0.000
GRP75	1.006	0.000
NOSIP	1.015	0.000
IGHG1	1.014	0.001
ATPG	1.015	0.001
SMC1B	1.012	0.001
ENOG	1.003	0.001
GRP78	1.006	0.001
DPYL3	1.005	0.001
TKT	1.005	0.001
ALBU	1.000	0.001
TBA1B	1.001	0.001
FBX2	1.012	0.002
TTC25	1.026	0.002
HS90A	1.002	0.002
VDAC2	1.007	0.002
AL7A1	1.012	0.003
NDRG2	1.014	0.003
IGKC	1.011	0.003
PP1A	1.014	0.003
MK01	1.007	0.003
GLNA	1.006	0.004
DPYL1	1.005	0.004
UCHL1	1.004	0.004
E41L3	1.003	0.004
PGAM1	1.002	0.004
ELAV2	1.049	0.004
MYPR	0.991	0.005
HAUS8	1.081	0.005
ATIF1	1.055	0.005
SPB5	1.021	0.006
CG033	1.019	0.006
DHYS	0.954	0.006

CNTN1	1.003	0.006
CATD	1.006	0.007
MT3	1.017	0.007
LSAMP	1.005	0.007
GBB2	1.006	0.007
FIBG	1.015	0.007
PRP19	1.048	0.008
UAP1L	0.974	0.008
SC65	1.038	0.008
KCC2D	1.007	0.008
KPCG	1.013	0.009
AT2B4	1.004	0.010
1433G	1.003	0.010
LMNB1	1.013	0.010
MIDN	1.034	0.010
RHOA	1.011	0.011
CLH1	1.001	0.011
AT1A4	1.004	0.011
RINI	1.008	0.012
HPT	1.037	0.012
PRDX3	1.008	0.012
ACTB	1.005	0.012
H90B2	1.005	0.012
HPLN2	0.982	0.013
A1AT	1.016	0.013
KAD1	1.007	0.014
CS052	0.980	0.014
PA1B2	1.010	0.014
TOP1M	0.994	0.014
ADT1	1.004	0.014
PPIA	1.007	0.014
ANKS3	1.032	0.015
HBD	1.007	0.015
TBB8	1.003	0.015
H90B3	1.008	0.015
RAP1B	1.010	0.016
RAC1	1.007	0.016
ACTC	0.994	0.016
DHE3	1.003	0.016
2AAA	1.004	0.017
PAI1	1.075	0.017
POTEE	1.002	0.018
EIF3A	1.015	0.018
H2B1C	1.030	0.019
SEP14	1.007	0.019
K2C5	1.012	0.019

ANR44	0.952	0.019
ACBP	0.992	0.019
AL4A1	1.014	0.019
FA81A	0.989	0.019
UBA1	1.004	0.020
MBP	0.998	0.020
UBP25	1.016	0.021
GFAP	0.999	0.021
GBB4	1.007	0.021
GT251	0.975	0.021
AQP4	1.021	0.021
RAB1C	1.010	0.023
DYN1	1.002	0.024
MPCP	1.005	0.024
SERA	1.008	0.024
NEGR1	1.020	0.024
TFAP4	1.016	0.025
RAB1B	1.006	0.029
NCK1	1.014	0.029
KAP3	1.011	0.029
RAB10	1.004	0.030
K1C10	0.997	0.030
GBB1	1.002	0.030
HCD2	1.017	0.031
CADM4	1.021	0.031
TRIM1	1.027	0.031
GLOD4	1.013	0.031
SYN1	1.002	0.032
PHB2	1.008	0.032
H2B1A	1.010	0.033
TAU	1.002	0.034
H2A2C	1.025	0.034
EPN3	1.025	0.035
TRAF3	0.940	0.036
HS904	1.007	0.036
ANR16	1.027	0.036
GDIB	1.003	0.036
SNP25	1.002	0.037
APT	1.021	0.037
SELR1	1.057	0.037
VATA	1.003	0.038
SYWC	1.005	0.038
CH60	1.002	0.038
KCC2B	1.002	0.038
KPYM	1.003	0.038
LDHA	1.002	0.039

TPPP3	1.015	0.039
RNB3L	0.990	0.041
AUXI	1.018	0.041
VGFR1	1.023	0.041
KCC2A	1.002	0.041
NPS3A	1.008	0.041
OLFM4	1.031	0.042
SEP7	1.003	0.042
STIP1	1.008	0.043
AT2B1	1.003	0.043
RTN4	1.004	0.043
HSP7C	1.002	0.043
TAGL3	0.993	0.044
YI023	0.983	0.044
ANXA6	1.002	0.045
PRS33	1.014	0.045
HNRPQ	1.030	0.046
HOME2	0.994	0.046
CALL3	1.004	0.047
OXR1	1.009	0.047
GSTM3	1.008	0.047
MDHC	1.002	0.047
DPYL5	0.990	0.047
SCND1	1.049	0.048
LAC2	1.015	0.049
BACHL	1.017	0.049
AP2B1	1.002	0.049
IQEC2	1.009	0.050

Table A.8.2 Proteins significantly (p<0.05) altered in acute PCP model brain tissue compared to control animals.

Protein Name	Fold Change	P Value
MBP	0.989	0.000
GLSK	1.006	0.001
MTAP2	0.997	0.001
CMC1	1.005	0.001
EDN2	0.972	0.001
VDAC3	0.985	0.001
VPP1	1.005	0.002
EF1A2	0.996	0.003
RPR1B	1.025	0.004
H2B1P	1.008	0.005
RRFM	0.985	0.009
QCR1	1.004	0.010

ARRB1	1.015	0.011
GEPH	1.012	0.011
SBSN	1.006	0.011
NCALD	1.003	0.011
POC1B	1.011	0.011
CN37	0.996	0.012
PFKAP	1.009	0.013
BASP1	0.997	0.015
ODPB	1.003	0.015
PSMD3	1.013	0.016
RAB13	0.991	0.021
H33	0.986	0.022
MCA3	0.987	0.025
AP2A1	1.006	0.026
SPG7	1.010	0.028
DPYL4	0.991	0.030
RAB10	1.004	0.030
H2B1K	0.984	0.032
VDAC2	0.996	0.032
RAB26	1.010	0.033
RAB3C	1.008	0.034
KLC3	1.006	0.035
SIR2	0.991	0.035
RAB35	0.985	0.036
GNAL	0.993	0.038
AINX	0.997	0.038
PDXK	1.006	0.039
SEGN	0.992	0.039
STX1A	1.004	0.042
DYN1	1.001	0.043
KPYR	1.004	0.044
H4	0.996	0.046
H2B1	0.997	0.047
ERAL1	1.012	0.049
HSP7C	0.998	0.050

Table A.8.3 Proteins significantly (p<0.05) altered in ketamine model brain tissue compared to control animals.

Protein Name	Fold Change	P Value
NEUG	0.960	0.000
GSTM4	1.016	0.000
KCRB	1.009	0.000
NDUA9	1.014	0.000
CAMKV	1.011	0.000
MYPR	0.989	0.000
GBB4	1.022	0.000
TIF1B	1.057	0.001
K2C1	1.014	0.001
GLNA	1.012	0.001
ATP4A	1.009	0.001
HS71L	1.015	0.001
STPG2	1.047	0.001
CLH1	1.006	0.001
SYT13	1.040	0.001
K2C8	0.973	0.001
NELFE	1.009	0.002
KPYM	1.006	0.002
STX17	1.019	0.002
FGFR2	1.048	0.003
AT1A4	1.008	0.003
RBM25	1.022	0.005
DHX8	1.023	0.005
SH3G1	1.064	0.006
ATP5J	1.022	0.006
BPNT1	1.020	0.006
DPYL3	1.012	0.007
RAB3C	0.984	0.007
COTL1	1.050	0.008
RUSD2	1.031	0.009
ACTS	1.052	0.009
GDIB	1.008	0.009
RL40	1.011	0.010
KPYR	1.015	0.010
AINX	0.995	0.010
CANB1	1.034	0.010
UCHL1	1.008	0.011
CH10	0.991	0.011
CAZA2	0.986	0.011
ALBU	1.004	0.013
ACTG	1.011	0.013
K2C73	1.011	0.014

MT3	0.993	0.014
CX6B1	0.986	0.015
GBB1	0.994	0.015
G3P	1.006	0.016
PHAR2	1.013	0.016
PDIA3	1.008	0.016
PP2BB	1.010	0.016
HXK1	1.007	0.016
GPM6B	1.014	0.017
GNA13	1.018	0.017
CMC1	1.006	0.017
DYN1	1.004	0.017
ARF3	1.010	0.017
MBP	0.995	0.017
RAC1	0.991	0.018
TBB2B	1.029	0.018
NRX1A	1.013	0.019
AT2B2	1.014	0.019
DHB2	1.023	0.019
HMGN2	1.021	0.020
OXR1	1.027	0.020
TT23L	1.025	0.021
SEP14	0.995	0.022
OPTN	1.017	0.022
KLOT	1.011	0.023
PGRC1	1.027	0.023
ACBP	0.989	0.024
PP2BA	1.007	0.024
CY1	1.019	0.026
CO1A1	1.020	0.026
K1C15	0.981	0.026
AP2A2	1.012	0.026
ODP2	1.005	0.027
NCALD	1.014	0.028
KCC2A	0.998	0.028
RHOF	1.024	0.029
CC175	1.015	0.030
ZN365	0.989	0.034
SYN2	1.008	0.034
PCSK1	1.013	0.035
HPCL4	1.012	0.036
MTPN	0.991	0.038
ZBT42	1.019	0.038
KAD1	1.007	0.038
TRIM1	0.987	0.039
VA0D1	0.991	0.040

NDUBA	1.012	0.042
F221B	1.021	0.045
ACADS	1.037	0.045
FSCN1	1.019	0.046
ENOG	1.004	0.048

Table A.8.4 Proteins significantly (p<0.05) altered in chronic PCP model brain tissue compared to control animals.

Protein Name	Fold Change	P Value
HBB1	1.012	0.000
ATPB	0.989	0.000
HBA	1.011	0.000
AT1A3	0.995	0.000
MTAP2	0.993	0.000
ENOG	0.991	0.000
1433Z	0.994	0.000
TBB2A	0.992	0.000
DPYL2	0.993	0.000
CH60	0.986	0.000
SEPT6	0.978	0.000
CN37	0.993	0.000
GNAO	0.994	0.001
HNRPD	0.978	0.001
HA2B	0.977	0.002
RAB6B	0.986	0.002
PP2BA	0.991	0.002
PICAL	0.977	0.002
CNTN1	0.989	0.003
LDHB	0.995	0.003
ENOB	0.983	0.003
GRP78	0.992	0.004
NDRG2	0.979	0.005
HSP7C	0.995	0.005
G6PI	0.994	0.006
KPYM	0.996	0.006
AL2SB	0.972	0.006
MT3	0.979	0.007
PARK7	0.989	0.008
SNAB	0.988	0.008
STX1B	0.994	0.009
NCAM1	0.990	0.009
RAB3B	0.981	0.010
AT2B1	0.994	0.010
CALX	0.990	0.010

TPM3	0.985	0.011
UMPS	0.965	0.011
HS90A	0.994	0.012
RAB12	1.053	0.012
EF2	0.987	0.014
AT1A2	0.996	0.014
AP3S2	0.969	0.015
SYUB	0.989	0.015
ACTC	0.966	0.015
DYN3	0.990	0.016
PACN1	0.987	0.017
VATG2	0.982	0.018
ATP5H	0.989	0.019
PP2BC	0.986	0.020
CX7A2	0.977	0.021
PSMD3	0.975	0.022
OSCP1	0.978	0.023
THY1	0.990	0.023
KS6A2	0.986	0.024
TAU	0.991	0.025
ST2A2	0.975	0.026
BACD3	0.982	0.026
KLD8A	0.964	0.027
PGAM1	0.994	0.028
AINX	0.992	0.029
SYPH	0.991	0.030
UCHL1	0.995	0.033
PROF1	0.989	0.033
RAC1	1.020	0.034
STIP1	0.973	0.035
API5	0.979	0.037
ALBU	1.004	0.038
MYH6	0.985	0.040
ASTL	0.982	0.040
KAD1	0.964	0.041
PRDX6	0.991	0.042
GSTM4	0.991	0.043
GNAI1	0.989	0.043
2AAA	0.993	0.044
CLCA	0.975	0.044
HSP72	0.990	0.045
IDH3B	0.985	0.046
MRRP1	0.986	0.046
VDAC3	1.008	0.046
RPGF5	0.976	0.047
VDAC2	0.988	0.047

1433E	0.994	0.048
AMPH	0.989	0.049
COX5B	0.984	0.050

Table A.8.5 Proteins significantly (p<0.05) altered in NR1 knockdown model brain tissue compared to wild-type animals.

Protein Name	Fold Change	P Value
LDHB	0.987	0.000
GNAO	0.985	0.000
TBB4A	0.981	0.000
KCRB	0.990	0.000
ENOB	0.972	0.000
HSP7C	0.990	0.000
DPYL2	0.992	0.000
DYN1	0.988	0.000
VISL1	0.982	0.000
G6PI	0.984	0.000
CAZA2	0.979	0.000
H2B1	1.009	0.000
GBB3	0.964	0.000
H14	1.018	0.001
ADM2	0.961	0.001
CLH1	0.993	0.001
THY1	0.962	0.001
ARF2	0.985	0.001
SNP25	0.989	0.002
1433T	0.991	0.002
MBP	0.992	0.002
LIAS	1.029	0.002
STXB1	0.994	0.002
LRC34	0.971	0.003
HS90A	0.990	0.003
AT1A1	0.988	0.003
TBB3	0.984	0.004
ACON	0.994	0.004
CETN2	0.960	0.004
SYUB	1.016	0.005
1433E	0.990	0.007
AT1A4	0.984	0.007
SEPT5	0.981	0.007
CI172	1.020	0.009
ZSC21	0.985	0.010
EF1A1	0.986	0.012
RLA1	0.981	0.013

CACB2	1.042	0.014
KCRS	0.968	0.015
SNAB	0.976	0.015
DHB3	1.013	0.016
GNAI1	0.985	0.017
MDHM	0.994	0.017
NSF	0.995	0.017
ACTS	0.984	0.018
GPSM1	0.976	0.018
DC1L1	1.011	0.018
TWST1	0.973	0.018
DCDC2	1.010	0.019
ODPB	0.982	0.020
AT1A2	0.993	0.021
TRPV6	0.944	0.021
LDHA	0.989	0.022
GNAI2	0.982	0.022
TNNC2	0.949	0.022
ATPA	0.995	0.022
PRDX6	0.972	0.023
KPYM	0.995	0.025
HMGCL	0.972	0.026
TBB4B	0.994	0.027
WDR5	0.964	0.027
TBA1B	0.992	0.027
CN37	0.983	0.028
2A5G	0.973	0.030
GNAI3	0.986	0.032
KCRU	0.977	0.035
TBB1	0.987	0.035
ALDOA	0.995	0.036
MAGIX	0.969	0.036
K2C72	0.972	0.038
1433Z	0.992	0.041
CDR2L	0.978	0.041
SNAPN	0.955	0.042
RRP15	1.034	0.043
TAGL2	0.991	0.043
RL40	0.985	0.044
TM143	0.965	0.045
TBAL3	1.020	0.047
ENOA	0.995	0.048
TBB6	0.989	0.049

Table A.8.6 The 162 proteins which were quantified across all human and rodent model tissue samples, as indicated in Figure 8.2. Arrows indicate the subsets of these proteins which were significantly differentially expressed in each comparison, and the direction of change.

Protein Name	Schizophrenia	aPCP	cPCP	Ketamine	NR1
AT1A2	↑		↓		↓
HBA	↑		↑		
HS90B	↑				
ALDOA	↑				↓
ENOG	↑		↓	↑	
DPYL3	↑			↑	
ALBU	↑		↑	↑	
TBA1B	↑				↓
HS90A	↑		↓		↓
VDAC2	↑	↓	↓		
GLNA	↑			↑	
DPYL1	↑				
PGAM1	↑		↓		
MYPR	↓			↓	
GBB2	↑				
KCC2D	↑				
1433G	↑				
CLH1	↑			↑	↓
AT1A4	↑			↑	↓
ACTB	↑				
ADT1	↑				
PPIA	↑				
MBP	↓	↓		↓	↓
GBB4	↑			↑	
DYN1	↑	↑		↑	↓
MPCP	↑				
RAB1B	↑				
RAB10	↑	↑			
GBB1	↑			↓	
SYN1	↑				
TAU	↑		↓		
GDIB	↑			↑	
SNP25	↑				↓
VATA	↑				
CH60	↑		↓		
KCC2B	↑				
KPYM	↑		↓	↑	↓
LDHA	↑				↓
KCC2A	↑			↓	
HSP7C	↑	↓	↓		↓

TAGL3	↓				
MDHC	↑				
AP2B1	↑				
HXK1				↑	
PRDX2					
EF1A1					↓
SAP					
POTEF					
TPIS					
ACTN1					
COF1					
SYN2				↑	
PRDX6			↓		↓
1433S					
ZC3HD					
RAB3A					
GNAO			↓		↓
PRDX5					
KCRU					↓
ENOA					↓
ALDOC					
GPM6A					
ADT2					
HSP72			↓		
STX1B			↓		
SYUA					
SH3G2					
DPYL2			↓		↓
EF1A2		↓			
PGK1					
ODPB		↑			↓
PP2BA			↓	↑	
1433Z			↓		↓
ACTBL					
AT5F1					
TBB3					↓
PEBP1					
ATPB			↓		
THY1			↓		↓
GNAI2					↓
CEND					
GBB3					↓
ATP4A				↑	
HBE					
KCC2G					
ENOB			↓		↓

SNAA					
AT1A1					↓
TBB4A					↓
NSF					↓
MDHM					↓
COX2					
ALDOB					
KPYR		↑		↑	
CN37		↓	↓		↓
1433E			↓		↓
EAA2					
AT12A					
ADT4					
G3P				↑	
AATC					
RAB37					
HS71L				↑	
H2A1H					
ACTS				↑	↓
GNAI1			↓		↓
FKB1A					
SH3G1				↑	
AT1B1					
GNAT3					
TBB1					↓
GNAT2					
VATH					
LDHB			↓		↓
KCRB				↑	↓
SYUB			↓		↑
RL40				↑	↓
AATM					
STXB1					↓
GNAS1					
TBB2B				↑	
1433T					↓
AP2M1					
G6PI			↓		↓
CALM					
CY561					
GDIA					
NP1L4					
GNAT1					
RAB14					
COX5A					
ENTP5					

VATB2					
RAB3D					
TBB2A			↓		
TBB6					↓
AT1A3			↓		
VDAC3		↓	↑		
ATPA					↓
TBA1A					
1433B					
H4		↓			
VDAC1					
BASP1		↓			
ACTG				↑	
SNAB			↓		↓
ARF5					
NDUA4					
TBB4B					↓
TBA4A					
NELFE				↑	
IDH3A					
QCR2					
RAB3C		↑		↓	
NDKA					
ATPO					
TBA1C					
ACON					↓
TBAL3					↑
RAB1A					
VISL1					↓
TBB5					

References

1. Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, “just the facts” what we know in 2008. 2. Epidemiology and etiology. Schizophr Res [Internet]. 2008 Jul [cited 2014 Jul 18];102(1–3):1–18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18514488

2. Hennekens CH, Hennekens AR, Hollar D, Casey DE. Schizophrenia and increased risks of cardiovascular disease. Am Heart J [Internet]. 2005 Dec [cited 2017 Mar 10];150(6):1115–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16338246

3. Laursen TM, Wahlbeck K, Hällgren J, Westman J, Ösby U, Alinaghizadeh H, et al. Life Expectancy and Death by Diseases of the Circulatory System in Patients with Bipolar Disorder or Schizophrenia in the Nordic Countries. Mazza M, editor. PLoS One [Internet]. 2013 Jun 24 [cited 2017 Mar 10];8(6):e67133. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23826212

4. Ringen PA, Engh JA, Birkenaes AB, Dieset I, Andreassen OA. Increased Mortality in Schizophrenia Due to Cardiovascular Disease â€“ A Non-Systematic Review of Epidemiology, Possible Causes, and Interventions. Front Psychiatry [Internet]. 2014 Sep 26 [cited 2017 Mar 10];5:137. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25309466

5. Thara R, Kamath S. Women and schizophrenia. Indian J Psychiatry [Internet]. 2015 Jul [cited 2017 Mar 6];57(Suppl 2):S246-51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26330642

6. Addington J, Addington D, Maticka-Tyndale E. Cognitive functioning and positive and negative symptoms in schizophrenia. Schizophr Res [Internet]. 1991 Sep [cited 2017 Mar 6];5(2):123–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1931805

7. Andreasen NC, Flaum M, Swayze VW, Tyrrell G, Arndt S. Positive and negative symptoms in schizophrenia. A critical reappraisal. Arch Gen Psychiatry [Internet]. 1990 Jul [cited 2017 Mar 6];47(7):615–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2360855

8. DE Hert M, Correll CU, Bobes J, Cetkovich-Bakmas M, Cohen D, Asai I, et al. Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. World Psychiatry [Internet]. 2011 Feb [cited 2017 Mar 6];10(1):52–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21379357

9. Rössler W, Salize HJ, van Os J, Riecher-Rössler A. Size of burden of schizophrenia and psychotic disorders. Eur Neuropsychopharmacol [Internet]. 2005 Aug [cited 2016 Feb 1];15(4):399–409. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15925493

10. Wu EQ, Birnbaum HG, Shi L, Ball DE, Kessler RC, Moulis M, et al. The economic burden of schizophrenia in the United States in 2002. J Clin Psychiatry [Internet]. 2005 Sep [cited 2017 Mar 6];66(9):1122–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16187769

11. Metastasio A, Bahn S. Utility of biomarkers to improve the diagnosis and treatment of schizophrenia. Future Neurol [Internet]. 2008 Nov [cited 2017 Mar 6];3(6):619–22. Available from: http://www.futuremedicine.com/doi/10.2217/14796708.3.6.619

12. Chan MK, Krebs M-O, Cox D, Guest PC, Yolken RH, Rahmoune H, et al. Development of a blood-based molecular biomarker test for identification of schizophrenia before disease onset. Transl Psychiatry [Internet]. 2015 Jul 14 [cited 2017 Jan 5];5:e601. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26171982

13. Bromet EJ, Kotov R, Fochtmann LJ, Carlson GA, Tanenberg-Karant M, Ruggero C, et al. Diagnostic Shifts During the Decade Following First Admission for Psychosis. Am J Psychiatry [Internet]. 2011 Nov [cited 2017 Aug 8];168(11):1186–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21676994

14. Fusar-Poli P, Bonoldi I, Yung AR, Borgwardt S, Kempton MJ, Valmaggia L, et al. Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. Arch Gen Psychiatry [Internet]. 2012 Mar [cited 2017 Jan 5];69(3):220–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22393215

15. Miyamoto S, Miyake N, Jarskog LF, Fleischhacker WW, Lieberman JA. Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. Mol Psychiatry [Internet]. 2012 Dec 15 [cited 2017 Mar 6];17(12):1206–27. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22584864

16. Mailman RB, Murthy V. Third generation antipsychotic drugs: partial agonism or receptor functional selectivity? Curr Pharm Des [Internet]. 2010 [cited 2017 Aug 8];16(5):488–501. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19909227

17. Emsley R, Chiliza B, Asmal L, Harvey BH. The nature of relapse in schizophrenia. BMC Psychiatry [Internet]. 2013 Dec 8 [cited 2017 Mar 6];13(1):50. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23394123

18. Chakos M, Lieberman J, Hoffman E, Bradford D, Sheitman B. Effectiveness of Second-Generation Antipsychotics in Patients With Treatment-Resistant Schizophrenia: A Review and Meta-Analysis of Randomized Trials. Am J Psychiatry [Internet]. 2001 Apr [cited 2017 Mar 6];158(4):518–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11282684

