UNIVERSITY OF MEDICINE AND HEALTH SCIENCES

Royal College of Surgeons in Ireland

repository@rcsi.com

Investigating the Molecular Biology of Psychosis: Evidence and Implications for Pathophysiology, Prediction and Prevention

AUTHOR(S)

David Mongan

CITATION

Mongan, David (2022): Investigating the Molecular Biology of Psychosis: Evidence and Implications for Pathophysiology, Prediction and Prevention. Royal College of Surgeons in Ireland. Thesis. https://doi.org/10.25419/rcsi.16550073.v1

DOI

10.25419/rcsi.16550073.v1

LICENCE

CC BY-NC-SA 4.0

This work is made available under the above open licence by RCSI and has been printed from https://repository.rcsi.com. For more information please contact repository@rcsi.com

URL

https://repository.rcsi.com/articles/thesis/Investigating_the_Molecular_Biology_of_Psychosis_Evidence_and_Implications_for_Pathophysiology_Prediction_and_Prevention/16550073/1

Investigating the molecular biology of psychosis: evidence and implications for pathophysiology, prediction and prevention

David Paul Mongan MB BCh BAO MRCPsych

Department of Psychiatry
Royal College of Surgeons in Ireland

Supervisors: Professor David Cotter[1] and Professor Mary Cannon[1]

[1] Department of Psychiatry, Royal College of Surgeons in Ireland

A thesis submitted to the School of Postgraduate Studies, Faculty of Medicine and Health Sciences, Royal College of Surgeons in Ireland in fulfilment of the degree of Doctor of Philosophy.

June 2021

Thesis Declaration

I declare that this thesis, which I submit to RCSI for examination in consideration of the award of Doctor of Philosophy, is my own personal effort. Where any of the content presented is the result of input or data from a related collaborative research programme this is duly acknowledged in the text such that it is possible to ascertain how much of the work is my own. I have not already obtained a degree in RCSI or elsewhere on the basis of this work. Furthermore, I took reasonable care to ensure that the work is original, and, to the best of my knowledge, does not breach copyright law, and has not been taken from other sources except where such work has been cited and acknowledged within the text.

Signed:

Student number: 18141081

Date: 9th June 2021

Contents

Thesis Declaration 2

Intellectual Property Declaration ..... **Error! Bookmark not defined.**

List of Abbreviations 8

List of Tables 10

List of Figures 11

Acknowledgements 12

Dedication 14

Summary 15

List of scientific papers for consideration in this thesis 17

Related scientific papers not in consideration 18

Conference presentations 19

Academic prizes and awards 19

Chapter 1. Introduction 20

1.1 Psychotic disorders 21

1.2 The phenotypic continuum of early psychosis 22

1.2.1 First episode psychosis 24

1.2.2 The psychosis prodrome 24

1.2.3 Clinical high-risk criteria 25

1.2.4 Psychotic experiences 27

1.3 Inflammation and the complement system in the pathophysiology of psychosis 28

1.3.1 Inflammatory markers in psychosis 29

1.3.2 The complement system in psychosis 30

1.4 Biomarkers of psychosis risk 32

1.4.1 Electroencephalography 33

1.4.2 Neurocognition 33

1.4.3 Neuroimaging 34

1.4.4 Cortisol 34

1.4.5 Peripheral blood markers 34

1.4.6 Predictive models 35

1.5 Polyunsaturated fatty acids and prevention 35

1.6 Thesis objectives and research questions 38

1.7 Data sources for this thesis 39

1.7.1 Objective 1 39
1.7.2 Objective 2 39
1.7.3 Objective 3 40

Chapter 2. Peripheral complement proteins in schizophrenia: A systematic review and meta-analysis of serological studies 41

2.1 Introduction 44
2.1.1 Rationale 44
2.1.2 Objective 46

2.2 Methods 46
2.2.1 Protocol and registration 46
2.2.2 Eligibility criteria 46
2.2.3 Information sources and search strategy 47
2.2.4 Study selection 47
2.2.5 Data collection process 47
2.2.6 Risk of bias assessment 48
2.2.7 Meta-analyses 48

2.3 Results 49
2.3.1 Study selection 49
2.3.2 Study characteristics 51
2.3.3 Qualitative synthesis of results 61
2.3.4 Meta-analyses 64

2.4 Discussion 67
2.4.1 Methodological diversity 67
2.4.2 Clinical diversity 68
2.4.3 Evidence in context 69
2.4.4 Limitations 72
2.4.5 Conclusions 72

Chapter 2 References 74

Chapter 2 Appendices 85
Appendix 2.1: Search terms used for each database 86
Appendix 2.2: Risk of bias assessment based on the Newcastle-Ottawa scale for case-control studies 88
Appendix 2.3: Assessment of publication bias and small-study effects 89
Appendix 2.4: Results of meta-regression analyses 90

Chapter 3. Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence 91

3.1 Introduction 97

3.2 Methods 97

3.2.1 Study 1: CHR sample 97
3.2.2 Study 2: general population sample 99
3.2.3 Data and statistical analyses 100
3.2.4 Predictive models 101
3.2.5 Supplementary analyses 102

3.3 Results 102
3.3.1 Study 1: CHR sample 102
3.3.2 Study 2: general population sample 116
3.3.3 Supplementary analyses 117

3.4 Discussion 117

Chapter 3 References 126

Chapter 3 Appendices 132
Appendix 3.1: Supplementary Methods 136
Appendix 3.2: Supplementary Tables 149
Appendix 3.3: Supplementary Figures 216

Chapter 4. Plasma polyunsaturated fatty acids and mental disorders in adolescence and early adulthood: cross-sectional and longitudinal associations in a general population cohort 242

4.1 Introduction 246

4.2 Methods 248
4.2.1 Participants and study design 248
4.2.2 Exposures 248
4.2.3 Outcomes 249
4.2.4 Confounders 249
4.2.5 Statistical analyses 250
4.2.6 Ethical approval and consent 252

4.3 Results 252
4.3.1 Sample characteristics and descriptive analyses 252
4.3.2 Logistic regression analyses 257
4.3.3 Sensitivity analyses 264

4.4 Discussion 264
4.4.1 Strengths and limitations 268
4.4.2 Conclusions 269

Chapter 4 References 271

Chapter 4 Appendices 277
Appendix 4.1: Supplementary Methods 278
Appendix 4.2: Supplementary Tables 283
Appendix 4.3: Supplementary Figure 299

Chapter 5. General Discussion 300

5.1 Review of main findings 301

5.1.1 Peripheral complement proteins in schizophrenia: a systematic review and meta-analysis of serological studies 301

5.1.2 Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence 301

5.1.3 Plasma polyunsaturated fatty acids and mental disorders in adolescence and early adulthood: cross-sectional and longitudinal associations in a general population cohort 302

5.2 Evidence and implications 302

5.2.1 Theme 1: The complement system in the pathophysiology of psychosis 303

5.2.2 Theme 2: Proteomic biomarkers for individualised prediction of psychosis risk 307

5.2.3 Theme 3: Polyunsaturated fatty acids and prevention 309

5.3 An integrated view of findings and implications from this thesis 311

5.4 Future directions 314

5.5 Methodological considerations 321

5.6 Conclusion 323

References for Chapter 1 and Chapter 5 324

List of Abbreviations

A2M	Alpha-2-macroglobulin
ALA	Alpha-linolenic acid
ALSPAC	Avon Longitudinal Study of Parents and Children
ANCOVA	Analysis of covariance
AUC	Area under the receiver-operating characteristic curve
AUDIT	Alcohol Use Disorders Identification Test
AUDIT-C	Alcohol Use Disorders Identification Test Concise
BLIPS	Brief Limited Intermittent Psychotic Symptoms
BMI	Body mass index
BPRS	Brief Psychiatric Rating Scale
C3	Complement component 3
C4	Complement component 4
C4BPA	C4b-binding protein alpha chain
C6	Complement component 6
CAARMS	Comprehensive Assessment of At Risk Mental States
CHR	Clinical high-risk for psychosis
CHR-NT	Clinical high-risk individuals who do not transition to psychosis
CHR-T	Clinical high-risk individuals who transition to psychosis
CI	Confidence Interval
CIS-R	Clinical Interview Schedule Revised
CRP	C-reactive protein
CV	Coefficient of variation
DDA	Data-dependent acquisition
DHA	Docosahexaenoic acid
DIA	Data-independent acquisition
DSM	Diagnostic and Statistical Manual

EEG	Electroencephalography
ELISA	Enzyme-linked immunosorbent assay
EPA	Eicosapentaenoic acid
EU-GEI	European Network of National Schizophrenia Networks studying Gene-Environment Interactions
FDR	False discovery rate
FEP	First-episode psychosis
GAD	Generalised anxiety disorder
GAF	General Assessment of Functioning
ICD	International Classification of Diseases
IGHM	Immunoglobulin heavy constant mu
IgM	Immunoglobulin M
IL-6	Interleukin-6
LFQ	Label-free quantification
MADRS	Montgomery-Åsberg Depression Rating Scale
MBL	Mannose-binding lectin
MS	Mass spectrometry
n-3	Omega- 3
n-6	Omega-6
NAPLS	North American Prodrome Longitudinal Study
NPV	Negative predictive value
OR	Odds ratio
PE	Psychotic experience
PLIKSi	Psychosis-Like Symptoms Interview
PPV	Positive predictive value
PUFA	Polyunsaturated fatty acid
QC	Quality control
SANS	Scale for Assessment of Negative Symptoms
TNF-Ŭ	Tumour necrosis factor alpha

List of Tables

Chapter 2

Table 2.1 Summary of included studies using activity-based assays é é é é é é .. 52

Table 2.2 Summary of included studies using concentration-based assays é é é .. 54

Chapter 3

Table 3.1 Sample characteristics for CHR-T and CHR-NT groups in the initial experiment é 104

Table 3.2 Performance metrics for unadjusted support vector machine models é . 109

Table 3.3 Ten percent highest-weighted features for Model 1a, Model 3 and Model 4 (ranked according to mean feature weight for models selected in cross-validation inner loop) é . 112

Chapter 4

Table 4.1 Characteristics of age 17 and age 24 samples according to mental disorder status é . 254

Table 4.2 Cross-sectional associations between plasma polyunsaturated fatty acid (PUFA) measures and odds of mental disorders at 17 years é é é é é é é é é é . 258

Table 4.3 Cross-sectional associations between plasma polyunsaturated fatty acid measures and odds of mental disorders at 24 years é é é é é é é é é é é é é 260

Table 4.4 Longitudinal associations between plasma polyunsaturated fatty acid measures at 17 years and odds of incident mental disorders at 24 years é é é é 226

List of Figures

Chapter 1

Figure 1.1 A psychosis continuum: an example of a p 22

Chapter 2

Figure 2.1 PRISMA flow chart showing search strategy and results é é é é é é 50

Figure 2.2 Forest plot for studies measuring peripheral complement component 3 concentration é . 65

Figure 2.3 Forest plot for studies measuring peripheral complement component 4 concentration é . 66

Chapter 3

Figure 3.1 Derivation of participants included in the initial EU-GEI mass spectrometry experiment and their provision of plasma samples é é é é é é é é . 98

Figure 3.2 Predicting transition to psychosis using clinical and proteomic data (model 1a) é 108

Chapter 4

Figure 4.1 Odds ratios for cross-sectional associations between plasma polyunsaturated fatty acid measures and mental disorders at 17 years é é é é é 259

Figure 4.2 Odds ratios for cross-sectional associations between plasma polyunsaturated fatty acid measures and mental disorders at 24 years é é é é é 261

Figure 4.3 Odds ratios for longitudinal associations between plasma polyunsaturated fatty acid measures at 17 years and incident mental disorders at 24 years é 263

Chapter 5

Figure 5.1 Major themes of evidence and associated implications arising from the studies included in this thesis é . . . 303

Figure 5.2 An hypothesised longitudinal trajectory of complement activity in psychosis development with examples of risk factors and preventative opportunities é 313

Acknowledgements

I firstly wish to thank my inspirational supervisors **Professor David Cotter** and **Professor Mary Cannon.** I am eternally grateful for the kindness, encouragement and guidance they have shown me over the last three years (and no doubt further into the future!). I could not ask for better supervisors and I thank you from the bottom of my heart for everything you have done for me.

I have been so fortunate to be based in a truly excellent research team and this work is the reflection of a collective effort. I owe a debt of gratitude to **Dr Melanie Föcking** for her expertise, thoughtfulness and advice at all stages. I also thank **Subash Raj Susai** whose wisdom, dedication and, above all, friendship were sincerely appreciated.

I owe special thanks to **Dr Colm Healy** (my statistical psychotherapist) without whom I would be lost! He always answered every query with a calm clarity of thought for which I am very grateful. I also thank **Dr Emmet Power** for his patient advice and listening ear. Every chat and opportunity to discuss ideas (or simply ventilate!) was always greatly appreciated.

I thank **Catriona O'Donoghue** for her support over the past three years. From the moment I first contacted the team she has been so helpful with matters big and small.

I wish to thank our collaborators, including those in University College Dublin (especially **Dr Gerard Cagney** and **Dr Kieran Wynne**), the University of Bristol (**Professor Stan Zammit** and **Dr Hannah Jones**) and Cardiff University (**Dr Meike Heurich-Sevcenco**), for their immensely helpful assistance and advice. Thank you also to **Dr Ciaran Mulholland** who set me on this path and ignited my passion for research.

I thank the anonymous reviewers who critically appraised the studies in this thesis for their insightful comments which helped to improve the presented manuscripts.

I thank my friends and colleagues in RCSI including **Sophie Sabherwal, Niamh Dooley, Ross Brannigan, Eleanor Carey, Helen Coughlan, Aisling O'Neil, Donal Campbell, Kathryn Yates, Ulla Lång, Lorna Staines, Jonah Byrne** and **Austin Lacey.** Thank you all for your support and for making my time in RCSI so enjoyable.

Thank you also to all the staff in the **Department of Psychiatry** in RCSI and Beaumont Hospital. I have learned so much from your tutelage and will seek to always emulate your professional, caring and patient-centred approach.

I thank everyone in the **Irish Clinical Academic Training (ICAT) programme** including the directors for their advice, feedback and guidance at all stages, and the other fellows for their invaluable peer support. I am so grateful for the opportunities that have opened to me through this wonderful programme. Special thanks to **Dr Karen Misstear** who has been an indefatigable source of help from the beginning of my ICAT journey. Thank you also to the staff in the **Royal College of Surgeons in Ireland**, particularly the School of Postgraduate Studies, for making the process so smooth.

I thank the **Wellcome Trust and the Health Research Board (Grant Number 203930/B/16/Z), the Health Service Executive National Doctors Training and Planning and the Health and Social Care, Research and Development Division, Northern Ireland** who provide funding and support for the ICAT programme.

I thank all the clinical and research staff who contributed to the **European Network of National Schizophrenia Networks studying Gene-Environment Interactions (EU-GEI)** study. I am also extremely grateful to the **Avon Longitudinal Study of Parents and Children (ALSPAC)** team which includes midwives, interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

I wish to thank in a very special way all the **participants and families** who contributed their precious data, samples and time to the EU-GEI and ALSPAC studies. This research is only possible because of their generosity and their contributions are truly invaluable.

Finally, I wish to thank my family and friends. I could not have undertaken this journey without your help. To my dad **Paul**, my mum **Vera** and my sister **Bria,** thank you for your love, care and support all my life. I am so fortunate to have such a loving family.

Dedication

This thesis is dedicated to my grandmothers Josie Cassidy and Sadie Mongan, who passed away during its completion.

ñBelieve that further shore
Is reachable from here.
Believe in miracles
And cures and healing wells.ò

The Cure at Troy, Seamus Heaney

Summary

Background: Psychotic disorders such as schizophrenia are severe mental illnesses associated with significant burden for individuals and society. Investigation of the molecular pathophysiology of psychosis may inform prediction and prevention strategies. This thesis focuses on the potential roles of the complement system and polyunsaturated fatty acids (PUFAs).

Methods: Firstly, a systematic literature review and meta-analyses were performed to synthesise results of studies measuring peripheral complement system protein activities or concentrations in patients with schizophrenia compared to controls.

Secondly, plasma proteomic analyses were performed in two nested case-control studies. The first comprised a subsample of 133 clinical high-risk (CHR) for psychosis participants in the European Network of National Schizophrenia Networks studying Gene-Environment Interactions (EU-GEI) longitudinal study, of whom 49 (36.8%) developed psychosis. The second comprised a subsample of 121 participants in the Avon Longitudinal Study of Parents and Children (ALSPAC) who did not report psychotic experiences (PEs) at age 12 years, of whom 55 (45.4%) had PEs at age 18 years. Machine learning algorithms were used to predict transition to psychosis in EU-GEI and development of PEs in ALSPAC.

Thirdly, plasma PUFA measures were assessed using nuclear magnetic spectroscopy in samples provided by ALSPAC cohort participants at age 17 years and age 24 years. Cross-sectional and longitudinal associations with mental disorders (psychotic disorder, depressive disorder and generalised anxiety disorder [GAD]) were evaluated using logistic regression, adjusting for key confounders.

Results: There was little evidence of a consistent pattern of complement dysregulation in patients with schizophrenia relative to controls in the existing literature, although there was marked heterogeneity between studies.

In the EU-GEI subsample, a model based on baseline clinical and proteomic data demonstrated excellent performance for prediction of transition to psychosis (area under

the receiver-operating characteristic curve [AUC] 0.95). Functional analysis of differentially expressed proteins implicated the complement and coagulation cascade. A model based on the 10 most predictive proteins accurately predicted transition status in training (AUC 0.99) and test data (AUC 0.92). In the ALSPAC subsample, a model using age 12 proteomic data predicted PEs at age 18 years with AUC 0.74.

Based on cohort-wide data in ALSPAC, at age 24 years the omega-6:omega-3 PUFA ratio was positively associated with psychotic disorder, depressive disorder and GAD, while docosahexaenoic acid (DHA) was inversely associated with psychotic disorder. In longitudinal analyses, there was evidence of an inverse association between DHA at age 17 years and incident psychotic disorder at age 24 years (adjusted odds ratio 0.44, 95% confidence interval 0.22 ï 0.87) with little evidence for depressive disorder or GAD.

Conclusions: Complement system dysregulation is detectable prior to the development of early psychosis phenotypes. Proteomic biomarkers may augment individualised prediction of outcomes in the clinical high-risk subpopulation. In the general population, higher levels of DHA in adolescence may be associated with reduced odds of psychosis in early adulthood, suggesting a potentially preventative role. These findings have implications regarding the pathophysiology, prediction and prevention of psychotic disorders.

List of scientific papers for consideration in this thesis

#	Reference	Impact factor	Quartile*
1.	**Mongan D**, Sabherwal S, Susai SR, Föcking M, Cannon M, Cotter DR. Peripheral complement proteins in schizophrenia: A systematic review and meta-analysis of serological studies. *Schizophrenia Research* 2020 Aug;222:58-72. doi: 10.1016/j.schres.2020.05.036.	3.8	Q1
2.	**Mongan D**, Föcking M, Healy C, Susai SR, Heurich M, Wynne K, Nelson B, McGorry PD, Amminger GP, Nordentoft M, Krebs MO, Riecher-Rössler A, Bressan RA, Barrantes-Vidal N, Borgwardt S, Ruhrmann S, Sachs G, Pantelis C, van der Gaag M, de Haan L, Valmaggia L, Pollak TA, Kempton MJ, Rutten BPF, Whelan R, Cannon M, Zammit S, Cagney G, Cotter DR, McGuire P, European Network of National Schizophrenia Networks Studying Gene-Environment Interactions (EU-GEI) High Risk Study Group. Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence. *JAMA Psychiatry* 2021;78(1):77-90. doi: 10.1001/jamapsychiatry.2020.2459.	17.5	Q1
3.	**Mongan D**, Healy C, Jones HJ, Zammit S, Cannon M, Cotter DR. Plasma polyunsaturated fatty acids and mental disorders in adolescence and early adulthood: cross-sectional and longitudinal associations in a general population cohort. *Translational Psychiatry* 2021;11:321. doi: 10.1038/s41398-021-01425-4	5.3	Q1

*Quartiles are based on Scrimago Journal Ranking (https://www.scimagojr.com/).

Related scientific papers not in consideration

#	Reference	Impact factor	Quartile*
1.	**Mongan D,** Ramesar M, Föcking M, Cannon M, Cotter DR. Role of inflammation in the pathogenesis of schizophrenia: A review of the evidence, proposed mechanisms and implications for treatment. *Early Intervention in Psychiatry* 2020 Aug;14(4):385-397. doi: 10.1111/eip.12859.	2.3	Q2
2.	**Mongan D**, Cannon M, Cotter DR. COVID-19, hypercoagulation and what it could mean for patients with psychotic disorders. *Brain, Behavior, and Immunity*. 2020 Aug; 88: 9ï 10. doi: 10.1016/j.bbi.2020.05.067.	6.6	Q1
3.	**Mongan D,** Yin X, Brennan L, Cotter DR. Dysregulation of lipid metabolism precedes psy metabolic disorder? *Biological Psychiatry* 2021 Feb 1; 89(3):209-211. doi: 10.1016/j.biopsych.2020.10.017.	12.1	Q1
4.	Susai SR, Sabherwal S, **Mongan D**, Föcking M, Cotter DR. Omega-3 fatty acid in ultra-high-risk psychosis: A systematic review based on functional outcome. *Early Intervention in Psychiatry* 2021 Mar 02. doi: 10.1111/eip.13133.	2.3	Q2
5.	Heurich M, Föcking M, **Mongan D**, Cagney G, Cotter DR. Dysregulation of complement and coagulation pathways: emerging mechanisms in the development of psychosis. *Molecular Psychiatry* (accepted 10th June 2021)	10.2	Q1

*Quartiles are based on Scrimago Journal Ranking (https://www.scimagojr.com/).