19. Lally J, MacCabe JH. Antipsychotic medication in schizophrenia: a review. Br Med Bull [Internet]. 2015 Jun [cited 2017 Aug 8];114(1):169–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25957394

20. Miyamoto S, Duncan GE, Marx CE, Lieberman JA. Treatments for schizophrenia: a critical review of pharmacology and mechanisms of action of antipsychotic drugs. Mol Psychiatry [Internet]. 2005 Jan 3 [cited 2017 Mar 6];10(1):79–104. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15289815

21. Lieberman JA, Tollefson G, Tohen M, Green AI, Gur RE, Kahn R, et al. Comparative Efficacy and Safety of Atypical and Conventional Antipsychotic Drugs in First-Episode Psychosis: A Randomized, Double-Blind Trial of Olanzapine Versus Haloperidol. Am J Psychiatry [Internet]. 2003 Aug [cited 2017 Mar 6];160(8):1396–404. Available from: http://psychiatryonline.org/doi/abs/10.1176/appi.ajp.160.8.1396

22. Carpenter WT, Koenig JI. The evolution of drug development in schizophrenia: past issues and future opportunities. Neuropsychopharmacology [Internet]. 2008 Aug [cited 2017 Mar 6];33(9):2061–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18046305

23. O'Brien PL, Thomas CP, Hodgkin D, Levit KR, Mark TL. The diminished pipeline for medications to treat mental health and substance use disorders. Psychiatr Serv [Internet]. 2014 Dec 1 [cited 2017 Mar 6];65(12):1433–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25178309

24. Sethi S, Brietzke E. Omics-Based Biomarkers: Application of Metabolomics in Neuropsychiatric Disorders. Int J Neuropsychopharmacol [Internet]. 2015 Oct 9 [cited 2017 Aug 8];19(3):pyv096. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26453695

25. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. Nature [Internet]. 2016 Jan 27 [cited 2017 Mar 6];530(7589):177–83. Available from: http://www.nature.com/doifinder/10.1038/nature16549

26. Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, Holmans PA, et al. Biological insights from

108 schizophrenia-associated genetic loci. Nature [Internet]. 2014 Jul 22 [cited 2017 Mar 9];511(7510):421–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25056061

27. Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kähler AK, Akterin S, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet [Internet]. 2013 Aug 25 [cited 2017 Mar 9];45(10):1150–9. Available from: http://www.nature.com/doifinder/10.1038/ng.2742

28. Kocerha J, Kauppinen S, Wahlestedt C. microRNAs in CNS Disorders. NeuroMolecular Med [Internet]. 2009 Sep 18 [cited 2017 Aug 8];11(3):162–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19536656

29. Pickard BS. Schizophrenia biomarkers: Translating the descriptive into the diagnostic. Pratt J, editor. J Psychopharmacol [Internet]. 2015 Feb [cited 2017 Mar 14];29(2):138–43. Available from: http://journals.sagepub.com/doi/10.1177/0269881114566631

30. Lindon JC, Holmes E, Nicholson JK. So what's the deal with metabonomics? Anal Chem [Internet]. 2003 Sep 1 [cited 2017 Aug 8];75(17):384A–391A. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14632032

31. McEvoy JP, Meyer JM, Goff DC, Nasrallah HA, Davis SM, Sullivan L, et al. Prevalence of the metabolic syndrome in patients with schizophrenia: Baseline results from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial and comparison with national estimates from NHANES III. Schizophr Res [Internet]. 2005 Dec 1 [cited 2017 Mar 13];80(1):19–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16137860

32. Levin Y, Wang L, Schwarz E, Koethe D, Leweke FM, Bahn S. Global proteomic profiling reveals altered proteomic signature in schizophrenia serum. Mol Psychiatry [Internet]. 2010 Nov [cited 2016 Jan 29];15(11):1088–100. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19546861

33. Jaros JAJ, Martins-de-Souza D, Rahmoune H, Rothermundt M, Leweke FM, Guest PC, et al. Protein phosphorylation patterns in serum from schizophrenia patients and healthy controls. J Proteomics [Internet]. 2012 Dec 5 [cited 2017 Mar 9];43–55. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1874391912003405

34. Domenici E, Willé DR, Tozzi F, Prokopenko I, Miller S, McKeown A, et al. Plasma protein biomarkers for depression and schizophrenia by multi analyte profiling of case-control collections. PLoS One [Internet]. 2010 Feb 11 [cited 2017 Jan 12];5(2):e9166. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/20161799

35. Martins-De-Souza D, Wobrock T, Zerr I, Schmitt A, Gawinecka J, Schneider-Axmann T, et al. Different apolipoprotein E, apolipoprotein A1 and prostaglandin-H2 D-isomerase levels in cerebrospinal fluid of schizophrenia patients and healthy controls. World J Biol Psychiatry [Internet]. 2010 Aug 7 [cited 2017 Mar 9];11(5):719–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20446881

36. Albertini V, Benussi L, Paterlini A, Glionna M, Prestia A, Bocchio-Chiavetto L, et al. Distinct cerebrospinal fluid amyloid-beta peptide signatures in cognitive decline associated with Alzheimer's disease and schizophrenia. Electrophoresis [Internet]. 2012 Dec [cited 2017 Mar 9];33(24):3738–44. Available from: http://doi.wiley.com/10.1002/elps.201200307

37. Huang DW, Sherman BT, Tan Q. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007;8(9):R183.

38. Johnston-Wilson NL, Sims CD, Hofmann JP, Anderson L, Shore AD, Torrey EF, et al. Disease-specific alterations in frontal cortex brain proteins in schizophrenia, bipolar disorder, and major depressive disorder. The Stanley Neuropathology Consortium. Mol Psychiatry [Internet]. 2000 Mar [cited 2017 Mar 9];5(2):142–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10822341

39. Martins-de-Souza D, Maccarrone G, Wobrock T, Zerr I, Gormanns P, Reckow S, et al. Proteome analysis of the thalamus and cerebrospinal fluid reveals glycolysis dysfunction and potential biomarkers candidates for schizophrenia. J Psychiatr Res [Internet]. 2010 Dec [cited 2017 Mar 9];44(16):1176–89. Available from: http://linkinghub.elsevier.com/retrieve/pii/S002239561000124X

40. Clark D, Dedova I, Cordwell S, Matsumoto I. A proteome analysis of the anterior cingulate cortex gray matter in schizophrenia. Mol Psychiatry [Internet]. 2006 May 21 [cited 2017 Mar 9];11(5):459–70, 423. Available from: http://www.nature.com/doifinder/10.1038/sj.mp.4001806

41. Sivagnanasundaram S, Crossett B, Dedova I, Cordwell S, Matsumoto I. Abnormal pathways in the genu of the corpus callosum in schizophrenia pathogenesis: a proteome study. Proteomics Clin Appl [Internet]. 2007 Oct [cited 2017 Mar 9];1(10):1291–305. Available from: http://doi.wiley.com/10.1002/prca.200700230

42. Föcking M, Dicker P, English JA, Schubert KO, Dunn MJ, Cotter DR. Common proteomic changes in the hippocampus in schizophrenia and bipolar disorder and particular evidence for involvement of cornu ammonis regions 2 and 3. Arch Gen Psychiatry [Internet]. 2011 May 2 [cited 2017 Mar 9];68(5):477–88. Available from: http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/archgenpsychiatry.2011.43

43. Nascimento JM, Martins-de-Souza D, Levin Y, Koutroukides T, Rahmoune H, Pietsch S. The proteome of schizophrenia. npj Schizophr [Internet]. 2015 Dec 4 [cited 2017 Mar 6];1(1):14003. Available from: http://www.nature.com/articles/npjschz20143

44. Davalieva K, Maleva Kostovska I, Dwork AJ. Proteomics Research in Schizophrenia. Front Cell Neurosci [Internet]. 2016 [cited 2017 Mar 6];10:18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26909022

45. Guest PC, Guest FL, Martins-de Souza D. Making Sense of Blood-Based Proteomics and Metabolomics in Psychiatric Research. Int J Neuropsychopharmacol [Internet]. 2015 Dec 30 [cited 2017 Aug 8];19(6). Available from: http://www.ncbi.nlm.nih.gov/pubmed/26721951

46. Owen MJ, Craddock N, O'Donovan MC. Suggestion of Roles for Both Common and Rare Risk Variants in Genome-wide Studies of Schizophrenia. Arch Gen Psychiatry [Internet]. 2010 Jul 1 [cited 2017 Mar 9];67(7):667. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20603448

47. Baloyianni N, Tsangaris GT. The audacity of proteomics: a chance to overcome current challenges in schizophrenia research. Expert Rev Proteomics [Internet]. 2009 Dec 9 [cited 2017 Mar 9];6(6):661–74. Available from: http://www.tandfonline.com/doi/full/10.1586/epr.09.85

48. Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet [Internet]. 2012 Mar 13 [cited 2017 Mar 9];13(4):227–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22411467

49. English JA, Pennington K, Dunn MJ, Cotter DR. The Neuroproteomics of Schizophrenia. Biol Psychiatry [Internet]. 2011 Jan 15 [cited 2017 Mar 9];69(2):163–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20887976

50. Guest PC, Chan MK, Gottschalk MG, Bahn S. The use of proteomic biomarkers for improved diagnosis and stratification of schizophrenia patients. Biomark Med [Internet]. 2014 Jan [cited 2017 Mar 9];8(1):15–27. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24325222

51. Ho B-C, Andreasen NC, Nopoulos P, Arndt S, Magnotta V, Flaum M, et al. Progressive Structural Brain Abnormalities and Their Relationship to Clinical Outcome. Arch Gen Psychiatry [Internet]. 2003 Jun 1 [cited 2017 Mar 10];60(6):585. Available from: http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/archpsyc.60.6.585

52. Wright IC, Rabe-Hesketh S, Woodruff PWR, David AS, Murray RM, Bullmore ET. Meta-Analysis of Regional Brain Volumes in Schizophrenia. Am J Psychiatry [Internet]. 2000 Jan [cited 2017 Mar 10];157(1):16–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10618008

53. Olabi B, Ellison-Wright I, McIntosh AM, Wood SJ, Bullmore E, Lawrie SM. Are There Progressive Brain Changes in Schizophrenia? A Meta-Analysis of Structural Magnetic Resonance Imaging Studies. Biol Psychiatry [Internet]. 2011 Jul 1 [cited 2017 Mar 10];70(1):88–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21457946

54. Ho B-C, Andreasen NC, Ziebell S, Pierson R, Magnotta V. Long-term Antipsychotic Treatment and Brain Volumes. Arch Gen Psychiatry [Internet]. 2011 Feb 7 [cited 2017 Mar 10];68(2):128. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21300943

55. Cannon TD, van Erp TG, Huttunen M, Lönnqvist J, Salonen O, Valanne L, et al. Regional gray matter, white matter, and cerebrospinal fluid distributions in schizophrenic patients, their siblings, and controls. Arch Gen Psychiatry [Internet]. 1998 Dec [cited 2017 Mar 10];55(12):1084–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9862551

56. Harrison PJ. The neuropathology of schizophrenia. A critical review of the data and their interpretation. Brain [Internet]. 1999 Apr [cited 2017 Mar 10];593–624. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10219775

57. Bunney WE, Bunney BG. Evidence for a compromised dorsolateral prefrontal cortical parallel circuit in schizophrenia. Brain Res Brain Res Rev [Internet]. 2000 Mar [cited 2017 Mar 10];31(2–3):138–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10719142

58. Narayan S, Tang B, Head SR, Gilmartin TJ, Sutcliffe JG, Dean B, et al. Molecular profiles of schizophrenia in the CNS at different stages of illness. Brain Res [Internet]. 2008 Nov 6 [cited 2017 Mar 10];1239:235–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18778695

59. Pennington K, Beasley CL, Dicker P, Fagan A, English J, Pariante CM, et al. Prominent synaptic and metabolic abnormalities revealed by proteomic analysis of the dorsolateral prefrontal cortex in schizophrenia and bipolar disorder. Mol Psychiatry [Internet]. 2008 Dec 16 [cited 2017 Mar 10];13(12):1102–17. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/17938637

60. Martins-de-Souza D, Gattaz WF, Schmitt A, Rewerts C, Maccarrone G, Dias-Neto E, et al. Prefrontal cortex shotgun proteome analysis reveals altered calcium homeostasis and immune system imbalance in schizophrenia. Eur Arch Psychiatry Clin Neurosci [Internet]. 2009 Apr 22 [cited 2017 Mar 10];259(3):151–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19165527

61. Arion D, Unger T, Lewis DA, Levitt P, Mirnics K. Molecular Evidence for Increased Expression of Genes Related to Immune and Chaperone Function in the Prefrontal Cortex in Schizophrenia. Biol Psychiatry [Internet]. 2007 Oct 1 [cited 2017 Mar 10];62(7):711–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17568569

62. Potvin S, Stip E, Sepehry AA, Gendron A, Bah R, Kouassi E. Inflammatory Cytokine Alterations in Schizophrenia: A Systematic Quantitative Review. Biol Psychiatry [Internet]. 2008 Apr 15 [cited 2017 Mar 10];63(8):801–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18005941

63. Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-Analysis of Cytokine Alterations in Schizophrenia: Clinical Status and Antipsychotic Effects. Biol Psychiatry [Internet]. 2011 Oct 1 [cited 2017 Mar 10];70(7):663–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21641581

64. Guest PC, Schwarz E, Krishnamurthy D, Harris LW, Leweke FM, Rothermundt M, et al. Altered levels of circulating insulin and other neuroendocrine hormones associated with the onset of schizophrenia. Psychoneuroendocrinology [Internet]. 2011 Aug [cited 2017 Mar 10];36(7):1092–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21251762

65. Green MJ, Matheson SL, Shepherd A, Weickert CS, Carr VJ. Brain-derived neurotrophic factor levels in schizophrenia: a systematic review with meta-analysis. Mol Psychiatry [Internet]. 2011 Sep 24 [cited 2017 Mar 10];16(9):960–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20733577

66. Raison CL, Miller AH. Is Depression an Inflammatory Disorder? Curr Psychiatry Rep [Internet]. 2011 Dec 17 [cited 2017 Mar 10];13(6):467–75. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21927805

67. Laan W, Grobbee DE, Selten J-P, Heijnen CJ, Kahn RS, Burger H. Adjuvant Aspirin Therapy Reduces Symptoms of Schizophrenia Spectrum Disorders. J Clin Psychiatry [Internet]. 2010

May 15 [cited 2017 Mar 10];71(05):520–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20492850

68. Müller N, Krause D, Dehning S, Musil R, Schennach-Wolff R, Obermeier M, et al. Celecoxib treatment in an early stage of schizophrenia: Results of a randomized, double-blind, placebo-controlled trial of celecoxib augmentation of amisulpride treatment. Schizophr Res [Internet]. 2010 Aug [cited 2017 Mar 10];121(1–3):118–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20570110

69. Palacio JR, Markert UR, Martínez P. Anti-inflammatory properties of N-acetylcysteine on lipopolysaccharide-activated macrophages. Inflamm Res [Internet]. 2011 Jul 20 [cited 2017 Mar 10];60(7):695–704. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21424515

70. Oja SS, Janáky R, Varga V, Saransaari P. Modulation of glutamate receptor functions by glutathione. Neurochem Int [Internet]. [cited 2017 Mar 10];37(2–3):299–306. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10812215

71. Spelman LM, Walsh PI, Sharifi N, Collins P, Thakore JH. Impaired glucose tolerance in first-episode drug-naïve patients with schizophrenia. Diabet Med [Internet]. 2007 May [cited 2017 Mar 10];24(5):481–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17381506

72. Amminger GP, Schäfer MR, Papageorgiou K, Klier CM, Cotton SM, Harrigan SM, et al. Long-Chain ω-3 Fatty Acids for Indicated Prevention of Psychotic Disorders. Arch Gen Psychiatry [Internet]. 2010 Feb 1 [cited 2017 Mar 10];67(2):146. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20124114

73. Ehrenreich H, Nave K-A. Phenotype-Based Genetic Association Studies (PGAS)—Towards Understanding the Contribution of Common Genetic Variants to Schizophrenia Subphenotypes. Genes (Basel) [Internet]. 2014 Feb 27 [cited 2017 Mar 10];5(1):97–105. Available from: http://www.mdpi.com/2073-4425/5/1/97/

74. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature [Internet]. 2009 Jul 1 [cited 2017 Mar 12];460(7256):748–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19571811

75. Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA, et al. Genome-wide association study identifies five new schizophrenia loci. Nat Genet [Internet]. 2011 Sep 18 [cited 2017 Mar 12];43(10):969–76. Available from:

http://www.nature.com/doifinder/10.1038/ng.940

76. Lee SH, DeCandia TR, Ripke S, Yang J, Sullivan PF, Goddard ME, et al. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. Nat Genet [Internet]. 2012 Feb 19 [cited 2017 Mar 12];44(3):247–50. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22344220

77. Loh P-R, Bhatia G, Gusev A, Finucane HK, Bulik-Sullivan BK, Pollack SJ, et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. Nat Genet [Internet]. 2015 Nov 2 [cited 2017 Mar 12];47(12):1385–92. Available from: http://www.nature.com/doifinder/10.1038/ng.3431

78. Ohi K, Shimada T, Yasuyama T, Uehara T, Kawasaki Y. Variability of 128 schizophrenia-associated gene variants across distinct ethnic populations. Transl Psychiatry [Internet]. 2017 Jan 3 [cited 2018 May 4];7(1):e988–e988. Available from: http://www.nature.com/articles/tp2016260

79. Benros ME, Nielsen PR, Nordentoft M, Eaton WW, Dalton SO, Mortensen PB. Autoimmune Diseases and Severe Infections as Risk Factors for Schizophrenia: A 30-Year Population-Based Register Study. Am J Psychiatry [Internet]. 2011 Dec [cited 2017 Mar 13];168(12):1303–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22193673

80. Smyth AM, Lawrie SM. The neuroimmunology of schizophrenia. Clin Psychopharmacol Neurosci [Internet]. 2013 Dec [cited 2017 Mar 13];11(3):107–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24465246

81. Hwang Y, Kim J, Shin J-Y, Kim J-I, Seo J-S, Webster MJ, et al. Gene expression profiling by mRNA sequencing reveals increased expression of immune/inflammation-related genes in the hippocampus of individuals with schizophrenia. Transl Psychiatry [Internet]. 2013 Oct 29 [cited 2017 Mar 13];3(10):e321. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24169640

82. Hanson DR, Gottesman II. Theories of schizophrenia: a genetic-inflammatory-vascular synthesis. BMC Med Genet [Internet]. 2005 Dec 11 [cited 2017 Mar 13];6(1):7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15707482

83. Chan MK, Guest PC, Levin Y, Umrania Y, Schwarz E, Bahn S, et al. Converging evidence of blood-based biomarkers for schizophrenia. In: International review of neurobiology [Internet]. 2011 [cited 2017 Mar 13]. p. 95–144. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/22050850

84. Dickerson F, Stallings C, Origoni A, Vaughan C, Khushalani S, Yang S, et al. C-reactive protein is elevated in schizophrenia. Schizophr Res [Internet]. 2013 Jan [cited 2017 Mar 13];143(1):198–202. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0920996412006159

85. Miller BJ, Culpepper N, Rapaport MH. C-Reactive Protein Levels in Schizophrenia. Clin Schizophr Relat Psychoses [Internet]. 2014 Jan [cited 2017 Mar 13];7(4):223–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23428789

86. Drexhage RC, Weigelt K, van Beveren N, Cohen D, Versnel MA, Nolen WA, et al. Immune and Neuroimmune Alterations in Mood Disorders and Schizophrenia. In: International review of neurobiology [Internet]. 2011 [cited 2017 Mar 13]. p. 169–201. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22050852

87. Murray RM, Lewis SW. Is schizophrenia a neurodevelopmental disorder? Br Med J (Clin Res Ed) [Internet]. 1987 Sep 19 [cited 2017 Mar 13];295(6600):681–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3117295

88. Brown AS, Begg MD, Gravenstein S, Schaefer CA, Wyatt RJ, Bresnahan M, et al. Serologic Evidence of Prenatal Influenza in the Etiology of Schizophrenia. Arch Gen Psychiatry [Internet]. 2004 Aug 1 [cited 2017 Mar 13];61(8):774. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15289276

89. Brown AS, Cohen P, Greenwald S, Susser E. Nonaffective Psychosis After Prenatal Exposure to Rubella. Am J Psychiatry [Internet]. 2000 Mar 1 [cited 2017 Mar 13];157(3):438–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10698821

90. Suvisaari J, Haukka J, Tanskanen A, Hovi T, Lönnqvist J. Association between prenatal exposure to poliovirus infection and adult schizophrenia. Am J Psychiatry [Internet]. 1999 Jul [cited 2017 Mar 13];156(7):1100–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10401461

91. Khandaker GM, Zimbron J, Dalman C, Lewis G, Jones PB. Childhood infection and adult schizophrenia: A meta-analysis of population-based studies. Schizophr Res [Internet]. 2012 Aug [cited 2017 Mar 13];139(1–3):161–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22704639

92. Koponen H, Rantakallio P, Veijola J, Jones P, Jokelainen J, Isohanni M. Childhood central

nervous system infections and risk for schizophrenia. Eur Arch Psychiatry Clin Neurosci [Internet]. 2004 Feb 1 [cited 2017 Mar 13];254(1):9–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14991373

93. Karlsson H, Bachmann S, Schroder J, McArthur J, Torrey EF, Yolken RH. Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. Proc Natl Acad Sci [Internet]. 2001 Apr 10 [cited 2017 Mar 13];98(8):4634–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11296294

94. Mortensen PB, Nørgaard-Pedersen B, Waltoft BL, Sørensen TL, Hougaard D, Torrey EF, et al. Toxoplasma gondii as a Risk Factor for Early-Onset Schizophrenia: Analysis of Filter Paper Blood Samples Obtained at Birth. Biol Psychiatry [Internet]. 2007 Mar 1 [cited 2017 Mar 13];61(5):688–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16920078

95. Brown AS, Schaefer CA, Quesenberry CP, Liu L, Babulas VP, Susser ES. Maternal Exposure to Toxoplasmosis and Risk of Schizophrenia in Adult Offspring. Am J Psychiatry [Internet]. 2005 Apr [cited 2017 Mar 13];162(4):767–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15800151

96. Webster JP, Kaushik M, Bristow GC, McConkey GA. Toxoplasma gondii infection, from predation to schizophrenia: can animal behaviour help us understand human behaviour? J Exp Biol [Internet]. 2013 Jan 1 [cited 2017 Mar 13];216(1):99–112. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23225872

97. Niebuhr DW, Millikan AM, Cowan DN, Yolken R, Li Y, Weber NS. Selected Infectious Agents and Risk of Schizophrenia Among U.S. Military Personnel. Am J Psychiatry [Internet]. 2008 Jan [cited 2017 Mar 13];165(1):99–106. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18086751

98. Amminger GP, McGorry PD, Berger GE, Wade D, Yung AR, Phillips LJ, et al. Antibodies to Infectious Agents in Individuals at Ultra-High Risk for Psychosis. Biol Psychiatry [Internet]. 2007 May 15 [cited 2017 Mar 13];61(10):1215–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17207471

99. Birkenaes AB, Opjordsmoen S, Brunborg C, Engh JA, Jonsdottir H, Ringen PA, et al. The level of cardiovascular risk factors in bipolar disorder equals that of schizophrenia: a comparative study. J Clin Psychiatry [Internet]. 2007 Jun [cited 2017 Mar 13];68(6):917–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17592917