Conference presentations

Oral presentations	
April 2021	Polyunsaturated fatty acids and mental disorders: cross-sectional and longitudinal associations in a large general population cohort ï Schizophrenia International Research Society 2021 Congress
April 2021	Development of proteomic prediction models for transition to psychotic disorder from the clinical high-risk state and psychotic experiences in adolescence ï Schizophrenia International Research Society 2021 Congress
March 2021	Polyunsaturated fatty acids and mental disorders: cross-sectional and longitudinal associations in a large general population cohort ï Rank Prize Symposium
October 2020	Development of proteomic prediction models for transition to psychotic disorder from the clinical high-risk state ï Beaumont Hospital Translational Research Awards
April 2019	Proteomic biomarkers of psychosis ï College of Psychiatrists in Ireland Spring Conference
Poster presentations	
March 2021	Polyunsaturated fatty acids and mental disorders: cross-sectional and longitudinal associations in a large general population cohort ï Longitudinal Studies 2021
April 2019	Complement proteins in schizophrenia: a systematic review protocol ï British Neuroscience Association Festival of Neuroscience

Academic prizes and awards

Date	**Award and organisation**
April 2021	Academic Trainees Conference best poster prize ï Royal College of Psychiatrists
April 2021	Early Career Award ï Schizophrenia International Research Society Congress 2021
March 2021	Best Early Career Research Presentation ï Rank Prize Symposium
December 2020	Student Engagement Programme funding award ï Royal College of Surgeons in Ireland
December 2020	CEO Innovation Prize ï Royal College of Surgeons in Ireland
November 2020	Best Research Presentation Prize ï Irish Clinical Academic Training (ICAT) programme annual retreat 2020
October 2020	Sheppard Prize for Translational Research ï Beaumont Hospital Translational Research Awards
November 2019	Best Research Presentation Prize ï Irish Clinical Academic Training (ICAT) programme annual retreat 2019

Chapter 1. Introduction

Chapter overview

This chapter introduces and defines the major topics of this thesis including the phenotypic continuum of psychosis. The text contains extracts from a narrative review entitled ó Role of inflammation in the pathoge evidence, proposed mechanisms and implications *Early Intervention in Psychiatry.* Dr Mongan was the first author of this article, drafted the initial submission and composed the final version of the manuscript following peer review.

Reference: Mongan D, Ramesar M, Föcking M, Cannon M, Cotter DR. Role of inflammation in the pathogenesis of schizophrenia: A review of the evidence, proposed mechanisms and implications for treatment. *Early Intervention in Psychiatry* 2020 Aug;14(4):385-397. doi: 10.1111/eip.12859.

1.1 Psychotic disorders

Psychotic disorders are severe mental illnesses characterised by loss of contact with reality. The cardinal symptoms of psychosis are conceptualised as three main groups: positive symptoms (such as thought disorder, hallucinations and delusions); negative symptoms (such as anhedonia and social withdrawal); and cognitive symptoms (such as attention and memory difficulties). These disorders are often associated with long-term functional impairment and disability for the affected individual (1).

Psychotic disorders give rise to combinations of these types of symptoms which vary in onset, duration, intensity and severity both between and within individuals over time. Examples of psychotic disorders as defined by the International Classification of Diseases (ICD) and Diagnostic and Statistical Manual (DSM) include schizophrenia, schizoaffective disorder, acute and transient psychotic disorder and delusional disorder (2, 3). In these ó primary psychosis is considered a dersô, t primary component of the disorder itself (4). These conditions form the focus of this thesis. It should however be noted that a state of psychosis can also occur secondary to other factors, including other psychiatric syndromes (such as in a manic episode of bipolar affective disorder, or in an episode of severe depression); non-psychiatric disorders (for example stroke, epilepsy or brain neoplasm); and use of certain medications or substances (including amphetamines and cannabis).

A meta-analysis of international studies reported a pooled median point prevalence for psychotic disorders of 3.9 per 1000 persons and median lifetime prevalence of 7.5 per 1000 (5). However, prevalence estimates vary by study design, assessment procedures and geographic location (6). For schizophrenia, males may be slightly more likely to be affected than females with a rate ratio of 1.4:1 (7), although evidence of no differences between sexes has also been reported (8). From a public health perspective, psychotic disorders account for substantial disability at individual and population levels, particularly given their often chronic and relapsing-remitting course. In the 2016 Global Burden of Disease Study, schizophrenia alone accounted for 13.4 million years of life lived with disability (8). There is also evidence that life expectancy in patients with schizophrenia compared to the general population is quantitatively reduced by

approximately 20 years (9). This may be related to environmental factors such as tobacco smoking and other substance use; co-morbid conditions such as obesity and cardiovascular disease; metabolic side-effects of antipsychotic medications; or as a result of biological processes inherent to the disorder itself (10-12).

In primary psychotic disorders, reduction in psychotic symptoms is often the main therapeutic goal in the initial acute stages of treatment. The most common pharmacological interventions are atypical or typical antipsychotic medications which generally (though not exclusively) target dopaminergic pathways in the brain that are thought to underlie the symptoms of psychosis. These are among the most effective available treatments for positive psychotic symptoms, but do not adequately treat negative symptoms or functional deficits (13). Psychological interventions such as cognitive behavioural therapy, family therapy and cognitive remediation, as well as social support for housing, education, substance use, employment and training, are also important components of management (14). All management strategies should be specifically tailored to the strengths and needs of the individual, their family and carers, and their wider social environment.

1.2 The phenotypic continuum of early psychosis

Given the long-term adverse consequences associated with psychosis, there has been a focus on characterising its longitudinal development with particular emphasis on early phenotypes to facilitate early detection, intervention and, potentially, prevention of these severe disorders. The longitudinal development of psychosis is complex, not least because initial presentations, illness trajectories and longitudinal outcomes are markedly heterogeneous (15, 16). For example, a meta-analysis of outcomes following first episode of psychosis found a pooled remission rate of 58% and a recovery rate of 38% (17). While there is often initial uncertainty regarding the underlying diagnosis following the first episode of psychosis, there is evidence that schizophrenia-spectrum and affective psychoses have high prospective diagnostic stability (18). Greater diagnostic variation has been found for other ICD or DSM psychotic disorders including schizoaffective disorder, substance-induced psychotic disorder, delusional disorder,

brief psychotic disorder, psychosis not otherwise specified and schizophreniform disorder (18).

Despite this heterogeneity in terms of outcomes, it is nevertheless possible to conceptualise a generalised framework that represents one possible early psychosis phenotypic trajectory as represented in Figure 1.1**.** While this risk trajectory is conceptualised as a continuum, it is also possible to distil certain discrete phases of phenotypic development. This continuum can also be viewed as occurring over time, with gradual progression from one phase to another, but not necessarily so. This should not be viewed as a deterministic inevitable march towards chronic psychotic disorder, nor as the only possible óvesas one possible o psychosis. probabilistic pathway of escalating risk associated with increasing phenotypic intensity. This model is necessarily a heuristic simplification but is a useful conceptual framework by which to explore several phases of early psychosis phenotypes. There follows an exploration of these individual phases in greater detail.

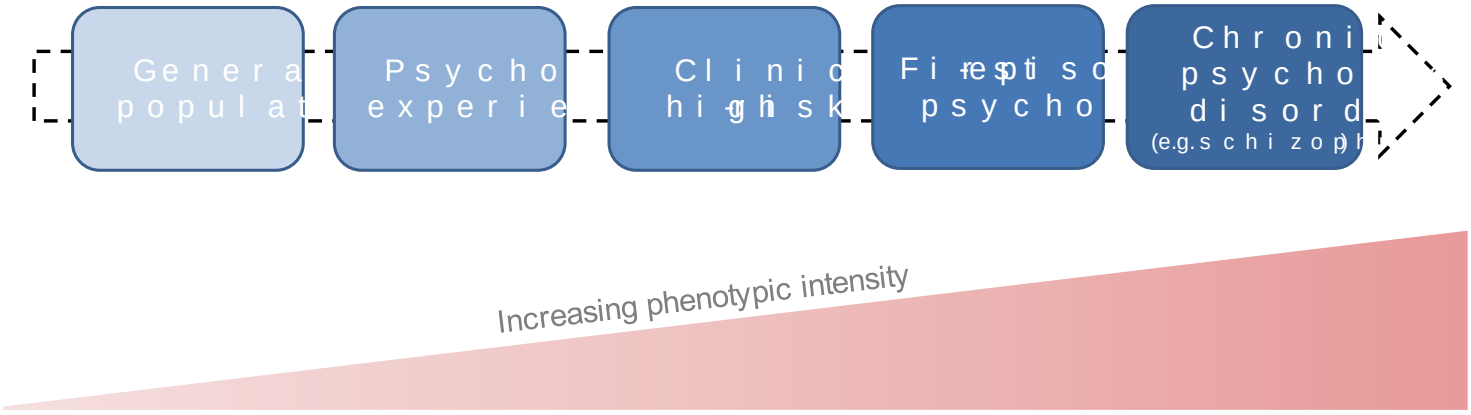


Figure 1.1 A psychosis continuum: an example of a psychosis risk trajectory.
This model is a representation of one possible phenotypic spectrum of psychosis. It serves as a conceptual framework and not necessarily as a deterministic progression of illness. The underlying spectrum represents increasing phenotypic intensity. The arrow can be thought of as representing time, although this is not always the case. For example, an individual may move dynamically from having subclinical psychotic experiences to fulfilling the clinical high-risk criteria, but may remit from fulfilling the criteria after several years without developing a first episode of psychosis (indeed most people who meet clinical high-risk criteria will not develop a first episode of psychosis on follow-up).

1.2.1 First episode psychosis

The first episode of psychosis (FEP) most commonly occurs in late adolescence or early adulthood, often before the age of 25 years (19). This is an important period of normative biopsychosocial development which includes, among multiple processes, the persistence of synaptic pruning in the brain, cortical maturation, establishment of personal identity, formation of relationships, educational attainment and initial steps in employment (20). The disability associated with psychosis and its attendant positive, negative and cognitive symptoms can hence significantly derail an individualôs long-term developmental trajectory, compounding the impact of adverse psychosocial effects and enduring functional impairment (21, 22).

A key finding of FEP research is that the duration of untreated psychosis is associated with adverse outcomes: the longer the duration of untreated psychosis, the worse the clinical and functional outcomes for patients (23, 24), even over relatively short timeframes. In a review of meta-analyses, a period of four weeks of untreated psychosis was associated with 20% more severe symptoms at follow-up compared to one week (25). Such findings have spurred the development of early intervention services with the aim of identifying and treating FEP as quickly as possible, to shorten the duration of untreated psychosis and optimise outcomes for patients (26). Early intervention in psychosis is now considered a standard of modern mental healthcare (27).

1.2.2 The psychosis prodrome

It has been recognised since at least the time of Emil Kraepelin (28) that the first psychotic episode is typically preceded by a prodromal period of psychotic symptoms that are of attenuated intensity, quality, frequency, duration or functional impact. Identifying patients FEP nôt phhi assep rwoodurl odm, ali nort hóeporr secondary preventative strategies focused on detecting the development of disorder at the earliest possible stage. However, a prodromal state can only be confidently diagnosed in retrospect, after the diagnosis of FEP is made. If the aim therefore is for secondary prevention of psychosis, it is instead necessary to characterise a subpopulation at increased risk for psychosis on a prospective basis.

1.2.3 Clinical high-risk criteria

Beginning initially in Australia, longitudinal research in early psychosis phenotypes led to the development of criteria to define a subpopulation of help-seeking individuals presenting with early psychotic symptoms and/or other risk factors who are at enhanced risk of developing FEP (29). These criteria are variously kn Stateô, -Ôrisk ô a o Hi ó-Ôrisk ô c (a CHR) High criteria for psy consistency, the author refers to the latter term throughout this thesis. Measurement tools based on semi-structured clinical interview were subsequently developed and validated to confirm the presence of the CHR state. Tools in most common use include the Comprehensive Assessment of At Risk Mental States (CAARMS) (30), the Structured Interview for Psychosis-risk Syndrome (31) and the Basel Screening Instrument for Psychosis (32).

The precise criteria defining those at CHR vary slightly depending on the specific instrument used, but generally comprise three main subgroups of help-seeking individuals aged 14 to 35 years old (33):

1. An óattenuated psychotic symptomsô subgrou that are reduced in intensity and/or duration compared to the symptoms of FEP;
2. A ógenetic risk and deteriorationô subgrou schizophrenia or personal history of schizotypal personality disorder and functional impairment;
3. A óbrief limited intmesrômi(tBtLeInPtS)pssyucbhgortoiucp swi symptoms of similar intensity to those in FEP, but of diminished duration (less than one week at a time).

In most studies, the attenuated psychosis subgroup makes up the largest proportion of CHR individuals (approximately 85%) followed by the BLIPS subgroup (approximately 10%) and the genetic risk and deterioration subgroup (approximately 5%), although the BLIPS subgroup may have the highest risk of progression to psychotic disorder (34, 35).

With time, the utility of the CHR criteria for identifying individuals at risk of psychosis was evidenced in further international studies (36, 37). Using CHR clinical instruments,

transition to psychosis is typically defined as the occurrence of at least one positive psychotic symptom of full-threshold intensity, occurring several times a week for a period of more than one week (38). An umbrella review of meta-analyses reported an overall risk of development of FEP among CHR individuals of 22% at three years (36). While transition to FEP remains a focal point for prevention, there has also been increasing recognition that the CHR subpopulation are at risk of other adverse outcomes (even in those who do not transition to FEP), including persistent attenuated psychotic symptoms, affective disorders, anxiety disorders, substance use disorders and functional impairment (39, 40).

A major research objective relates to identifying specific predictors of transition to improve prediction of psychosis in CHR individuals. The CHR assessment alone has excellent ability to órule outô psychosis risk ó r u l i n g i n ô p s y c h o s i s r i s k i n i n d i v i d u a l s t e s t p o s i t i v e (i.e. high sensitivity but only moderate specificity) (41). At group level, several studies have previously identified a range of predictors of transition. These include clinical variables such as impairments in cognition and poor baseline social functioning, as well as environmental factors such as psychological trauma, bullying and cannabis use (42). This research has been supplemented by examination of potential biomarkers including electroencephalography, neuroimaging and blood markers (43). However, group-level differences are less useful for individualised prediction. In keeping with a move towards precision medicine, which advocates for prognostication and treatment tailored to the individual, studies have begun to focus on developing prediction algorithms based on a variety of clinical and neurobiological data modalities. These have had varying degrees of predictive accuracy (44-47) but none are yet in routine clinical use.

Regarding interventions to reduce the risk of transition, a meta-analysis of randomised clinical trials reported that cognitive behavioural therapy was associated with a reduction in incidence at 12 months (risk ratio [RR] 0.52, 95% confidence interval [CI] 0.33 ï 0.82) and 18 ï 48 months (RR 0.60, 95% CI 0.42 ï 0.84) (48). Pharmacological interventions studied to-date in randomised trials have included low-dose antipsychotics, D-serine and omega-3 fatty acids, but evidence for their effects on

transition risk is inconsistent (49). Several reviews have concluded that there is currently insufficient high-quality evidence to recommend the routine use of any specific intervention in place of or as adjuncts to needs-based or supportive interventions (49-51). Further evidence is thus required to inform recommendations for preventative interventions in the CHR population.

Several criticisms of the CHR concept have been advanced, including the lack of focus on symptoms other than positive psychotic phenomena, such as depressive and anxiety symptoms (52). The transition rate has fallen over time, with lower rates of transition generally reported in more recent studies (33). The reasons for the decline in transition rates are unclear, but could include improved detection and prevention, differences in the precise definition of CHR or ótransitionô (53). Further criticism relates to findings from a retrospective survey in South London, where an established early detection and psychosis prevention service exists, which found that most young adults presenting with FEP were not previously seen by the CHR service (54). However, this may reflect problems related to availability, accessibility and awareness of CHR services as currently constructed, rather than a limitation of the CHR concept itself (55). Indeed, one study found that more than 50% of individuals with FEP, on retrospective interview, reported symptoms consistent with a CHR state prior to the onset of full-threshold psychosis (56). The CHR subpopulation remains a help-seeking group that requires clinical care, and the concept serves as an important paradigm to enable clinical and research efforts focused on prevention of psychosis.

1.2.4 Psychotic experiences

Historically, positive psychotic symptoms were generally thought of as definitively abnormal phenomena. However, the CHR concept challenges this notion and implies a more fluid, continual model of psychosis liability. More recent research has found that experiences of thought disorder, delusions and hallucinations are in fact common in the general population (57). These phenomena have been variously described as ósubclinicalô, ósubthresholdô or ópsychotic-likeô symptoms or experiences. For consistency, the author refers to ópsychotic

In a meta-analysis, the median prevalence of PEs in the general population was estimated to be approximately 5% (58). This figure is often higher in younger age groups and especially in childhood and adolescence (59). PEs can be associated with varying degrees of distress or functional impairment. In approximately 75 ï 90% of cases, PEs are transient phenomena (58). However, a subset of individuals experience recurrence or persistence of PEs, which may be associated with greater risk of development of psychotic disorder (60).

There is now a body of research showing that PEs in childhood, adolescence or early adulthood are associated with an increased risk of adverse homotypic and heterotypic psychiatric outcomes (59). These relate not only to development of psychotic disorders but also of non-psychotic disorders such as depression and substance use as well as self-harm and suicide (61). In one meta-analysis of longitudinal community studies, individuals with PEs had four-fold increased risk of subsequently developing psychosis and three-fold risk of developing any mental disorder (59). Thus, rather than being confined to psychosis risk, PEs may represent transdiagnostic clinical indicators of severity of psychopathology more generally (57).

1.3 Inflammation and the complement system in the pathophysiology of psychosis

Investigating the molecular pathophysiology of psychosis may provide insights to inform detection, prediction and prevention strategies. In recent decades, the role of inflammation and the immune system has come into increasing focus. Inflammation is a complex biological cascade activated by a noxious stimulus. The process of inflammation will generally signal and activate the immune system to defend against such threats, with the aims of minimising tissue damage and preventing systemic spread. In the acute context, for example physical trauma or local infection, inflammation results in the classical, familiar symptoms of heat, redness, swelling and pain (62). When the activating stimulus is successfully withdrawn or treated, the acute inflammatory response appropriately subsides. In contrast, in the context of psychotic disorders, there is evidence for a chronic, low-grade activation of inflammation (63).

1.3.1 Inflammatory markers in psychosis

Measurement of peripheral inflammation and immune activation requires the use of markers such as cytokines. These are key signalling proteins that modulate inflammation and the immune response, several of which have been studied as potential biomarkers of schizophrenia. To date, research in psychosis has focused on measurement of acute inflammatory markers. For example, raised serum interleukin-6 (IL-6) and tumour necrosis factor-Ŭ (TŬNF), both inflammatory cytokines, are raised in acute psychotic relapses in patients with schizophrenia, and IL-6 levels have been found to reduce after anti-psychotic treatment of acute illness (64). Meta-analyses have demonstrated increased pro-inflammatory cytokines in patients with schizophrenia compared to controls (65). Similar findings have been noted in medication-naïve patients with first-episode psychosis (66) and in some studies of people at risk of psychosis (67, 68).

It is possible that increased inflammation may occur early in the course of the disorder (at least in a subset of affected patients) and is not entirely a result of later factors such as medication use. However, studies measuring inflammatory biomarkers in psychosis have often been cross-sectional in nature with the limitation that the temporal relationship between inflammation and psychosis cannot be directly inferred. Longitudinal studies have attempted to address this. For example, elevated serum C-reactive protein (CRP) at age 15 years has been associated with increased risk of developing schizophrenia by age 27 years (69). In a general population study, higher serum levels of IL-6 at age 9 years were associated with a twofold greater risk of psychotic disorder at age 18 years (70). Further longitudinal studies are required, but these findings suggest that the increase in inflammatory tone can occur at an early phase, preceding the onset of clinical-threshold overt psychotic symptoms.

Biomarker investigations add to the evidence regarding inflammation and immune dysregulation in psychosis, but may also have translational utility in terms of diagnosis or prognosis. For example, Perkins et al (71) identified a panel of 15 analytes (including several immunomodulatory cytokines) that together distinguished between people at

high risk of psychosis who did or did not transition to FEP. Similarly, raised IL12p40 has been associated with increased risk of transition in the CHR subpopulation (72).

1.3.2 The complement system in psychosis

The complement system is a network of plasma proteins representing a key component of the innate immune system for which evidence is accumulating to implicate a potential role in psychosis pathophysiology. The system functions as a cascade activated by foreign or damaged cells or debris along one of three molecular pathways (the classical, alternative and lectin pathways) that converge on a terminal pathway leading to formation of a membrane attack complex capable of lysing pathogenic surfaces (73). Complement system proteins have a range of further immune functions including opsonisation, promotion of phagocyte function, B cell activation, maintenance and elimination (74). The complement network also displays intimate functional connections and cross-talk with other related systems including the acute phase response and coagulation cascade (75).

While older in evolutionary terms compared to adaptive immune responses (76), the consequences of inappropriate activation, deficiency or dysregulation of the complement system can be no less deleterious in humans. When control of the system becomes dysregulated, or the response is inappropriately directed toward self-antigens, the complement system can contribute to disease states. Complement system dysregulation has been associated with several neurological disorders that have psychiatric manifestations including systemic lupus erythematosus (77), multiple sclerosis (78, 79) and epilepsy (80). The complement system is also relevant to the pathogenesis of neurodegenerative conditions such as Alzheimer's (81), where mouse models have shown that complement proteins and microglial cells contribute to synaptic loss (82).

There are several lines of evidence linking the complement system to the pathophysiology of psychosis (83). Several individual complement gene polymorphisms have been associated with schizophrenia (84). Variations in the major histocompatibility complex (a region which encodes several immune proteins, including proteins of the complement pathway) have been significantly associated with the disorder in genome-

wide association studies (85-87). In a landmark study, Sekar and colleagues showed that allelic variation of the complement component 4 (C4) gene was associated with schizophrenia risk (88). Moreover, the authors demonstrated that this correlated with increased levels of C4A mRNA in post-mortem brain tissue from patients with schizophrenia and that C4 protein was expressed at synapses. Studies have shown complement proteins complement component 1q (C1q) and complement component 3 (C3) to be expressed at synapses in the rodent brain and, together with microglia, these proteins mediate synaptic elimination in post-natal neurodevelopment (89, 90). It is thus plausible that aberration of the complement system at least partially explains the observation of reduced synaptic density in schizophrenia (91).

From a serological perspective, individual complement proteins (including C1, C1q, C2, C3, C4 and C5) in sera from patients with schizophrenia have shown increased haemolytic activity compared to controls (92). However, other studies have provided inconsistent findings. For example, a further study showed increased haemolytic activity of C1, C3 and C4 but decreased activity of C2 compared to controls and no mean difference in total complement haemolytic activity between patient and control samples (93). Studies have also measured complement proteins directly using techniques such as nephelometry and enzyme-linked immunosorbent assay. Several such studies have demonstrated increased levels of complement proteins such as C3 and C4 (94-96) although contradictory results have also been reported (97-99). Serological studies measuring complement system proteins have used multiple methods and examined patients at various stages of disorder or clinical status (for example acute psychosis compared to remission), which may account for some of these inconsistencies. Overall, it is unclear from serological studies whether peripheral complement markers are consistently dysregulated in patients with psychosis compared to controls.

Rather than patients with chronic psychosis, Kopczynska and colleagues measured 11 complement proteins by ELISA in serum samples from patients with FEP compared to controls (100). A logistic regression model implicated several complement analytes (including, among others, C3, C4 and C5) and the model demonstrated moderate performance in distinguishing cases and controls. However, when measured

individually, only one (the terminal complement complex) was found to differ significantly in patients compared to controls. In a study that used discovery-based proteomic techniques to measure proteins in blood samples taken at age 11 from young people who went on to develop psychotic disorder at age 18, several proteins were significantly altered in this group compared to controls who did not (101). The complement and coagulation systems were most significantly implicated. These findings suggest that dysregulation of the complement system occurs early and may be predictive of increased risk for disorder in later life. Indeed, increased levels of C1q have been found in mothers whose children developed psychosis compared to mothers whose offspring had no psychiatric disorder, suggesting that prenatal exposure may be relevant (102). However, it is not yet known whether complement biomarkers may helpfully inform clinical prediction of psychosis in the general population or high-risk subpopulations.

1.4 Biomarkers of psychosis risk

There is clear evidence that early intervention in psychosis is associated with improved outcomes (103). However, given the heterogeneity in terms of suspected aetiology and clinical course, markers of future risk (beyond early clinical symptoms alone) would be helpful to improve prognostication. Biomarker ectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic response (104). Biomarkers to a th may be used to inform diagnosis, prognosis, monitoring or prediction of illness course or therapeutic response. Conceptually, biomarkers could enable improved prognosis in early psychosis and stratification of interve profile and level of associated risk. This is already the case in other areas of clinical medicine where biomarkers are used to tailor treatment strategies and prognosis for the individual patient (105).

Accurate prediction of outcomes in early psychosis is a major goal of current research. The CHR paradigm represents a window of opportunity for prediction of FEP, but only approximately 20 ï 30% of CHR individuals will go on to develop psychosis (33). Individualised prediction may be improved with consideration of other data modalities

besides clinical information alone. As well as potentially providing pathophysiological insights as above, biomarkers that accurately predict onset of psychosis would have important clinical implications, potentially enabling stratification strategies based on psychosis risk. This could in turn facilitate research into (and eventually clinical implementation of) interventions that are personalised to the individual and informed by their underlying pathophysiology. The first necessary step is to demonstrate that biomarkers can usefully predict risk of development of FEP in at-risk individuals. To this end, previous studies in the existing literature have investigated markers from several different domains.

1.4.1 Electroencephalography

Differences in electroencephalography (EEG) microstates, spectral power and alpha peak frequencies have been detected in CHR individuals who transition to psychosis (CHR-T) compared to those who do not (CHR-NT) (106, 107). Other studies have found no evidence of quantitative EEG differences in CHR or FEP compared to controls (108). Differences in EEG-based event-related potential measures such as P300 and mismatch negativity have also been reported in CHR-T compared to CHR-NT (109). However, EEG studies are often limited by sample size. Difficulties related to data acquisition, confounding effects and interpretability may further limit clinical application. Furthermore, most EEG studies to-date have reported group-level differences rather than modelling predictive performance (109).

1.4.2 Neurocognition

Multiple studies have reported global and specific cognitive deficits in CHR individuals compared to controls at group level (110-113). Spatial memory significantly predicted transition in CHR in one study (114). However, a larger study of CHR individuals within the North American Prodrome Longitudinal Study (NAPLS) cohort found neurocognitive data did not usefully inform prediction of psychosis beyond clinical criteria (although worse verbal memory predicted more rapid conversion) (110). A meta-analysis of nine neurocognitive studies concluded that CHR-T performed significantly worse than CHR-

NT on working memory and visual learning (115), but whether these group-level differences translate into improved prediction of transition outcome remains uncertain.

1.4.3 Neuroimaging

Imaging studies provide evidence of neuroanatomical abnormalities in CHR-T compared to CHR-NT, including grey matter volume reductions in areas such as the anterior cingulate, frontal cortex and temporal cortex; differences in white matter integrity; and larger pituitary volume (116-119). Functional magnetic resonance imaging studies have reported evidence of aberrant network connectivity in CHR (120). Machine learning techniques have been used in combination with neuroimaging data to generate outcome prediction models in CHR with relatively high balanced accuracy (121-123). However, imaging techniques produce millions of variables per participant which can lead to risk of overfitting and limited generalisability, notwithstanding concerns regarding cost and accessibility.

1.4.4 Cortisol

Previous studies have reported increased activity of the hypothalamic-pituitary axis in schizophrenia (124). Blunted cortisol awakening responses have been reported in CHR compared to healthy controls (125). Initial findings from the NAPLS cohort suggested higher baseline levels of salivary cortisol in CHR-T compared to CHR-NT (126). However, subsequent studies examining salivary or blood cortisol levels have been inconsistent (127). A meta-analysis found that CHR cortisol levels were increased compared to controls, but did not discriminate between CHR-T and CHR-NT (128) suggesting limited utility as a predictive biomarker of psychosis risk.

1.4.5 Peripheral blood markers

Blood-based tests have a potential advantage over other investigations in terms of accessibility in the clinical setting. Informed by the evidence implicating inflammation in schizophrenia and other psychotic disorders (129), investigations of blood markers of transition risk in CHR have focused on measurement of markers of acute inflammation including cytokines such as C-reactive protein and the interleukin family. A meta-analysis found that blood IL-6 levels were higher and interleukin-1β levels were lower in

CHR compared to controls (68). With respect to prediction of transition, previous studies have used blood-based molecular analytes reflecting a range of biological processes (typically related to inflammation and the immune system) to build predictive models with area under the receiver-operating characteristic curve (AUC) greater than 0.80 (130, 131).

1.4.6 Predictive models

In an effort to improve individualised prediction of psychosis, several studies have advanced upon analysis of group-level differences by developing predictive models using classical statistical methods or machine learning techniques. A systematic review of models for psychosis in CHR reported results of 38 primary studies (45). These included 11 studies of clinical models, five studies of biological models (including blood-based, EEG and neuroimaging markers), five studies of neurocognitive models, five studies of environmental models and 18 studies using combinations of these. The PPV ranged from 24% to 88% for clinical models; 78% to 83% for biological models; 57% to 83% for neurocognitive models; 26% to 63% for environmental models; and 49% to 86% for combined models. To date, no biomarker-based predictive models are in routine use in clinical practice for prediction of transition to psychosis and further longitudinal investigations are warranted.

1.5 Polyunsaturated fatty acids and prevention

Accurate prediction of onset and outcomes of mental disorders is clearly a laudable aim. However, investigation of risk and preventative factors is also crucial to better understand their early pathophysiology and to inform prevention strategies. Such strategies can be aimed at the level of the general population (universal prevention), subpopulations with identified risk factors (selected prevention) or individuals displaying early subclinical symptoms of the disorder (indicated prevention) (132). Taking schizophrenia as an example, universal strategies might include public awareness campaigns and legislation regarding control of substances such as cannabis; selected prevention could include active monitoring of individuals with positive family history; while indicated prevention is exemplified by the CHR paradigm.

In line with a growing recognition of the bidirectional interplay between nutrition and mental health (133, 134), the role of bioactive dietary factors is increasingly being investigated as a potential mechanism by which risk of mental disorders may be modulated at population, subgroup and individual levels. For example, prenatal and early childhood dietary quality has been associated with risk of internalising and externalising problems in childhood (135). Such insights can helpfully inform prevention strategies by, for example, adapting universal or age-specific dietary guidelines for the general population or recommending dietary interventions in high-risk groups.

Among macronutrients examined in the existing literature, polyunsaturated fatty acids (PUFAs) have received much attention in the field of prevention of mental disorders. Omega-6 (n-6) PUFAs generally produce pro-inflammatory mediators, while omega-3 (n-3) PUFAs produce broadly anti-inflammatory mediators (136). Given the evidence implicating inflammation in the pathophysiology of several mental disorders, interest in PUFAs as risk-modulating agents in this context arises in part because of their roles in altering inflammatory tone (136). However, in addition to modulation of inflammation, n-3 fatty acids have multiple further potentially relevant neurobiological effects including antioxidant properties, modulation of synapse formation and neuroprotective effects (137). Omega-3 PUFAs must be obtained from the diet (for example in flaxseed, oily fish and supplements) because they cannot be synthesised in adequate amounts by humans. Of the n-3 PUFAs, docosahexaenoic acid (DHA) is the most abundant in the human brain (138) and as such is among the most-studied in relation to mental health.

Multiple previous cross-sectional studies have demonstrated lower peripheral levels of n-3 PUFAs in people with psychotic disorders, as well as depressive disorder and anxiety disorders, compared to controls (139-141). Cross-sectional associations with depressive and anxiety symptoms in women in the general population have been previously noted in relation to self-reported DHA intake (142). However, an important limitation of cross-sectional studies is that the temporal relationship between exposure and outcome is not definitively known. Given the relative sparsity of longitudinal research, it is unclear whether PUFA abnormalities precede the onset of mental disorders in the general population. Several cohort studies have reported associations

between self-reported fish intake and subsequent depression (143). A study of young adults found an inverse association between fish consumption and subsequent depression in women but not in men (144). Generally, longitudinal studies analysing PUFA biomarkers (rather than self-reported intake) have not provided consistent evidence for associations with depression in the general population (145, 146). Furthermore, trials of PUFA supplements as therapies for mental disorders have produced mixed results (147).

Focusing on psychosis, the use of n-3 PUFA supplementation for prevention has been studied in two prominent randomised controlled trials. The Vienna High Risk trial found that CHR individuals randomised to receive omega-3 PUFAs had lower risk of transition to FEP compared to those randomised to placebo (148). However, in a subsequent larger multicentre trial (the NEURAPRO study) these findings were not replicated (149). The investigators reported that participant compliance was poor, which might partially explain this discrepancy. A subsequent analysis of the NEURAPRO cohort found that plasma n-3 PUFA levels significantly predicted clinical and functional outcomes, though not transition to psychosis (150). However, at present, the randomised controlled trial evidence for n-3 PUFAs for prevention of psychosis remains insufficient to recommend their routine use in CHR individuals to prevent transition.

Associations between biomarkers of PUFA status and mental disorders have generally not been extensively studied in adolescence and early adulthood, despite the fact that the risk of onset of mental disorders, including psychosis, is highest during this period of development (151). Considering psychosis specifically, a previous study in the Avon Longitudinal Study of Parents and Children (ALSPAC) found minimal evidence of cross-sectional associations between plasma PUFA markers and PEs at age 17 years or longitudinally between PUFA markers at age 15 years and PEs at age 17 years (152). A further study in the same cohort examined associations between plasma PUFA markers in childhood (age 7 years) and mid-adolescence (age 16 years) in relation to PEs or psychotic disorder in late adolescence (age 18 years), but found little evidence to support such associations (153). One longitudinal study in a CHR for psychosis cohort found that a higher n-6:n-3 fatty acid ratio at baseline predicted mood disorder (but not

other psychiatric disorders) on follow-up (154). There have been calls to clarify whether PUFA abnormalities precede the onset of psychosis and other mental disorders in longitudinal studies (155).

1.6 Thesis objectives and research questions

The main body of this thesis comprises three studies each focusing on molecular pathophysiology in different phenotypic phases of psychosis. The corresponding objectives and principal research questions are summarised below.

Objective 1. To systematically review the evidence from serological studies examining complement proteins in schizophrenia.

Research question: In patients with schizophrenia, are peripheral complement protein concentrations or complement activity consistently altered in comparison to controls without the disorder?

Objective 2. To develop models incorporating plasma proteomic data for prediction of transition to psychotic disorder in individuals at clinical high-risk for psychosis and for prediction of adolescent psychotic experiences.

Research question: Are proteomic data predictive of transition to psychotic disorder in people at clinical high-risk of psychosis, and of adolescent psychotic experiences in a general population sample?

Objective 3. To evaluate cross-sectional and longitudinal associations between plasma polyunsaturated fatty acids and psychotic disorder, depressive disorder and generalised anxiety disorder in late adolescence and early adulthood in the general population.

Research question: In the general population, are plasma polyunsaturated fatty acids associated cross-sectionally and/or longitudinally with mental disorders in late adolescence and early adulthood?

1.7 Data sources for this thesis

1.7.1 Objective 1

For objective 1, data were drawn from the existing literature by means of a systematic review of peer-reviewed published serological studies that measured activities or concentrations of complement system proteins in patients with schizophrenia compared to controls without the disorder.

1.7.2 Objective 2

For objective 2, data were acquired from two longitudinal studies.

Firstly, a CHR sample was drawn from the European Network of National Schizophrenia Networks studying Gene-Environment Interactions (EU-GEI) study. The EU-GEI study includes a prospective cohort of 344 CHR participants recruited across 11 international sites (156, 157). Individuals with CHR symptoms, referred by local mental health services, were eligible to participate if they met CHR criteria as per the Comprehensive Assessment of At-Risk Mental States (30). Accrual began in September 2010 and the last baseline assessment was performed in July 2015.

Secondly, a general population sample was drawn from the Avon Longitudinal Study of Parents and Children (ALSPAC). The ALSPAC study is a prospective birth cohort (158, 159). Pregnant women in Avon, United Kingdom with delivery dates between 1st April 1991 to 31st December 1992 were invited and 14,541 pregnancies were enrolled. When the oldest children were approximately 7 years old, an attempt was made to bolster the sample with cases who did not join originally. The sample size for analyses using data from age 7 is 15,454 pregnancies (14,901 children alive at 1 year). For this objective, data were collected via clinics conducted when the participants were aged approximately 12 years old and subsequently when approximately 18 years old.

1.7.3 Objective 3

For objective 3, data were derived from the ALSPAC birth cohort, primarily based on clinics conducted when participants were aged approximately 17 years old and subsequently when approximately 24 years old.

Chapter 2. Peripheral complement proteins in schizophrenia: A systematic review and meta-analysis of serological studies

Chapter overview

This chapter comprises a manuscript published in *Schizophrenia Research* of a systematic review and meta-analysis focusing on serological studies of peripheral complement proteins in schizophrenia.

Role of the thesis author

Dr Mongan contributed to the conception and design of the study, formulated the research question and database search strategy, performed data extraction and statistical analyses, drafted the initial manuscript and produced the final version of the text following peer review. He was the first and corresponding author on the published manuscript.

Reference: Mongan D, Sabherwal S, Susai SR, Föcking M, Cannon M, Cotter DR. Peripheral complement proteins in schizophrenia: A systematic review and meta-analysis of serological studies. *Schizophrenia Research* 2020 Aug;222:58-72. doi: 10.1016/j.schres.2020.05.036.

Peripheral complement proteins in schizophrenia: a systematic review and meta-analysis of serological studies

David Mongan*[1], Sophie Sabherwal[1], Subash Raj Susai[1], Melanie Föcking[1], Mary Cannon[1], David R Cotter[1]

[1]Department of Psychiatry, Royal College of Surgeons in Ireland, Dublin, Ireland.

***Corresponding author:** David Mongan
Address: Department of Psychiatry, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin 9, Ireland.

Number of tables: 2

Number of figures: 2

Word count: 5669

Appendices:

Appendix 2.1: Search terms used for each database
Appendix 2.2: Risk of bias assessment
Appendix 2.3: Assessment of publication bias and small-study effects
Appendix 2.4: Results of meta-regression analyses

Declarations of interest: none

Abstract

Background: There is renewed focus on the complement system in the pathogenesis of schizophrenia. In addition to providing aetiological insights, consistently dysregulated complement proteins in serum or plasma may have clinical utility as biomarkers.

Methods: We performed a systematic literature review searching PubMed, Embase and PsycINFO for studies measuring complement system activity or complement protein concentrations in serum or plasma from patients with schizophrenia compared to controls. Random-effects meta-analyses were performed to calculate pooled effect estimates *g* (standardised mean difference [SMD]) for complement proteins whose concentrations were measured in three or more studies. The review was pre-registered on the PROSPERO database (CRD42018109012).

Results: Database searching identified 1146 records. Fifty-eight full-text articles were assessed for eligibility and 24 studies included. Seven studies measured complement system activity. Activity of the classical pathway did not differ between cases and controls in four of six studies, and conflicting results were noted in two studies of alternative pathway activity. Twenty studies quantified complement protein concentrations of which complement components 3 (C3) and 4 (C4) were measured in more than three studies. Meta-analyses showed no evidence of significant differences between cases and controls for 11 studies of C3 (SMD 0.04, 95% confidence interval [CI] -0.29ï 0.36) and 10 studies of C4 (SMD 0.10, 95% CI -0.21ï 0.41).

Conclusions: Serological studies provide mixed evidence regarding dysregulation of the complement system in schizophrenia. Larger studies of a longitudinal nature, focusing on early phenotypes, could provide further insights regarding the potential role of the complement system in psychotic disorders.

2.1 Introduction

2.1.1 Rationale

Schizophrenia is a severe mental disorder, among the most disabling conditions worldwide (Collaborators, 2017). While the precise pathogenesis remains unknown, hypotheses for involvement of the immune system and inflammation have developed from several lines of enquiry. For example, infants born to mothers exposed to the influenza virus in pregnancy are at increased risk of developing schizophrenia later in life (Limosin et al., 2003). Epidemiological studies suggest associations between several autoimmune diseases and psychosis (Benros et al., 2014; Cullen et al., 2019). Several studies have reported raised serological markers of inflammation in patients with schizophrenia compared to controls. In their meta-analysis, Miller et al identified multiple inflammatory cytokines consistently raised in patients with schizophrenia compared to controls (Miller et al., 2011). In a subsequent meta-analysis of patients with first-episode psychosis (FEP), several cytokines were again found to be raised in comparison to controls (Upthegrove et al., 2014).

Recently there has been renewed focus in schizophrenia research on the complement system, an important component of innate immu in conjunction with adaptive immune responses. Holers (2014) provides a detailed discussion of the components and functions of the complement system. To summarise, the complement system is a complex network of approximately 60 interacting proteins that, upon activation, augments immune and inflammatory responses through opsonisation of antigens, chemotaxis and, ultimately, formation of a membrane attack complex (C5b-9) capable of rupturing invading bacterial cells (Mayilyan, Weinberger, et al., 2008; Nimgaonkar et al., 2017). Activation of the system occurs via three main routes known as the classical, alternative and mannose-binding lectin (MBL) pathways. While each pathway is initiated by different molecular mechanisms, all three lead to activation of the central complement protein known as complement component 3 (C3). C3 and its activation products promote opsonisation and chemotaxis as well as contributing to formation of the membrane attack complex. A range of control proteins,

such as complement factor H, complement factor I and C4b binding protein, regulate activity of the system (Zipfel & Skerka, 2009).

The complement system serves vital immune functions in health by augmenting the clearance of micro-organisms and necrotic cells. However, dysregulation or inappropriate activation in the presence of self-antigens can contribute to disease states, as is associated with disorders such as systemic lupus erythematosus (Popescu & Kao, 2011), multiple sclerosis (Ingram et al., 2009; Michailidou et al., 2016) and Alzheimer's (Hong et al., 2016; VanItallie, 2017). Genome-wide association studies have associated schizophrenia susceptibility with polygenic variation in the major histocompatibility complex (Consortium, 2014), a region containing a set of genes that code for multiple immune-related proteins, including several complement components. Sekar et al (2016) showed that this association was partly related to allelic variation of the complement component 4 (C4) gene. The same authors also found that C4 RNA expression was increased in post-mortem brain tissue from patients with schizophrenia compared to controls. These findings have spurred interest regarding the role of the complement system in schizophrenia. In the brain, complement proteins are thought to play a role in tagging synapses for elimination by microglia (Stephan et al., 2012) which could represent an underlying biological mechanism for the observations of reduced synaptic density in schizophrenia (Osimo et al., 2018).