100. Heiskanen T, Niskanen L, Lyytikäinen R, Saarinen PI, Hintikka J. Metabolic syndrome in patients with schizophrenia. J Clin Psychiatry [Internet]. 2003 May [cited 2017 Mar 13];64(5):575–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12755662

101. Saari KM, Lindeman SM, Viilo KM, Isohanni MK, Järvelin M-R, Laurén LH, et al. A 4-fold risk of metabolic syndrome in patients with schizophrenia: the Northern Finland 1966 Birth Cohort study. J Clin Psychiatry [Internet]. 2005 May [cited 2017 Mar 13];66(5):559–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15889940

102. Osborn DPJ, Levy G, Nazareth I, Petersen I, Islam A, King MB. Relative Risk of Cardiovascular and Cancer Mortality in People With Severe Mental Illness From the United Kingdom's General Practice Research Database. Arch Gen Psychiatry [Internet]. 2007 Feb 1 [cited 2017 Mar 13];64(2):242. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17283292

103. Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. Lancet [Internet]. 2005 Apr [cited 2017 Mar 13];365(9468):1415–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15836891

104. Leucht S, Cipriani A, Spineli L, Mavridis D, Örey D, Richter F, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet [Internet]. 2013 Sep 14 [cited 2017 Mar 13];382(9896):951–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23810019

105. Thakore JH. Metabolic disturbance in first-episode schizophrenia. Br J Psychiatry Suppl [Internet]. 2004 Apr [cited 2017 Mar 13];47:S76-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15056598

106. Newcomer JW. Second-generation (atypical) antipsychotics and metabolic effects: a comprehensive literature review. CNS Drugs [Internet]. 2005 [cited 2017 Mar 13];19 Suppl 1:1–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15998156

107. van Beveren NJM, Schwarz E, Noll R, Guest PC, Meijer C, de Haan L, et al. Evidence for disturbed insulin and growth hormone signaling as potential risk factors in the development of schizophrenia. Transl Psychiatry [Internet]. 2014 Aug 26 [cited 2017 Mar 13];4(8):e430. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25158005

108. Guest PC, Wang L, Harris LW, Burling K, Levin Y, Ernst A, et al. Increased levels of circulating insulin-related peptides in first-onset, antipsychotic naïve schizophrenia patients. Mol Psychiatry [Internet]. 2010 Feb 1 [cited 2015 Nov 9];15(2):118–9. Available from:

http://www.nature.com/mp/journal/v15/n2/full/mp200981a.html#bib8

109. Sham PC, O'Callaghan E, Takei N, Murray GK, Hare EH, Murray RM. Schizophrenia following pre-natal exposure to influenza epidemics between 1939 and 1960. Br J Psychiatry [Internet]. 1992 Apr [cited 2017 Mar 13];160:461–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1294066

110. O'Callaghan E, Sham P, Takei N, Glover G, Murray RM. Schizophrenia after prenatal exposure to 1957 A2 influenza epidemic. Lancet (London, England) [Internet]. 1991 May 25 [cited 2017 Mar 13];337(8752):1248–50. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1674062

111. Castle D, Sham P, Murray R. Differences in distribution of ages of onset in males and females with schizophrenia. Schizophr Res [Internet]. 1998 Oct 9 [cited 2017 Mar 13];33(3):179–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9789910

112. Pérez-Iglesias R, Tordesillas-Gutiérrez D, Barker GJ, McGuire PK, Roiz-Santiañez R, Mata I, et al. White matter defects in first episode psychosis patients: A voxelwise analysis of diffusion tensor imaging. Neuroimage [Internet]. 2010 Jan 1 [cited 2017 Mar 13];49(1):199–204. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19619664

113. STEEN RG, Mull C, McClure R, Hamer RM, Lieberman JA. Brain volume in first-episode schizophrenia: Systematic review and meta-analysis of magnetic resonance imaging studies. Br J Psychiatry [Internet]. 2006 Jun 1 [cited 2017 Mar 13];188(6):510–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16738340

114. Witthaus H, Kaufmann C, Bohner G, Özgürdal S, Gudlowski Y, Gallinat J, et al. Gray matter abnormalities in subjects at ultra-high risk for schizophrenia and first-episode schizophrenic patients compared to healthy controls. Psychiatry Res Neuroimaging [Internet]. 2009 Sep 30 [cited 2017 Mar 13];173(3):163–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19616415

115. Grace AA. Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. Nat Rev Neurosci [Internet]. 2016 Jun 3 [cited 2017 Mar 13];17(8):524–32. Available from: http://www.nature.com/doifinder/10.1038/nrn.2016.57

116. Howes O, McCutcheon R, Stone J. Glutamate and dopamine in schizophrenia: An update for the 21 st century. Pratt J, editor. J Psychopharmacol [Internet]. 2015 Feb [cited 2017 Mar 13];29(2):97–115. Available from: http://journals.sagepub.com/doi/10.1177/0269881114563634

117. Kambeitz J, Abi-Dargham A, Kapur S, Howes OD. Alterations in cortical and extrastriatal subcortical dopamine function in schizophrenia: systematic review and meta-analysis of imaging studies. Br J Psychiatry [Internet]. 2014 Jun 1 [cited 2017 Mar 13];204(6):420–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25029687

118. Allen P, Chaddock CA, Howes OD, Egerton A, Seal ML, Fusar-Poli P, et al. Abnormal Relationship Between Medial Temporal Lobe and Subcortical Dopamine Function in People With an Ultra High Risk for Psychosis. Schizophr Bull [Internet]. 2012 Sep 1 [cited 2017 Mar 13];38(5):1040–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21536784

119. Howes O, Bose S, Turkheimer F, Valli I, Egerton A, Stahl D, et al. Progressive increase in striatal dopamine synthesis capacity as patients develop psychosis: a PET study. Mol Psychiatry [Internet]. 2011 Sep 1 [cited 2017 Mar 13];16(9):885–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21358709

120. Egerton A, Chaddock CA, Winton-Brown TT, Bloomfield MAP, Bhattacharyya S, Allen P, et al. Presynaptic Striatal Dopamine Dysfunction in People at Ultra-high Risk for Psychosis: Findings in a Second Cohort. Biol Psychiatry [Internet]. 2013 Jul 15 [cited 2017 Mar 13];74(2):106–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23312565

121. Howes OD, Montgomery AJ, Asselin M-C, Murray RM, Valli I, Tabraham P, et al. Elevated Striatal Dopamine Function Linked to Prodromal Signs of Schizophrenia. Arch Gen Psychiatry [Internet]. 2009 Jan 1 [cited 2017 Mar 13];66(1):13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19124684

122. Howes OD, Kapur S. A neurobiological hypothesis for the classification of schizophrenia: type A (hyperdopaminergic) and type B (normodopaminergic). Br J Psychiatry [Internet]. 2014 Jul 1 [cited 2017 Mar 13];205(1):1–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24986384

123. Stone JM, Morrison PD, Pilowsky LS. Review: Glutamate and dopamine dysregulation in schizophrenia — a synthesis and selective review. J Psychopharmacol [Internet]. 2007 Jun [cited 2017 Mar 13];21(4):440–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17259207

124. Sokolov BP. Expression of NMDAR1, GluR1, GluR7, and KA1 glutamate receptor mRNAs is decreased in frontal cortex of "neuroleptic-free" schizophrenics: evidence on reversible up-regulation by typical neuroleptics. J Neurochem [Internet]. 1998 Dec [cited

2017 Mar 13];71(6):2454–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9832144

125. Olney JW, Farber NB. Glutamate receptor dysfunction and schizophrenia. Arch Gen Psychiatry [Internet]. 1995 Dec [cited 2015 Dec 30];52(12):998–1007. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7492260

126. Marsman A, van den Heuvel MP, Klomp DWJ, Kahn RS, Luijten PR, Hulshoff Pol HE. Glutamate in Schizophrenia: A Focused Review and Meta-Analysis of 1H-MRS Studies. Schizophr Bull [Internet]. 2013 Jan 1 [cited 2017 Mar 13];39(1):120–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21746807

127. Javitt DC. Glutamatergic theories of schizophrenia. Isr J Psychiatry Relat Sci [Internet]. 2010 [cited 2017 Mar 13];47(1):4–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20686195

128. Lisman JE, Coyle JT, Green RW, Javitt DC, Benes FM, Heckers S, et al. Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. Trends Neurosci [Internet]. 2008 May [cited 2017 Mar 13];31(5):234–42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18395805

129. Eggers AE. A serotonin hypothesis of schizophrenia. Med Hypotheses [Internet]. 2013 Jun [cited 2017 Mar 13];80(6):791–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23557849

130. Takahashi M, Hayashi H, Watanabe Y, Sawamura K, Fukui N, Watanabe J, et al. Diagnostic classification of schizophrenia by neural network analysis of blood-based gene expression signatures. Schizophr Res [Internet]. 2010 Jun [cited 2017 Mar 14];119(1–3):210–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20083392

131. Kumarasinghe N, Tooney PA, Schall U. Finding the needle in the haystack: A review of microarray gene expression research into schizophrenia. Aust New Zeal J Psychiatry [Internet]. 2012 Jul [cited 2017 Mar 14];46(7):598–610. Available from: http://journals.sagepub.com/doi/10.1177/0004867412442405

132. Schwarz E, Guest PC, Rahmoune H, Harris LW, Wang L, Leweke FM, et al. Identification of a biological signature for schizophrenia in serum. Mol Psychiatry [Internet]. 2012 May [cited 2015 Oct 26];17(5):494–502. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21483431

133. Nishioka Masaki , Bundo Miki, Kasai Kiyoto IK. . DNA methylation in schizophrenia: progress and challenges of epigenetic studies. Genome Med. 2012;4(12):96.

134. Zarogianni E, Moorhead TWJ, Lawrie SM. Towards the identification of imaging biomarkers in schizophrenia, using multivariate pattern classification at a single-subject level. NeuroImage Clin [Internet]. 2013 [cited 2017 Mar 14];3:279–89. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24273713

135. McGorry PD, Mihalopoulos C, Henry L, Dakis J, Jackson HJ, Flaum M, et al. Spurious precision: procedural validity of diagnostic assessment in psychotic disorders. Am J Psychiatry [Internet]. 1995 Feb [cited 2017 Jan 2];152(2):220–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7840355

136. Tandon R, Gaebel W, Barch DM, Bustillo J, Gur RE, Heckers S, et al. Definition and description of schizophrenia in the DSM-5. Schizophr Res [Internet]. 2013 Oct [cited 2017 Mar 14];150(1):3–10. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0920996413002831

137. Zhou N, Wang J, Yu Y, Shi J, Li X, Xu B, et al. Mass spectrum analysis of serum biomarker proteins from patients with schizophrenia. Biomed Chromatogr [Internet]. 2014 May 1 [cited 2017 Aug 6];28(5):654–9. Available from: http://doi.wiley.com/10.1002/bmc.3084

138. Rosenhan DL. On being sane in insane places. Science [Internet]. 1973 Jan 19 [cited 2017 Aug 6];179(4070):250–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/4683124

139. Gonzalez-Pinto A, Gutierrez M, Mosquera F, Ballesteros J, Lopez P, Ezcurra J, et al. First episode in bipolar disorder: misdiagnosis and psychotic symptoms. J Affect Disord [Internet]. 1998 Jul [cited 2017 Jan 2];50(1):41–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9716278

140. Pearlson GD. Etiologic, Phenomenologic, and Endophenotypic Overlap of Schizophrenia and Bipolar Disorder. Annu Rev Clin Psychol [Internet]. 2015 Mar 28 [cited 2017 Jan 2];11(1):251–81. Available from: http://www.annualreviews.org/doi/10.1146/annurev-clinpsy-032814-112915

141. Koutsouleris N, Meisenzahl EM, Borgwardt S, Riecher-Rössler A, Frodl T, Kambeitz J, et al. Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. Brain [Internet]. 2015 Jul [cited 2017 Jan 2];138(Pt 7):2059–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25935725

142. Csernansky JG, Schuchart EK. Relapse and rehospitalisation rates in patients with schizophrenia: effects of second generation antipsychotics. CNS Drugs [Internet]. 2002 [cited 2017 Apr 16];16(7):473–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12056922

143. Sorkin A, Weinshall D, Modai I, Peled A. Improving the Accuracy of the Diagnosis of Schizophrenia by Means of Virtual Reality. Am J Psychiatry [Internet]. 2006 Mar [cited 2017 Jan 5];163(3):512–20. Available from: http://psychiatryonline.org/doi/abs/10.1176/appi.ajp.163.3.512

144. Perkins D, Lieberman J, Gu H, Tohen M, McEvoy J, Green A, et al. Predictors of antipsychotic treatment response in patients with first-episode schizophrenia, schizoaffective and schizophreniform disorders. Br J Psychiatry [Internet]. 2004 Jul [cited 2017 Jan 8];185:18–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15231551

145. Borrebaeck CAK. Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. Nat Rev Cancer [Internet]. 2017 Feb 3 [cited 2017 Apr 16];17(3):199–204. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28154374

146. Agranoff D, Fernandez-Reyes D, Papadopoulos MC, Rojas SA, Herbster M, Loosemore A, et al. Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum. Lancet [Internet]. 2006 Sep 16 [cited 2017 Apr 16];368(9540):1012–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16980117

147. Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam C-P, et al. Blood-Based Protein Biomarkers for Diagnosis of Alzheimer Disease. Arch Neurol [Internet]. 2012 Oct 1 [cited 2017 Apr 16];69(10):1318. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22801742

148. Chan T, Gu F. Early diagnosis of sepsis using serum biomarkers. Expert Rev Mol Diagn [Internet]. 2011 Jun 9 [cited 2017 Apr 16];11(5):487–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21707457

149. World Health Organization. Cancer control: early detection: WHO guide for effective programmes [Internet]. WHO. 2007. Available from: http://www.who.int/cancer/publications/cancer_control_detection/en/

150. Soares-Weiser K, Maayan N, Bergman H, Davenport C, Kirkham AJ, Grabowski S, et al. First rank symptoms for schizophrenia. In: Soares-Weiser K, editor. Cochrane Database of Systematic Reviews [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2015 [cited 2017 Sep 26]. p. CD010653. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25879096

151. Eskandari F, Webster JI, Sternberg EM. Neural immune pathways and their connection to inflammatory diseases. Arthritis Res Ther [Internet]. 2003 [cited 2017 Apr 16];5(6):251–65. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14680500

152. Weickert CS, Weickert TW, Pillai A, Buckley PF. Biomarkers in Schizophrenia: A Brief Conceptual Consideration. Dis Markers [Internet]. 2013 [cited 2017 Jan 5];35(1):3–9. Available from: http://www.hindawi.com/journals/dm/2013/510402/

153. Kurian SM, Le-Niculescu H, Patel SD, Bertram D, Davis J, Dike C, et al. Identification of blood biomarkers for psychosis using convergent functional genomics. Mol Psychiatry [Internet]. 2011 Jan 24 [cited 2017 May 1];16(1):37–58. Available from: http://www.nature.com/doifinder/10.1038/mp.2009.117

154. Shi W, Du J, Qi Y, Liang G, Wang T, Li S, et al. Aberrant expression of serum miRNAs in schizophrenia. J Psychiatr Res [Internet]. 2012 Feb [cited 2017 May 1];46(2):198–204. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0022395611002159

155. Sun X, Lu J, Zhang L, Song H, Zhao L, Fan H, et al. Aberrant microRNA expression in peripheral plasma and mononuclear cells as specific blood-based biomarkers in schizophrenia patients. J Clin Neurosci [Internet]. 2015 Mar [cited 2017 May 1];22(3):570–4. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0967586814005967

156. Schwarz E, Izmailov R, Spain M, Barnes A, Mapes JP, Guest PC, et al. Validation of a blood-based laboratory test to aid in the confirmation of a diagnosis of schizophrenia. Biomark Insights [Internet]. 2010 May 12 [cited 2017 Jan 12];5:39–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20520744

157. Tomasik J, Schwarz E, Guest PC, Bahn S. Blood test for schizophrenia. Eur Arch Psychiatry Clin Neurosci [Internet]. 2012 Nov 26 [cited 2017 Aug 6];262(S2):79–83. Available from: http://link.springer.com/10.1007/s00406-012-0354-3

158. Niculescu AB, Levey D, Le-Niculescu H, Niculescu E, Kurian SM, Salomon D. Psychiatric blood biomarkers: avoiding jumping to premature negative or positive conclusions. Mol Psychiatry [Internet]. 2015 Mar 13 [cited 2017 May 1];20(3):286–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25582618

159. Cadenhead KS, Addington J, Cannon T, Cornblatt B, McGlashan T, Perkins D, et al. Treatment history in the psychosis prodrome: characteristics of the North American Prodrome Longitudinal Study Cohort. Early Interv Psychiatry [Internet]. 2010 Aug [cited 2017 Jan

8];4(3):220–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20712727

160. Tandon R. Definition of psychotic disorders in the DSM-5 too radical, too conservative, or just right! Schizophr Res [Internet]. 2013 Oct [cited 2017 Jan 7];150(1):1–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23978774

161. Fusar-Poli P, Carpenter WT, Woods SW, McGlashan TH. Attenuated psychosis syndrome: ready for DSM-5.1? Annu Rev Clin Psychol [Internet]. 2014 [cited 2017 Jan 5];10:155–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24471375

162. Klosterkötter J, Hellmich M, Steinmeyer EM, Schultze-Lutter F, G G, RQ B, et al. Diagnosing Schizophrenia in the Initial Prodromal Phase. Arch Gen Psychiatry [Internet]. 2001 Feb 1 [cited 2017 Jan 5];58(2):158. Available from: http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/archpsyc.58.2.158

163. Riecher-Rossler A, Gschwandtner U, Borgwardt S, Aston J, Pfluger M, Rossler W. Early detection and treatment of schizophrenia: how early? Acta Psychiatr Scand [Internet]. 2006 Feb [cited 2017 Apr 17];113(s429):73–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16445487

164. Perkins DO. Evaluating and treating the prodromal stage of schizophrenia. Curr Psychiatry Rep [Internet]. 2004 Aug [cited 2017 Apr 17];6(4):289–95. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15260945

165. Leucht S, Corves C, Arbter D, Engel RR, Li C, Davis JM. Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. Lancet (London, England) [Internet]. 2009 Jan 3 [cited 2017 Jan 8];373(9657):31–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19058842

166. Perkins D Jeffries C Addington J Bearden C Cadenhead K et. al. Towards a psychosis risk blood diagnostic for persons experiencing high-risk symptoms: preliminary results from the NAPLS project. Schizophr Bull. 2015;41(2):419–28.

167. McDermott JE, Wang J, Mitchell H, Webb-Robertson B-J, Hafen R, Ramey J, et al. Challenges in Biomarker Discovery: Combining Expert Insights with Statistical Analysis of Complex Omics Data. Expert Opin Med Diagn [Internet]. 2013 Jan [cited 2017 Jan 12];7(1):37–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23335946

168. Barla A, Jurman G, Riccadonna S, Merler S, Chierici M, Furlanello C. Machine learning

methods for predictive proteomics. Brief Bioinform [Internet]. 2007 Sep 28 [cited 2017 Apr 17];9(2):119–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18310105

169. O'Bryant SE, Xiao G, Barber R, Huebinger R, Wilhelmsen K, Edwards M, et al. A Blood-Based Screening Tool for Alzheimer's Disease That Spans Serum and Plasma: Findings from TARC and ADNI. Bush AI, editor. PLoS One [Internet]. 2011 Dec 7 [cited 2017 Apr 17];6(12):e28092. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22163278

170. van Dijk SJ, Feskens EJM, Heidema AG, Bos MB, van de Rest O, Geleijnse JM, et al. Plasma Protein Profiling Reveals Protein Clusters Related to BMI and Insulin Levels in Middle-Aged Overweight Subjects. Sorensen TIA, editor. PLoS One [Internet]. 2010 Dec 23 [cited 2017 Apr 17];5(12):e14422. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21203453

171. Zhang Z, Yu Y, Xu F, Berchuck A, van Haaften-Day C, Havrilesky LJ, et al. Combining multiple serum tumor markers improves detection of stage I epithelial ovarian cancer. Gynecol Oncol [Internet]. 2007 Dec [cited 2017 Apr 17];107(3):526–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17920110

172. Bloemberg TG, Wessels HJCT, van Dael M, Gloerich J, van den Heuvel LP, Buydens LMC, et al. Pinpointing Biomarkers in Proteomic LC/MS Data by Moving-Window Discriminant Analysis. Anal Chem [Internet]. 2011 Jul 1 [cited 2017 Apr 17];83(13):5197–206. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21557614

173. Wilkinson DJ. Bayesian methods in bioinformatics and computational systems biology. Brief Bioinform [Internet]. 2006 Dec 19 [cited 2017 Apr 17];8(2):109–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17430978

174. Harris K, Girolami M, Mischak H. Definition of Valid Proteomic Biomarkers: A Bayesian Solution. In Springer, Berlin, Heidelberg; 2009 [cited 2017 Apr 17]. p. 137–49. Available from: http://link.springer.com/10.1007/978-3-642-04031-3_13

175. Wu B, Abbott T, Fishman D, McMurray W, Mor G, Stone K, et al. Comparison of statistical methods for classification of ovarian cancer using mass spectrometry data. Bioinformatics [Internet]. 2003 Sep 1 [cited 2017 Apr 17];19(13):1636–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12967959

176. Stanley E, Delatola EI, Nkuipou-Kenfack E, Spooner W, Kolch W, Schanstra JP, et al. Comparison of different statistical approaches for urinary peptide biomarker detection in the context of coronary artery disease. BMC Bioinformatics [Internet]. 2016 Dec 6 [cited 2017 Apr

17];17(1):496. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27923348

177. Sampson DL, Parker TJ, Upton Z, Hurst CP. A comparison of methods for classifying clinical samples based on proteomics data: a case study for statistical and machine learning approaches. PLoS One [Internet]. 2011 [cited 2017 Apr 17];6(9):e24973. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21969867

178. Frantzi M, Bhat A, Latosinska A. Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. Clin Transl Med [Internet]. 2014 Mar 29 [cited 2017 Apr 17];3(1):7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24679154

179. Dakna M, Harris K, Kalousis A, Carpentier S, Kolch W, Schanstra JP, et al. Addressing the Challenge of Defining Valid Proteomic Biomarkers and Classifiers. BMC Bioinformatics [Internet]. 2010 Dec 10 [cited 2017 Apr 17];11(1):594. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21208396

180. Skates SJ, Gillette MA, LaBaer J, Carr SA, Anderson L, Liebler DC, et al. Statistical Design for Biospecimen Cohort Size in Proteomics-based Biomarker Discovery and Verification Studies. J Proteome Res [Internet]. 2013 Dec 6 [cited 2017 Apr 17];12(12):5383–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24063748

181. Surinova S, Schiess R, Hüttenhain R, Cerciello F, Wollscheid B, Aebersold R. On the Development of Plasma Protein Biomarkers. J Proteome Res [Internet]. 2011 Jan 7 [cited 2017 Jan 12];10(1):5–16. Available from: http://pubs.acs.org/doi/abs/10.1021/pr1008515

182. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol [Internet]. 2006 Aug [cited 2017 Jan 12];24(8):971–83. Available from: http://www.nature.com/doifinder/10.1038/nbt1235

183. Misek DE, Kim EH. Protein biomarkers for the early detection of breast cancer. Int J Proteomics [Internet]. 2011 [cited 2017 Aug 6];2011:343582. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22084684

184. Diamandis EP. The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem? BMC Med [Internet]. 2012 Dec 9 [cited 2017 Jan 12];10(1):87. Available from: http://bmcmedicine.biomedcentral.com/articles/10.1186/1741-7015-10-87

185. Oon SF, Pennington SR, Fitzpatrick JM, Watson RWG. Biomarker research in prostate cancer--towards utility, not futility. Nat Rev Urol [Internet]. 2011 Mar [cited 2017 Jan 12];8(3):131–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21394176

186. Hernández B, Parnell A, Pennington SR. Why have so few proteomic biomarkers "survived" validation? (Sample size and independent validation considerations). Proteomics [Internet]. 2014 Jul [cited 2017 Jan 12];14(13–14):1587–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24737731

187. Hernández B, Pennington SR, Parnell AC. Bayesian methods for proteomic biomarker development. EuPA Open Proteomics. 2015;9:54–64.

188. Drucker E, Krapfenbauer K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. EPMA J [Internet]. 2013 Feb 25 [cited 2017 Aug 7];4(1):7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23442211

189. Vittinghoff E, McCulloch CE. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression. Am J Epidemiol [Internet]. 2007 Jan 12 [cited 2017 Aug 1];165(6):710–8. Available from: https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwk052

190. Perkins DO, Jeffries CD, Addington J, Bearden CE, Cadenhead KS, Cannon TD, et al. Towards a psychosis risk blood diagnostic for persons experiencing high-risk symptoms: preliminary results from the NAPLS project. Schizophr Bull [Internet]. 2015 Mar [cited 2017 Aug 1];41(2):419–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25103207

191. Chan MK, Cooper JD, Bot M, Steiner J, Penninx BWJH, Bahn S. Identification of an Immune-Neuroendocrine Biomarker Panel for Detection of Depression: A Joint Effects Statistical Approach. Neuroendocrinology [Internet]. 2016 [cited 2017 Jan 12];103(6):693–710. Available from: http://www.karger.com/?doi=10.1159/000442208

192. Lausser L, Müssel C, Maucher M, Kestler HA. Measuring and visualizing the stability of biomarker selection techniques. Comput Stat [Internet]. 2013 Feb 5 [cited 2017 Jan 12];28(1):51–65. Available from: http://link.springer.com/10.1007/s00180-011-0284-y

193. Jelizarow M, Guillemot V, Tenenhaus A, Strimmer K, Boulesteix A-L. Over-optimism in bioinformatics: an illustration. Bioinformatics [Internet]. 2010 Aug 15 [cited 2017 Jan 12];26(16):1990–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20581402

194. Hill SM, Neve RM, Bayani N, Kuo W-L, Ziyad S, Spellman PT, et al. Integrating biological

knowledge into variable selection: an empirical Bayes approach with an application in cancer biology. BMC Bioinformatics [Internet]. 2012 May 11 [cited 2017 Apr 18];13:94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22578440

195. Hill SM, Lu Y, Molina J, Heiser LM, Spellman PT, Speed TP, et al. Bayesian inference of signaling network topology in a cancer cell line. Bioinformatics [Internet]. 2012 Nov 1 [cited 2017 Apr 18];28(21):2804–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22923301

196. Wesseling H, Guest PC, Lago SG, Bahn S, A C, AJR H, et al. Technological advances for deciphering the complexity of psychiatric disorders: merging proteomics with cell biology. Int J Neuropsychopharmacol [Internet]. 2014 Aug 14 [cited 2017 Aug 8];17(08):1327–41. Available from: https://academic.oup.com/ijnp/article-lookup/doi/10.1017/S146114571400008X

197. Herberth M, Koethe D, Cheng TMK, Krzyszton ND, Schoeffmann S, Guest PC, et al. Impaired glycolytic response in peripheral blood mononuclear cells of first-onset antipsychotic-naive schizophrenia patients. Mol Psychiatry [Internet]. 2011 Aug 29 [cited 2017 Aug 8];16(8):848–59. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20585325

198. Gladkevich A, Kauffman HF, Korf J. Lymphocytes as a neural probe: potential for studying psychiatric disorders. Prog Neuropsychopharmacol Biol Psychiatry [Internet]. 2004 May [cited 2017 Aug 8];28(3):559–76. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15093964

199. Brennand KJ, Simone A, Jou J, Gelboin-Burkhart C, Tran N, Sangar S, et al. Modelling schizophrenia using human induced pluripotent stem cells. Nature [Internet]. 2011 May 12 [cited 2017 Aug 8];473(7346):221–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21490598

200. Howes OD, McCutcheon R. Inflammation and the neural diathesis-stress hypothesis of schizophrenia: a reconceptualization. Transl Psychiatry [Internet]. 2017 Feb 7 [cited 2017 Aug 8];7(2):e1024. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28170004

201. Mondelli V, Vernon AC, Turkheimer F, Dazzan P, Pariante CM. Brain microglia in psychiatric disorders. The Lancet Psychiatry [Internet]. 2017 Jul [cited 2017 Aug 8];4(7):563–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28454915

202. Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, Holmans PA, et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature [Internet]. 2014 Jul 22 [cited 2017 Mar

12];511(7510):421–7. Available from: http://www.nature.com/doifinder/10.1038/nature13595

203. Kirkpatrick B, Miller BJ. Inflammation and schizophrenia. Schizophr Bull [Internet]. 2013 Nov [cited 2017 May 2];39(6):1174–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24072812

204. McCusker RH, Kelley KW. Immune-neural connections: how the immune system's response to infectious agents influences behavior. J Exp Biol [Internet]. 2013 Jan 1 [cited 2017 May 2];216(1):84–98. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23225871

205. Barry S, Clarke G, Scully P, Dinan T. Kynurenine pathway in psychosis: evidence of increased tryptophan degradation. J Psychopharmacol [Internet]. 2009 May 1 [cited 2017 May 2];23(3):287–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18562404

206. Nilsson LK, Linderholm KR, Engberg G, Paulson L, Blennow K, Lindström LH, et al. Elevated levels of kynurenic acid in the cerebrospinal fluid of male patients with schizophrenia. Schizophr Res [Internet]. 2005 Dec 15 [cited 2017 May 2];80(2–3):315–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16125901

207. Ravikumar A, Deepadevi K V, Arun P, Manojkumar V, Kurup PA. Tryptophan and tyrosine catabolic pattern in neuropsychiatric disorders. Neurol India [Internet]. 2000 Sep [cited 2017 May 2];48(3):231–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11025626

208. Sathyasaikumar K V., Stachowski EK, Wonodi I, Roberts RC, Rassoulpour A, McMahon RP, et al. Impaired Kynurenine Pathway Metabolism in The Prefrontal Cortex of Individuals With Schizophrenia. Schizophr Bull [Internet]. 2011 Nov 1 [cited 2017 May 2];37(6):1147–56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21036897

209. Kettenmann H, Hanisch U-K, Noda M, Verkhratsky A. Physiology of Microglia. Physiol Rev [Internet]. 2011 Apr 1 [cited 2017 May 2];91(2):461–553. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21527731

210. Nimmerjahn A, Kirchhoff F, Helmchen F. Resting Microglial Cells Are Highly Dynamic Surveillants of Brain Parenchyma in Vivo. Science (80-) [Internet]. 2005 May 27 [cited 2017 May 2];308(5726):1314–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15831717

211. Nayak D, Roth TL, McGavern DB. Microglia Development and Function. Annu Rev Immunol [Internet]. 2014 Mar 21 [cited 2017 May 2];32(1):367–402. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/24471431

212. Munn NA. Microglia dysfunction in schizophrenia: an integrative theory. Med Hypotheses [Internet]. 2000 Feb [cited 2017 May 2];54(2):198–202. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10790752

213. Monji A, Kato TA, Mizoguchi Y, Horikawa H, Seki Y, Kasai M, et al. Neuroinflammation in schizophrenia especially focused on the role of microglia. Prog Neuro-Psychopharmacology Biol Psychiatry [Internet]. 2013 Apr 5 [cited 2017 May 2];42:115–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22192886

214. Réus GZ, Fries GR, Stertz L, Badawy M, Passos IC, Barichello T, et al. The role of inflammation and microglial activation in the pathophysiology of psychiatric disorders. Neuroscience [Internet]. 2015 Aug 6 [cited 2017 May 2];300:141–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25981208

215. Bodhankar S, Lapato A, Chen Y, Vandenbark AA, Saugstad JA, Offner H. Role for microglia in sex differences after ischemic stroke: importance of M2. Metab Brain Dis [Internet]. 2015 Dec 6 [cited 2017 May 2];30(6):1515–29. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26246072

216. Han L, Cai W, Mao L, Liu J, Li P, Leak RK, et al. Rosiglitazone Promotes White Matter Integrity and Long-Term Functional Recovery After Focal Cerebral Ischemia. Stroke [Internet]. 2015 Sep [cited 2017 May 2];46(9):2628–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26243225

217. Napoli I, Neumann H. Protective effects of microglia in multiple sclerosis. Exp Neurol [Internet]. 2010 Sep [cited 2017 May 2];225(1):24–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19409897

218. Bora E, Fornito A, Radua J, Walterfang M, Seal M, Wood SJ, et al. Neuroanatomical abnormalities in schizophrenia: A multimodal voxelwise meta-analysis and meta-regression analysis. Schizophr Res [Internet]. 2011 Apr [cited 2017 May 2];127(1–3):46–57. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21300524

219. Mileaf MI, Byne W. Neuronal deficit in medial pulvinar from right but not left hemisphere in schizophrenia. Schizophr Res [Internet]. 2012 Feb [cited 2017 May 2];134(2–3):291–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22079946

220. Kubicki M, Shenton ME. Diffusion Tensor Imaging findings and their implications in schizophrenia. Curr Opin Psychiatry [Internet]. 2014 May [cited 2017 May 2];27(3):179–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24613986

221. Uranova NA, Vikhreva OV, Rachmanova VI, Orlovskaya DD. Microglial activation in white matter in schizophrenia: Findings from a postmortem electron microscopic morphometric study. Neurol Psychiatry Brain Res [Internet]. 2014 Feb [cited 2017 May 2];20(1):25. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0941950014001754

222. Orsini F, De Blasio D, Zangari R, Zanier ER, De Simoni M-G. Versatility of the complement system in neuroinflammation, neurodegeneration and brain homeostasis. Front Cell Neurosci [Internet]. 2014 [cited 2017 May 2];8:380. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25426028

223. Stephan AH, Barres BA, Stevens B. The Complement System: An Unexpected Role in Synaptic Pruning During Development and Disease. Annu Rev Neurosci [Internet]. 2012 Jul 21 [cited 2017 May 2];35(1):369–89. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22715882

224. Patterson PH. Maternal infection and immune involvement in autism. Trends Mol Med [Internet]. 2011 Jul [cited 2017 May 2];17(7):389–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21482187

225. Radewicz K, Garey LJ, Gentleman SM, Reynolds R. Increase in HLA-DR immunoreactive microglia in frontal and temporal cortex of chronic schizophrenics. J Neuropathol Exp Neurol [Internet]. 2000 Feb [cited 2017 May 2];59(2):137–50. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10749103

226. Wierzba-Bobrowicz T, Lewandowska E, Lechowicz W, Stepień T, Pasennik E. Quantitative analysis of activated microglia, ramified and damage of processes in the frontal and temporal lobes of chronic schizophrenics. Folia Neuropathol [Internet]. 2005 [cited 2017 May 2];43(2):81–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16012909

227. Fillman SG, Cloonan N, Catts VS, Miller LC, Wong J, McCrossin T, et al. Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol Psychiatry [Internet]. 2013 Feb 7 [cited 2017 May 2];18(2):206–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22869038

228. Rao JS, Kim H-W, Harry GJ, Rapoport SI, Reese EA. Increased neuroinflammatory and arachidonic acid cascade markers, and reduced synaptic proteins, in the postmortem frontal

cortex from schizophrenia patients. Schizophr Res [Internet]. 2013 Jun [cited 2017 May 2];147(1):24–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23566496

229. Paris JJ, Singh HD, Carey AN, McLaughlin JP. Exposure to HIV-1 Tat in brain impairs sensorimotor gating and activates microglia in limbic and extralimbic brain regions of male mice. Behav Brain Res [Internet]. 2015 Sep 15 [cited 2017 May 2];291:209–18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26005128

230. Ribeiro BMM, do Carmo MRS, Freire RS, Rocha NFM, Borella VCM, de Menezes AT, et al. Evidences for a progressive microglial activation and increase in iNOS expression in rats submitted to a neurodevelopmental model of schizophrenia: Reversal by clozapine. Schizophr Res [Internet]. 2013 Dec [cited 2017 May 2];151(1–3):12–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24257517

231. Van den Eynde K, Missault S, Fransen E, Raeymaekers L, Willems R, Drinkenburg W, et al. Hypolocomotive behaviour associated with increased microglia in a prenatal immune activation model with relevance to schizophrenia. Behav Brain Res [Internet]. 2014 Jan 1 [cited 2017 May 2];258:179–86. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24129217

232. Zhu F, Liu Y, Zhao J, Zheng Y. Minocycline alleviates behavioral deficits and inhibits microglial activation induced by intrahippocampal administration of Granulocyte–Macrophage Colony-Stimulating Factor in adult rats. Neuroscience [Internet]. 2014 Apr 25 [cited 2017 May 2];266:275–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24486961

233. Kobayashi K, Imagama S, Ohgomori T, Hirano K, Uchimura K, Sakamoto K, et al. Minocycline selectively inhibits M1 polarization of microglia. Cell Death Dis [Internet]. 2013 Mar 7 [cited 2017 May 2];4(3):e525. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23470532

234. Miyaoka T, Yasukawa R, Yasuda H, Hayashida M, Inagaki T, Horiguchi J. Minocycline as Adjunctive Therapy for Schizophrenia. Clin Neuropharmacol [Internet]. 2008 Sep [cited 2017 May 2];31(5):287–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18836347

235. Levkovitz Y, Mendlovich S, Riwkes S, Braw Y, Levkovitch-Verbin H, Gal G, et al. A Double-Blind, Randomized Study of Minocycline for the Treatment of Negative and Cognitive Symptoms in Early-Phase Schizophrenia. J Clin Psychiatry [Internet]. 2010 Feb 15 [cited 2017 May 2];71(02):138–49. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19895780

236. Dean OM, Data-Franco J, Giorlando F, Berk M. Minocycline. CNS Drugs [Internet]. 2012 May 1

[cited 2017 May 2];26(5):391–401. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22486246

237. Kato TA, Monji A, Mizoguchi Y, Hashioka S, Horikawa H, Seki Y, et al. Anti-Inflammatory properties of antipsychotics via microglia modulations: are antipsychotics a “fire extinguisher” in the brain of schizophrenia? Mini Rev Med Chem [Internet]. 2011 Jun [cited 2017 May 2];11(7):565–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21699487

238. Maynard TM, Sikich L, Lieberman JA, LaMantia AS. Neural development, cell-cell signaling, and the "two-hit" hypothesis of schizophrenia. Schizophr Bull [Internet]. 2001 [cited 2017 May 2];27(3):457–76. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11596847

239. Bayer TA, Falkai P, Maier W. Genetic and non-genetic vulnerability factors in schizophrenia: the basis of the "two hit hypothesis". J Psychiatr Res [Internet]. [cited 2017 May 2];33(6):543–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10628531

240. van Os J, Rutten BP, Poulton R. Gene-Environment Interactions in Schizophrenia: Review of Epidemiological Findings and Future Directions. Schizophr Bull [Internet]. 2008 Aug 20 [cited 2017 May 2];34(6):1066–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18791076

241. Feigenson KA, Kusnecov AW, Silverstein SM. Inflammation and the two-hit hypothesis of schizophrenia. Neurosci Biobehav Rev [Internet]. 2014 Jan [cited 2017 May 2];38:72–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24247023

242. Brown AS, Derkits EJ. Prenatal Infection and Schizophrenia: A Review of Epidemiologic and Translational Studies. Am J Psychiatry [Internet]. 2010 Mar [cited 2017 May 2];167(3):261–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20123911

243. Brown AS, Patterson PH. Maternal Infection and Schizophrenia: Implications for Prevention. Schizophr Bull [Internet]. 2011 Mar 1 [cited 2017 May 2];37(2):284–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21134972

244. Susser E, Hoek HW, Brown A. Neurodevelopmental disorders after prenatal famine: The story of the Dutch Famine Study. Am J Epidemiol [Internet]. 1998 Feb 1 [cited 2017 May 2];147(3):213–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9482494

245. Hoek HW, Brown AS, Susser E. The Dutch famine and schizophrenia spectrum disorders. Soc

Psychiatry Psychiatr Epidemiol [Internet]. 1998 Aug [cited 2017 May 2];33(8):373–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9708024

246. Hickie IB, Banati R, Stewart CH, Stewart CH, Lloyd AR. Are common childhood or adolescent infections risk factors for schizophrenia and other psychotic disorders? Med J Aust [Internet]. 2009 Feb 16 [cited 2017 May 2];190(4 Suppl):S17-21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19220168

247. Bland ST, Beckley JT, Young S, Tsang V, Watkins LR, Maier SF, et al. Enduring consequences of early-life infection on glial and neural cell genesis within cognitive regions of the brain. Brain Behav Immun [Internet]. 2010 Mar [cited 2017 May 2];24(3):329–38. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19782746

248. Juckel G, Manitz MP, Brüne M, Friebe A, Heneka MT, Wolf RJ. Microglial activation in a neuroinflammational animal model of schizophrenia — a pilot study. Schizophr Res [Internet]. 2011 Sep [cited 2017 May 2];131(1–3):96–100. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21752601

249. Frank MG, Thompson BM, Watkins LR, Maier SF. Glucocorticoids mediate stress-induced priming of microglial pro-inflammatory responses. Brain Behav Immun [Internet]. 2012 Feb [cited 2017 May 2];26(2):337–45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22041296

250. Beppu K, Kosai Y, Kido MA, Akimoto N, Mori Y, Kojima Y, et al. Expression, subunit composition, and function of AMPA-type glutamate receptors are changed in activated microglia; possible contribution of GluA2 (GluR-B)-deficiency under pathological conditions. Glia [Internet]. 2013 Jun [cited 2017 May 2];61(6):881–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23468421

251. Rosenberg L, Franzén B, Auer G, Lehtiö J, Forshed J. Multivariate meta-analysis of proteomics data from human prostate and colon tumours. BMC Bioinformatics [Internet]. 2010 Sep 17 [cited 2017 May 2];11(1):468. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20849579

252. Nuechterlein KH, Robbins TW, Einat H. Distinguishing Separable Domains of Cognition in Human and Animal Studies: What Separations Are Optimal for Targeting Interventions? A Summary of Recommendations From Breakout Group 2 at the Measurement and Treatment Research to Improve Cognition in Schizophrenia New Approaches Conference. Schizophr Bull

[Internet]. 2005 Sep 15 [cited 2017 May 4];31(4):870–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16150960

253. Keefe RSE, Bilder RM, Davis SM, Harvey PD, Palmer BW, Gold JM, et al. Neurocognitive Effects of Antipsychotic Medications in Patients With Chronic Schizophrenia in the CATIE Trial. Arch Gen Psychiatry [Internet]. 2007 Jun 1 [cited 2017 May 4];64(6):633. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17548746

254. Mintz J, Kopelowicz A. CUtLASS Confirms CATIE. Arch Gen Psychiatry [Internet]. 2007 Aug 1 [cited 2017 May 4];64(8):978. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17679644

255. Jones CA, Watson DJG, Fone KCF. Animal models of schizophrenia. Br J Pharmacol [Internet]. 2011 Oct [cited 2016 Apr 14];164(4):1162–94. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3229756&tool=pmcentrez&rendertype=abstract

256. Bondi C, Matthews M, Moghaddam B. Glutamatergic animal models of schizophrenia. Curr Pharm Des [Internet]. 2012 Jan [cited 2016 Feb 1];18(12):1593–604. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22280432

257. LUBY ED, COHEN BD, ROSENBAUM G, GOTTLIEB JS, KELLEY R. Study of a new schizophrenomimetic drug; sernyl. AMA Arch Neurol Psychiatry [Internet]. 1959 Mar [cited 2017 May 4];81(3):363–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/13626287

258. Stanley Burns R, Lerner SE. Perspectives. Clin Toxicol [Internet]. 1976 Jan 25 [cited 2017 May 4];9(4):477–501. Available from: http://www.ncbi.nlm.nih.gov/pubmed/789004

259. Pearlson GD. Psychiatric and medical syndromes associated with phencyclidine (PCP) abuse. Johns Hopkins Med J [Internet]. 1981 Jan [cited 2017 May 4];148(1):25–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6109037

260. Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, Bremner JD, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. Arch Gen Psychiatry [Internet]. 1994 Mar [cited 2017 May 4];51(3):199–214. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8122957

261. Javitt DC, Zukin SR. Recent advances in the phencyclidine model of schizophrenia. Am J

Psychiatry [Internet]. 1991 Oct [cited 2016 Jan 22];148(10):1301–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1654746

262. BAKKER CB, AMINI FB. Observations on the psychotomimetic effects of Sernyl. Compr Psychiatry [Internet]. 1961 Oct [cited 2017 May 4];2:269–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/13864199

263. Rainey JM, Crowder MK. Prolonged psychosis attributed to phencyclidine: report of three cases. Am J Psychiatry [Internet]. 1975 Oct [cited 2017 May 4];132(10):1076–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1166881

264. Greenberg BD, Segal DS. Acute and chronic behavioral interactions between phencyclidine (PCP) and amphetamine: evidence for a dopaminergic role in some PCP-induced behaviors. Pharmacol Biochem Behav [Internet]. 1985 Jul [cited 2017 May 4];23(1):99–105. Available from: http://www.ncbi.nlm.nih.gov/pubmed/4041046

265. Steinpreis RE, Sokolowski JD, Papanikolaou A, Salamone JD. The effects of haloperidol and clozapine on PCP- and amphetamine-induced suppression of social behavior in the rat. Pharmacol Biochem Behav [Internet]. 1994 Mar [cited 2017 May 4];47(3):579–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8208777

266. Verma A, Moghaddam B. NMDA receptor antagonists impair prefrontal cortex function as assessed via spatial delayed alternation performance in rats: modulation by dopamine. J Neurosci [Internet]. 1996 Jan [cited 2017 May 4];16(1):373–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8613804

267. Mansbach RS, Geyer MA. Effects of phencyclidine and phencyclidine biologs on sensorimotor gating in the rat. Neuropsychopharmacology [Internet]. 1989 Dec [cited 2017 May 4];2(4):299–308. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2692589

268. Buonanno A, Ozaki M, Sasner M, Yano R, Lu HS. Neuregulin-beta induces expression of an NMDA-receptor subunit. Nature [Internet]. [cited 2017 May 4];390(6661):691–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9414162

269. Mohn AR, Gainetdinov RR, Caron MG, Koller BH. Mice with reduced NMDA receptor expression display behaviors related to schizophrenia. Cell [Internet]. 1999 Aug 20 [cited 2016 Jan 18];98(4):427–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10481908

270. Nestler EJ, Hyman SE. Animal models of neuropsychiatric disorders. Nat Neurosci [Internet].

2010 Oct [cited 2014 Jul 11];13(10):1161–9. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3750731&tool=pmcentrez&rendertype=abstract

271. Powell CM, Miyakawa T. Schizophrenia-relevant behavioral testing in rodent models: a uniquely human disorder? Biol Psychiatry [Internet]. 2006 Jun 15 [cited 2015 Oct 12];59(12):1198–207. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3928106&tool=pmcentrez&rendertype=abstract

272. Kluge W, Alsaif M, Guest PC, Schwarz E, Bahn S. Translating potential biomarker candidates for schizophrenia and depression to animal models of psychiatric disorders. Expert Rev Mol Diagn [Internet]. 2011 Sep [cited 2016 Jan 30];11(7):721–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21902534

273. Sevimoglu T, Arga KY. The role of protein interaction networks in systems biomedicine. Comput Struct Biotechnol J [Internet]. 2014 Aug [cited 2017 May 4];11(18):22–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25379140

274. Smedley D, Oellrich A, Köhler S, Ruef B, Sanger Mouse Genetics Project SMG, Westerfield M, et al. PhenoDigm: analyzing curated annotations to associate animal models with human diseases. Database (Oxford) [Internet]. 2013 [cited 2017 May 4];2013:bat025. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23660285

275. Washington NL, Haendel MA, Mungall CJ, Ashburner M, Westerfield M, Lewis SE. Linking Human Diseases to Animal Models Using Ontology-Based Phenotype Annotation. Buetow KH, editor. PLoS Biol [Internet]. 2009 Nov 24 [cited 2017 May 4];7(11):e1000247. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19956802

276. Stelzhammer V, Guest PC, Rothermundt M, Sondermann C, Michael N, Schwarz E, et al. Electroconvulsive therapy exerts mainly acute molecular changes in serum of major depressive disorder patients. Eur Neuropsychopharmacol [Internet]. 2013 Oct [cited 2016 Nov 17];23(10):1199–207. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23183131

277. Arnold SE, Xie SX, Leung Y-Y, Wang L-S, Kling MA, Han X, et al. Plasma biomarkers of depressive symptoms in older adults. Transl Psychiatry [Internet]. 2012 Jan [cited 2016 Nov 17];2(1):e65. Available from: http://www.nature.com/doifinder/10.1038/tp.2011.63

278. Hall JR, Johnson LA, Barber RC, Vo HT, Winter AS, O'Bryant SE, et al. Biomarkers of basic

activities of daily living in Alzheimer's disease. J Alzheimers Dis [Internet]. 2012 [cited 2016 Nov 17];31(2):429–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22571981

279. Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam C-P, et al. Blood-based protein biomarkers for diagnosis of Alzheimer disease. Arch Neurol [Internet]. 2012 Oct [cited 2016 Nov 17];69(10):1318–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22801742

280. Lee HJ, Kim YT, Park PJ, Shin YS, Kang KN, Kim Y, et al. A novel detection method of non-small cell lung cancer using multiplexed bead-based serum biomarker profiling. J Thorac Cardiovasc Surg [Internet]. 2012 Feb [cited 2016 Nov 17];143(2):421–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22104668

281. Bertenshaw GP, Yip P, Seshaiah P, Zhao J, Chen T-H, Wiggins WS, et al. Multianalyte profiling of serum antigens and autoimmune and infectious disease molecules to identify biomarkers dysregulated in epithelial ovarian cancer. Cancer Epidemiol Biomarkers Prev [Internet]. 2008 Oct [cited 2016 Nov 17];17(10):2872–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18843033

282. Gurbel PA, Kreutz RP, Bliden KP, DiChiara J, Tantry US. Biomarker analysis by fluorokine multianalyte profiling distinguishes patients requiring intervention from patients with long-term quiescent coronary artery disease: a potential approach to identify atherosclerotic disease progression. Am Heart J [Internet]. 2008 Jan [cited 2017 Aug 1];155(1):56–61. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0002870307007077

283. Ozcan S, Cooper JD, Lago SG, Kenny D, Rustogi N, Stocki P, et al. Towards reproducible MRM based biomarker discovery using dried blood spots. Sci Rep. 2017 Mar;7:1–10.