Aside from providing aetiopathogenic insights, if peripheral complement proteins are consistently dysregulated in schizophrenia, then they may have clinical utility as biomarkers. Investigators have generally used two broad categories of assays to study this area: activity-based assays and concentration-based assays. Activity-based assays typically assess haemolytic activity by measuring the ability of serum to lyse sheep erythrocytes sensitised with rabbit immunoglobulin M for classical pathway activity, or unsensitised rabbit erythrocytes for alternative pathway activity under conditions that block functioning of the classical pathway (Costabile, 2010; Oppermann & Wurzner, 2010). Assays are also available that measure functional (rather than haemolytic) activity, such as using enzyme-linked immunosorbent assay (ELISA) techniques to measure the formation of membrane attack complexes after complement activation

(Seelen et al., 2005). Concentration-based assays, rather than providing a measure of activity, quantify the concentration of individual complement proteins in the blood. This may be performed using several methods including ELISA, nephelometry, radial immunodiffusion and other techniques (Kirschfink & Mollnes, 2003; Nilsson & Ekdahl, 2012).

2.1.2 Objective

The objective of this review was to systematically synthesise and evaluate the available evidence from serological studies to answer the question: in patients with schizophrenia, are peripheral complement protein concentrations or complement activity consistently altered in comparison to controls without the disorder?

2.2 Methods

2.2.1 Protocol and registration

The protocol for this review is registered on the PROSPERO database (https://www.crd.york.ac.uk/prospero/; registration number CRD42018109012).

2.2.2 Eligibility criteria

Inclusion criteria for studies were: case-control or cohort studies; comparing serum or plasma complement protein concentrations or activity in patients diagnosed with schizophrenia to controls without schizophrenia; published in peer-reviewed journals; available in English. Review articles that did not present original data were included in the primary search so that their reference lists could be screened for potential additional records.

Exclusion criteria for studies were: genetic or RNA expression studies; non-human studies; no healthy control comparison group; poster or conference abstracts; studies measuring the same proteins in a previously-presented sample; studies focusing exclusively on psychiatric disorders other than schizophrenia (for example depression or bipolar disorder). We chose to focus on studies that used targeted approaches measuring complement proteins or their activity. Thus, studies that relied upon proteome-wide or discovery-based approaches (such as óshotgu

spectrometry) were not included. Evidence derived from such methods in schizophrenia (Sabherwal et al., 2016) and other severe mental disorders (Comes et al., 2018) has been synthesised in recent systematic reviews.

2.2.3 Information sources and search strategy

We searched PubMed, Embase and PsycINFO from inception to current date using search strategies optimised for each database. The search strategy was developed and performed in collaboration with an information specialist (the full list of search terms for each database is included in Appendix 2.1). The search was last performed on 17th June 2019. Reference lists of included studies were also examined for further potential records.

2.2.4 Study selection

Titles and abstracts were screened independently by two authors (DM and SS). Full-text articles were obtained for the remaining records, which were assessed for inclusion according to the eligibility criteria above. Differences were resolved by discussion and involvement of a third author where necessary.

2.2.5 Data collection process

For each included study, data were extracted by two authors independently (DM and SS) using a custom-developed and pre-piloted data extraction form. Key data items extracted included country of origin; case definition and method of diagnosis; hospitalisation status of cases; number of cases and controls; type of assay used; complement pathways or proteins measured and their mean values and standard deviations (or other summary measures where appropriate) in cases and controls. Where data were unavailable or unclear, we attempted to contact the study authors. Where data were presented in graphical form only (and raw data unavailable from authors), data extraction software was used to estimate the numerical values (WebPlotDigitizer version 4.2; https://automeris.io/WebPlotDigitizer/).

2.2.6 Risk of bias assessment

Risk of bias was assessed at the study level using the Newcastle-Ottawa quality assessment scale (Wells et al., 2013). This measure consists of three subscales used to assess study quality, with a maximum score of nine points: four points for selection of cases and controls, two points for comparability between cases and controls and three points for determination of exposure.

2.2.7 Meta-analyses

We proposed to perform meta-analyses for complement proteins measured using concentration-based assays in at least three studies. Among all proteins measured across the included studies, C3 and C4 met this criterion. Sample size, mean and standard deviation values in cases and controls for concentrations of these proteins were recorded. Where studies presented the standard error of the mean, we converted this to the standard deviation. Where studies presented the median and interquartile range, we attempted to contact authors for the mean and standard deviation values.

Statistical analyses were performed using Stata version 15 (StataCorp). Predicting significant between-study heterogeneity, we used random-effects models to calculate pooled effect size estimates (Hedges ô *g*, standardised mean difference [SMD]). Statistical heterogeneity was assessed using the I^2 statistic. For all analyses, the threshold for statistical significance was p=0.05.

Publication bias

Evidence for publication bias was assessed via inspection of funnel plots. Evidence for small-study effects was assessed using Eggerôs test (Egger et al., 1997).

Meta-regression analyses

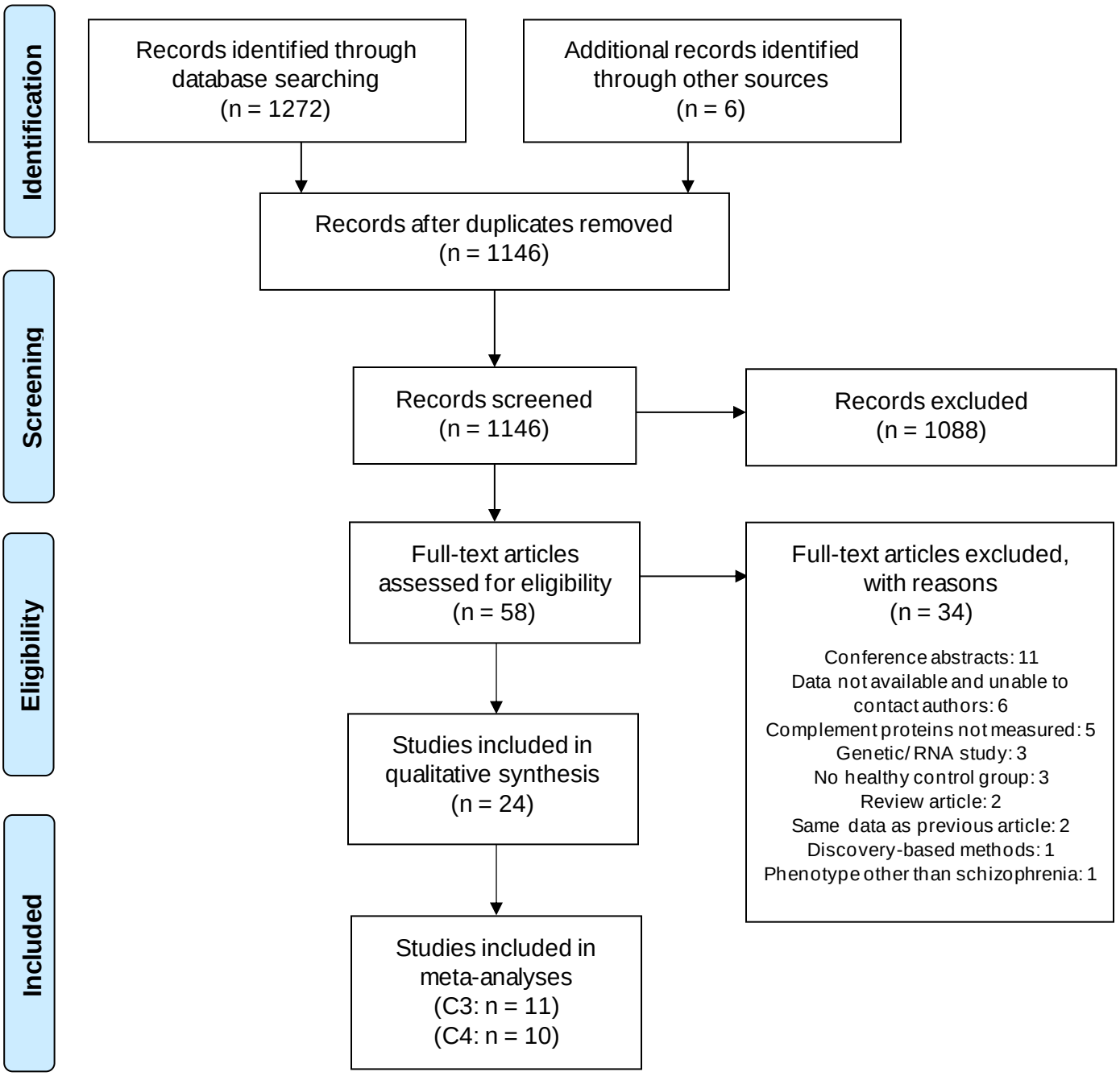
Post-hoc, we performed univariate meta-regression analyses to explore heterogeneity by evaluating the effects of three study characteristics upon the effect estimates: medication status at time of blood sampling (concurrent use of antipsychotics vs. no concurrent use), hospitalisation status of cases (inpatient vs. outpatient) and assay methodology (ELISA vs. radial immunodiffusion or other; radial immunodiffusion vs. ELISA or other; and other vs. ELISA or radial immunodiffusion).

2.3 Results

2.3.1 Study selection

Figure 2.1 shows the results of the search strategy and study selection process, which was performed and is reported in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). Database searching identified 1272 records (462 from PubMed, 373 from Embase and 437 from PsycINFO) and 6 records were identified from reference screening. Following removal of duplicates, 1146 title and abstract records were screened of which 1088 were excluded. Full-text articles were obtained for the remaining 58 records, which were assessed for eligibility according to the inclusion and exclusion criteria. Thirty-four articles were excluded for multiple reasons (detailed in Figure 2.1) resulting in 24 included studies (Ali et al., 2017; Boyajyan et al., 2010; Cazzullo et al., 1998; Foldager et al., 2012; Fontana et al., 1980; Hakobyan et al., 2005; Hong et al., 2016; Idonije et al., 2012; Ji et al., 2019; Kucharska-Mazur et al., 2014; Laskaris et al., 2018; Li et al., 2016; Li et al., 2012; Maes et al., 1997; Mayilyan et al., 2006; Mayilyan et al., 2008; Ramsey et al., 2013; Santos Soria et al., 2012; Sasaki et al., 1994; Schwarz et al., 2012; Spivak et al., 1993; Spivak et al., 1989; Walss-Bass et al., 2019; Wong et al., 1996).

Figure 2.1 PRISMA flow chart showing search strategy and results



C3: complement component 3; C4: complement component 4

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

2.3.2 Study characteristics

Characteristics of studies included in this review are presented in Table 2.1 (studies using activity-based assays) and Table 2.2 (studies using concentration-based assays). Unadjusted values for assay results are presented unless otherwise stated.

Table 2.1: Summary of included studies using activity-based assays

Study	Country	Hospitalisation status of cases	Diagnostic criteria	Medication status	Sample size	Measure	Sample type	Assay	Main findings	Notes
Boyajyan et al, 2010	Armenia	Inpatient	ICD-10	Medicated and drug-free	36 medicated cases, 25 drug-free cases vs. 26 controls	Alternative pathway activity, C3 activity	Serum	Haemolytic activity	Mean alternative pathway activity in medicated cases (228.3 AH50 units/ml, p=0.001) and drug-free SZ (155.6 AH50 units/ml, p=0.001) controls (93.7 AH50 units/ml) Mean C3 activity in medicated cases (294.1 C3H50 units/ml, p=0.001) and drug-free SZ (138.5 C3H50 units/ml, p=0.002) significan (99.5 C3H50 units/ml)	All cases had multiple-episode SZ of paranoid subtype No difference between smokers and non-smokers or males and females
Hakobyan et al, 2005	Armenia	Inpatient	ICD-10	Medicated and drug-free (for mean of 10 weeks)	24 cases (13 drug-free) vs. 28 controls	Classical pathway activity and activities of C1, C2, C3 and C4	Serum	Haemolytic activity	Mean classical pathway activity not significantly different between cases (39.24 CH50 units) vs. controls (37.31 CH50 units) Mean C1 activity si (39.10 CH50 units) vs. controls (22.90 CH50 units), p<0.002 Mean C2 activity si (3.52 CH50 units) vs. controls (6.82 CH50 units), p<0.001 Mean C3 activity si (204.83 CH50 units) vs. controls (101.88 CH50 units), p<0.004 Mean C4 activity si (944.57 CH50 units) vs. controls (686.37), p<0.001	Cases were in acute relapse of multiple-episode paranoid SZ and had first- or second-degree family history of SZ No difference between smokers and non-smokers or males and females Only mean C3 activity significantly differed between medicated (289.53 CH50 units) and drug-free cases (133.16 CH50 units), p=0.001
Li et al, 2012	China	Outpatient	DSM-IV	Drug-free	47 cases (29 antipsychotic naïve, 18 drug-	Classical pathway, alternative pathway and MBL	Serum	Qualitative ELISA	Alternative pathway activity significantly decreased in drug-free cases (75%) vs. medicated cases (82%), p=0.001; and drug-free cases vs. controls (86%), p=0.001	

					free for 8 weeks) vs. 53 controls	pathway activity			No evidence of significant difference between drug-free or drug-treated cases and controls for classical or MBL pathways	
Mayilyan et al, 2006	Armenia	Outpatient	DSM-IV	Medicated and drug-free (>3 months)	45 cases (9 drug-free) vs. 29 controls	Classical pathway activity (CH50), C4 activity, MBL-bound MASP1 activity, MBL-bound MASP2 activity	Serum	Haemolytic activity	Median total complement activity significantly ŷ in vs. controls (1087 units/ml), p<0.02 Median C4 activity cases (143 000 U/ml) vs. controls (109 000 U/ml), p<0.02 Mean MBL-bound MASP-2 activity significantly ŷ in controls (5.66 U/ml), p<0.01 No significant difference in MBL-bound MASP1 activity between cases and controls	All cases diagnosed with paranoid subtype and had first- or second-degree family history of SZ
Sasaki et al, 1994	Japan	Inpatient	DSM-III-R	Medicated and drug-free	14 cases (3 drug-free) vs. 20 controls	Classical pathway activity (CH50)	Serum	Haemolytic activity	Mean CH50 not significantly different between cases (36.7 units/ml) and controls (36.3 units/ml)	Cases had chronic SZ, in acute relapse
Spivak et al, 1989	Israel	Inpatient	DSM-III	Drug-free (Ó 3 mon	20 cases vs. 18 controls	Classical pathway activity (CH100)	Serum	Haemolytic activity	Mean CH100 not significantly different between cases (69.1 units/ml) and controls (75.9 units/ml)	Controls recruited from medical staff
Spivak et al, 1993	Israel	Inpatient	DSM-III-R	Drug-naïve, drug-free months) and medicated	91 cases (20 drug-naïve, 37 drug-free, 34 medicated) vs. 37 controls	Classical pathway activity (CH100)	Serum	Haemolytic activity	Mean CH100 signific units/ml) vs. controls (85 units/ml), p=0.007 No differences between cases vs. controls when stratified by medication status	Cases had chronic SZ Controls recruited from medical staff

SZ: schizophrenia; ELISA: Enzyme-linked immunosorbent assay; ICD: International Classification of Diseases; DSM: Diagnostic and Statistical Manual; MBL: mannose-binding lectin; MASP: mannose-binding lectin serine protease

Table 2.2: Summary of included studies using concentration-based assays

Study	Country	Hospitalisation status of cases	Diagnostic criteria	Medication status of cases	Sample size	Complement proteins	Sample type	Assay	Main findings	Notes
Ali et al, 2017	Egypt	Outpatient	DSM-IV	Drug-free	44 cases vs. 50 controls	C3, C4	Serum	ADVIA Chemistry XPT System	Mean C3 in cases (169.1 mg/dl) significic controls (133.5 mg/dl), p<0.001 Mean C4 in cases (33.9 mg/dl) not significantly different vs. controls (31.3 mg/dl)	C3 and C4 were not significantly correlated with PANSS score
Cazzullo et al, 1998	Italy	Not known	DSM-III-R	Medicated	29 cases vs. 20 controls	C3, C4	Serum	Nephelometry	Median C3 in cases (155 mg/dl) not significantly different vs. controls (134 mg/dl) Median C4 in cases (33 mg/dl) not significantly different vs. controls (28 mg/dl)	Cases were treatment-resistant with active SZ symptoms C4 values estimated from graph
Foldager et al, 2012	Denmark	Inpatient and outpatient	ICD-10, DSM-IV	Medicated	100 cases vs. 350 controls	MBL, MASP-2	Serum	Time-resolved immuno-fluorometric assay	Median MBL in cases (584 ng/ml) not significantly different vs. controls (578 ng/ml); significantly higher in cases when adjusted for MBL2 haplotypes Median MASP-2 in cases (425 ng/ml) not significantly different vs. controls (417 ng/ml)	Samples for cases obtained from previous genetic studies and acutely psychotic at time of sampling Controls were blood donors, not psychiatrically screened
Fontana et al, 1980	Switz-erland	Inpatient	ICD	Medicated	48 cases vs. 100 controls	C3, C4, C1-inhibitor	Serum	C3 by radial immunodiffusion C4 and C1-inhibitor by immuno-electro-phoresis	Mean C3 not significantly different between cases (130.2 mg/ml) vs. controls (150.0 mg/ml) Mean C4 not significantly different between cases (16.7 mg/ml) vs. controls (18.6 m/ml)	Controls were blood donors

									Mean C1-inhibitor not significantly different between cases (17.0 mg/ml) vs. controls (21.5 mg/ml)	
Hong et al, 2016	China	Inpatient	ICD-10	Drug-naïve and drug-free (Ć weeks)	41 cases vs. 33 controls	C3	Plasma	ELISA	Mean C3 not significantly different between cases (161.19 εg/ml) (164.20 εg/ml)	Controls recruited from medical staff
Idonije et al, 2012	Nigeria	Inpatient	DSM-IV	Medicated and drug-naïve	20 chronic medicated cases, 15 first episode drug-naïve cases vs. 20 controls	C1q, C3c, C4, C5, C1 inhibitor, C3 activator	Serum	Immunoplates	Mean C1q signif chronic cases (3.61 g/l) vs. controls (9.40 g/l), p<0.001; significantly Ź cases (3.09 g/l) vs. controls, p<0.001; no significant difference between chronic and first episode cases Mean C3c not significantly different between chronic cases (1.30 g/l) vs. controls (1.41 g/l); sig episode cases (0.68 g/l, p<0.001) vs. controls; significantly ŷ vs. first episode cases (p<0.001) Mean C4 not significantly different between chronic cases (0.17 g/l) vs. controls (0.26 g/l) or first episode cases (0.23 g/l) vs. controls; no significant difference between chronic cases vs. first episode cases Mean C5 not significantly different between chronic cases (1.55 g/l) vs. controls (1.50 g/l) or first episode cases (1.20) vs. controls; significantly ŷ in chronic ca episode cases, p=0.04	

									Mean C1 inhibitor not significantly different between chronic cases (1.21 g/l) vs. controls (0.53 g/l) or first episode cases (0.49 g/l) vs. controls or chronic cases vs. first episode cases Mean C3 activator not significantly different between chronic cases (0.75 g/l) vs. controls (0.89 g/l); or first episode cases (0.66 g/l) vs. controls; or chronic cases vs. first episode cases	
Ji et al, 2019	China	Inpatient	ICD-10	Not known	40 cases vs. 40 controls	C4	Serum	ELISA	Mean C4 signifi cases (154.2 ug/ml) vs. controls (216.2 ug/ml), p<0.001	Cases had first episode SZ
Kucharska-Mazur et al, 2014	Poland	Not known	ICD-10	Drug-naïve	22 cases vs. 35 controls	C3a, C5a, C5b-9	Plasma	ELISA	Mean C3a signif cases (494.33 ng/ml) vs. controls (591.13 ng/ml), p 0.03 Mean C5a not significantly different between cases (52.94 ng/ml) and controls (53.24 ng/ml) Mean C5b-9 not significantly different between cases (355.03 ng/ml) and controls (220.73 ng/ml)	Cases had first episode SZ (22 of 28 recruited patients with first episode psychosis)
Laskaris et al, 2018	Australia	Outpatient	DSM-IV	Medicated	50 cases vs. 54 controls	C1q, C3 and C4	Serum	ELISA multiplex	Mean C1q not significantly different between in cases (48.9 εg/ml) vs. controls (45.8 εg/ml) Mean C3 not significantly different between cases (16.9 εg/ml) vs. cont εg/ml) Mean C4 signifi cases (292.2 εg controls (252.0	Cases had chronic SZ Values adjusted for BMI