284. Hoofnagle AN, Wener MH. The fundamental flaws of immunoassays and potential solutions using tandem mass spectrometry. J Immunol Methods [Internet]. 2009 Aug 15 [cited 2017 Sep 1];347(1–2):3–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19538965

285. Chan MK, Cooper JD, Bahn S. Commercialisation of Biomarker Tests for Mental Illnesses: Advances and Obstacles. Trends Biotechnol [Internet]. 2015 Dec [cited 2017 Sep 1];33(12):712–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26549771

286. Li G-Z, Vissers JPC, Silva JC, Golick D, Gorenstein M V., Geromanos SJ. Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. Proteomics [Internet]. 2009 Mar [cited 2017 Sep 1];9(6):1696–719. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19294629

287. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, et al. Comparison of Label-free Methods for Quantifying Human Proteins by Shotgun Proteomics. Mol Cell Proteomics [Internet]. 2005 Oct [cited 2017 Sep 1];4(10):1487–502. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15979981

288. Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. Nat Biotechnol [Internet]. 2010 Jul 9 [cited 2017 Sep 1];28(7):710–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20622845

289. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. Quantitative, Multiplexed Assays for Low Abundance Proteins in Plasma by Targeted Mass Spectrometry and Stable Isotope Dilution. Mol Cell Proteomics [Internet]. 2007 Dec [cited 2017 Sep 1];6(12):2212–29. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17939991

290. Cohen Freue G V., Borchers CH. Multiple Reaction Monitoring (MRM): Principles and Application to Coronary Artery Disease. Circ Cardiovasc Genet [Internet]. 2012 Jun 1 [cited 2017 Sep 1];5(3):378–378. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22715283

291. Wesseling H, Gottschalk MG, Bahn S. Targeted multiplexed selected reaction monitoring analysis evaluates protein expression changes of molecular risk factors for major psychiatric disorders. Int J Neuropsychopharmacol [Internet]. 2015 Jan [cited 2016 Jan 20];18(1). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4368865&tool=pmcentrez&rendertype=abstract

292. Surinova S, Hüttenhain R, Chang C-Y, Espona L, Vitek O, Aebersold R. Automated selected reaction monitoring data analysis workflow for large-scale targeted proteomic studies. Nat Protoc [Internet]. 2013 Jul 25 [cited 2017 Sep 1];8(8):1602–19. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23887179

293. Holman SW, Sims PFG, Eyers CE. The use of selected reaction monitoring in quantitative proteomics. Bioanalysis [Internet]. 2012 Jul [cited 2017 Sep 1];4(14):1763–86. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22877222

294. Ronsein GE, Pamir N, von Haller PD, Kim DS, Oda MN, Jarvik GP, et al. Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. J Proteomics [Internet]. 2015 Jan 15 [cited 2017 Sep 1];113:388–99. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/25449833

295. Valet G. Cytomics as a new potential for drug discovery. Drug Discov Today [Internet]. 2006 Sep [cited 2017 Sep 1];11(17–18):785–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16935745

296. Lago SG. Drug discovery in neuropsychiatric disorders using high-content single-cell screening of signaling network responses ex vivo. submitted. 2017;

297. Krutzik PO, Nolan GP. Fluorescent cell barcoding in flow cytometry allows high-throughput drug screening and signaling profiling. Nat Methods. 2006;3(5):361–8.

298. Nestler, E. J. and PG. Protein phosphorylation is of fundamental importance in biological regulation. Basic Neurochemistry: Molecular, Cellular and Medical Aspects, 6th edition. Lippincott-Raven, Philadelphia, PA; 1999.

299. Tang Z, Wang K, Tan W, Ma C, Li J, Liu L, et al. Real-time investigation of nucleic acids phosphorylation process using molecular beacons. Nucleic Acids Res [Internet]. 2005 Jun 16 [cited 2017 Sep 1];33(11):e97–e97. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15961728

300. Clutter MR, Heffner GC, Krutzik PO, Sachen KL, Nolan GP. Tyramide signal amplification for analysis of kinase activity by intracellular flow cytometry. Cytom Part A [Internet]. 2010 Nov [cited 2017 Sep 1];77A(11):1020–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20824632

301. Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP. Causal Protein-Signaling Networks Derived from Multiparameter Single-Cell Data. Science (80-) [Internet]. 2005 [cited 2017 Sep 1];308(5721). Available from: http://science.sciencemag.org/content/308/5721/523

302. van der Doef TF, de Witte LD, Sutterland AL, Jobse E, Yaqub M, Boellaard R, et al. In vivo (R)-[(11)C]PK11195 PET imaging of 18kDa translocator protein in recent onset psychosis. NPJ Schizophr [Internet]. 2016 [cited 2017 Sep 1];2:16031. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27602389

303. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. R Foundation for Statistical Computing, Vienna, Austria. 2013.

304. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open

software development for computational biology and bioinformatics. Genome Biol [Internet]. 2004 [cited 2016 Nov 20];5(10):R80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15461798

305. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res [Internet]. 2003 Nov [cited 2016 Nov 20];13(11):2498–504. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14597658

306. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemom Intell Lab Syst [Internet]. 1987 Aug [cited 2016 Nov 20];2(1–3):37–52. Available from: http://linkinghub.elsevier.com/retrieve/pii/0169743987800849

307. Lander ES. Array of hope. Nat Genet [Internet]. 1999 Jan [cited 2016 Nov 21];21(1 Suppl):3–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9915492

308. Fare TL, Coffey EM, Dai H, He YD, Kessler DA, Kilian KA, et al. Effects of Atmospheric Ozone on Microarray Data Quality. Anal Chem [Internet]. 2003 Sep [cited 2016 Nov 21];75(17):4672–5. Available from: http://pubs.acs.org/doi/abs/10.1021/ac034241b

309. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, et al. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proc Natl Acad Sci U S A [Internet]. 2004 Jun 22 [cited 2016 Nov 21];101(25):9309–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15184677

310. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics [Internet]. 2012 Mar 15 [cited 2016 Nov 21];28(6):882–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22257669

311. Burnham KP, Anderson DR. Multimodel Inference. Sociol Methods Res [Internet]. 2004 Nov 30 [cited 2017 Sep 1];33(2):261–304. Available from: http://journals.sagepub.com/doi/10.1177/0049124104268644

312. Hothorn T, Hornik K, van de Wiel M, Zeileis A. Implementing a class of permutation tests: The coin package. J Stat Softw. 2008;28(8):1–23.

313. Haynes W. Wilcoxon Rank Sum Test. In: Encyclopedia of Systems Biology [Internet]. New York, NY: Springer New York; 2013 [cited 2017 Sep 29]. p. 2354–5. Available from:

http://link.springer.com/10.1007/978-1-4419-9863-7_1185

314. Hastie T, Tibshirani R, Friedman JH (Jerome H. The elements of statistical learning : data mining, inference, and prediction. 745 p.

315. Seber GAF (George AF, Lee AJ. Linear regression analysis. Wiley-Interscience; 2003. 557 p.

316. Poole MA, O 'farrell PN. The assumptions of the linear regression model. [cited 2017 Sep 1]; Available from: http://people.uleth.ca/~towni0/PooleOfarrell71.pdf

317. Fox J. Regression diagnostics. Sage Publications; 1991. 92 p.

318. Menard SW. Applied logistic regression analysis [Internet]. Sage Publications; 2002 [cited 2017 Sep 1]. 111 p. Available from: https://books.google.co.uk/books?hl=en&lr=&id=EAI1QmUUsbUC&oi=fnd&pg=PR5&dq=S.+Menard,+Applied+Logistic+Regression+Analysis&ots=4UBNKWrWHR&sig=iInPAaVmHACFORW6HHpfhpZ0SMw#v=onepage&q=S. Menard%2C Applied Logistic Regression Analysis&f=false

319. Stoltzfus JC. Logistic Regression: A Brief Primer. Acad Emerg Med [Internet]. 2011 Oct [cited 2017 Sep 1];18(10):1099–104. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21996075

320. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw [Internet]. 2010 Feb 2 [cited 2017 Sep 1];33(1):1–22. Available from: http://www.jstatsoft.org/v33/i01/

321. Tibshirani R. Regression Shrinkage and Selection via the Lasso [Internet]. Vol. 58, Journal of the Royal Statistical Society. Series B (Methodological). WileyRoyal Statistical Society; 1996 [cited 2017 Sep 1]. p. 267–88. Available from: http://www.jstor.org/stable/2346178

322. Chan MK, Cooper JD, Bot M, Steiner J, Penninx BWJH, Bahn S. Identification of an Immune-Neuroendocrine Biomarker Panel for Detection of Depression: A Joint Effects Statistical Approach. Neuroendocrinology [Internet]. 2016 [cited 2017 Aug 1];103(6):693–710. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26580065

323. Cooper JD, Ozcan S, Gardner RM, Rustogi N, Wicks S, van Rees GF, et al. Schizophrenia-risk and urban birth are associated with proteomic changes in neonatal dried blood spots. Transl Psychiatry [Internet]. 2017 Dec 18 [cited 2018 May 1];7(12):1290. Available from: http://www.nature.com/articles/s41398-017-0027-0

324. IZMIRLIAN G. Application of the Random Forest Classification Algorithm to a SELDI-TOF Proteomics Study in the Setting of a Cancer Prevention Trial. Ann N Y Acad Sci [Internet]. 2004 May [cited 2017 Sep 1];1020(1):154–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15208191

325. Breiman L, Leo. Random Forests. Mach Learn [Internet]. 2001 [cited 2017 Apr 30];45(1):5–32. Available from: http://link.springer.com/10.1023/A:1010933404324

326. Kuhn M. Building Predictive Models in *R* Using the **caret** Package. J Stat Softw [Internet]. 2008 [cited 2017 Apr 30];28(5):1–26. Available from: http://www.jstatsoft.org/v28/i05/

327. Agarwal R, Ranjan P, Chipman H. A new Bayesian ensemble of trees classifier for identifying multi-class labels in satellite images. 2013 Apr 15 [cited 2017 Sep 1]; Available from: http://arxiv.org/abs/1304.4077

328. Sampson DL, Parker TJ, Upton Z, Hurst CP, Eilers P. A Comparison of Methods for Classifying Clinical Samples Based on Proteomics Data: A Case Study for Statistical and Machine Learning Approaches. Brusic V, editor. PLoS One [Internet]. 2011 Sep 28 [cited 2017 May 25];6(9):e24973. Available from: http://dx.plos.org/10.1371/journal.pone.0024973

329. Li J, Das K, Fu G, Li R, Wu R. The Bayesian lasso for genome-wide association studies. Bioinformatics [Internet]. 2011 Feb 15 [cited 2017 Apr 25];27(4):516–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21156729

330. Park T, Casella G. The Bayesian Lasso. [cited 2017 Sep 1]; Available from: http://www.stat.ufl.edu/archived/casella/Papers/Lasso.pdf

331. Gramacy RB, Polson NG. Simulation-based Regularized Logistic Regression. [cited 2017 Sep 1]; Available from: https://arxiv.org/pdf/1005.3430.pdf

332. Chipman HA, George EI, McCulloch RE. BART: Bayesian additive regression trees. Ann Appl Stat [Internet]. 2010 Mar [cited 2017 Sep 1];4(1):266–98. Available from: http://projecteuclid.org/euclid.aoas/1273584455

333. Bleich J, Kapelner A, George EI, Jensen ST. Variable selection for BART: An application to gene regulation. 2013 Oct 17 [cited 2017 May 1]; Available from: http://arxiv.org/abs/1310.4887

334. Adam Kapelner A, Bleich Maintainer Adam Kapelner J. Package “bartMachine.” 2016 [cited 2017 Sep 1]; Available from: https://cran.r-project.org/web/packages/bartMachine/bartMachine.pdf

335. Pinheiro JC, Bates DM. lme and nlme: Mixed-effects Methods and Classes for S and S-plus. [cited 2017 Sep 1]; Available from: http://sun.cwru.edu/~jiayang/nlme.pdf

336. Pinheiro J, Bates D, DebRoy S SD. nlme: Linear and Nonlinear Mixed Effects Models. R Core Team. 2017;

337. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett [Internet]. 2006 Jun [cited 2017 Sep 1];27(8):861–74. Available from: http://linkinghub.elsevier.com/retrieve/pii/S016786550500303X

338. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biom J [Internet]. 2005 Aug [cited 2017 Apr 30];47(4):458–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16161804

339. Lüdemann L, Grieger W, Wurm R, Wust P, Zimmer C. Glioma assessment using quantitative blood volume maps generated by T1-weighted dynamic contrast-enhanced magnetic resonance imaging: a receiver operating characteristic study. Acta Radiol [Internet]. 2006 Apr [cited 2017 Apr 30];47(3):303–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16613313

340. Obuchowski NA. Receiver Operating Characteristic Curves and Their Use in Radiology. Radiology [Internet]. 2003 Oct [cited 2017 Apr 30];229(1):3–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14519861

341. Metz CE. Basic principles of ROC analysis. Semin Nucl Med [Internet]. 1978 Oct [cited 2017 Apr 30];8(4):283–98. Available from: http://www.ncbi.nlm.nih.gov/pubmed/112681

342. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics [Internet]. 2005 Oct 15 [cited 2017 Apr 30];21(20):3940–1. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bti623

343. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics [Internet]. 2004 Jan 1 [cited 2017 Sep 1];20(1):93–9. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btg382

344. Jung K, Dihazi H, Bibi A, Dihazi GH, Beissbarth T. Adaption of the global test idea to proteomics data with missing values. Bioinformatics [Internet]. 2014 May 15 [cited 2017 Sep 1];30(10):1424–30. Available from: https://academic.oup.com/bioinformatics/article-

lookup/doi/10.1093/bioinformatics/btu062

345. Liu Q, Dinu I, Adewale AJ, Potter JD, Yasui Y. Comparative evaluation of gene-set analysis methods. BMC Bioinformatics [Internet]. 2007 Nov 7 [cited 2017 Sep 29];8(1):431. Available from: http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-431

346. Houwing-Duistermaat JJ, Derkx BH, Rosendaal FR, van Houwelingen HC. Testing familial aggregation. Biometrics [Internet]. 1995 Dec [cited 2017 Sep 29];51(4):1292–301. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8589223

347. le Cessie S, van Houwelingen HC. Testing the fit of a regression model via score tests in random effects models. Biometrics [Internet]. 1995 Jun [cited 2017 Sep 29];51(2):600–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7662848

348. Goeman J, Oosting J, Finos L, Solari A. The Global Test and the globaltest R package. 2017 [cited 2017 Sep 1]; Available from: https://www.bioconductor.org/packages/devel/bioc/vignettes/globaltest/inst/doc/GlobalTest.pdf

349. Hermjakob H, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, et al. IntAct: an open source molecular interaction database. Nucleic Acids Res [Internet]. 2004;32(Database issue):D452-5. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=308786&tool=pmcentrez&rendertype=abstract

350. Zanzoni A, Montecchi-Palazzi L, Quondam M, Ausiello G, Helmer-Citterich M, Cesareni G. MINT: a Molecular INTeraction database. FEBS Lett [Internet]. 2002;513(1):135–40. Available from: http://www.sciencedirect.com/science/article/pii/S0014579301032938

351. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, et al. UniProt: the Universal Protein knowledgebase. Nucleic Acids Res. 2004;32(Database issue):D115–9.

352. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics [Internet]. 2009;25(8):1091–3. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2666812&tool=pmcentrez&rendertype=abstract

353. Cox D, Chan MK, Bahn S. The potential of immune biomarkers to advance personalized

medicine approaches for schizophrenia. J Nerv Ment Dis [Internet]. 2015 May [cited 2015 Oct 6];203(5):393–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25919386

354. He Q-Y, Chiu J-F. Proteomics in biomarker discovery and drug development. J Cell Biochem [Internet]. 2003 Aug 1 [cited 2017 Jan 12];89(5):868–86. Available from: http://doi.wiley.com/10.1002/jcb.10576

355. Schwarz E, Izmailov R, Spain M, Barnes A, Mapes JP, Guest PC, et al. Validation of a blood-based laboratory test to aid in the confirmation of a diagnosis of schizophrenia. Biomark Insights [Internet]. 2010 Jan [cited 2015 Oct 26];5:39–47. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2879227&tool=pmcentrez&rendertype=abstract

356. Levin Y, Wang L, Schwarz E, Koethe D, Leweke FM, Bahn S. Global proteomic profiling reveals altered proteomic signature in schizophrenia serum. Mol Psychiatry [Internet]. 2010 Nov 23 [cited 2017 Mar 9];15(11):1088–100. Available from: http://www.nature.com/doifinder/10.1038/mp.2009.54

357. Whiteaker JR, Zhao L, Zhang HY, Feng L-C, Piening BD, Anderson L, et al. Antibody-based enrichment of peptides on magnetic beads for mass-spectrometry-based quantification of serum biomarkers. Anal Biochem [Internet]. 2007 Mar 1 [cited 2017 Aug 1];362(1):44–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17241609

358. Delaleu N, Immervoll H, Cornelius J, Jonsson R. Biomarker profiles in serum and saliva of experimental Sjögren's syndrome: associations with specific autoimmune manifestations. Arthritis Res Ther [Internet]. 2008 [cited 2017 Aug 1];10(1):R22. Available from: http://arthritis-research.biomedcentral.com/articles/10.1186/ar2375

359. Gurbel PA, Kreutz RP, Bliden KP, DiChiara J, Tantry US. Biomarker analysis by fluorokine multianalyte profiling distinguishes patients requiring intervention from patients with long-term quiescent coronary artery disease: a potential approach to identify atherosclerotic disease progression. Am Heart J [Internet]. 2008 Jan [cited 2016 Nov 17];155(1):56–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18082490

360. Bot M, Chan MK, Jansen R, Lamers F, Vogelzangs N, Steiner J, et al. Serum proteomic profiling of major depressive disorder. Transl Psychiatry [Internet]. 2015 Jul 14 [cited 2017 Aug 1];5(7):e599. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26171980

361. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why

small sample size undermines the reliability of neuroscience. Nat Rev Neurosci [Internet]. 2013 Apr 10 [cited 2017 Aug 1];14(5):365–76. Available from: http://www.nature.com/doifinder/10.1038/nrn3475

362. Haenisch F, Cooper JD, Reif A, Kittel-Schneider S, Steiner J, Leweke FM, et al. Towards a blood-based diagnostic panel for bipolar disorder. Brain Behav Immun [Internet]. 2016 Feb [cited 2017 Aug 1];52:49–57. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0889159115300271

363. Grissa D, Pétéra M, Brandolini M, Napoli A, Comte B, Pujos-Guillot E. Feature Selection Methods for Early Predictive Biomarker Discovery Using Untargeted Metabolomic Data. Front Mol Biosci [Internet]. 2016 [cited 2017 Apr 25];3:30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27458587

364. Scott IM, Lin W, Liakata M, Wood JE, Vermeer CP, Allaway D, et al. Merits of random forests emerge in evaluation of chemometric classifiers by external validation. Anal Chim Acta [Internet]. 2013 Nov 1 [cited 2017 Apr 25];801:22–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24139571

365. McDermott JE, Wang J, Mitchell H, Webb-Robertson B-J, Hafen R, Ramey J, et al. Challenges in Biomarker Discovery: Combining Expert Insights with Statistical Analysis of Complex Omics Data. Expert Opin Med Diagn [Internet]. 2013 Jan [cited 2017 Apr 17];7(1):37–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23335946

366. de Witte L, Tomasik J, Schwarz E, Guest PC, Rahmoune H, Kahn RS, et al. Cytokine alterations in first-episode schizophrenia patients before and after antipsychotic treatment. Schizophr Res [Internet]. 2014 Apr [cited 2017 Apr 25];154(1–3):23–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24582037

367. Magaud E, Morvan Y, Rampazzo A, Alexandre C, Willard D, Gaillard R, et al. Subjects at Ultra High Risk for psychosis have “heterogeneous” intellectual functioning profile: a multiple-case study. Schizophr Res [Internet]. 2014 Feb [cited 2017 Jan 13];152(2–3):415–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24365404

368. Swan AL, Mobasheri A, Allaway D, Liddell S, Bacardit J. Application of machine learning to proteomics data: classification and biomarker identification in postgenomics biology. OMICS [Internet]. 2013 Dec [cited 2017 May 25];17(12):595–610. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24116388

369. Chen T, Cao Y, Zhang Y, Liu J, Bao Y, Wang C, et al. Random Forest in Clinical Metabolomics for Phenotypic Discrimination and Biomarker Selection. Evidence-Based Complement Altern Med [Internet]. 2013 [cited 2017 Apr 25];2013:1–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23573122

370. Hapfelmeier A, Hothorn T, Ulm K, Strobl C. A new variable importance measure for random forests with missing data. Stat Comput [Internet]. 2014 Jan 28 [cited 2017 Apr 25];24(1):21–34. Available from: http://link.springer.com/10.1007/s11222-012-9349-1

371. Menze BH, Kelm BM, Masuch R, Himmelreich U, Bachert P, Petrich W, et al. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. BMC Bioinformatics [Internet]. 2009 Jul 10 [cited 2017 Apr 25];10(1):213. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19591666

372. Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML, et al. A tutorial review: Metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding. Anal Chim Acta [Internet]. 2015 Jun 16 [cited 2017 Apr 25];879:10–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26002472

373. Gromski PS, Xu Y, Correa E, Ellis DI, Turner ML, Goodacre R. A comparative investigation of modern feature selection and classification approaches for the analysis of mass spectrometry data. Anal Chim Acta [Internet]. 2014 Jun 4 [cited 2017 Apr 25];829:1–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24856395

374. Bishop CM, Tipping ME. Bayesian Regression and Classification. NATO Sci Ser III Comput Syst Sci [Internet]. [cited 2017 May 25];190. Available from: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/bishop-nato-bayes.pdf

375. Li J, Das K, Fu G, Li R, Wu R. The Bayesian lasso for genome-wide association studies. Bioinformatics [Internet]. 2011 Feb 15 [cited 2017 Apr 30];27(4):516–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21156729

376. de los Campos G, Naya H, Gianola D, Crossa J, Legarra A, Manfredi E, et al. Predicting Quantitative Traits With Regression Models for Dense Molecular Markers and Pedigree. Genetics [Internet]. 2009 May 1 [cited 2017 Apr 25];182(1):375–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19293140

377. Cai X, Huang A, Xu S. Fast empirical Bayesian LASSO for multiple quantitative trait locus mapping. BMC Bioinformatics [Internet]. 2011 May 26 [cited 2017 Apr 25];12(1):211.