Li et al, 2016	China	Inpatient	ICD-10	Drug-f r e e 4 weeks)	40 cases vs. 40 controls	C3	Plasma	ELISA	M e d i a n C 3 s i g n i cases (118742.49 ng/ml) vs. controls (160853.17 ng/ml), p=0.04	Controls recruited from medical staff C3 negatively correlated with PANSS score (r ï 0.37, p=0.03)
Maes et al, 1997	USA	Inpatient	DSM-III-R	Drug-free (Ó 1 w e e and medicated	27 cases (17 drug-free, 10 medicated) vs. 21 controls	C3c, C4	Plasma	Nephelometry	M e a n C 3 c s i g n i f drug-free cases (92 mg/dl) vs. controls (74 mg/dl), p=0.03; not significantly different between medicated cases (89 mg/dl) vs. controls M e a n C 4 s i g n i f i drug-free cases (41 mg/dl) vs. controls (29 mg/dl), p=0.02; not significantly different between medicated cases (37 mg/dl) vs. controls	Values adjusted for age and gender
Mayilyan et al, 2006	Armenia	Outpatient	DSM-IV	Drug-free (>3 months) and medicated	45 cases (9 drug-free) vs. 29 controls	MBL	Serum	ELISA	Mean MBL serum concentration not significantly different between cases (1.74 ug/ml) vs. controls (2.55 ug/ml)	Cases had chronic SZ in remission and had positive family history of SZ
Mayilyan et al, 2008	Armenia	Outpatient	ICD-10, DSM-IV	Drug-f r e e 2 months) and medicated	45 cases (9 drug-free) vs. 51 controls	C4B	Serum	ELISA	M e d i a n C 4 B s i g n cases (159.1 mg/l) vs. controls (180.0 mg/l), p<0.01	Cases had chronic SZ in remission
Ramsey et al, 2013	Germany	Not known	DSM-IV	Drug-naïve	133 cases vs. 133 controls	C3	Serum	ELISA multiplex	C 3 s i g n i f i c a n t l (ratio cases:controls 1.07), p=0.002 (p=0.013 corrected for multiple comparisons)	Cases had first episode SZ
Santos-Soria et al, 2012	Brazil	Outpatient	DSM-IV	Medicated	53 cases vs. 80 controls	C3, C4	Serum	Immuno-turbidimetry	M e a n C 3 s i g n i f i cases (190.3 mg/dl) vs. controls (162.6 mg/dl), p<0.01 Mean C4 not significantly different between cases (40.3 mg/dl) vs. controls (38.4 mg/dl)	Cases had chronic SZ in remission

Schwarz et al, 2012	Germany, Holland, UK	Inpatient	DSM-IV	Drug-naïve	71 cases vs. 59 controls	C3	Serum	ELISA multiplex	Mean C3 not significantly different between cases (0.97 mg/ml) and controls (0.91 mg/ml)	Cases had first episode paranoid SZ
Spivak et al, 1989	Israel	Inpatient	DSM-III	Drug-free (Ó 3 mo n	20 cases vs. 18 controls	C3, C4	Serum	Radial immuno-diffusion	Mean C3 not significantly different between cases (146.1 mg/100ml) vs. controls (136.2 mg/100ml) Mean C4 not significantly different between cases (35.3 mg/100ml) and controls (36.7 mg/100ml)	Cases had chronic SZ Controls recruited from medical staff
Spivak et al, 1993	Israel	Inpatient	DSM-III-R	Drug-naïve, drug-f r e e months) and medicated	91 cases (20 drug-naïve, 37 drug-free, 34 medicated) vs. 37 controls	C3, C4	Serum	Radial immuno-diffusion	Mean C3 not significantly different between cases (128 mg/100ml) vs. controls (132 mg/100ml) Mean C4 not significantly different between cases (33 mg/100ml) and controls (35 mg/100ml) No significant differences for C3 or C4 in cases vs. controls when stratified by medication status	Cases had chronic SZ Controls recruited from medical staff
Walss-Bass et al, 2018	USA	Outpatient	DSM-IV-TR	Medicated	60 cases vs. 20 controls	C4	Plasma	ELISA	Mean C4 significantly increased in cases (242.5 ug/ml) vs. controls (191.8 ug/ml), p=0.0097	Cases had chronic SZ
Wong et al, 1996	Singapore	Inpatient	ICD-9	Medicated and drug-free (15 drug-free)	Series 1: 44 recovered (neither positive nor negative symptoms), 30 chronic (only negative symptoms), 15 acute (positive and negative	C3	Serum	Immuno-electro-phoresis	Series 1: Mean C3 significantly Ź i n c h r o n i c c a controls (0.88 g/l), p<0.01 No significant differences for acute or recovered patients compared to controls	All participants were males Controls in series 1 were blood donors Controls in series 2 were healthy Chinese male civil servants attending compulsory

	symptoms) vs. 41 controls Series 2: 50 acute cases (positive and negative symptoms) vs. 50 controls		Series 2: mean C3 not significantly different between acute cases (0.81 g/l) vs. controls (0.70 g/l)	medical examination No significant differences in C3 in cases on or off antipsychotics

SZ: schizophrenia; ELISA: Enzyme-linked immunosorbent assay; ICD: International Classification of Diseases; DSM: Diagnostic and Statistical Manual; MBL: mannose-binding lectin; MASP: mannose-binding lectin serine protease

All included studies were of a case-control design and defined the diagnosis of schizophrenia in cases according to contemporaneous International Classification of Diseases (ICD) or Diagnostic and Statistical Manual (DSM) criteria. Sample sizes ranged from 34 to 450 participants. In the majority of included studies assays were performed using serum samples; five studies instead used plasma (Hong et al., 2016; Kucharska-Mazur et al., 2014; Li et al., 2016; Maes et al., 1997; Walss-Bass et al., 2019). Thirteen studies recruited inpatient cases (Boyajyan et al., 2010; Fontana et al., 1980; Hakobyan et al., 2005; Hong et al., 2016; Idonije et al., 2012; Ji et al., 2019; Li et al., 2016; Maes et al., 1997; Sasaki et al., 1994; Schwarz et al., 2012; Spivak et al., 1993; Spivak et al., 1989; Wong et al., 1996), seven studies recruited outpatients (Ali et al., 2017; Laskaris et al., 2018; Li et al., 2012; Mayilyan et al., 2006; Mayilyan et al., 2008; Santos Soria et al., 2012; Walss-Bass et al., 2019), one recruited both inpatient and outpatient cases (Foldager et al., 2012) and in the remainder the hospitalisation status of patients was not determined (Cazzullo et al., 1998; Kucharska-Mazur et al., 2014; Ramsey et al., 2013).

In many studies, the clinical status of patients at the time of blood sampling (i.e. in remission or acute psychosis) was unclear. In at least four studies (Cazzullo et al., 1998; Foldager et al., 2012; Hakobyan et al., 2005; Sasaki et al., 1994) sampling was performed when patients were experiencing acute exacerbations or active symptoms of schizophrenia, while two studies stated that cases were stable or in remission (Mayilyan, et al., 2008; Santos Soria et al., 2012). Most studies recruited cases with chronic schizophrenia, although in five studies all (Ji et al., 2019; Kucharska-Mazur et al., 2014; Ramsey et al., 2013; Schwarz et al., 2012) or a proportion (Idonije et al., 2012) of cases were in their first episode of illness. Most studies included participants who were prescribed antipsychotics at the time of sampling, of which five reported data stratified by medication status (Boyajyan et al., 2010; Hakobyan et al., 2005; Idonije et al., 2012; Maes et al., 1997; Spivak et al., 1993). In others, cases were drug-naïve (never exposed to antipsychotic therapy) (Hong et al., 2016; Kucharska-Mazur et al., 2014; Ramsey et al., 2013; Schwarz et al., 2012) or drug-free for a period of time (Ali et al., 2017; H. Li et al., 2016; Y. Li et al., 2012; Spivak et al., 1989) although in the latter

case, the duration of time off medication prior to blood sampling was variable (from one week to more than 3 months).

Assays used

The included studies utilised several different methodologies to measure complement proteins. Seven studies used activity-based assays measuring haemolytic or functional activity of complement pathways or proteins (Table 2.1). Twenty studies quantified complement protein concentrations rather than activity (Table 2.2). Of these, ten used enzyme-linked immunoassay (ELISA), four used radial immunodiffusion, two used nephelometry and the remainder used other techniques.

Risk of bias assessment

The median quality score according to the Newcastle-Ottawa scale was three and the range of scores was from one to seven (see Appendix 2.2).

2.3.3 Qualitative synthesis of results

Activity-based assays

Of the seven studies assessing complement activity (Table 2.1), six studies measured activity of the classical pathway. One study found total complement activity to be significantly lower in cases compared to controls (Spivak et al., 1993), one found total complement activity significantly increased in cases (Mayilyan et al., 2006) and four found no evidence of a significant difference (Hakobyan et al., 2005; Li et al., 2012; Sasaki et al., 1994; Spivak et al., 1989).

Two studies measured alternative pathway activity, which was found to be significantly increased in medicated and non-medicated cases compared to controls in one study (Boyajyan et al., 2010) though significantly reduced in cases compared to controls in the other study (Li et al., 2012).

Activity of the MBL pathway was measured in one study (Li et al., 2012) which found no evidence of a significant difference between cases and controls.

Regarding activity of individual proteins, C3 activity was measured in two studies and found to be significantly increased relative to controls (Boyajyan et al., 2010; Hakobyan

et al., 2005) as well as in patients on antipsychotic therapy compared to non-medicated patients (Hakobyan et al., 2005). C4 activity was measured in two studies, both of which found significantly increased activity compared to controls (Hakobyan et al., 2005; Mayilyan et al., 2006). C1 and C2 activity were measured in one study (Hakobyan et al., 2005) which found that C1 activity was significantly increased and C2 activity significantly decreased in cases compared to controls. One study found MBL-bound mannose-binding lectin serine peptidase 2 (MASP2) activity was significantly increased in cases relative to controls, although the same study reported no significant difference in MBL-bound MASP1 activity (Mayilyan et al., 2006).

Concentration-based assays

Twenty studies measured serum or plasma concentrations of specific complement proteins in schizophrenia patients and controls (Table 2.2). In total, concentrations of 13 different complement proteins were assessed across all included studies.

Twelve studies reported peripheral concentrations of C3, of which three studies found significantly increased levels in serum from cases compared to controls (Ali et al., 2017; Ramsey et al., 2013; Santos Soria et al., 2012). One study found significantly decreased levels in cases (Li et al., 2016). Another study reported a similar finding in one of two case-control series, but this was not replicated in the other series within the same study (Wong et al., 1996). The remaining studies reported no evidence of a significant difference between cases and controls (Cazzullo et al., 1998; Fontana et al., 1980; W. Hong et al., 2016; Laskaris et al., 2018; Schwarz et al., 2012; Spivak et al., 1993; Spivak et al., 1989).

Ten studies reported C4 concentrations, of which two found significantly increased levels in cases (Laskaris et al., 2018; Walss-Bass et al., 2019), one found significantly decreased levels in cases (Ji et al., 2019) and seven studies found no significant difference (Ali et al., 2017; Fontana et al., 1980; Idonije et al., 2012; Maes et al., 1997; Santos Soria et al., 2012; Spivak et al., 1993; Spivak et al., 1989). One of these studies stratified cases by medication status (Maes et al., 1997) and found increased serum C4 in drug-free patients compared to controls without schizophrenia (Maes et al., 1997).

However, the minimal specified duration off antipsychotics was short (at least one week) and the sample size was small (ten medicated and 17 drug-free patients).

Two studies reported C3c concentrations. One study (Idonije et al., 2012) found no significant difference between cases and controls. The other (Maes et al., 1997) reported no difference between medicated schizophrenia patients and controls, despite evidence of a significant increase in drug-free patients compared to controls. However, the same caveats apply as described for the same study in relation to C4.

Two studies reported C1 inhibitor concentrations, both of which found no evidence of a significant difference between cases and controls (Fontana et al., 1980; Idonije et al., 2012). Two studies reported C1q concentrations. In one study, concentration of C1q was significantly lower in both medicated and antipsychotic-free cases compared to controls (Idonije et al., 2012) although no difference was found in the other study measuring this protein (Laskaris et al., 2018).

Two studies reported concentrations of MBL (Foldager et al., 2012; Mayilyan et al., 2006). In both studies, no significant difference was found on primary analysis between cases and controls. However, in one study (Foldager et al., 2012), a significant increase in cases was reported following adjustment for MBL2 haplotype.

Concentrations of C4b, C5, C3 activator, MASP2, C3a, C5a and C5b-9 were measured in individual studies. C4b was found to be significantly reduced in cases (Mayilyan et al., 2008). C5 was unchanged between cases and controls but was increased in chronically unwell medicated patients compared to newly diagnosed drug-free patients (Idonije et al., 2012). For both C3 activator (Idonije et al., 2012) and MASP2 (Foldager et al., 2012) no significant differences were reported. In a study measuring C3a, C5a and C5b-9 in patients with first-episode schizophrenia, C3a was significantly reduced in cases compared to controls, but no significant differences were reported for C5a or C5b-9 (Kucharska-Mazur et al., 2014).

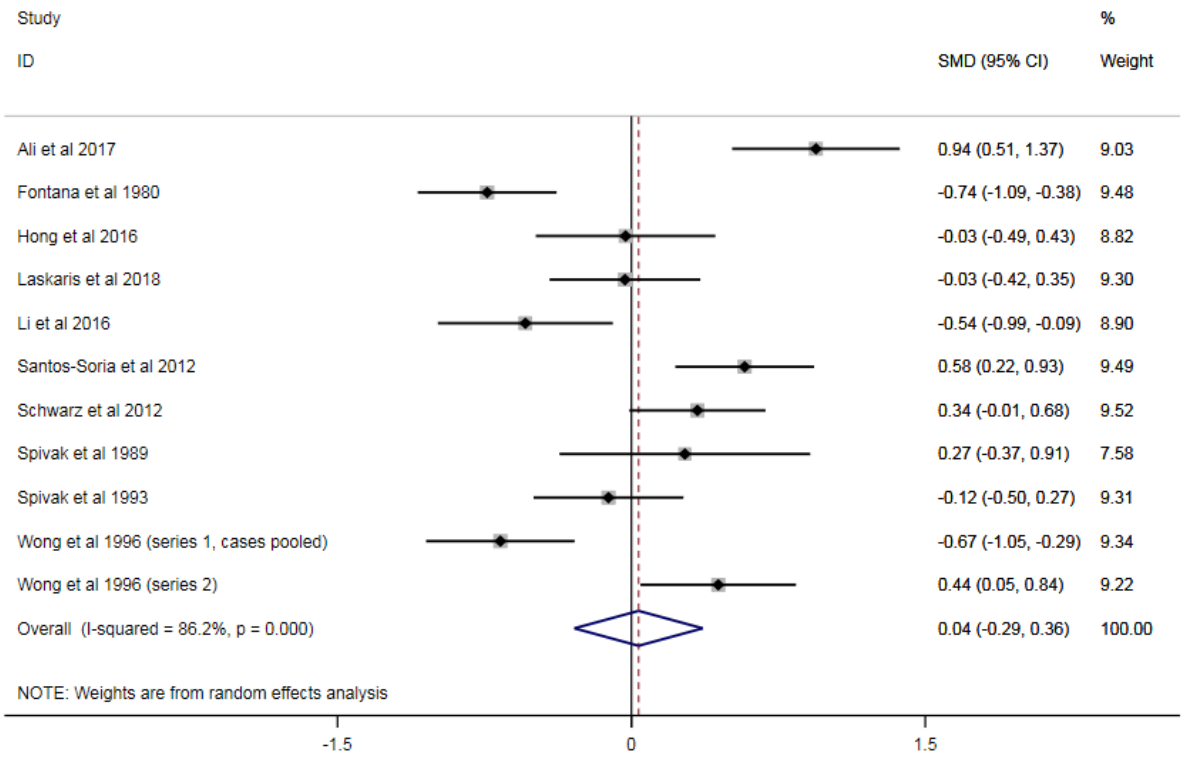
2.3.4 Meta-analyses

Complement component 3

For C3, sufficient data were available to perform meta-analysis of 11 studies (Ali et al., 2017; Fontana et al., 1980; Hong et al., 2016; Laskaris et al., 2018; Li et al., 2016; Santos Soria et al., 2012; Schwarz et al., 2012; Spivak et al., 1993; Spivak et al., 1989; Wong et al., 1996). One study (Wong et al., 1996) presented data for two different case-control series which were considered as separate studies in the meta-analysis. In one of these series, data for cases were presented stratified by symptomatology into three groups (positive and negative symptoms; only negative symptoms; no symptoms) which were compared separately to the same group of healthy controls. To facilitate comparison with the other studies, we calculated the pooled mean and standard deviation across the three groups of cases and used these values in the meta-analysis.

Results of the meta-analysis are presented in the forest plot in Figure 2.2. The pooled effect estimate showed no evidence for a significant difference between cases and controls (SMD 0.04, 95% confidence interval [CI] -0.29 ï 0.36). There was evidence of significant between-study statistical heterogeneity (I^2 86.2%, p<0.001).

Figure 2.2: Forest plot for studies measuring peripheral complement component 3 concentration



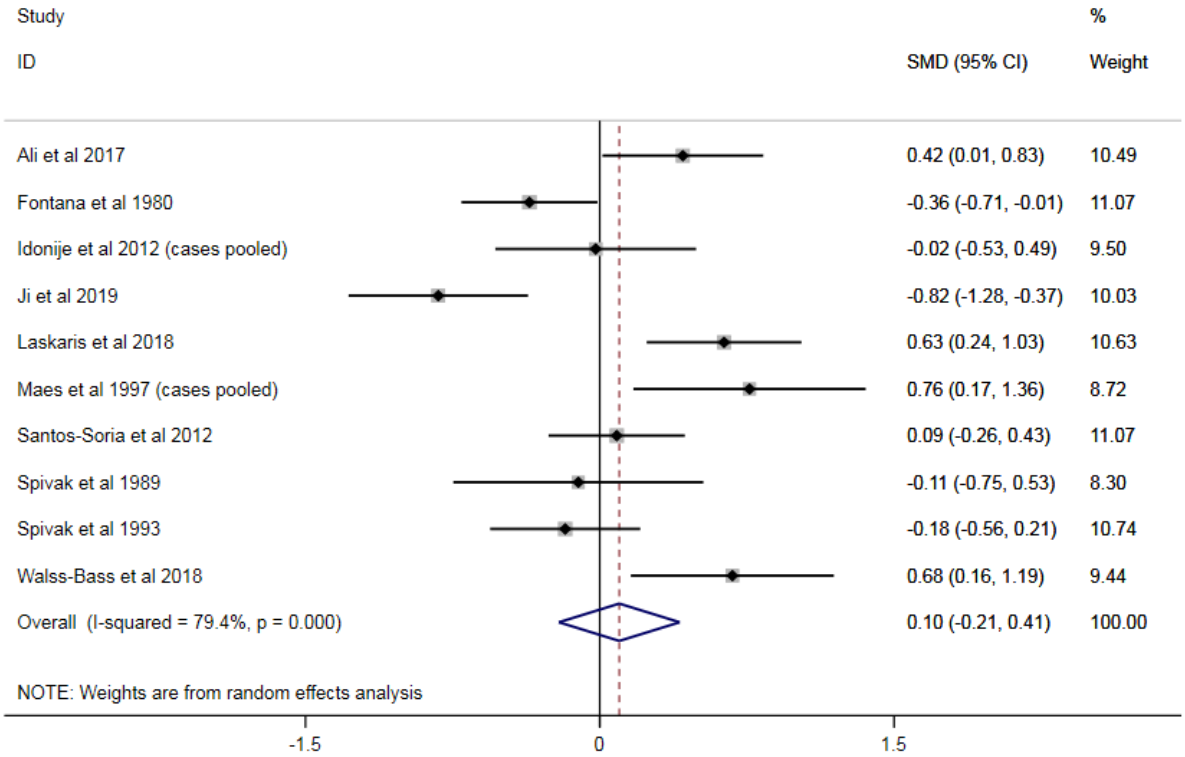
SMD: standardised mean difference; 95% CI: 95% confidence interval

Complement component 4

For C4, sufficient data were available to perform meta-analysis of ten studies (Ali et al., 2017; Fontana et al., 1980; Idonije et al., 2012; Ji et al., 2019; Laskaris et al., 2018; Maes et al., 1997; Santos Soria et al., 2012; Spivak et al., 1993; Spivak et al., 1989; Walss-Bass et al., 2019). In one study (Idonije et al., 2012) data for cases were presented stratified by stage of illness (chronic vs. FEP) and the two groups were compared separately to the same group of healthy controls. Another study (Maes et al., 1997) presented data for cases stratified by medication status (drug-free vs. on antipsychotics) and the two groups were compared separately to the same group of healthy controls. In both cases, we calculated the pooled means and standard deviations across groups of cases and used these values in the meta-analysis.

Results of the meta-analysis are presented in the forest plot in Figure 2.3. The pooled effect estimate showed no evidence for a significant difference between cases and controls (SMD 0.10, 95% CI -0.21 ï 0.41). There was evidence of significant between-study statistical heterogeneity (I^2 79.4%, p<0.001).

Figure 2.3: Forest plot for studies measuring peripheral complement component 4 concentration



SMD: standardised mean difference; 95% CI: 95% confidence interval

Publication bias

For both C3 and C4, inspection of funnel plots did not indicate asymmetry suggestive of significant publication bias. Egger's test (see Appendix 2.3).

Meta-regression analyses

For both C3 and C4, none of the characteristics examined (antipsychotic use, hospitalisation status and assay methodology) showed evidence of significant effects upon the pooled estimates (see Appendix 2.4).

2.4 Discussion

To our knowledge, this is the first systematic literature review and meta-analysis of serological studies measuring complement system activity or protein concentrations in patients with schizophrenia compared to controls. Among studies measuring activity of the classical pathway, activity did not significantly differ between cases and controls in four of six studies. Conflicting results were noted in two studies of alternative pathway activity. Individual studies measuring the classical and MBL pathways reported no evidence of a significant difference. Among studies measuring specific complement protein concentrations, C3 and C4 were the most frequently examined proteins. In meta-analyses, we found no evidence of significant differences between cases and controls for C3 or C4. Regarding other proteins, C1 inhibitor concentration did not differ in two studies. C1q concentration was reduced in one study but did not differ in a second study. No difference was found in two studies of MBL levels. For proteins measured in individual studies, a significant decrease was noted for C4b and C3a, but no significant changes were described for C3 activator, C5, C5a, C5b-9 or MASP2.

Collectively, evidence from these serological studies presents a mixed picture regarding changes in key peripheral complement proteins in schizophrenia. Whereas individual studies report significant differences in activity or concentrations of certain proteins in cases compared to controls, there is little evidence of a consistent pattern of change across studies. However, for most proteins, the available evidence is scarce, limiting our ability to draw broad inferences. Furthermore, the available studies are methodologically and clinically diverse.

2.4.1 Methodological diversity

Regarding methods of measurement, the majority of studies examining activity of the complement system used haemolytic activity assays. However, studies that quantified complement protein concentrations used a range of methods with varying degrees of sensitivity (Bouzek & Mancal, 1989; Koelle & Bartholomew, 1982; Mali et al., 2009), limiting comparisons between studies. Effect sizes for individual proteins may be small and potentially undetectable by relatively insensitive methods. Differences with respect to sample medium, handling, processing and storage between studies are also relevant.

For example, thawing effects are exaggerated for some biomarkers of complement activation in serum and citrated plasma, but not in plasma treated with ethylenediaminetetraacetic acid (Yang et al., 2015). It is also possible that specific isoforms of certain complement proteins are changed or dysregulated in cases relative to controls, but outside of mass spectrometric methods there is limited availability of assays that accurately differentiate between multiple isoforms in this way.

Differences in study design may further contribute to heterogeneity in results. For example, some studies introduced bias by recruiting controls from medical staff or other potentially non-representative sources such as blood donors. Many studies did not adequately control for potential confounders such as age (Stephan et al., 2013), body mass index (Cominetti et al., 2018; Yang et al., 2003) or disorders of physical health such as diabetes mellitus (Engstrom et al., 2005). Finally, in most studies the sample size was small, and possibly under-powered to detect meaningful differences for multiple proteins between cases and controls.

2.4.2 Clinical diversity

Differences related to clinical characteristics of the patient samples in the included studies also limit generalisable inferences. For example, studies differed on whether patients were never, previously or currently prescribed antipsychotic medication. The effects of antipsychotic medication on the complement system have not been widely investigated. However, antipsychotics have a range of immunomodulatory effects including acting upon cytokine networks (Drzyzga et al., 2006; Pollmacher et al., 2000) and through these may exert influence on the complement system. Several of the included studies stratified by medication status, with conflicting results. Among studies assessing complement activity, Boyajyan et al (2010) found increased alternative pathway and C3 activity in medicated and drug-free patients compared to controls. Hakobyan et al (2005) found increased C3 activity in medicated compared to non-medicated patients. Li et al (2012) found decreased alternative pathway activity in drug-free compared to medicated patients. Spivak et al (1993) reported no difference in total complement activity between medicated or non-medicated cases and controls. Among studies measuring complement protein concentrations, Maes et al (1997) reported

increased C3c and C4 in drug-free patients versus controls, but not in patients currently taking antipsychotics. Spivak et al (1993) reported no difference in medicated compared to non-medicated cases with respect to C3 and C4 concentrations. Idonije et al (2012) compared controls with two groups of cases: patients with chronic schizophrenia who were prescribed antipsychotics, and patients with first-episode schizophrenia who were drug-free. As such, it is not possible when interpreting the results from this study to separate the effects of stage of illness and medication. Nevertheless, C1q was found to be reduced in both sets of patients compared to controls; C3c was reduced only in the first-episode drug-free group; and no differences relative to controls were found for C4, C5, C1 inhibitor or C3 activator. To summarise, there is some limited evidence that antipsychotic medication may influence the complement system, but the precise nature of these effects requires further study.

With regard to stage of illness, relatively few studies have investigated complement proteins in patients at different stages of disorder within the same study. There is however value in investigating biomarkers in patients at different stages of illness, and especially in early stages, to gain aetiopathogenic insights and, potentially, to inform early intervention strategies. One included study (Laskaris et al., 2018) compared C1q, C3 and C4 levels in healthy controls with three groups of cases: patients with chronic schizophrenia, patients with FEP and individuals at clinical high-risk (CHR) for psychosis (Yung et al., 2005). Compared to controls, C1q levels did not differ for each of the three groups of cases, C3 was higher in CHR and C4 was higher in CHR and chronic schizophrenia. The sample size for the CHR group was small (ten participants), but these results preliminarily suggest that complement protein changes may be present in early psychosis phenotypes.

2.4.3 Evidence in context

There have been mixed findings regarding complement polymorphisms in candidate gene studies as reviewed previously by Mayilyan et al (2008) and more recently by Woo et al (2019), although genome-wide association studies in schizophrenia provide strong evidence for implication of the major histocompatibility complex (Consortium, 2014; Purcell et al., 2009; Ripke et al., 2013). This association was later partly explained by

allelic variation of C4 (Sekar et al., 2016). C4 RNA expression was found to be increased in post-mortem brain samples from patients with schizophrenia compared to controls, and on immunohistochemical analysis was noted to localise to neurons and synapses (Sekar et al., 2016). In concert with previous data (Stevens et al., 2007), these findings suggest a role for complement in contributing to synaptic pruning. This process is thought to be excessive in schizophrenia, given evidence of reduced synaptic density in association with the disorder (Glausier & Lewis, 2013; Osimo et al., 2018). Complement proteins may serve other functions in the central nervous system such as influencing neuronal migration in neurodevelopment (Gorelik et al., 2017) and neural plasticity after ischaemic injury (Stokowska et al., 2017). Most complement proteins do not cross the blood-brain barrier, and thus it is possible that concentrations of complement proteins in serum or plasma do not reflect complement activity in the brain. Indeed, this may also in part explain the lack of consensus identified from serological studies in schizophrenia. However, the complement system and blood-brain barrier interact with one another (Alexander, 2018; Jacob et al., 2010) and this relationship might be especially relevant in psychosis, given evidence of associated blood-brain barrier dysfunction (Pollak et al., 2018) that might facilitate passage of peripherally-derived complement proteins into the central nervous system, or vice-versa.

Notwithstanding the methodological and clinical diversity between studies, the findings of this review may be in keeping with the hypothesis that significant immune dysregulation occurs in only a sub-population of patients with schizophrenia (Miller & Goldsmith, 2017; Schwarz et al., 2014). In potential support of this view, the high statistical heterogeneity in our meta-analyses of C3 and C4 was not fully explained by between-study variation in antipsychotic use, hospitalisation status or assay methodology. Some of this heterogeneity may reflect biological differences among patients diagnosed with schizophrenia. Higher baseline levels of pro-inflammatory cytokines predict treatment non-response in FEP (Mondelli et al., 2015) suggesting that patients with an óinflammatory subtypeô may b Such patients may be less likely to respond optimally to usual treatments and may potentially benefit from adjunctive anti-inflammatory or immunomodulatory agents (Cho et al., 2019). Whether there are one or several such immunophenotypes in

schizophrenia, potentially characterised by differences in the precise nature and extent of immune disturbances, remains to be determined (Miller & Goldsmith, 2017).

It may also be the case that significant complement dysregulation occurs early in the course of disorder but becomes quiescent by the time of schizophrenia diagnosis (potentially several years after the initial onset of psychosis), and thus is not detectable relative to controls in cross-sectional studies. In support, a recent study in the earlier and broader phenotype of FEP found higher levels of the membrane attack complex in patients compared to controls (Kopczynska et al., 2017). Further evidence that certain complement changes occur early in the psychosis spectrum comes from mass spectrometry-based proteomic studies in non-clinical general population samples showing differential expression of complement proteins in age 12 plasma samples from individuals who go on to report psychotic experiences at age 18 compared to controls who do not (English et al., 2018; Föcking et al., 2019). Another prospective study found that maternal immunoglobulin G markers of C1q were higher in mothers whose children later developed schizophrenic or affective psychosis compared to mothers whose children did not develop mental disorder (Severance et al., 2014) suggesting that these disturbances may even span generations.

Dysregulation of the complement system may also be associated with phenotypes outside of the schizophrenia or psychosis spectrum. Complement protein changes have been observed in other mental disorders (Comes et al., 2018; Druart & Le Magueresse, 2019) including depression (Chen et al., 2015; Stelzhammer et al., 2014), bipolar disorder (Reginia et al., 2018) and autism spectrum disorder (Corbett et al., 2007; Fagan et al., 2017). Thus these changes may reflect shared molecular pathology and aetiological mechanisms between these clinical phenotypes (Druart & Le Magueresse, 2019). The possible primary and perpetuating causes of complement dysregulation and inflammation in schizophrenia and other mental disorders are unclear, but maternal immune activation (Conway & Brown, 2019), exposure to infections (Kneeland & Fatemi, 2013), the microbiome (Severance & Yolken, 2019) and other environmental factors such as childhood adversity (Danese & J Lewis, 2017) and substance use (Miller et al., 2018) are potential contributing factors. Drugs targeting the complement system

are available and others in active development (Carpanini et al., 2019; Druart & Le Magueresse, 2019; Ricklin et al., 2018; Zelek et al., 2019), but whether they may prove useful in the treatment or prevention of psychotic disorders will require extensive preclinical testing before human trials.

2.4.4 Limitations

At the study level, as previously discussed, most studies included in this review were of small sample size and potentially inadequately powered. Studies were generally prone to risk of bias, and reporting of data was inconsistent. In particular, in many studies it was unclear whether the case definition was independently validated; few studies reported the non-response rate in terms of recruitment of cases and controls; and several studies did not adequately control for possible confounders (such as BMI, tobacco and illicit drug use) in the design or analysis.

At the review level, some limitations should be noted. Firstly, our search strategy was limited to studies published in English. Secondly, our focus was on targeted methods of complement protein measurement and thus this synthesis does not include results from proteome-wide investigations. Thirdly, we only included studies published in peer-reviewed academic journals, potentially at the expense of including an increased breadth of literature as derived from grey literature sources such as posters and conference abstracts.

2.4.5 Conclusions

The available evidence from serological studies is inconsistent regarding dysregulation of the complement system in schizophrenia. Studies have generally been small with methodological and clinical heterogeneity. Further studies in larger samples would help to clarify the potential role of complement proteins as biomarkers. Recent evidence from genetic and proteomic studies has suggested that complement changes are detectable and may occur early in the course of disorder. Large-scale prospective studies with repeated biosampling would be helpful to examine longitudinal changes in complement proteins in early and later development of psychotic and other mental disorders. Studies have begun to focus on measurement of complement proteins in early phenotypes,

such as patients in their first episode of psychosis and in clinical high-risk individuals. Pursuing these promising lines of investigation will provide illuminating insights regarding the aetiopathogenic role of the complement system and potential attendant opportunities for early intervention.

Acknowledgements: The authors wish to acknowledge the advice and support provided by Mr Paul Murphy (Information Specialist, Royal College of Surgeons in Ireland) in the planning and execution of the database search.

Disclosure of conflicts of interest: All authors declare no conflicts of interest in relation to this work.

Contributions: DM, SRS, MF and DRC contributed to the conception and design of the review. DM and SS performed the database search and data extraction. DM, SS, SRS, MF, MC and DRC contributed to the drafting and review of the manuscript.

Role of the funding source: This work was performed within the Irish Clinical Academic Training (ICAT) Programme, which is supported by the Wellcome Trust and the Health Research Board (Grant Number 203930/B/16/Z), the Health Service Executive National Doctors Training and Planning and the Health and Social Care, Research and Development Division, Northern Ireland. The funders had no involvement in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Chapter 2 References

Alexander, J. J. (2018). Blood-brain barrier (BBB) and the complement landscape. *Mol Immunol, 102*, 26-31. doi:10.1016/j.molimm.2018.06.267

Ali, F. T., Abd El-Azeem, E. M., Hamed, M. A., Ali, M. A. M., Abd Al-Kader, N. M., & Hassan, E. A. (2017). Redox dysregulation, immuno-inflammatory alterations and genetic variants of BDNF and MMP-9 in schizophrenia: Pathophysiological and phenotypic implications. *Schizophr Res, 188*, 98-109. doi:10.1016/j.schres.2017.01.016

Benros, M. E., Eaton, W. W., & Mortensen, P. B. (2014). The epidemiologic evidence linking autoimmune diseases and psychosis. *Biol Psychiatry, 75*(4), 300-306. doi:10.1016/j.biopsych.2013.09.023

Bouzek, J., & Mancal, P. (1989). Comparability of some serum protein determinations by radial immunodiffusion, laser nephelometric and turbidimetric assays employing Q-antisera SEVAC. *J Hyg Epidemiol Microbiol Immunol, 33*(1), 105-111.

Boyajyan, A., Khoyetsyan, A., & Chavushyan, A. (2010). Alternative complement pathway in schizophrenia. *Neurochem Res, 35*(6), 894-898. doi:10.1007/s11064-010-0126-2

Carpanini, S. M., Torvell, M., & Morgan, B. P. (2019). Therapeutic Inhibition of the Complement System in Diseases of the Central Nervous System. *Front Immunol, 10*(362). doi:10.3389/fimmu.2019.00362

Cazzullo, C. L., Saresella, M., Roda, K., Calvo, M. G., Bertrando, P., Doria, S., . . . Ferrante, P. (1998). Increased levels of CD8+ and CD4+ 45RA+ lymphocytes in schizophrenic patients. *Schizophr Res, 31*(1), 49-55.

Chen, J., Huang, C., Song, Y., Shi, H., Wu, D., Yang, Y., . . . Xie, P. (2015). Comparative proteomic analysis of plasma from bipolar depression and depressive disorder: identification of proteins associated with immune regulatory. *Protein Cell, 6*(12), 908-911. doi:10.1007/s13238-015-0218-5

Cho, M., Lee, T. Y., Kwak, Y. B., Yoon, Y. B., Kim, M., & Kwon, J. S. (2019). Adjunctive use of anti-inflammatory drugs for schizophrenia: A meta-analytic investigation of

randomized controlled trials. *Aust N Z J Psychiatry, 53*(8), 742-759. doi:10.1177/0004867419835028

Comes, A. L., Papiol, S., Mueller, T., Geyer, P. E., Mann, M., & Schulze, T. G. (2018). Proteomics for blood biomarker exploration of severe mental illness: pitfalls of the past and potential for the future. *Transl Psychiatry, 8*(1), 160. doi:10.1038/s41398-018-0219-2

Cominetti, O., Núñez Galindo, A., Corthésy, J., Valsesia, A., Irincheeva, I., Kussmann, M., . . . Dayon, L. (2018). Obesity shows preserved plasma proteome in large independent clinical cohorts. *Sci Rep, 8*(1), 16981. doi:10.1038/s41598-018-35321-7

Consortium, Schizpohrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature, 511*(7510), 421-427. doi:10.1038/nature13595

Conway, F., & Brown, A. S. (2019). Maternal Immune Activation and Related Factors in the Risk of Offspring Psychiatric Disorders. *Front Psychiatry, 10*, 430. doi:10.3389/fpsyt.2019.00430

Corbett, B. A., Kantor, A. B., Schulman, H., Walker, W. L., Lit, L., Ashwood, P., . . . Sharp, F. R. (2007). A proteomic study of serum from children with autism showing differential expression of apolipoproteins and complement proteins. *Mol Psychiatry, 12*(3), 292-306. doi:10.1038/sj.mp.4001943

Costabile, M. (2010). Measuring the 50% haemolytic complement (CH50) activity of serum. *Journal of visualized experiments : JoVE*(37), 1923. doi:10.3791/1923

Cullen, A. E., Holmes, S., Pollak, T. A., Blackman, G., Joyce, D. W., Kempton, M. J., . . . Mondelli, V. (2019). Associations Between Non-neurological Autoimmune Disorders and Psychosis: A Meta-analysis. *Biol Psychiatry, 85*(1), 35-48. doi:10.1016/j.biopsych.2018.06.016

Danese, A., & J Lewis, S. (2017). Psychoneuroimmunology of Early-Life Stress: The Hidden Wounds of Childhood Trauma? *Neuropsychopharmacology, 42*(1), 99-114. doi:10.1038/npp.2016.198

Collaborators, GBD 2016 Disease and Injury Incidence and Prevalence (2017). Global, regional, and national incidence, prevalence, and years lived with disability for

328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet, 390*(10100), 1211-1259. doi:10.1016/S0140-6736(17)32154-2

Druart, M., & Le Magueresse, C. (2019). Emerging Roles of Complement in Psychiatric Disorders. *Front Psychiatry, 10*, 573. doi:10.3389/fpsyt.2019.00573

Drzyzga, L., Obuchowicz, E., Marcinowska, A., & Herman, Z. S. (2006). Cytokines in schizophrenia and the effects of antipsychotic drugs. *Brain Behav Immun, 20*(6), 532-545. doi:10.1016/j.bbi.2006.02.002

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634. doi:10.1136/bmj.315.7109.629

English, J. A., Lopez, L. M., O'Gorman, A., Focking, M., Hryniewiecka, M., Scaife, C., . . . Cotter, D. R. (2018). Blood-Based Protein Changes in Childhood Are Associated With Increased Risk for Later Psychotic Disorder: Evidence From a Nested Case-Control Study of the ALSPAC Longitudinal Birth Cohort. *Schizophr Bull, 44*(2), 297-306. doi:10.1093/schbul/sbx075

Engstrom, G., Hedblad, B., Eriksson, K. F., Janzon, L., & Lindgarde, F. (2005). Complement C3 is a risk factor for the development of diabetes: a population-based cohort study. *Diabetes, 54*(2), 570-575. doi:10.2337/diabetes.54.2.570

Fagan, K., Crider, A., Ahmed, A. O., & Pillai, A. (2017). Complement C3 Expression Is Decreased in Autism Spectrum Disorder Subjects and Contributes to Behavioral Deficits in Rodents. *Mol Neuropsychiatry, 3*(1), 19-27. doi:10.1159/000465523

Föcking, M., Sabherwal, S., Cates, H. M., Scaife, C., Dicker, P., Hryniewiecka, M., . . . Cotter, D. R. (2019). Complement pathway changes at age 12 are associated with psychotic experiences at age 18 in a longitudinal population-based study: evidence for a role of stress. *Mol Psychiatry*. doi:10.1038/s41380-018-0306-z

Foldager, L., Steffensen, R., Thiel, S., Als, T. D., Nielsen, H. J., Nordentoft, M., . . . Jensenius, J. C. (2012). MBL and MASP-2 concentrations in serum and MBL2 promoter polymorphisms are associated to schizophrenia. *Acta Neuropsychiatrica, 24*(4), 199-207. doi:10.1111/j.1601-5215.2011.00618.x

Fontana, A., Storck, U., Angst, J., Dubs, R., Abegg, A., & Grob, P. J. (1980). An immunological basis of schizophrenia and affective disorders? *Neuropsychobiology, 6*(5), 284-289. doi:10.1159/000117771

Glausier, J. R., & Lewis, D. A. (2013). Dendritic spine pathology in schizophrenia. *Neuroscience, 251*, 90-107. doi:10.1016/j.neuroscience.2012.04.044

Gorelik, A., Sapir, T., Haffner-Krausz, R., Olender, T., Woodruff, T. M., & Reiner, O. (2017). Developmental activities of the complement pathway in migrating neurons. *Nat Commun, 8*, 15096. doi:10.1038/ncomms15096

Hakobyan, S., Boyajyan, A., & Sim, R. B. (2005). Classical pathway complement activity in schizophrenia. *Neurosci Lett, 374*(1), 35-37. doi:10.1016/j.neulet.2004.10.024

Holers, V. M. (2014). Complement and its receptors: new insights into human disease. *Annu Rev Immunol, 32*, 433-459. doi:10.1146/annurev-immunol-032713-120154

Hong, S., Beja-Glasser, V. F., Nfonoyim, B. M., Frouin, A., Li, S., Ramakrishnan, S., . . . Stevens, B. (2016). Complement and microglia mediate early synapse loss in Alzheimer mouse models. *Science, 352*(6286), 712-716. doi:10.1126/science.aad8373

Hong, W., Zhao, M., Li, H., Peng, F., Wang, F., Li, N., . . . Xie, B. (2016). Higher Plasma S100B Concentrations in Schizophrenia Patients, and Dependently Associated with Inflammatory Markers. *Sci Rep, 6*, 27584. doi:10.1038/srep27584

Idonije, O. B., Akinlade, K. S., Ihenyen, O., & Arinola, O. G. (2012). Complement factors in newly diagnosed Nigerian schizoprenic patients and those on antipsychotic therapy. *Niger J Physiol Sci, 27*(1), 19-21.

Ingram, G., Hakobyan, S., Robertson, N. P., & Morgan, B. P. (2009). Complement in multiple sclerosis: its role in disease and potential as a biomarker. *Clinical & Experimental Immunology, 155*(2), 128-139. doi:10.1111/j.1365-2249.2008.03830.x

Jacob, A., Hack, B., Chiang, E., Garcia, J. G. N., Quigg, R. J., & Alexander, J. J. (2010). C5a alters blood-brain barrier integrity in experimental lupus. *The FASEB Journal, 24*(6), 1682-1688. doi:10.1096/fj.09-138834

Ji, R. N., Zhang, L. L., Zhao, M. F., He, H. F., Bai, W., Duan, R. X., & Kou, C. G. (2019). Decreased serum complement component 4 levels in patients with

schizophrenia. *Psychiatr Genet, 29*(4), 127-129. doi:10.1097/ypg.0000000000000226

Kirschfink, M., & Mollnes, T. E. (2003). Modern complement analysis. *Clinical and diagnostic laboratory immunology, 10*(6), 982-989. doi:10.1128/cdli.10.6.982-989.2003

Kneeland, R. E., & Fatemi, S. H. (2013). Viral infection, inflammation and schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry, 42*, 35-48. doi:10.1016/j.pnpbp.2012.02.001

Koelle, M., & Bartholomew, W. R. (1982). Comparison of a fluorometric method with radial immunodiffusion assays for determination of complement components C3 and C4. *J Clin Microbiol, 16*(2), 271-275.