Available from: http://www.ncbi.nlm.nih.gov/pubmed/21615941

378. Bonato V, Baladandayuthapani V, Broom BM, Sulman EP, Aldape KD, Do K-A. Bayesian ensemble methods for survival prediction in gene expression data. Bioinformatics [Internet]. 2011 Feb 1 [cited 2017 Apr 25];27(3):359–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21148161

379. WHITTINGHAM MJ, STEPHENS PA, BRADBURY RB, FRECKLETON RP. Why do we still use stepwise modelling in ecology and behaviour? J Anim Ecol [Internet]. 2006 Sep [cited 2017 May 25];75(5):1182–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16922854

380. Xi B, Gu H, Baniasadi H, Raftery D. Statistical analysis and modeling of mass spectrometry-based metabolomics data. Methods Mol Biol [Internet]. 2014 [cited 2017 May 25];1198:333–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25270940

381. El Khouli RH, Macura KJ, Barker PB, Habba MR, Jacobs MA, Bluemke DA, et al. Relationship of temporal resolution to diagnostic performance for dynamic contrast enhanced MRI of the breast. J Magn Reson Imaging [Internet]. 2009 Nov [cited 2017 Apr 30];30(5):999–1004. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19856413

382. Guyon+ I, Weston+ J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. [cited 2017 Apr 28]; Available from: http://www.thespermwhale.com/jaseweston/GENESEL.PDF

383. Kursa MB. Robustness of Random Forest-based gene selection methods. BMC Bioinformatics [Internet]. 2014 Jan 13 [cited 2017 Apr 30];15(1):8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24410865

384. Ma L, Fu T, Blaschke T, Li M, Tiede D, Zhou Z, et al. Evaluation of Feature Selection Methods for Object-Based Land Cover Mapping of Unmanned Aerial Vehicle Imagery Using Random Forest and Support Vector Machine Classifiers. ISPRS Int J Geo-Information [Internet]. 2017 Feb 18 [cited 2017 Apr 30];6(2):51. Available from: http://www.mdpi.com/2220-9964/6/2/51

385. Domenici E, Willé DR, Tozzi F, Prokopenko I, Miller S, McKeown A, et al. Plasma protein biomarkers for depression and schizophrenia by multi analyte profiling of case-control collections. Domschke K, editor. PLoS One [Internet]. 2010 Feb 11 [cited 2017 Mar 9];5(2):e9166. Available from: http://dx.plos.org/10.1371/journal.pone.0009166

386. Nieuwenhuis M, van Haren NEM, Hulshoff Pol HE, Cahn W, Kahn RS, Schnack HG.

Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. Neuroimage [Internet]. 2012 Jul 2 [cited 2017 Sep 7];61(3):606–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22507227

387. Boja ES, Fehniger TE, Baker MS, Marko-Varga G, Rodriguez H. Analytical validation considerations of multiplex mass-spectrometry-based proteomic platforms for measuring protein biomarkers. J Proteome Res [Internet]. 2014 Dec 5 [cited 2017 Sep 28];13(12):5325–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25171765

388. Kuschner KW, Malyarenko DI, Cooke WE, Cazares LH, Semmes OJ, Tracy ER. A Bayesian network approach to feature selection in mass spectrometry data. BMC Bioinformatics [Internet]. 2010 Apr 8 [cited 2017 Sep 28];11:177. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20377906

389. Sakurai K, Migita O, Toru M, Arinami T. An association between a missense polymorphism in the close homologue of L1 (CHL1, CALL) gene and schizophrenia. Mol Psychiatry [Internet]. 2002 Apr 24 [cited 2017 Sep 28];7(4):412–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11986985

390. Katic J, Loers G, Kleene R, Karl N, Schmidt C, Buck F, et al. Interaction of the Cell Adhesion Molecule CHL1 with Vitronectin, Integrins, and the Plasminogen Activator Inhibitor-2 Promotes CHL1-Induced Neurite Outgrowth and Neuronal Migration. J Neurosci [Internet]. 2014 Oct 29 [cited 2017 Sep 28];34(44):14606–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25355214

391. Jiang L, Lindpaintner K, Li H-F, Gu N-F, Langen H, He L, et al. Proteomic analysis of the cerebrospinal fluid of patients with schizophrenia. Amino Acids [Internet]. 2003 Jul 17 [cited 2017 Aug 12];25(1):49–57. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12836058

392. Elliott DA, Weickert CS, Garner B. Apolipoproteins in the brain: implications for neurological and psychiatric disorders. Clin Lipidol [Internet]. 2010 Aug 1 [cited 2017 Aug 30];51(4):555–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21423873

393. Wan C, La Y, Zhu H, Yang Y, Jiang L, Chen Y, et al. Abnormal changes of plasma acute phase proteins in schizophrenia and the relation between schizophrenia and haptoglobin (Hp) gene. Amino Acids [Internet]. 2007 Jan 10 [cited 2017 Aug 30];32(1):101–8. Available from: http://link.springer.com/10.1007/s00726-005-0292-8

394. Carrizo E, Fernández V, Quintero J, Connell L, Rodríguez Z, Mosquera M, et al. Coagulation

and inflammation markers during atypical or typical antipsychotic treatment in schizophrenia patients and drug-free first-degree relatives. Schizophr Res [Internet]. 2008 Aug [cited 2017 Aug 11];103(1–3):83–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18436434

395. Mayilyan KR, Weinberger DR, Sim RB. The complement system in schizophrenia. Drug News Perspect [Internet]. 2008 May [cited 2017 Aug 30];21(4):200–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18560619

396. Maes M, Delanghe J, Bocchio Chiavetto L, Bignotti S, Tura GB, Pioli R, et al. Haptoglobin polymorphism and schizophrenia: genetic variation on chromosome 16. Psychiatry Res [Internet]. 2001 Oct 10 [cited 2017 Aug 30];104(1):1–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11600184

397. Chow TJ, Chern Loh H. Altered Levels of Serum Haptoglobin and Apo A-I in Schizophrenia (Perubahan Paras Haptoglobin dan Apo A-I Serum dalam Skizofrenia). Sains Malaysiana [Internet]. 2011 [cited 2017 Aug 30];40(11):1319–23. Available from: http://www.ukm.my/jsm/pdf_files/SM-PDF-40-11-2011/18 Tze Jen Chow.pdf

398. Ohi K, Shimada T, Nitta Y, Kihara H, Okubo H, Uehara T, et al. Schizophrenia risk variants in ITIH4 and CALN1 regulate gene expression in the dorsolateral prefrontal cortex. Psychiatr Genet [Internet]. 2016 Jun [cited 2017 Aug 30];26(3):142–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26991396

399. Chen P-Y, Huang M-C, Chiu C-C, Liu H-C, Lu M-L, Chen C-H. Association of plasma retinol-binding protein-4, adiponectin, and high molecular weight adiponectin with metabolic adversities in patients with schizophrenia. Prog Neuro-Psychopharmacology Biol Psychiatry [Internet]. 2011 Dec 1 [cited 2017 Aug 30];35(8):1927–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21840365

400. Maes M, Delange J, Ranjan R, Meltzer HY, Desnyder R, Cooremans W, et al. Acute phase proteins in schizophrenia, mania and major depression: modulation by psychotropic drugs. Psychiatry Res [Internet]. 1997 Jan 15 [cited 2017 Sep 19];66(1):1–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9061799

401. Bloomfield PS, Selvaraj S, Veronese M, Rizzo G, Bertoldo A, Owen DR, et al. Microglial Activity in People at Ultra High Risk of Psychosis and in Schizophrenia: An [^{11} C]PBR28 PET Brain Imaging Study. Am J Psychiatry [Internet]. 2016 Jan [cited 2017 Aug 9];173(1):44–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26472628

402. Trépanier MO, Hopperton KE, Mizrahi R, Mechawar N, Bazinet RP. Postmortem evidence of cerebral inflammation in schizophrenia: a systematic review. Mol Psychiatry [Internet]. 2016 Aug 7 [cited 2017 Aug 9];21(8):1009–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27271499

403. Fillman SG, Weickert TW, Lenroot RK, Catts S V, Bruggemann JM, Catts VS, et al. Elevated peripheral cytokines characterize a subgroup of people with schizophrenia displaying poor verbal fluency and reduced Broca's area volume. Mol Psychiatry [Internet]. 2016 Aug 21 [cited 2017 Aug 9];21(8):1090–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26194183

404. Nakagawa Y, Chiba K. Role of Microglial M1/M2 Polarization in Relapse and Remission of Psychiatric Disorders and Diseases. Pharmaceuticals [Internet]. 2014 Nov 25 [cited 2017 Aug 9];7(12):1028–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25429645

405. Dean B. Understanding the role of inflammatory-related pathways in the pathophysiology and treatment of psychiatric disorders: evidence from human peripheral studies and CNS studies. Int J Neuropsychopharmacol [Internet]. 2011 Aug [cited 2017 Aug 9];14(7):997–1012. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21156092

406. Hu X, Leak RK, Shi Y, Suenaga J, Gao Y, Zheng P, et al. Microglial and macrophage polarization—new prospects for brain repair. Nat Rev Neurol [Internet]. 2015 Nov 11 [cited 2017 May 2];11(1):56–64. Available from: http://www.nature.com/doifinder/10.1038/nrneurol.2014.207

407. Ransohoff RM. A polarizing question: do M1 and M2 microglia exist? Nat Neurosci [Internet]. 2016 Jul 26 [cited 2017 May 2];19(8):987–91. Available from: http://www.nature.com/doifinder/10.1038/nn.4338

408. Ransohoff RM, Perry VH. Microglial Physiology: Unique Stimuli, Specialized Responses. Annu Rev Immunol [Internet]. 2009 Apr [cited 2017 Aug 9];27(1):119–45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19302036

409. Cherry JJD, Olschowka JJA, O'Banion M, Rio-Hortega P Del, McGeer P, McGeer E, et al. Neuroinflammation and M2 microglia: the good, the bad, and the inflamed. J Neuroinflammation. 2014;11(98):1–15.

410. Nakagawa Y, Chiba K. Role of microglial m1/m2 polarization in relapse and remission of psychiatric disorders and diseases. Pharmaceuticals (Basel) [Internet]. 2014 Jan [cited 2015

Nov 20];7(12):1028–48. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4276905&tool=pmcentrez&rendertype=abstract

411. Hollander Z, Lazárová M, Lam KKY, Ignaszewski A, Oudit GY, Dyck JRB, et al. Proteomic biomarkers of recovered heart function. Eur J Heart Fail [Internet]. 2014 May 1 [cited 2017 Sep 8];16(5):551–9. Available from: http://doi.wiley.com/10.1002/ejhf.65

412. Curatolo P, Moavero R, de Vries PJ. Neurological and neuropsychiatric aspects of tuberous sclerosis complex. Lancet Neurol [Internet]. 2015 Jul [cited 2017 Aug 9];14(7):733–45. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1474442215000691

413. Hu X, Leak RK, Shi Y, Suenaga J, Gao Y, Zheng P, et al. Microglial and macrophage polarization—new prospects for brain repair. Nat Rev Neurol [Internet]. 2014 Nov 11 [cited 2017 Aug 9];11(1):56–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25385337

414. Knijnenburg TA, Wessels LFA, Reinders MJT, Shmulevich I. Fewer permutations, more accurate P-values. Bioinformatics [Internet]. 2009 Jun 15 [cited 2018 Apr 30];25(12):i161–8. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp211

415. Yu K, Liang F, Ciampa J, Chatterjee N. Efficient p-value evaluation for resampling-based tests. Biostatistics [Internet]. 2011 Jul [cited 2018 Apr 30];12(3):582–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21209154

416. Benjamini Y, Hochberg Y. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B [Internet]. 1995;57(1):289–300. Available from: http://www.stat.purdue.edu/~doerge/BIOINFORM.D/FALL06/Benjamini and Y FDR.pdf%5Cnhttp://engr.case.edu/ray_soumya/mlrg/controlling_fdr_benjamini95.pdf

417. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? Schizophr Res [Internet]. 2005 Nov 15 [cited 2015 Sep 26];79(2–3):231–8. Available from: http://www.sciencedirect.com/science/article/pii/S0920996405001611

418. Qin H, Yeh W-I, De Sarno P, Holdbrooks AT, Liu Y, Muldowney MT, et al. Signal transducer and activator of transcription-3/suppressor of cytokine signaling-3 (STAT3/SOCS3) axis in myeloid cells regulates neuroinflammation. Proc Natl Acad Sci [Internet]. 2012 Mar 27 [cited 2017 Aug 11];109(13):5004–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22411837

419. Li D, Wang C, Yao Y, Chen L, Liu G, Zhang R, et al. mTORC1 pathway disruption ameliorates brain inflammation following stroke via a shift in microglia phenotype from M1 type to M2 type. FASEB J [Internet]. 2016 Oct 24 [cited 2017 Aug 11];30(10):3388–99. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27342766

420. Dello Russo C, Lisi L, Tringali G, Navarra P. Involvement of mTOR kinase in cytokine-dependent microglial activation and cell proliferation. Biochem Pharmacol [Internet]. 2009 Nov 1 [cited 2017 Aug 11];78(9):1242–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19576187

421. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. Nature [Internet]. 2016 Jan 27 [cited 2017 Aug 8];530(7589):177–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26814963

422. Yeh FL, Wang Y, Tom I, Gonzalez LC, Sheng M. TREM2 Binds to Apolipoproteins, Including APOE and CLU/APOJ, and Thereby Facilitates Uptake of Amyloid-Beta by Microglia. Neuron [Internet]. 2016 Jul 20 [cited 2017 Aug 11];91(2):328–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27477018

423. Borda JT, Alvarez X, Mohan M, Hasegawa A, Bernardino A, Jean S, et al. CD163, a marker of perivascular macrophages, is up-regulated by microglia in simian immunodeficiency virus encephalitis after haptoglobin-hemoglobin complex stimulation and is suggestive of breakdown of the blood-brain barrier. Am J Pathol [Internet]. 2008 Mar [cited 2017 Aug 11];172(3):725–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18276779

424. Doens D, Fernández PL. Microglia receptors and their implications in the response to amyloid β for Alzheimer's disease pathogenesis. J Neuroinflammation [Internet]. 2014 Mar 13 [cited 2017 Aug 11];11(1):48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24625061

425. Knöchel C, Kniep J, Cooper JD, Stäblein M, Wenzler S, Sarlon J, et al. Altered apolipoprotein C expression in association with cognition impairments and hippocampus volume in schizophrenia and bipolar disorder. Eur Arch Psychiatry Clin Neurosci [Internet]. 2017 Apr 22 [cited 2017 Aug 11];267(3):199–212. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27549216

426. Saetre P, Emilsson L, Axelsson E, Kreuger J, Lindholm E, Jazin E. Inflammation-related genes up-regulated in schizophrenia brains. BMC Psychiatry [Internet]. 2007 Dec 6 [cited 2017 Aug

11];7(1):46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17822540

427. Tomasik J, Rahmoune H, Guest PC, Bahn S. Neuroimmune biomarkers in schizophrenia. Schizophr Res [Internet]. 2014 Aug 11 [cited 2016 Feb 1]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/25124519

428. Montesinos-Rongen M, Brunn A, Bentink S, Basso K, Lim WK, Klapper W, et al. Gene expression profiling suggests primary central nervous system lymphomas to be derived from a late germinal center B cell. Leukemia [Internet]. 2008 Feb 8 [cited 2017 Aug 13];22(2):400–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17989719

429. Law RH, Zhang Q, McGowan S, Buckle AM, Silverman GA, Wong W, et al. An overview of the serpin superfamily. Genome Biol [Internet]. 2006 [cited 2017 Aug 13];7(5):216. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16737556

430. Yifeng Yang †,‡,⊥, Chunling Wan †,‡,⊥, Huafang Li §, Hui Zhu †,‡, Yujuan La †,‡, Zhengrui Xi †,‡, et al. Altered Levels of Acute Phase Proteins in the Plasma of Patients with Schizophrenia. 2006 [cited 2017 Aug 11]; Available from: http://pubs.acs.org/doi/abs/10.1021/ac051916x

431. Cudaback E, Li X, Yang Y, Yoo T, Montine KS, Craft S, et al. Apolipoprotein C-I is an APOE genotype-dependent suppressor of glial activation. J Neuroinflammation [Internet]. 2012 Aug 10 [cited 2017 Aug 13];9:192. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22883744

432. Rudduck C, Franzén G, Fröhlander N, Lindström L. Haptoglobin and transferrin types in schizophrenia. Hum Hered [Internet]. 1985 [cited 2017 Aug 12];35(2):65–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3857216

433. Yee JY, Nurjono M, Ng WY, Teo SR, Lee T-S, Lee J. Peripheral blood gene expression of acute phase proteins in people with first episode psychosis. Brain Behav Immun [Internet]. 2017 Oct [cited 2017 Aug 12];65:337–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28627459

434. Zhao X, Song S, Sun G, Strong R, Zhang J, Grotta JC, et al. Neuroprotective Role of Haptoglobin after Intracerebral Hemorrhage. J Neurosci [Internet]. 2009 Dec 16 [cited 2017 Aug 13];29(50):15819–27. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20016097

435. Skurkovich S V, Aleksandrovsky YA, Chekhonin VP, Ryabukhin IA, Chakhava KO, Skurkovich B. Improvement in negative symptoms of schizophrenia with antibodies to tumor necrosis

factor-alpha and to interferon-gamma: a case report. J Clin Psychiatry [Internet]. 2003 Jun [cited 2017 Aug 13];64(6):734–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12823095

436. Kim Y-K, Myint A-M, Lee B-H, Han C-S, Lee H-J, Kim D-J, et al. Th1, Th2 and Th3 cytokine alteration in schizophrenia. Prog Neuro-Psychopharmacology Biol Psychiatry [Internet]. 2004 Nov [cited 2017 Aug 13];28(7):1129–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15610925

437. Kato T, Monji A, Hashioka S, Kanba S. Risperidone significantly inhibits interferon-γ-induced microglial activation in vitro. Schizophr Res [Internet]. 2007 May [cited 2017 Aug 13];92(1–3):108–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17363222

438. Li Y, Zhou K, Zhang Z, Sun L, Yang J, Zhang M, et al. Label-free quantitative proteomic analysis reveals dysfunction of complement pathway in peripheral blood of schizophrenia patients: evidence for the immune hypothesis of schizophrenia. Mol Biosyst [Internet]. 2012 Oct [cited 2017 Apr 30];8(10):2664. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22797129

439. Athanas KM, Mauney SL, Woo T-UW. Increased extracellular clusterin in the prefrontal cortex in schizophrenia. Schizophr Res [Internet]. 2015 Dec [cited 2017 Aug 13];169(1–3):381–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26482819

440. Chen X, Rivard L, Naqvi S, Nakada S, Padbury JF, Sanchez-Esteban J, et al. Expression and localization of Inter-alpha Inhibitors in rodent brain. Neuroscience [Internet]. 2016 Jun 2 [cited 2017 Aug 13];324:69–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26964679

441. Paglinawan R, Malipiero U, Schlapbach R, Frei K, Reith W, Fontana A. TGF? directs gene expression of activated microglia to an anti-inflammatory phenotype strongly focusing on chemokine genes and cell migratory genes. Glia [Internet]. 2003 Dec [cited 2017 Aug 13];44(3):219–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14603463

442. Butovsky O, Jedrychowski MP, Moore CS, Cialic R, Lanser AJ, Gabriely G, et al. Identification of a unique TGF-β–dependent molecular and functional signature in microglia. Nat Neurosci [Internet]. 2013 Dec 8 [cited 2017 Aug 13];17(1):131–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24316888

443. You W, Wang Z, Li H, Shen H, Xu X, Jia G, et al. Inhibition of mammalian target of rapamycin attenuates early brain injury through modulating microglial polarization after experimental

subarachnoid hemorrhage in rats. J Neurol Sci. 2016 Aug;367:224–31.

444. Felger JC, Lotrich FE. Inflammatory cytokines in depression: neurobiological mechanisms and therapeutic implications. Neuroscience [Internet]. 2013 Aug 29 [cited 2018 May 5];246:199–229. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23644052

445. Song Q, Xie D, Pan S, Xu W. Rapamycin protects neurons from brain contusion-induced inflammatory reaction via modulation of microglial activation. Mol Med Rep. 2015;12(5):7203–10.

446. Srivastava IN, Shperdheja J, Baybis M, Ferguson T, Crino PB. mTOR pathway inhibition prevents neuroinflammation and neuronal death in a mouse model of cerebral palsy. Neurobiol Dis. 2016;85:144–54.

447. Xie L, Sun F, Wang J, Mao X, Xie L, Yang S-H, et al. mTOR signaling inhibition modulates macrophage/microglia-mediated neuroinflammation and secondary injury via regulatory T cells after focal ischemia. J Immunol. 2014 Jun;192(12):6009–19.

448. Verheijden S, Beckers L, Casazza A, Butovsky O, Mazzone M, Baes M. Identification of a chronic non-neurodegenerative microglia activation state in a mouse model of peroxisomal ??-oxidation deficiency. Glia. 2015;63(9):1606–20.

449. Harris H, Rubinsztein DC. Control of autophagy as a therapy for neurodegenerative disease. Nat Rev Neurol. 2011;8(2):108–17.

450. Merenlender-Wagner A, Malishkevich A, Shemer Z, Udawela M, Gibbons A, Scarr E, et al. Autophagy has a key role in the pathophysiology of schizophrenia. Mol Psychiatry. 2015;20(1):126–32.

451. Gassen NC, Hartmann J, Schmidt M V, Rein T. FKBP5/FKBP51 enhances autophagy to synergize with antidepressant action. Autophagy. 2015 Jan;11(3):578–80.

452. Gassen NC, Hartmann J, Zschocke J, Stepan J, Hafner K, Zellner A, et al. Association of FKBP51 with Priming of Autophagy Pathways and Mediation of Antidepressant Treatment Response: Evidence in Cells, Mice, and Humans. Nestler E, editor. PLoS Med. 2014 Nov;11(11):e1001755.

453. Qin H, Yeh W, Sarno P De, Holdbrooks AT, Liu Y, Muldowney MT, et al. Signal transducer and activator of transcription- axis in myeloid cells regulates neuroin fl ammation. 2012;3:1–6.

454. Murase S, McKay RD. Neuronal activity-dependent STAT3 localization to nucleus is dependent on Tyr-705 and Ser-727 phosphorylation in rat hippocampal neurons. Eur J Neurosci. 2014 Feb;39(4):557–65.

455. Inta D, Lang UE, Borgwardt S, Meyer-Lindenberg A, Gass P. Microglia Activation and Schizophrenia: Lessons From the Effects of Minocycline on Postnatal Neurogenesis, Neuronal Survival and Synaptic Pruning. Schizophr Bull. 2016;sbw088.

456. Kwon S-H, Han J-K, Choi M, Kwon Y-J, Kim SJ, Yi EH, et al. Dysfunction of Microglial STAT3 Alleviates Depressive Behavior via Neuron-Microglia Interactions. Neuropsychopharmacology. 2017 May;

457. Insel TR. Rethinking schizophrenia. Nature. 2010;468(7321):187–93.

458. Gururajan A, van den Buuse M. Is the mTOR-signalling cascade disrupted in Schizophrenia? J Neurochem [Internet]. 2014 May [cited 2017 Sep 17];129(3):377–87. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24266366

459. Pham X, Song G, Lao S, Goff L, Zhu H, Valle D, et al. The DPYSL2 gene connects mTOR and schizophrenia. Transl Psychiatry [Internet]. 2016 Nov 1 [cited 2018 May 6];6(11):e933–e933. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27801893

460. Nicolas CS, Amici M, Bortolotto ZA, Doherty A, Csaba Z, Fafouri A, et al. The role of JAK-STAT signaling within the CNS. JAK-STAT [Internet]. 2013 Jan 1 [cited 2018 May 6];2(1):e22925. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24058789

461. Hakobyan S, Boyajyan A, Sim RB. Classical pathway complement activity in schizophrenia. Neurosci Lett. 2005;374(1):35–7.

462. Stevens B, Allen NJ, Vazquez LE, Howell GR, Christopherson KS, Nouri N, et al. The Classical Complement Cascade Mediates CNS Synapse Elimination. Cell. 2007;131(6):1164–78.

463. Schafer DP, Lehrman EK, Kautzman AG, Koyama R, Mardinly AR, Yamasaki R, et al. Microglia Sculpt Postnatal Neural Circuits in an Activity and Complement-Dependent Manner. Neuron. 2012 May;74(4):691–705.

464. Bilbo SD, Schwarz JM. Early-life programming of later-life brain and behavior: a critical role for the immune system. Front Behav Neurosci. 2009;3(14):1–14.

465. Selemon LD, Zecevic N. Schizophrenia: a tale of two critical periods for prefrontal cortical

development. Transl Psychiatry. 2015 Aug;5:1–11.

466. Knöchel C, Kniep J, Cooper JD, Stäblein M, Wenzler S, Sarlon J, et al. Altered apolipoprotein C expression in association with cognition impairments and hippocampus volume in schizophrenia and bipolar disorder. Eur Arch Psychiatry Clin Neurosci. 2016 Aug;1–14.

467. Zhang L, Zhao J. Profile of minocycline and its potential in the treatment of schizophrenia. Neuropsychiatr Dis Treat. 2014;10:1103–11.

468. Oya K, Kishi T, Iwata N. Efficacy and tolerability of minocycline augmentation therapy in schizophrenia: a systematic review and meta-analysis of randomized controlled trials. Hum Psychopharmacol. 2014 Aug;29:483–91.

469. Venneti S, Lopresti BJ, Wiley CA. Molecular imaging of microglia/macrophages in the brain. Glia [Internet]. 2013 Jan [cited 2017 Sep 15];61(1):10–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22615180

470. van Berckel BN, Bossong MG, Boellaard R, Kloet R, Schuitemaker A, Caspers E, et al. Microglia Activation in Recent-Onset Schizophrenia: A Quantitative (R)-[11C]PK11195 Positron Emission Tomography Study. Biol Psychiatry [Internet]. 2008 Nov 1 [cited 2017 Sep 15];64(9):820–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18534557

471. Doorduin J, de Vries EFJ, Willemsen ATM, de Groot JC, Dierckx RA, Klein HC. Neuroinflammation in Schizophrenia-Related Psychosis: A PET Study. J Nucl Med [Internet]. 2009 Nov 1 [cited 2017 Sep 15];50(11):1801–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19837763

472. Cotel M-C, Lenartowicz EM, Natesan S, Modo MM, Cooper JD, Williams SCR, et al. Microglial activation in the rat brain following chronic antipsychotic treatment at clinically relevant doses. Eur Neuropsychopharmacol. 2015;25(11):2098–107.

473. Holmes SE, Hinz R, Drake RJ, Gregory CJ, Conen S, Matthews JC, et al. In vivo imaging of brain microglial activity in antipsychotic-free and medicated schizophrenia: a [11C](R)-PK11195 positron emission tomography study. Mol Psychiatry. 2016 Dec;21(12):1672–9.

474. Laskaris LE, Di Biase MA, Everall I, Chana G, Christopoulos A, Skafidas E, et al. Microglial activation and progressive brain changes in schizophrenia. Br J Pharmacol [Internet]. 2016 Feb [cited 2017 Sep 16];173(4):666–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26455353

475. Kenk M, Selvanathan T, Rao N, Suridjan I, Rusjan P, Remington G, et al. Imaging neuroinflammation in gray and white matter in schizophrenia: an in-vivo PET study with [18F]-FEPPA. Schizophr Bull [Internet]. 2015 Jan [cited 2017 Sep 15];41(1):85–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25385788

476. Qin H, Yeh W, Sarno P De, Holdbrooks AT, Liu Y, Muldowney MT, et al. Signal transducer and activator of transcription- axis in myeloid cells regulates neuroinflammation. PNAS. 2012;109(13):5004–9.

477. Batarseh A, Papadopoulos V. Regulation of translocator protein 18kDa (TSPO) expression in health and disease states☆. Mol Cell Endocrinol [Internet]. 2010 Oct 7 [cited 2017 Sep 15];327(1–2):1–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20600583

478. Hill MJ, Donocik JG, Nuamah RA, Mein CA, Sainz-Fuertes R, Bray NJ. Transcriptional consequences of schizophrenia candidate miR-137 manipulation in human neural progenitor cells. Schizophr Res [Internet]. 2014 Mar [cited 2017 Sep 15];153(1–3):225–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24556472

479. Luo X, Huang L, Han L, Luo Z, Hu F, Tieu R, et al. Systematic prioritization and integrative analysis of copy number variations in schizophrenia reveal key schizophrenia susceptibility genes. Schizophr Bull [Internet]. 2014 Nov [cited 2017 Sep 15];40(6):1285–99. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24664977

480. Harry GJ. Microglia during development and aging. Pharmacol Ther [Internet]. 2013 Sep [cited 2017 Sep 15];139(3):313–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23644076

481. Huntoon KM, Wang Y, Eppolito CA, Barbour KW, Berger FG, Shrikant PA, et al. The acute phase protein haptoglobin regulates host immunity. J Leukoc Biol [Internet]. 2008 Apr 3 [cited 2017 Sep 15];84(1):170–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18436583

482. Vila-Rodriguez F, Honer WG, Innis SM, Wellington CL, Beasley CL. ApoE and cholesterol in schizophrenia and bipolar disorder: comparison of grey and white matter and relation with APOE genotype. J Psychiatry Neurosci [Internet]. 2011 Jan [cited 2017 Sep 15];36(1):47–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20964956

483. Deleidi M, Jäggle M, Rubino G. Immune aging, dysmetabolism, and inflammation in

neurological diseases. Front Neurosci [Internet]. 2015 [cited 2017 Sep 16];9:172. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26089771

484. Guo L, LaDu MJ, Van Eldik LJ. A Dual Role for Apolipoprotein E in Neuroinflammation: Anti- and Pro-Inflammatory Activity. J Mol Neurosci [Internet]. 2004 [cited 2017 Sep 16];23(3):205–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15181248

485. Lavisse S, Guillermier M, Herard A-S, Petit F, Delahaye M, Van Camp N, et al. Reactive Astrocytes Overexpress TSPO and Are Detected by TSPO Positron Emission Tomography Imaging. J Neurosci [Internet]. 2012 Aug 8 [cited 2017 Sep 15];32(32):10809–18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22875916

486. Tóth M, Little P, Arnberg F, Häggkvist J, Mulder J, Halldin C, et al. Acute neuroinflammation in a clinically relevant focal cortical ischemic stroke model in rat: longitudinal positron emission tomography and immunofluorescent tracking. Brain Struct Funct [Internet]. 2016 Apr 20 [cited 2017 Sep 15];221(3):1279–90. Available from: http://link.springer.com/10.1007/s00429-014-0970-y

487. Nicolas CS, Peineau S, Amici M, Csaba Z, Fafouri A, Javalet C, et al. The Jak/STAT pathway is involved in synaptic plasticity. Neuron [Internet]. 2012 Jan 26 [cited 2018 May 6];73(2):374–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22284190

488. Bachstetter AD, Rowe RK, Kaneko M, Goulding D, Lifshitz J, Van Eldik LJ. The p38α MAPK regulates microglial responsiveness to diffuse traumatic brain injury. J Neurosci [Internet]. 2013 Apr 3 [cited 2017 Sep 15];33(14):6143–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23554495

489. Huang G, Wang Y, Vogel P, Kanneganti T-D, Otsu K, Chi H. Signaling via the kinase p38α programs dendritic cells to drive TH17 differentiation and autoimmune inflammation. Nat Immunol [Internet]. 2012 Jan 8 [cited 2017 Sep 15];13(2):152–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22231518

490. Dudley AC, Thomas D, Best J, Jenkins A. The STATs in cell stress-type responses. Cell Commun Signal [Internet]. 2004 Aug 6 [cited 2017 Sep 15];2(1):8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15296508

491. Kovarik P, Stoiber D, Eyers PA, Menghini R, Neininger A, Gaestel M, et al. Stress-induced phosphorylation of STAT1 at Ser727 requires p38 mitogen-activated protein kinase whereas IFN-gamma uses a different signaling pathway. Proc Natl Acad Sci U S A [Internet]. 1999 Nov

23 [cited 2017 Sep 15];96(24):13956–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10570180

492. Ramsauer K, Sadzak I, Porras A, Pilz A, Nebreda AR, Decker T, et al. p38 MAPK enhances STAT1-dependent transcription independently of Ser-727 phosphorylation. Proc Natl Acad Sci [Internet]. 2002 Oct 1 [cited 2017 Sep 15];99(20):12859–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12232043

493. Amato D, Beasley CL, Hahn MK, Vernon AC. Neuroadaptations to antipsychotic drugs: Insights from pre-clinical and human post-mortem studies. Neurosci Biobehav Rev [Internet]. 2017 May [cited 2017 Sep 15];76(Pt B):317–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27756689

494. Singh RK, Jia C, Garcia F, Carrasco GA, Battaglia G, Muma NA. Activation of the JAK-STAT pathway by olanzapine is necessary for desensitization of serotonin2A receptor-stimulated phospholipase C signaling in rat frontal cortex but not serotonin2A receptor-stimulated hormone release. J Psychopharmacol [Internet]. 2010 Jul [cited 2017 Sep 16];24(7):1079–88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19304867

495. Richetto J, Chesters R, Cattaneo A, Labouesse MA, Gutierrez AMC, Wood TC, et al. Genome-Wide Transcriptional Profiling and Structural Magnetic Resonance Imaging in the Maternal Immune Activation Model of Neurodevelopmental Disorders. Cereb Cortex [Internet]. 2016 Oct 23 [cited 2017 Sep 15];27(6):3397–413. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27797829

496. The influence of clozapine treatment and other antipsychotics on the 18 kDa translocator protein, formerly named the peripheral-type benzodiazepine receptor, and steroid production. Eur Neuropsychopharmacol [Internet]. 2008 Jan 1 [cited 2017 Sep 15];18(1):24–33. Available from: http://www.sciencedirect.com/science/article/pii/S0924977X0700096X?showall=true

497. Cornett EM, Novitch M, Kaye AD, Kata V, Kaye AM. Medication-Induced Tardive Dyskinesia: A Review and Update. Ochsner J [Internet]. 2017 [cited 2018 May 5];17(2):162–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28638290

498. Müller N, Weidinger E, Leitner B, Schwarz MJ. The role of inflammation in schizophrenia. Front Neurosci [Internet]. 2015 Oct 21 [cited 2018 May 5];9:372. Available from: http://journal.frontiersin.org/Article/10.3389/fnins.2015.00372/abstract

499. Stone JM. Glutamatergic antipsychotic drugs: a new dawn in the treatment of schizophrenia? Ther Adv Psychopharmacol [Internet]. 2011 Feb [cited 2018 May 5];1(1):5–18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23983922

500. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry [Internet]. 2006 Jan [cited 2016 Feb 25];163(1):28–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16390886

501. Charney DS, Manji HK. Life stress, genes, and depression: multiple pathways lead to increased risk and new opportunities for intervention. Sci STKE [Internet]. 2004 Mar 23 [cited 2016 Feb 18];2004(225):re5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15039492

502. Maier W. Common risk genes for affective and schizophrenic psychoses. Eur Arch Psychiatry Clin Neurosci [Internet]. 2008 Jun [cited 2016 Jan 29];258 Suppl:37–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18516516

503. Lohoff FW. Overview of the genetics of major depressive disorder. Curr Psychiatry Rep [Internet]. 2010 Dec [cited 2016 Jun 26];12(6):539–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20848240

504. Markou A, Chiamulera C, Geyer MA, Tricklebank M, Steckler T. Removing obstacles in neuroscience drug discovery: the future path for animal models. Neuropsychopharmacology [Internet]. 2009 Jan [cited 2016 Mar 28];34(1):74–89. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2651739&tool=pmcentrez&rendertype=abstract

505. Berton O, Hahn C-G, Thase ME. Are we getting closer to valid translational models for major depression? Science [Internet]. 2012 Oct 5 [cited 2016 Apr 8];338(6103):75–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23042886

506. Barkus C. Genetic mouse models of depression. Curr Top Behav Neurosci [Internet]. 2013 Jan [cited 2016 Apr 8];14:55–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22890963

507. Nestler EJ, Hyman SE. Animal models of neuropsychiatric disorders. Nat Neurosci [Internet]. 2010 Oct [cited 2014 Jul 11];13(10):1161–9. Available from: http://dx.doi.org/10.1038/nn.2647

508. Chan MK, Gottschalk MG, Haenisch F, Tomasik J, Ruland T, Rahmoune H, et al. Applications of blood-based protein biomarker strategies in the study of psychiatric disorders. Prog Neurobiol [Internet]. 2014 Nov [cited 2016 Jun 26];122:45–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25173695

509. Torrey EF, Webster M, Knable M, Johnston N, Yolken RH. The Stanley Foundation brain collection and Neuropathology Consortium. Schizophr Res. 2000;44(2):151–5.

510. Gottschalk MG, Wesseling H, Guest PC, Bahn S. Proteomic enrichment analysis of psychotic and affective disorders reveals common signatures in presynaptic glutamatergic signaling and energy metabolism. Int J Neuropsychopharmacol [Internet]. 2014;18(2):pyu019. Available from: http://ijnp.oxfordjournals.org/content/18/2/pyu019.abstract

511. Ma D, Chan MK, Lockstone HE, Pietsch SR, Jones DNC, Cilia J, et al. Antipsychotic treatment alters protein expression associated with presynaptic function and nervous system development in rat frontal cortex. J Proteome Res [Internet]. 2009 Jul [cited 2016 Oct 9];8(7):3284–97. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19400588

512. Berton O, McClung C a, Dileone RJ, Krishnan V, Renthal W, Russo SJ, et al. Essential role of BDNF in the mesolimbic dopamine pathway in social defeat stress. Science (80-). 2006;311(5762):864–8.

513. Krishnan V, Han M-H, Graham DL, Berton O, Renthal W, Russo SJ, et al. Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. Cell [Internet]. 2007;131(2):391–404. Available from: http://www.sciencedirect.com/science/article/pii/S0092867407012068

514. Moreau JL, Jenck F, Martin JR, Mortas P, Haefely WE. Antidepressant treatment prevents chronic unpredictable mild stress-induced anhedonia as assessed by ventral tegmentum self-stimulation behavior in rats. Eur Neuropsychopharmacol. 1992;2(1):43–9.

515. Willner P, Towell a., Sampson D, Sophokleous S, Muscat R. Reduction of sucrose preference by chronic unpredictable mild stress, and its restoration by a tricyclic antidepressant. Psychopharmacology (Berl). 1987;93(3):358–64.

516. Pothion S, Bizot J-C, Trovero F, Belzung C. Strain differences in sucrose preference and in the consequences of unpredictable chronic mild stress. Behav Brain Res [Internet]. 2004 Nov 5 [cited 2016 Oct 3];155(1):135–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15325787

517. Herrera-Ruiz M, García-Beltrán Y, Mora S, Díaz-Véliz G, Viana GSB, Tortoriello J, et al. Antidepressant and anxiolytic effects of hydroalcoholic extract from Salvia elegans. J Ethnopharmacol [Internet]. 2006 Aug 11 [cited 2016 Oct 3];107(1):53–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16530995

518. Borsini F, Meli A. Is the forced swimming test a suitable model for revealing antidepressant activity? Psychopharmacology (Berl) [Internet]. 1988 [cited 2016 Oct 3];94(2):147–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3127840

519. Pollak DD, Rey CE, Monje FJ. Rodent models in depression research: classical strategies and new directions. Ann Med [Internet]. 2010 May 6 [cited 2016 Oct 3];42(4):252–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20367120

520. Yang CR, Zhang ZG, Bai YY, Zhou HF, Zhou L, Ruan CS, et al. Foraging activity is reduced in a mouse model of depression. Neurotox Res [Internet]. 2014 Apr [cited 2016 Oct 3];25(3):235–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23873577

521. Sickmann HM, Arentzen TS, Dyrby TB, Plath N, Kristensen MP. Prenatal stress produces sex-specific changes in depression-like behavior in rats: implications for increased vulnerability in females. J Dev Orig Health Dis [Internet]. 2015 Oct 1 [cited 2015 Oct 2];6(5):462–74. Available from: http://journals.cambridge.org/abstract_S2040174415001282

522. Beniger JR, Barnett V, Lewis T. Outliers in Statistical Data. Contemp Sociol. 1980;9(4):560.

523. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc [Internet]. 2007;2(10):2366–82. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3685583&tool=pmcentrez&rendertype=abstract

524. Li G-L, Xu X-H, Wang B-A, Yao Y-M, Qin Y, Bai S-R, et al. Analysis of protein-protein interaction network and functional modules on primary osteoporosis. Eur J Med Res [Internet]. 2014 Mar 21 [cited 2018 May 1];19(1):15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24656062

525. Taylor NR. SMALL WORLD NETWORK STRATEGIES FOR STUDYING PROTEIN STRUCTURES AND BINDING. Comput Struct Biotechnol J [Internet]. 2013 Feb 1 [cited 2018 Apr 17];5(6). Available from: https://www.sciencedirect.com/science/article/pii/S2001037014600349

526. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One [Internet]. 2010;5(11):e13984. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2981572&tool=pmcentrez&rendertype=abstract

527. Li L, Ching W, Chan Y, Mamitsuka H. On network-based kernel methods for protein-protein interactions with applications in protein functions prediction. J Syst Sci Complex [Internet]. 2010 Nov 9 [cited 2015 Oct 19];23(5):917–30. Available from: http://link.springer.com/10.1007/s11424-010-0207-y

528. Zare H, Kaveh M, Khodursky A. Inferring a Transcriptional Regulatory Network from Gene Expression Data Using Nonlinear Manifold Embedding. Rogers S, editor. PLoS One [Internet]. 2011 Aug 12 [cited 2015 Oct 19];6(8):e21969. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0021969

529. Lerman G, Shakhnovich BE. Defining functional distance using manifold embeddings of gene ontology annotations. Proc Natl Acad Sci U S A. 2007;104(27):11334–9.

530. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. Science [Internet]. 2000 Dec 22 [cited 2014 Jul 12];290(5500):2323–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11125150

531. Fröhlich H, Speer N, Poustka A, Beissbarth T. GOSim--an R-package for computation of information theoretic GO similarities between terms and gene products. BMC Bioinformatics [Internet]. 2007 May 22 [cited 2018 Apr 17];8:166. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17519018

532. Fröhlich H, Speer N, Zell A. Kernel Based Functional Gene Grouping. In: Proc Int Joint Conf Neural Networks. 2006. p. 6886–91.

533. Speer N, Frohlich H, Spieth C, Zell A. Functional grouping of genes using spectral clustering and Gene Ontology. In: Neural Networks, 2005 IJCNN ’05 Proceedings 2005 IEEE International Joint Conference on [Internet]. 2005. p. 298–303. Available from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1555846

534. Schaefer MH, Serrano L, Andrade-Navarro MA. Correcting for the study bias associated with protein-protein interaction measurements reveals differences between protein degree distributions from different cancer types. Front Genet [Internet]. 2015 [cited 2018 Apr

18];6:260. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26300911

535. Krishnan V, Nestler EJ. The molecular neurobiology of depression. Nature [Internet]. 2008 Oct 16 [cited 2014 Jul 10];455(7215):894–902. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2721780&tool=pmcentrez&rendertype=abstract

536. McArthur R, Borsini F. Animal models of depression in drug discovery: a historical perspective. Pharmacol Biochem Behav [Internet]. 2006 Jul [cited 2016 Mar 22];84(3):436–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16844210

537. Krishnan V, Nestler EJ. Animal models of depression: molecular perspectives. Curr Top Behav Neurosci [Internet]. 2011 Jan [cited 2016 Apr 11];7:121–47. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3270071&tool=pmcentrez&rendertype=abstract

538. Bayés A, Grant SGN. Neuroproteomics: understanding the molecular organization and complexity of the brain. Nat Rev Neurosci [Internet]. 2009 Sep [cited 2016 Jun 26];10(9):635–46. Available from: http://www.nature.com/doifinder/10.1038/nrn2701

539. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry [Internet]. 2000 Oct [cited 2016 Jun 26];157(10):1552–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11007705

540. Schmidt HD, Duman RS. The role of neurotrophic factors in adult hippocampal neurogenesis, antidepressant treatments and animal models of depressive-like behavior. Behav Pharmacol [Internet]. 2007 Sep [cited 2016 Mar 22];18(5–6):391–418. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17762509

541. Farooq RK, Isingrini E, Tanti A, Le Guisquet A-M, Arlicot N, Minier F, et al. Is unpredictable chronic mild stress (UCMS) a reliable model to study depression-induced neuroinflammation? Behav Brain Res [Internet]. 2012 May 16 [cited 2016 Mar 10];231(1):130–7. Available from: http://www.sciencedirect.com/science/article/pii/S0166432812002124

542. Van Bokhoven P, Oomen CA, Hoogendijk WJG, Smit AB, Lucassen PJ, Spijker S. Reduction in hippocampal neurogenesis after social defeat is long-lasting and responsive to late antidepressant treatment. Eur J Neurosci [Internet]. 2011 May [cited 2016 Jul 5];33(10):1833–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21488984

543. Christoffel DJ, Golden SA, Dumitriu D, Robison AJ, Janssen WG, Ahn HF, et al. IκB kinase regulates social defeat stress-induced synaptic and behavioral plasticity. J Neurosci [Internet]. 2011 Jan 5 [cited 2016 Mar 20];31(1):314–21. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3219041&tool=pmcentrez&rendertype=abstract

544. Monje FJ, Cabatic M, Divisch I, Kim E-J, Herkner KR, Binder BR, et al. Constant darkness induces IL-6-dependent depression-like behavior through the NF-κB signaling pathway. J Neurosci [Internet]. 2011 Jun 22 [cited 2016 Apr 16];31(25):9075–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21697358

545. Huang Y, Chen H-C, Chiang C-W, Yeh C-T, Chen S-J, Chou C-K. Identification of a two-layer regulatory network of proliferation-related microRNAs in hepatoma cells. Nucleic Acids Res [Internet]. 2012 Nov 1 [cited 2016 Jan 30];40(20):10478–93. Available from: http://nar.oxfordjournals.org/content/40/20/10478.short

546. Plaisier CL, Pan M, Baliga NS. A miRNA-regulatory network explains how dysregulated miRNAs perturb oncogenic processes across diverse cancers. Genome Res [Internet]. 2012 Nov 1 [cited 2016 Jan 30];22(11):2302–14. Available from: http://genome.cshlp.org/content/22/11/2302.full

547. Schlicker A, Domingues FS, Rahnenführer J, Lengauer T. A new measure for functional similarity of gene products based on Gene Ontology. BMC Bioinformatics [Internet]. 2006;7:302. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-33748335463&partnerID=tZOtx3y1

548. Czéh B, Fuchs E, Wiborg O, Simon M. Animal models of major depression and their clinical implications. Prog Neuropsychopharmacol Biol Psychiatry [Internet]. 2016 Jan 4 [cited 2016 Mar 5];64:293–310. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25891248

549. Carboni L. Peripheral biomarkers in animal models of major depressive disorder. Dis Markers [Internet]. 2013 [cited 2016 Jun 26];35(1):33–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24167347

550. Brigitta B. Pathophysiology of depression and mechanisms of treatment. Dialogues Clin Neurosci [Internet]. 2002 Mar [cited 2016 Mar 8];4(1):7–20. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3181668&tool=pmcentrez&rendertype=abstract

551. Buchanan RW, Gold JM. Negative symptoms: diagnosis, treatment and prognosis. Int Clin Psychopharmacol [Internet]. 1996 May [cited 2016 Feb 1];11 Suppl 2:3–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8803654

552. Green MF. What are the functional consequences of neurocognitive deficits in schizophrenia? Am J Psychiatry [Internet]. 1996 Mar [cited 2015 Dec 29];153(3):321–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8610818

553. Greenwood KE, Landau S, Wykes T. Negative symptoms and specific cognitive impairments as combined targets for improved functional outcome within cognitive remediation therapy. Schizophr Bull [Internet]. 2005 Oct [cited 2016 Feb 1];31(4):910–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16049165

554. Konradi C, Heckers S. Molecular aspects of glutamate dysregulation: implications for schizophrenia and its treatment. Pharmacol Ther [Internet]. 2003 Feb [cited 2016 Feb 1];97(2):153–79. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4203361&tool=pmcentrez&rendertype=abstract

555. Tsai G, Coyle JT. Glutamatergic mechanisms in schizophrenia. Annu Rev Pharmacol Toxicol [Internet]. 2002 Jan [cited 2016 Jan 13];42:165–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11807169

556. Kim JS, Kornhuber HH, Schmid-Burgk W, Holzmüller B. Low cerebrospinal fluid glutamate in schizophrenic patients and a new hypothesis on schizophrenia. Neurosci Lett [Internet]. 1980 Dec [cited 2016 Apr 28];20(3):379–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6108541

557. Halene TB, Ehrlichman RS, Liang Y, Christian EP, Jonak GJ, Gur TL, et al. Assessment of NMDA receptor NR1 subunit hypofunction in mice as a model for schizophrenia. Genes Brain Behav [Internet]. 2009 Oct [cited 2016 Jan 18];8(7):661–75. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2757454&tool=pmcentrez&rendertype=abstract

558. Jentsch JD, Roth RH. The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. Neuropsychopharmacology [Internet]. 1999 Mar [cited 2015 Nov 11];20(3):201–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10063482

559. Kapur S, Seeman P. NMDA receptor antagonists ketamine and PCP have direct effects on the dopamine D(2) and serotonin 5-HT(2)receptors-implications for models of schizophrenia. Mol Psychiatry [Internet]. 2002 Jan [cited 2015 Nov 11];7(8):837–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12232776

560. Liu J, Moghaddam B. Regulation of glutamate efflux by excitatory amino acid receptors: evidence for tonic inhibitory and phasic excitatory regulation. J Pharmacol Exp Ther [Internet]. 1995 Sep [cited 2016 Apr 28];274(3):1209–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7562490

561. Moghaddam B, Adams B, Verma A, Daly D. Activation of glutamatergic neurotransmission by ketamine: a novel step in the pathway from NMDA receptor blockade to dopaminergic and cognitive disruptions associated with the prefrontal cortex. J Neurosci [Internet]. 1997 Apr 15 [cited 2016 Feb 25];17(8):2921–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9092613

562. Grayson B, Adamson L, Harte M, Leger M, Marsh S, Piercy C, et al. The involvement of distraction in memory deficits induced by NMDAR antagonism: relevance to cognitive deficits in schizophrenia. Behav Brain Res [Internet]. 2014 Jun 1 [cited 2016 Feb 1];266:188–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24632009

563. Neill JC, Harte MK, Haddad PM, Lydall ES, Dwyer DM. Acute and chronic effects of NMDA receptor antagonists in rodents, relevance to negative symptoms of schizophrenia: a translational link to humans. Eur Neuropsychopharmacol [Internet]. 2014 May [cited 2016 Feb 1];24(5):822–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24287012

564. Lahti AC, Koffel B, LaPorte D, Tamminga CA. Subanesthetic doses of ketamine stimulate psychosis in schizophrenia. Neuropsychopharmacology [Internet]. 1995 Aug [cited 2016 Jan 31];13(1):9–19. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8526975

565. Ramnani N, Owen AM. Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. Nat Rev Neurosci [Internet]. 2004 Mar [cited 2016 Mar 15];5(3):184–94. Available from: http://dx.doi.org/10.1038/nrn1343

566. Wesseling H, Rahmoune H, Tricklebank M, Guest PC, Bahn S. A Targeted Multiplexed Proteomic Investigation Identifies Ketamine-Induced Changes in Immune Markers in Rat Serum and Expression Changes in Protein Kinases/Phosphatases in Rat Brain. J Proteome Res [Internet]. 2015 Jan 2 [cited 2016 Feb 1];14(1):411–21. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/25363195

567. Gastambide F, Mitchell SN, Robbins TW, Tricklebank MD, Gilmour G. Temporally distinct cognitive effects following acute administration of ketamine and phencyclidine in the rat. Eur Neuropsychopharmacol [Internet]. 2013 Nov [cited 2016 Jan 12];23(11):1414–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23561394

568. Li J, Ishiwari K, Conway MW, Francois J, Huxter J, Lowry JP, et al. Dissociable effects of antipsychotics on ketamine-induced changes in regional oxygenation and inter-regional coherence of low frequency oxygen fluctuations in the rat. Neuropsychopharmacology [Internet]. 2014 Jun [cited 2016 Jan 12];39(7):1635–44. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4023136&tool=pmcentrez&rendertype=abstract

569. Littlewood CL, Jones N, O'Neill MJ, Mitchell SN, Tricklebank M, Williams SCR. Mapping the central effects of ketamine in the rat using pharmacological MRI. Psychopharmacology (Berl) [Internet]. 2006 May [cited 2016 Jan 12];186(1):64–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16550385

570. Smith JW, Gastambide F, Gilmour G, Dix S, Foss J, Lloyd K, et al. A comparison of the effects of ketamine and phencyclidine with other antagonists of the NMDA receptor in rodent assays of attention and working memory. Psychopharmacology (Berl) [Internet]. 2011 Sep [cited 2016 Jan 12];217(2):255–69. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21484239

571. Ernst A, Ma D, Garcia-Perez I, Tsang TM, Kluge W, Schwarz E, et al. Molecular validation of the acute phencyclidine rat model for schizophrenia: identification of translational changes in energy metabolism and neurotransmission. J Proteome Res [Internet]. 2012 Jul 6 [cited 2016 Jan 18];11(7):3704–14. Available from: http://dx.doi.org/10.1021/pr300197d

572. Palmowski P, Rogowska-Wrzesinska A, Williamson J, Beck HC, Mikkelsen JD, Hansen HH, et al. Acute phencyclidine treatment induces extensive and distinct protein phosphorylation in rat frontal cortex. J Proteome Res [Internet]. 2014 Mar 7 [cited 2016 Jan 18];13(3):1578–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24564430

573. Wesseling H, Want EJ, Guest PC, Rahmoune H, Holmes E, Bahn S. Hippocampal Proteomic and Metabonomic Abnormalities in Neurotransmission, Oxidative Stress, and Apoptotic Pathways in a Chronic Phencyclidine Rat Model. J Proteome Res [Internet]. 2015 Aug 7 [cited 2016 Jan 18];14(8):3174–87. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26043028

574. Wesseling H, Guest PC, Lee C-M, Wong EH, Rahmoune H, Bahn S. Integrative proteomic analysis of the NMDA NR1 knockdown mouse model reveals effects on central and peripheral pathways associated with schizophrenia and autism spectrum disorders. Mol Autism [Internet]. 2014 Jan [cited 2016 Feb 1];5:38. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4109791&tool=pmcentrez&rendertype=abstract

575. Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, et al. A human protein-protein interaction network: A resource for annotating the proteome. Cell. 2005;122(6):957–68.

576. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol [Internet]. 2007 Jan [cited 2015 Aug 25];8(9):R183. Available from: http://genomebiology.com/2007/8/9/R183

577. Lee PR, Brady DL, Shapiro RA, Dorsa DM, Koenig JI. Social interaction deficits caused by chronic phencyclidine administration are reversed by oxytocin. Neuropsychopharmacology [Internet]. 2005 Oct [cited 2016 Feb 1];30(10):1883–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15798779

578. Sams-Dodd F. Phencyclidine-induced stereotyped behaviour and social isolation in rats: a possible animal model of schizophrenia. Behav Pharmacol [Internet]. 1996 Jan [cited 2016 Feb 1];7(1):3–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11224390

579. Becker A, Peters B, Schroeder H, Mann T, Huether G, Grecksch G. Ketamine-induced changes in rat behaviour: A possible animal model of schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry [Internet]. 2003 Jun [cited 2016 Apr 28];27(4):687–700. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12787858

580. Dzirasa K, Ramsey AJ, Takahashi DY, Stapleton J, Potes JM, Williams JK, et al. Hyperdopaminergia and NMDA receptor hypofunction disrupt neural phase signaling. J Neurosci [Internet]. 2009 Jun 24 [cited 2016 Feb 1];29(25):8215–24. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2731697&tool=pmcentrez&rendertype=abstract

581. Chan MK, Gottschalk MG, Haenisch F, Tomasik J, Ruland T, Rahmoune H, et al. Applications of blood-based protein biomarker strategies in the study of psychiatric disorders. Prog

Neurobiol [Internet]. 2014 Nov [cited 2016 Jan 30];122:45–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25173695

582. Mäki P, Veijola J, Jones PB, Murray GK, Koponen H, Tienari P, et al. Predictors of schizophrenia--a review. Br Med Bull [Internet]. 2005 Jan [cited 2015 Dec 29];73–74:1–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15947217

583. Fonio E, Golani I, Benjamini Y. Measuring behavior of animal models: faults and remedies. Nat Methods [Internet]. 2012 Dec [cited 2015 Dec 15];9(12):1167–70. Available from: http://dx.doi.org/10.1038/nmeth.2252

584. Tordjman S, Drapier D, Bonnot O, Graignic R, Fortes S, Cohen D, et al. Animal models relevant to schizophrenia and autism: validity and limitations. Behav Genet [Internet]. 2007 Jan [cited 2016 Jan 30];37(1):61–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17160702

585. Lei G, Xia Y, Johnson KM. The role of Akt-GSK-3beta signaling and synaptic strength in phencyclidine-induced neurodegeneration. Neuropsychopharmacology [Internet]. 2008 May [cited 2016 Jan 29];33(6):1343–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17637606

586. Adams B, Moghaddam B. Corticolimbic dopamine neurotransmission is temporally dissociated from the cognitive and locomotor effects of phencyclidine. J Neurosci [Internet]. 1998 Jul 15 [cited 2016 Feb 1];18(14):5545–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9651235

587. López-Gil X, Babot Z, Amargós-Bosch M, Suñol C, Artigas F, Adell A. Clozapine and haloperidol differently suppress the MK-801-increased glutamatergic and serotonergic transmission in the medial prefrontal cortex of the rat. Neuropsychopharmacology [Internet]. 2007 Oct 14 [cited 2016 Jan 18];32(10):2087–97. Available from: http://dx.doi.org/10.1038/sj.npp.1301356

588. Jentsch JD. Enduring Cognitive Deficits and Cortical Dopamine Dysfunction in Monkeys After Long-Term Administration of Phencyclidine. Science (80-) [Internet]. 1997 Aug 15 [cited 2016 Feb 1];277(5328):953–5. Available from: http://science.sciencemag.org/content/277/5328/953.abstract

589. Olney JW, Labruyere J, Price MT. Pathological changes induced in cerebrocortical neurons by phencyclidine and related drugs. Science [Internet]. 1989 Jun 16 [cited 2016 Feb 1];244(4910):1360–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2660263

590. Prabakaran S, Swatton JE, Ryan MM, Huffaker SJ, Huang JT-J, Griffin JL, et al. Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. Mol Psychiatry [Internet]. 2004 Jul [cited 2015 Oct 1];9(7):684–97, 643. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15098003

591. Rajasekaran A, Venkatasubramanian G, Berk M, Debnath M. Mitochondrial dysfunction in schizophrenia: pathways, mechanisms and implications. Neurosci Biobehav Rev [Internet]. 2015 Jan [cited 2015 Dec 24];48:10–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25446950

592. Konradi C, Eaton M, MacDonald ML, Walsh J, Benes FM, Heckers S. Molecular evidence for mitochondrial dysfunction in bipolar disorder. Arch Gen Psychiatry [Internet]. 2004 Mar [cited 2015 Dec 24];61(3):300–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14993118

593. Gottschalk MG, Wesseling H, Guest PC, Bahn S. Proteomic enrichment analysis of psychotic and affective disorders reveals common signatures in presynaptic glutamatergic signaling and energy metabolism. Int J Neuropsychopharmacol [Internet]. 2015 Jan [cited 2016 Jan 11];18(2). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4368887&tool=pmcentrez&rendertype=abstract

594. Agarwal A, Zhang M, Trembak-Duff I, Unterbarnscheidt T, Radyushkin K, Dibaj P, et al. Dysregulated expression of neuregulin-1 by cortical pyramidal neurons disrupts synaptic plasticity. Cell Rep [Internet]. 2014 Aug 21 [cited 2018 May 5];8(4):1130–45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25131210

595. Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D, et al. Common variants conferring risk of schizophrenia. Nature [Internet]. 2009 Jul 1 [cited 2017 Mar 13];460(7256):744. Available from: http://www.nature.com/doifinder/10.1038/nature08186

596. Uhlhaas PJ, Singer W. Abnormal neural oscillations and synchrony in schizophrenia. Nat Rev Neurosci [Internet]. 2010 Feb [cited 2018 May 5];11(2):100–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20087360

597. BERTRAM I, BERNSTEIN H-G, LENDECKEL U, BUKOWSKA A, DOBROWOLNY H, KEILHOFF G, et al. Immunohistochemical Evidence for Impaired Neuregulin-1 Signaling in the Prefrontal Cortex in Schizophrenia and in Unipolar Depression. Ann N Y Acad Sci [Internet]. 2007 Jan 1 [cited 2018 May 5];1096(1):147–56. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/17405926

598. Law AJ, Lipska BK, Weickert CS, Hyde TM, Straub RE, Hashimoto R, et al. Neuregulin 1 transcripts are differentially expressed in schizophrenia and regulated by 5' SNPs associated with the disease. Proc Natl Acad Sci U S A [Internet]. 2006 Apr 25 [cited 2018 May 5];103(17):6747–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16618933

599. Wong AHC, Josselyn SA. Caution When Diagnosing Your Mouse with Schizophrenia: The Use and Misuse of Model Animals for Understanding Psychiatric Disorders. Biol Psychiatry [Internet]. 2015 May 6 [cited 2015 Nov 4];79(1):32–8. Available from: http://www.sciencedirect.com/science/article/pii/S0006322315003613

600. Gandal MJ, Anderson RL, Billingslea EN, Carlson GC, Roberts TPL, Siegel SJ. Mice with reduced NMDA receptor expression: more consistent with autism than schizophrenia? Genes Brain Behav [Internet]. 2012 Aug [cited 2016 May 4];11(6):740–50. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3808979&tool=pmcentrez&rendertype=abstract

601. Stewart AM, Kalueff A V. Developing better and more valid animal models of brain disorders. Behav Brain Res [Internet]. 2015 Jan 1 [cited 2016 Jan 29];276:28–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24384129

602. Keefe RSE, Bilder RM, Davis SM, Harvey PD, Palmer BW, Gold JM, et al. Neurocognitive effects of antipsychotic medications in patients with chronic schizophrenia in the CATIE Trial. Arch Gen Psychiatry [Internet]. 2007 Jun [cited 2015 Dec 29];64(6):633–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17548746

603. Sarnyai Z, Alsaif M, Bahn S, Ernst A, Guest PC, Hradetzky E, et al. Behavioral and molecular biomarkers in translational animal models for neuropsychiatric disorders. Int Rev Neurobiol [Internet]. 2011 Jan [cited 2016 Jan 29];101:203–38. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22050853

604. Tomasik J, Schultz TL, Kluge W, Yolken RH, Bahn S, Carruthers VB. Shared Immune and Repair Markers During Experimental Toxoplasma Chronic Brain Infection and Schizophrenia. Schizophr Bull [Internet]. 2015 Sep 20 [cited 2016 Feb 1]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/26392628

605. Kendler KS, Diehl SR. The genetics of schizophrenia: a current, genetic-epidemiologic perspective. Schizophr Bull [Internet]. 1993 Jan [cited 2016 Jan 29];19(2):261–85. Available

from: http://www.ncbi.nlm.nih.gov/pubmed/8322035

606. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch Gen Psychiatry [Internet]. 2003 Dec [cited 2015 Feb 26];60(12):1187–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14662550

607. Prata D, Mechelli A, Kapur S. Clinically meaningful biomarkers for psychosis: a systematic and quantitative review. Neurosci Biobehav Rev [Internet]. 2014 Sep [cited 2017 Sep 26];45:134–41. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0149763414001274

608. Larson MK, Walker EF, Compton MT. Early signs, diagnosis and therapeutics of the prodromal phase of schizophrenia and related psychotic disorders. Expert Rev Neurother [Internet]. 2010 Aug [cited 2017 Jun 2];10(8):1347–59. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20662758

609. Bak M, Delespaul P, Hanssen M, de Graaf R, Vollebergh W, van Os J. How false are "false" positive psychotic symptoms? Schizophr Res [Internet]. 2003 Jul 1 [cited 2017 Jun 2];62(1–2):187–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12765760

610. Bahn S, Noll R, Barnes A, Schwarz E, Guest PC. Challenges of Introducing New Biomarker Products for Neuropsychiatric Disorders into the Market. In: International review of neurobiology [Internet]. 2011 [cited 2017 Jun 2]. p. 299–327. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22050857

611. Kourou K, Exarchos TP, Exarchos KP, Karamouzis M V., Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J [Internet]. 2015 [cited 2017 Sep 26];13:8–17. Available from: http://linkinghub.elsevier.com/retrieve/pii/S2001037014000464

612. Yang J, Chen T, Sun L, Zhao Z, Qi X, Zhou K, et al. Potential metabolite markers of schizophrenia. Mol Psychiatry [Internet]. 2013 Jan [cited 2017 Sep 26];18(1):67–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22024767

613. Chen F, Xue J, Zhou L, Wu S, Chen Z. Identification of serum biomarkers of hepatocarcinoma through liquid chromatography/mass spectrometry-based metabonomic method. Anal Bioanal Chem [Internet]. 2011 Oct [cited 2017 Jan 12];401(6):1899–904. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21833635

614. Subramaniam M, Lam M, Guo ME, He VYF, Lee J, Verma S, et al. Body Mass Index, Obesity,

and Psychopathology in Patients With Schizophrenia. J Clin Psychopharmacol [Internet]. 2014 Feb [cited 2017 Jun 2];34(1):40–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24346756

615. Radhakrishnan R, Wilkinson ST, D'Souza DC. Gone to Pot - A Review of the Association between Cannabis and Psychosis. Front psychiatry [Internet]. 2014 [cited 2017 Jun 2];5:54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24904437

616. Rothschild AJ. Challenges in the Treatment of Major Depressive Disorder With Psychotic Features. Schizophr Bull [Internet]. 2013 Jul 1 [cited 2017 Sep 26];39(4):787–96. Available from: https://academic.oup.com/schizophreniabulletin/article-lookup/doi/10.1093/schbul/sbt046

617. Poyurovsky M, Koran LM. Obsessive–compulsive disorder (OCD) with schizotypy vs. schizophrenia with OCD: diagnostic dilemmas and therapeutic implications. J Psychiatr Res [Internet]. 2005 Jul [cited 2017 Sep 26];39(4):399–408. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15804390

618. Dalsgaard S, Mortensen PB, Frydenberg M, Maibing CM, Nordentoft M, Thomsen PH. Association between Attention-Deficit Hyperactivity Disorder in childhood and schizophrenia later in adulthood. Eur Psychiatry [Internet]. 2014 May [cited 2017 Sep 26];29(4):259–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24016863

619. Patel JP, Frey BN, Patel JP, Frey BN. Disruption in the Blood-Brain Barrier: The Missing Link between Brain and Body Inflammation in Bipolar Disorder? Neural Plast. 2015;2015:1–12.

620. Gururajan A, van den Buuse M. Is the mTOR-signalling cascade disrupted in Schizophrenia? J Neurochem [Internet]. 2014 May [cited 2015 Nov 10];129(3):377–87. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24266366

621. Powell JD, Pollizzi KN, Heikamp EB, Horton MR. Regulation of Immune Responses by mTOR. Annu Rev Immunol [Internet]. 2012 Apr 23 [cited 2017 Sep 17];30(1):39–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22136167

622. Bisht K, Sharma KP, Lecours C, Gabriela Sánchez M, El Hajj H, Milior G, et al. Dark microglia: A new phenotype predominantly associated with pathological states. Glia [Internet]. 2016 May [cited 2017 Aug 9];64(5):826–39. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26847266

623. Melief J, Koning N, Schuurman KG, Van De Garde MDB, Smolders J, Hoek RM, et al. Phenotyping primary human microglia: Tight regulation of LPS responsiveness. Glia [Internet]. 2012 Oct [cited 2017 Aug 9];60(10):1506–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22740309

624. Eßlinger M, Wachholz S, Manitz M-P, Plümper J, Sommer R, Juckel G, et al. Schizophrenia associated sensory gating deficits develop after adolescent microglia activation. Brain Behav Immun. 2016;58:99–106.

625. Sellgren CM, Sheridan SD, Gracias J, Xuan D, Fu T, Perlis RH. Patient-specific models of microglia-mediated engulfment of synapses and neural progenitors. Mol Psychiatry. 2017 Feb;22(2):170–7.

626. Brennand KJ, Gage FH. Modeling psychiatric disorders through reprogramming. Dis Model Mech. 2012 Jan;5(1):26–32.

627. Sarnyai Z, Guest PC. Connecting Brain Proteomics with Behavioural Neuroscience in Translational Animal Models of Neuropsychiatric Disorders. In Springer, Cham; 2017 [cited 2017 Sep 19]. p. 97–114. Available from: http://link.springer.com/10.1007/978-3-319-52479-5_6