Kopczynska, M., Zelek, W., Touchard, S., Gaughran, F., Di Forti, M., Mondelli, V., . . . Morgan, B. P. (2017). Complement system biomarkers in first episode psychosis. *Schizophr Res*. doi:10.1016/j.schres.2017.12.012

Kucharska-Mazur, J., Tarnowski, M., Dolegowska, B., Budkowska, M., Pedziwiatr, D., Jablonski, M., . . . Samochowiec, J. (2014). Novel evidence for enhanced stem cell trafficking in antipsychotic-naive subjects during their first psychotic episode. *J Psychiatr Res, 49*, 18-24. doi:10.1016/j.jpsychires.2013.10.016

Laskaris, L., Zalesky, A., Weickert, C. S., Di Biase, M. A., Chana, G., Baune, B. T., . . . Cropley, V. (2018). Investigation of peripheral complement factors across stages of psychosis. *Schizophr Res*. doi:10.1016/j.schres.2018.11.035

Li, H., Zhang, Q., Li, N., Wang, F., Xiang, H., Zhang, Z., . . . Hong, W. (2016). Plasma levels of Th17-related cytokines and complement C3 correlated with aggressive behavior in patients with schizophrenia. *Psychiatry Res, 246*, 700-706. doi:10.1016/j.psychres.2016.10.061

Li, Y., Zhou, K., Zhang, Z., Sun, L., Yang, J., Zhang, M., . . . Wan, C. (2012). 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, 8*(10), 2664-2671. doi:10.1039/c2mb25158b

Limosin, F., Rouillon, F., Payan, C., Cohen, J. M., & Strub, N. (2003). Prenatal exposure to influenza as a risk factor for adult schizophrenia. *Acta Psychiatr Scand, 107*(5), 331-335.

Maes, M., Delange, J., Ranjan, R., Meltzer, H. Y., Desnyder, R., Cooremans, W., & Scharpe, S. (1997). Acute phase proteins in schizophrenia, mania and major depression: modulation by psychotropic drugs. *Psychiatry Res, 66*(1), 1-11.

Mali, B., Armbruster, D., Serediak, E., & Ottenbreit, T. (2009). Comparison of immunoturbidimetric and immunonephelometric assays for specific proteins. *Clin Biochem, 42*(15), 1568-1571. doi:10.1016/j.clinbiochem.2009.06.016

Mayilyan, K. R., Arnold, J. N., Presanis, J. S., Soghoyan, A. F., & Sim, R. B. (2006). Increased complement classical and mannan-binding lectin pathway activities in schizophrenia. *Neurosci Lett, 404*(3), 336-341. doi:10.1016/j.neulet.2006.06.051

Mayilyan, K. R., Dodds, A. W., Boyajyan, A. S., Soghoyan, A. F., & Sim, R. B. (2008). Complement C4B protein in schizophrenia. *World J Biol Psychiatry, 9*(3), 225-230. doi:10.1080/15622970701227803

Mayilyan, K. R., Weinberger, D. R., & Sim, R. B. (2008). The complement system in schizophrenia. *Drug News Perspect, 21*(4), 200-210. doi:10.1358/dnp.2008.21.4.1213349

Michailidou, I., Naessens, D. M. P., Hametner, S., Guldenaar, W., Kooi, E.-J., Geurts, J. J. G., . . . Ramaglia, V. (2016). Complement C3 on microglial clusters in multiple sclerosis occur in chronic but not acute disease: Implication for disease pathogenesis. *GLIA, 65*(2), 264-277. doi:10.1002/glia.23090

Miller, B. J., Buckley, P., Seabolt, W., Mellor, A., & Kirkpatrick, B. (2011). Meta-Analysis of Cytokine Alterations in Schizophrenia: Clinical Status and Antipsychotic Effects. *Biol Psychiatry, 70*(7), 663-671. doi:10.1016/j.biopsych.2011.04.013

Miller, B. J., Buckley, P. F., & McEvoy, J. P. (2018). Inflammation, substance use, psychopathology, and cognition in phase 1 of the clinical antipsychotic trials of intervention effectiveness study. *Schizophr Res, 195*, 275-282. doi:10.1016/j.schres.2017.08.027

Miller, B. J., & Goldsmith, D. R. (2017). Towards an Immunophenotype of Schizophrenia: Progress, Potential Mechanisms, and Future Directions.

Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology, 42(1), 299-317. doi:10.1038/npp.2016.211

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med, 6*(7), e1000097. doi:10.1371/journal.pmed.1000097

Mondelli, V., Ciufolini, S., Belvederi Murri, M., Bonaccorso, S., Di Forti, M., Giordano, A., . . . Dazzan, P. (2015). Cortisol and Inflammatory Biomarkers Predict Poor Treatment Response in First Episode Psychosis. *Schizophrenia bulletin, 41*(5), 1162-1170. doi:10.1093/schbul/sbv028

Nilsson, B., & Ekdahl, K. N. (2012). Complement Diagnostics: Concepts, Indications, and Practical Guidelines. *Clinical and Developmental Immunology, 2012*, 11. doi:10.1155/2012/962702

Nimgaonkar, V. L., Prasad, K. M., Chowdari, K. V., Severance, E. G., & Yolken, R. H. (2017). The complement system: a gateway to gene-environment interactions in schizophrenia pathogenesis. *Mol Psychiatry, 22*(11), 1554-1561. doi:10.1038/mp.2017.151

Oppermann, M., & Wurzner, R. (2010). Modern determination of complement activation. *Semin Thromb Hemost, 36*(6), 611-619. doi:10.1055/s-0030-1262882

Osimo, E. F., Beck, K., Reis Marques, T., & Howes, O. D. (2018). Synaptic loss in schizophrenia: a meta-analysis and systematic review of synaptic protein and mRNA measures. *Mol Psychiatry*. doi:10.1038/s41380-018-0041-5

Pollak, T. A., Drndarski, S., Stone, J. M., David, A. S., McGuire, P., & Abbott, N. J. (2018). The bloodï brain barrier in psychosis. *The Lancet Psychiatry, 5*(1), 79-92. doi:10.1016/s2215-0366(17)30293-6

Pollmacher, T., Haack, M., Schuld, A., Kraus, T., & Hinze-Selch, D. (2000). Effects of antipsychotic drugs on cytokine networks. *J Psychiatr Res, 34*(6), 369-382. doi:10.1016/s0022-3956(00)00032-7

Popescu, A., & H. Kao, A. (2011). Neuropsychiatric Systemic Lupus Erythematosus. *Curr Neuropharmacol, 9*(3), 449-457. doi:10.2174/157015911796557984

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., & Sklar, P. (2009). Common polygenic variation contributes to risk of

schizophrenia and bipolar disorder. *Nature, 460*(7256), 748-752. doi:10.1038/nature08185

Ramsey, J. M., Schwarz, E., Guest, P. C., van Beveren, N. J., Leweke, F. M., Rothermundt, M., . . . Bahn, S. (2013). Distinct molecular phenotypes in male and female schizophrenia patients. *PLoS One, 8*(11), e78729. doi:10.1371/journal.pone.0078729

Reginia, A., Kucharska-Mazur, J., Jabğoœski, M., Budkowska Sagan, L., . . . Samochowiec, J. (2018). Assessment of Complement Cascade Components in Patients With Bipolar Disorder. *Front Psychiatry, 9*, 614-614. doi:10.3389/fpsyt.2018.00614

Ricklin, D., Mastellos, D. C., Reis, E. S., & Lambris, J. D. (2018). The renaissance of complement therapeutics. *Nature reviews. Nephrology, 14*(1), 26-47. doi:10.1038/nrneph.2017.156

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kahler, A. K., Akterin, S., . . . Sullivan, P. F. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet, 45*(10), 1150-1159. doi:10.1038/ng.2742

Sabherwal, S., English, J. A., Focking, M., Cagney, G., & Cotter, D. R. (2016). Blood biomarker discovery in drug-free schizophrenia: the contributionof proteomics and multiplex immunoassays. *Expert Rev Proteomics, 13*(12), 1141-1155. doi:10.1080/14789450.2016.1252262

Santos Soria, L., Moura Gubert, C., Cereser, K. M., Gama, C. S., & Kapczinski, F. (2012). Increased serum levels of C3 and C4 in patients with schizophrenia compared to eutymic patients with bipolar disorder and healthy. *Rev Bras Psiquiatr, 34*(1), 119-120.

Sasaki, T., Nanko, S., Fukuda, R., Kawate, T., Kunugi, H., & Kazamatsuri, H. (1994). Changes of immunological functions after acute exacerbation in schizophrenia. *Biol Psychiatry, 35*(3), 173-178.

Schwarz, E., Guest, P. C., Rahmoune, H., Harris, L. W., Wang, L., Leweke, F. M., . . . Bahn, S. (2012). Identification of a biological signature for schizophrenia in serum. *Mol Psychiatry, 17*(5), 494-502. doi:10.1038/mp.2011.42

Schwarz, E., van Beveren, N. J., Ramsey, J., Leweke, F. M., Rothermundt, M., Bogerts, B., . . . Bahn, S. (2014). Identification of subgroups of schizophrenia patients with changes in either immune or growth factor and hormonal pathways. *Schizophrenia bulletin, 40*(4), 787-795. doi:10.1093/schbul/sbt105

Seelen, M. A., Roos, A., Wieslander, J., Mollnes, T. E., Sjoholm, A. G., Wurzner, R., . . . Daha, M. R. (2005). Functional analysis of the classical, alternative, and MBL pathways of the complement system: standardization and validation of a simple ELISA. *J Immunol Methods, 296*(1-2), 187-198. doi:10.1016/j.jim.2004.11.016

Sekar, A., Bialas, A. R., de Rivera, H., Davis, A., Hammond, T. R., Kamitaki, N., . . . McCarroll, S. A. (2016). Schizophrenia risk from complex variation of complement component 4. *Nature, 530*(7589), 177-183. doi:10.1038/nature16549

Severance, E. G., Gressitt, K. L., Buka, S. L., Cannon, T. D., & Yolken, R. H. (2014). Maternal complement C1q and increased odds for psychosis in adult offspring. *Schizophr Res, 159*(1), 14-19. doi:10.1016/j.schres.2014.07.053

Severance, E. G., & Yolken, R. H. (2019). From Infection to the Microbiome: An Evolving Role of Microbes in Schizophrenia. *Curr Top Behav Neurosci*. doi:10.1007/7854_2018_84

Spivak, B., Radwan, M., Brandon, J., Baruch, Y., Stawski, M., Tyano, S., & Weizman, A. (1993). Reduced total complement haemolytic activity in schizophrenic patients. *Psychol Med, 23*(2), 315-318.

Spivak, B., Radwan, M., Elimelech, D., Baruch, Y., Avidan, G., & Tyano, S. (1989). A study of the complement system in psychiatric patients. *Biol Psychiatry, 26*(6), 640-642.

Stelzhammer, V., Haenisch, F., Chan, M. K., Cooper, J. D., Steiner, J., Steeb, H., . . . Bahn, S. (2014). Proteomic changes in serum of first onset, antidepressant drug-naive major depression patients. *Int J Neuropsychopharmacol, 17*(10), 1599-1608. doi:10.1017/s1461145714000819

Stephan, A. H., Barres, B. A., & Stevens, B. (2012). The Complement System: An Unexpected Role in Synaptic Pruning During Development and Disease. *Annu Rev Neurosci, 35*(1), 369-389. doi:10.1146/annurev-neuro-061010-113810

Stephan, A. H., Madison, D. V., Mateos, J. M., Fraser, D. A., Lovelett, E. A., Coutellier, L., . . . Barres, B. A. (2013). A dramatic increase of C1q protein in the CNS during normal aging. *J Neurosci, 33*(33), 13460-13474. doi:10.1523/jneurosci.1333-13.2013

Stevens, B., Allen, N. J., Vazquez, L. E., Howell, G. R., Christopherson, K. S., Nouri, N., . . . Barres, B. A. (2007). The classical complement cascade mediates CNS synapse elimination. *Cell, 131*(6), 1164-1178. doi:10.1016/j.cell.2007.10.036

Stokowska, A., Atkins, A. L., Moran, J., Pekny, T., Bulmer, L., Pascoe, M. C., . . . Pekna, M. (2017). Complement peptide C3a stimulates neural plasticity after experimental brain ischaemia. *Brain, 140*(2), 353-369. doi:10.1093/brain/aww314

Upthegrove, R., Manzanares-Teson, N., & Barnes, N. M. (2014). Cytokine function in medication-naive first episode psychosis: A systematic review and meta-analysis. *Schizophr Res, 155*(1-3), 101-108. doi:10.1016/j.schres.2014.03.005

VanItallie, T. B. (2017). Alzheimer's disease: Innate immunity gone awry? *Metabolism, 69*, S41-S49. doi:10.1016/j.metabol.2017.01.014

Walss-Bass, C., Lokesh, G. L. R., Dyukova, E., Gorenstein, D. G., Roberts, D. L., Velligan, D., & Volk, D. E. (2019). X-Aptamer Technology Identifies C4A and ApoB in Blood as Potential Markers for Schizophrenia. *Mol Neuropsychiatry, 5*(1), 52-59. doi:10.1159/000492331

Wells, G., Shea, B., Ô Connell, D., Losos, M., & Tugwell, P. (2013). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Retrieved from http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Wong, C. T., Tsoi, W. F., & Saha, N. (1996). Acute phase proteins in male Chinese schizophrenic patients in Singapore. *Schizophr Res, 22*(2), 165-171.

Woo, J. J., Pouget, J. G., Zai, C. C., & Kennedy, J. L. (2019). The complement system in schizophrenia: where are we now and what's next? *Mol Psychiatry*. doi:10.1038/s41380-019-0479-0

Yang, S., McGookey, M., Wang, Y., Cataland, S. R., & Wu, H. M. (2015). Effect of Blood Sampling, Processing, and Storage on the Measurement of Complement

Activation Biomarkers. *American Journal of Clinical Pathology, 143*(4), 558-565. doi:10.1309/AJCPXPD7ZQXNTIAL

Yang, Y., Chung, E. K., Zhou, B., Blanchong, C. A., Yu, C. Y., Fust, G., . . . Varga, L. (2003). Diversity in intrinsic strengths of the human complement system: serum C4 protein concentrations correlate with C4 gene size and polygenic variations, hemolytic activities, and body mass index. *J Immunol, 171*(5), 2734-2745. doi:10.4049/jimmunol.171.5.2734

Yung, A. R., Yuen, H. P., McGorry, P. D., Phillips, L. J., Kelly, D., Dell'Olio, M., . . . Buckby, J. (2005). Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. *Aust N Z J Psychiatry, 39*(11-12), 964-971. doi:10.1080/j.1440-1614.2005.01714.x

Zelek, W. M., Xie, L., Morgan, B. P., & Harris, C. L. (2019). Compendium of current complement therapeutics. *Molecular Immunology, 114*, 341-352. doi:https://doi.org/10.1016/j.molimm.2019.07.030

Zipfel, P. F., & Skerka, C. (2009). Complement regulators and inhibitory proteins. *Nat Rev Immunol, 9*(10), 729-740. doi:10.1038/nri2620

Chapter 2 Appendices

Appendix 2.1: Search terms used for each database é é é é é é é é é é é 8 6

Appendix 2.2: Risk of bias assessment based on the Newcastle Ottawa scale for case control studies é é é é é é é é é é é é é é é é é é 8 8

Appendix 2.3: Assessment of publication bias and small-study effects é é é . 89

Appendix 2.4: Results of meta-regression analyses é é é é é é é é é é é 9 0

Appendix 2.1: Search terms used for each database

PubMed Medline	
1	(("Schizophrenia"[Mesh]) OR "Psychotic Disorders"[Mesh]) OR ((((schizophrenia[Title/Abstract]) OR schizophrenic[Title/Abstract]) OR psychosis[Title/Abstract]) OR psychotic[Title/Abstract])
2	("Complement System Proteins"[Mesh]) OR complement[Title/Abstract]
3	**1 AND 2**

EMBASE Elsevier	
1	'schizophrenia'/exp OR schizoprenia:ti,ab OR schizophrenic:ti,ab OR 'psychosis'/exp OR psychosis:ti,ab OR psychotic:ti,ab
2	'complement'/exp OR complement:ti,ab,de
3	1 AND 2
4	**#3 AND [embase]/lim NOT ([embase]/lim AND [medline]/lim)** **LIMITED TO EMBASE ONLY RECORDS, EXCLUDING MEDLINE**

PSYCINFO Ebscohost	
1	DE "Schizophrenia" OR DE "Acute Schizophrenia" OR DE "Catatonic Schizophrenia" OR DE "Childhood Schizophrenia" OR DE "Paranoid Schizophrenia" OR DE "Process Schizophrenia" OR DE "Schizophrenia (Disorganized Type)" OR DE "Schizophreniform Disorder" OR DE "Undifferentiated Schizophrenia"
2	DE "Psychosis" OR DE "Acute Psychosis" OR DE "Affective Psychosis" OR DE "Alcoholic Psychosis" OR DE "Capgras Syndrome" OR DE "Childhood Psychosis" OR DE "Chronic Psychosis" OR DE "Experimental Psychosis" OR DE "Hallucinosis" OR DE "Paranoia (Psychosis)" OR DE "Postpartum Psychosis" OR DE "Reactive Psychosis" OR DE "Schizophrenia" OR DE "Senile Psychosis" OR DE "Toxic Psychoses"
3	(TI schizophrenia OR schizophrenic OR psychosis OR psychotic) OR (AB schizophrenia OR schizophrenic OR psychosis OR psychotic)
4	**1 OR 2 OR 3**
5	(TI complement) OR (AB complement)
6	**4 AND 5**

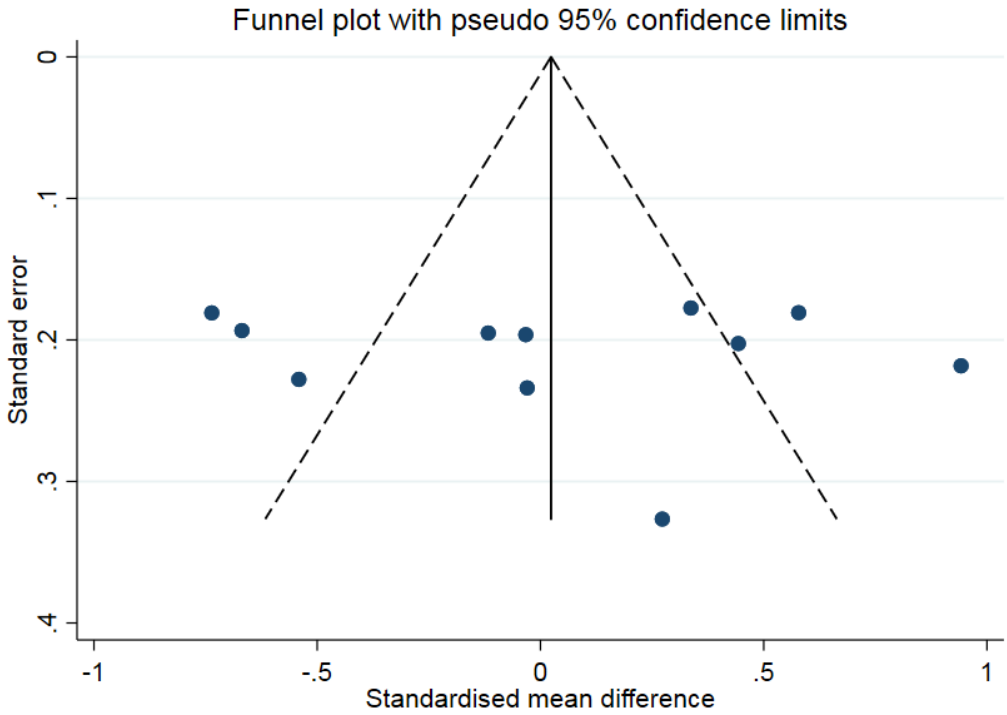
Appendix 2.2: Risk of bias assessment based on the Newcastle-Ottawa scale for case-control studies

Study	Selection score (maximum 4)	Comparability score (maximum 2)	Exposure score (maximum 3)	Total score (maximum 9)
Ali et al 2017	3	2	2	7
Boyajyan et al 2010	2	0	0	2
Cazzullo et al 1998	1	1	0	2
Foldager et al 2012	0	1	0	1
Fontana et al 1980	1	0	0	1
Hakobyan et al 2004	2	0	0	2
Hong et al 2016	2	2	0	4
Idonije et al 2012	1	2	0	3
Ji et al 2019	3	0	0	3
Kucharska-Mazur et al 2014	2	2	0	4
Laskaris et al 2018	1	2	1	4
Li et al 2012	2	1	0	3
Li et al 2016	1	2	1	4
Maes et al 1997	1	2	0	3
Mayilyan et al 2006	0	2	0	2
Mayiliyan et al 2008	3	1	0	4
Ramsey et al 2013	2	2	1	5
Santos-Soria et al 2012	1	2	0	3
Sasaki et al 1994	1	0	0	1
Schwarz et al 2012	3	2	0	5
Spivak et al 1989	0	1	0	1
Spivak et al 1993	1	1	0	2
Walls-Bass et al 2018	2	2	1	5
Wong et al 1996	0	1	0	1

Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

Appendix 2.3: Assessment of publication bias and small-study effects

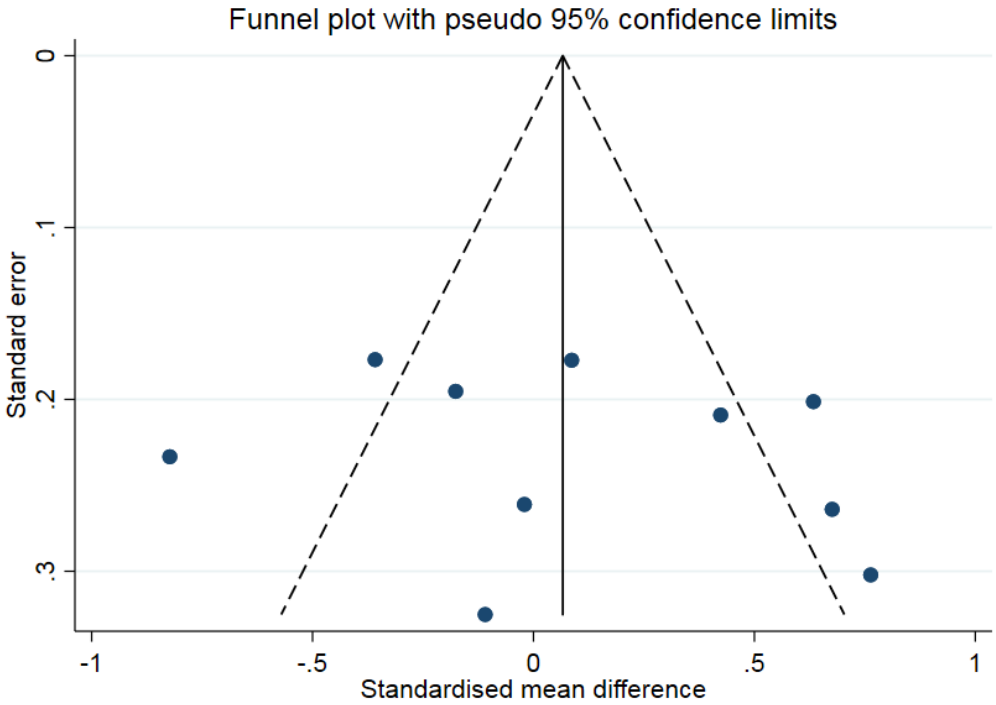
1) Funnel plot and Eggerôs test for complement



Eggerôs small-study effects:

t=0.28, *p*=0.789

2) Funnel plot and Eggerôs test for complement



Eggerôs t small-study effects:

t=0.67, *p*=0.524

Appendix 2.4: Results of meta-regression analyses

1) Meta-regression analyses for complement component 3

Characteristic	Number of studies	Coefficient	Standard error	t	*p*	Lower 95% confidence interval	Upper 95% confidence interval
Current antipsychotic use (yes vs. no)	11	-0.29	0.33	-0.86	0.412	-1.05	0.47
Hospitalisation status (inpatient vs. outpatient)	11	-0.63	0.32	-1.96	0.082	-1.36	0.10
Assay (ELISA vs radial immunodiffusion/other)	11	-0.15	0.36	-0.43	0.679	-0.96	0.65
Assay (radial immunodiffusion vs ELISA/other)	11	-0.36	0.37	-0.97	0.357	-1.20	0.48
Assay (other vs ELISA/ radial immunodiffusion)	11	0.45	0.33	1.38	0.200	-0.29	1.19

2) Meta-regression analyses for complement component 4

Characteristic	Number of studies	Coefficient	Standard error	t	*p*	Lower 95% confidence interval	Upper 95% confidence interval
Current antipsychotic use (yes vs. no) [a]	9	0.01	0.37	0.02	0.988	-0.86	0.87
Hospitalisation status (inpatient vs. outpatient)	10	-0.60	0.27	-2.20	0.059	-1.22	0.03
Assay (ELISA vs radial immunodiffusion/other)	10	0.09	0.37	0.23	0.821	-0.77	0.94
Assay (radial immunodiffusion vs ELISA/other)	10	-0.31	0.42	-0.73	0.488	-1.28	0.66
Assay (other vs ELISA/ radial immunodiffusion)	10	0.12	0.34	0.34	0.740	-0.67	0.90

[a] Medication status was unable to be confirmed for one study (Ji et al 2019) and thus this study was not included in this specific analysis. ELISA: enzyme-linked immunosorbent assay.

Chapter 3. Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence

Chapter overview

This chapter comprises a manuscript published in *JAMA Psychiatry* examining proteomic biomarkers for prediction of transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence.

Role of the thesis author

Dr Mongan contributed to the bioinformatics and statistical analyses, interpretation of data and drafting of the initial manuscript and final version of the text following peer review. He was the first author on the published manuscript.

Reference: Mongan D, Föcking M, Healy C, Susai SR, Heurich M, Wynne K, Nelson B, McGorry PD, Amminger GP, Nordentoft M, Krebs MO, Riecher-Rössler A, Bressan RA, Barrantes-Vidal N, Borgwardt S, Ruhrmann S, Sachs G, Pantelis C, van der Gaag M, de Haan L, Valmaggia L, Pollak TA, Kempton MJ, Rutten BPF, Whelan R, Cannon M, Zammit S, Cagney G, Cotter DR, McGuire P, European Network of National Schizophrenia Networks Studying Gene-Environment Interactions (EU-GEI) High Risk Study Group. Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence. *JAMA Psychiatry* 2021;78(1):77-90. doi: 10.1001/jamapsychiatry.2020.2459.

An animation explaining the findings of this study for a general audience is available to view online at https://www.youtube.com/watch?v=zcs8j8z_R7Q.

Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence

Authors: David Mongan[1], MB BCh BAO; Melanie Föcking[1], PhD; Colm Healy[1], MSc; Subash Raj Susai[1], MSc; Meike Heurich[2], PhD; Kieran Wynne[3], MSc; Barnaby Nelson[4], PhD; Pat McGorry[4], PhD; Paul Amminger[4], PhD; Merete Nordentoft[5], PhD; Marie-Odile Krebs[6], PhD; Anita Riecher-Rössler[7], PhD; Rodrigo A. Bressan[8], PhD; Neus Barrantes-Vidal[9], PhD; Stefan Borgwardt[7,10], MD; Stephan Ruhrmann[11], MD; Gabriele Sachs[12], PhD; Christos Pantelis[13], MD; Mark van der Gaag[14,15], PhD; Lieuwe de Haan[16], PhD; Lucia Valmaggia[17], PhD; Thomas A Pollak[18], PhD; Matthew J Kempton[18], PhD; Bart Rutten[19], PhD; Robert Whelan[20], PhD; Mary Cannon[1], PhD; Stan Zammit[21,22], PhD; Gerard Cagney[3], PhD; David R. Cotter[1*ƶ], PhD; Philip McGuire[18*], PhD for the EU-GEI High Risk Study GroupÀ.

*Joint senior authors ƶCorresponding author ÀSee end o
and affiliations

1. Department of Psychiatry, Royal College of Surgeons in Ireland, Dublin, Ireland
2. School of Pharmacy and Pharmaceutical Sciences, Cardiff University
3. School of Biomolecular and Biomedical Science, Conway Institute, University College Dublin (UCD)
4. Centre for Youth Mental Health, University of Melbourne, 35 Poplar Road (Locked Bag 10), Parkville, Victoria 485 3052, Australia
5. Mental Health Centre Copenhagen, Copenhagen University Hospital
6. University Paris Descartes, GHU Paris ï Sainte Anne, CôJAAD, Hospitalo-Universitaire, Inserm U1266, Institut de Psychiatrie (CNRS 3557) Paris, France
7. Department of Psychiatry, Medical Faculty, University of Basel, Switzerland
8. LiNC - Lab Interdisciplinar Neurociências Clínicas, Depto Psiquiatria, Escola Paulista de Medicina, Universidade Federal de São Paulo ï UNIFESP
9. Departament de Psicologia Clínica i de la Salut (Universitat Autònoma de Barcelona), Fundació Sanitària Sant Pere Claver (Spain), Spanish Mental Health Research Network (CIBERSAM)

10. Department of Psychiatry and Psychotherapy, Translational Psychiatry Unit, University zu Lübeck, Germany
11. Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany
12. Medical University of Vienna, Department of Psychiatry and Psychotherapy
13. Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne and Melbourne Health, Carlton South, Victoria, Australia
14. VU University, Faculty of Behavioural and Movement Sciences, Department of Clinical Psychology and EMGO+ Institute for Health and Care Research, van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands
15. Parnassia Psychiatric Institute, Department of Psychosis Research, Zoutkeetsingel 40, 2512 HN The Hague, The Netherlands
16. AMC, Academic Psychiatric Centre, Department Early Psychosis, Meibergdreef 5, 1105 AZ Amsterdam, The Netherlands
17. Department of Psychology, Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom.
18. Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom.
19. Department of Psychiatry & Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Centre, Maastricht, The Netherlands
20. Trinity Institute of Neuroscience, Trinity College Dublin, Ireland
21. MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK
22. Bristol Medical School, University of Bristol, Bristol, UK

Corresponding author: Professor David Cotter, Department of Psychiatry, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin 9, Ireland.

Keywords: plasma, proteomics, Ultra High-Risk, Clinical High-Risk, At Risk Mental State, psychotic disorder, psychotic experiences, machine learning, complement, coagulation, EU-GEI, ALSPAC

Word count: 3746

Number of tables: 3

Number of figures: 2

Appendices:

Appendix 3.1: Supplementary Methods

Appendix 3.2: Supplementary Tables 3.1 ï 3.21

Appendix 3.3: Supplementary Figures 3.1 ï 3.25

Abstract

Importance: Biomarkers that are predictive of outcomes in individuals at risk of psychosis would facilitate individualised prognosis and stratification strategies.

Objective: To investigate whether proteomic biomarkers may aid prediction of transition to psychotic disorder in the clinical high-risk (CHR) state and adolescent psychotic experiences (PEs) in the general population.

Design, setting and participants: This diagnostic study comprised two case-control studies nested within the European Network of National Schizophrenia Networks Studying Gene-Environment Interactions (EU-GEI) and the Avon Longitudinal Study of Parents and Children (ALSPAC). EU-GEI is an international multisite prospective study of participants at CHR referred from local mental health services. ALSPAC is a United Kingdomï based general population birth cohort. Included were EU-GEI participants who met CHR criteria at baseline and ALSPAC participants who did not report PEs at age 12 years. Data were analysed from September 2018 to April 2020.

Main outcomes and measures: In EU-GEI, transition status was assessed by the Comprehensive Assessment of At-Risk Mental States or contact with clinical services. In ALSPAC, PEs at age 18 years were assessed using the Psychosis-Like Symptoms Interview. Proteomic data were obtained from mass spectrometry of baseline plasma samples in EU-GEI and plasma samples at age 12 years in ALSPAC. Support vector machine learning algorithms were used to develop predictive models.

Results: The EU-GEI subsample (133 participants at CHR; mean [SD] age, 22.6 [4.5] years; 68 [51.1%] male) comprised 49 (36.8%) participants who developed psychosis and 84 (63.2%) who did not. A model based on baseline clinical and proteomic data demonstrated excellent performance for prediction of transition outcome (area under the receiver operating characteristic curve [AUC], 0.95; positive predictive value [PPV], 75.0%; and negative predictive value [NPV], 98.6%). Functional analysis of differentially expressed proteins implicated the complement and coagulation cascade. A model based on the 10 most predictive proteins accurately predicted transition status in training (AUC, 0.99; PPV, 76.9%; and NPV, 100%) and test (AUC, 0.92; PPV, 81.8%; and NPV, 96.8%) data. The ALSPAC subsample (121 participants from the general population with plasma samples

available at age 12 years; 61 [50.4%] male) comprised 55 participants (45.5%) with PEs at age 18 years and 61 (50.4%) without PEs at age 18 years. A model using proteomic data at age 12 years predicted PEs at age 18 years, with an AUC of 0.74 (PPV, 67.8%; and NPV, 75.8%).

Conclusions and Relevance: In individuals at risk of psychosis, proteomic biomarkers may contribute to individualised prognosis and stratification strategies. These findings implicate early dysregulation of the complement and coagulation systems in development of psychosis outcomes.

3.1 Introduction

Early detection of psychosis may improve clinical outcomes.[1] Clinical high-risk (CHR) criteria[2] enable identification of vulnerable groups with 3-year transition rates to first-episode psychosis (FEP) of 16ï 35%.[3] However, it is difficult to predict outcomes individually. Previous studies have also characterised an phenotypeô including individual psycho[4] These with subthreshold symptoms are associated with increased risk of psychotic and non-psychotic disorders[5] and reduced functioning.[6]

Biomarkers may augment prognosis and stratification strategies.[7] We aimed to compare plasma protein expression in individuals at CHR who do and do not develop psychosis and to develop models incorporating proteomic data for individualised prediction of transition to FEP. We also aimed to apply similar methods for prediction of PEs in a general population sample.

3.2 Methods

Ethical approval for this diagnostic study was granted by the Royal College of Surgeons in Ireland (REC1240bb). Ethics committees of participating sites granted approval for the European Network of National Schizophrenia Networks studying Gene-Environment Interactions (EU-GEI) study. Approval was also obtained from the Avon Longitudinal Study of Parents and Children (ALSPAC) Ethics and Law Committee and local research ethics committees. Informed consent for biological samples was obtained in accordance with the Human Tissue Act (2004). Informed consent for use of questionnaire and clinic data was obtained following recommendations of ALSPAC Ethics and Law Committee at the time.

3.2.1 Study 1: CHR sample

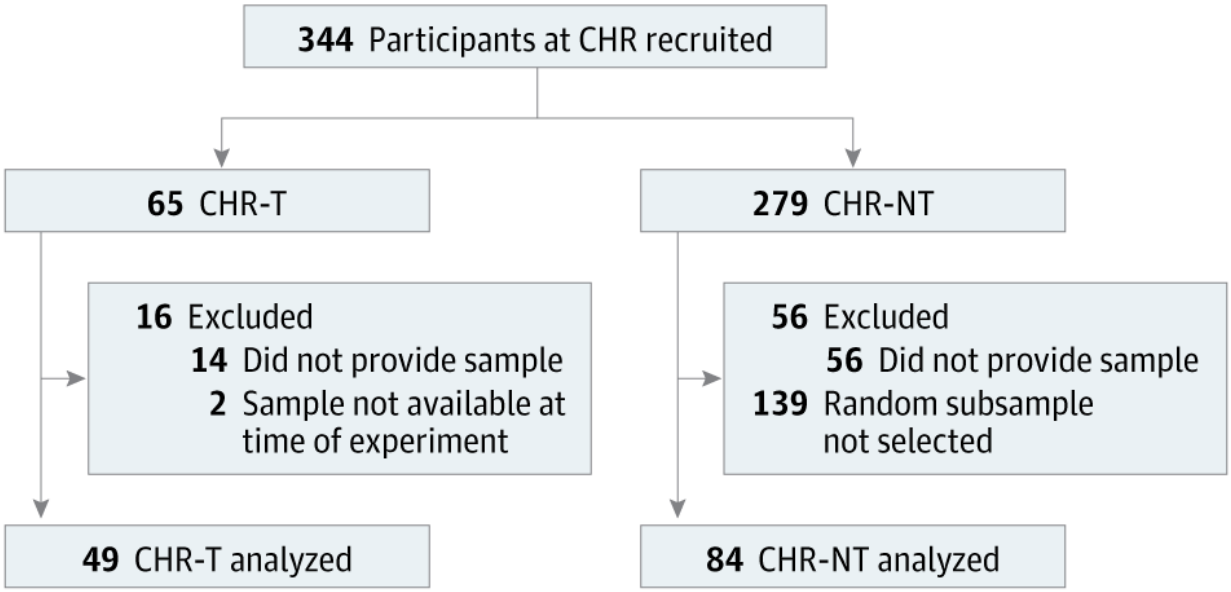
Participants and study design

The EU-GEI study includes a prospective cohort of 344 CHR participants recruited across 11 international sites.[8,9] Individuals with CHR symptoms who were referred by local mental health services were eligible to participate if they met CHR criteria as per the Comprehensive Assessment of At-Risk Mental States[10] (CAARMS) and provided written informed consent. Exclusion criteria were: current or past psychotic disorder; symptoms explained by a medical disorder or drug or alcohol use; and IQ less than 60.

Plasma samples were obtained at baseline, and clinical assessments were performed at baseline, 12 months and 24 months. After 24 months, or if a face-to-face interview was not possible, attempts were made to confirm transition status via the clinical team or records. Assessors were not systematically blind to transition status because, in some cases, clinical services contacted the research team in advance to advise that transition had occurred. Accrual began in September 2010. The last baseline assessment was performed in July 2015.

The present investigation comprised a nested case-control study comparing plasma proteins from CHR participants who transitioned to psychosis on follow-up (CHR-T, n=49) with a control group of randomly-selected participants who did not (CHR-NT, n=84) (Figure 3.1). Based on prior experience[11] the experiment was limited to this number to ensure optimal technical performance across mass spectrometry runs.

Figure 3.1. Derivation of participants included in the initial EU-GEI mass spectrometry experiment and their provision of plasma samples



CHR indicates clinical high risk; CHR-NT, participants at clinical high risk who did not transition to psychosis; CHR-T, participants at clinical high risk who transitioned to first-episode psychosis; and EU-GEI, European Network of National Schizophrenia Networks Studying Gene-Environment Interactions.

Outcome and clinical measures

Transition was defined as the onset of non-organic psychotic disorder as assessed either by CAARMS interview[10] or by contact with the clinical team or review of clinical records. Sixty-five of 344 participants at CHR (18.9%) developed psychosis on follow-up, 57 within 24 months and 8 after 24 months.

Baseline clinical measures were recorded. These included age, sex, body mass index (BMI; calculated as weight in kilograms divided by height in metres squared), years in education, General Assessment of Functioning (GAF) subscales for symptoms and disability,[12,13] the Scale for Assessment of Negative Symptoms (SANS),[14] the Brief Psychiatric Rating Scale (BPRS)[15] and the Montgomery- Åsberg Depression Rating Scale (MADRS).[16]

Sample preparation and proteomics

Laboratory staff were blind to case-control status. Protein depletion, digestion and peptide purification were performed using baseline plasma samples. Discovery-based proteomic methods were used.[11] Briefly, 5εl from each pre injected on a Thermo Scientific Q-Exactive mass spectrometer operated in data-dependent acquisition (DDA) mode for label-free liquid chromatography mass spectrometry[11,17-19] (see Supplementary Methods).

Enzyme-linked immunosorbent assay (ELISA) validation

Nine proteins in plasma samples from the same participants at CHR described above (Figure 3.1) were assessed using ELISA (see Supplementary Methods).

Replication experiment

In an effort to reproduce our findings, we conducted a partial replication of the initial mass spectrometry experiment by analysing baseline plasma from 49 CHR-T cases (two of these cases were different from the initial experiment) and an entirely new group of 86 CHR-NT controls (see Supplementary Methods).

3.2.2 Study 2: general population sample

Participants and study design

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a prospective birth cohort[20,21] (http://www.bristol.ac.uk/alspac/researchers/our-data/). Pregnant

women in Avon, United Kingdom with delivery dates between 1st April 1991 to 31st December 1992 were invited and 14,541 pregnancies were enrolled. When the oldest children were approximately age 7 years, an attempt was made to bolster the sample with cases who did not join originally. The sample size for analyses using data from age 7 years is 15,454 pregnancies (14,901 children alive at 1 year).

Plasma samples obtained at age 12 years from ALSPAC participants who did or did not report PEs at age 18 years were previously investigated.[11,19] In data-independent acquisition (DIA) analyses focused on proteins of the complement pathway, several were differentially expressed. Here, we performed DDA analyses (rather than DIA) in this sample to achieve broader proteome coverage.

Outcome

Psychotic experiences (PEs) were assessed at 12 and 18 years using the Psychosis-Like Symptom Interview[4] and rated as not present, suspected or definite. Of 4,060 participants assessed at both timepoints, 190 (4.7%) had suspected or definite PEs at age 18 years but not at age 12 years.[4] The present study was based on a subsample of case participants (who did not report PEs at age 12 years but reported at least one definite PE at age 18) and randomly-selected control participants (who did not report PEs at either age 12 or 18).

Sample preparation and proteomics

Age 12 plasma samples were prepared as previously described.[11] Data-dependent acquisition proteomic analyses performed as for Study 1.

3.2.3 Data and statistical analyses

Data were analysed from September 2018 to April 2020. Clinical data were tested for differences using the two-sided t-test for continuous and χ^2-test for categorical variables in SPSS v.25 (IBM). *P* values were corrected for multiple comparisons using the Benjamini-Hochberg procedure[22] with 5% false discovery rate (FDR). The threshold for statistical significance was FDR-corrected $p<0.05$.

Label-free quantification (LFQ) was performed in MaxQuant v.1.5.2.8.[23,24] Proteins identified with at least two peptides (one uniquely assignable to the protein) and quantified in more than 80% of plasma samples were taken forward for analysis and log2-transformed. Missing values were imputed using imputeLCMD v.2.0[25] in

RStudio (http://www.rstudio.com/). LFQ values were converted to z-scores and winsorised within ±3z.

Analysis of covariance (ANCOVA) was performed in Stata version 15 (StataCorp LLC), comparing mean LFQ for each protein in cases and controls. Covariates included age, sex, BMI and years in education in Study 1 and sex, BMI at age 12 years and maternal social class in Study 2. *P* values were corrected for multiple comparisons with 5% FDR.

3.2.4 Predictive models

Neurominer version 1.0 (https://www.pronia.eu/neurominer/) for MatLab 2018a (MathWorks Inc.) was used to develop support vector machine (SVM) models (see Supplementary Methods). Development of each model is summarised in Supplementary Table 3.1.

Models 1a-c: predicting transition using clinical and proteomic data

First, we developed a model predicting transition using clinical and proteomic data together (model 1a). Supplementary Table 3.2 lists the included clinical features. Geographical generalisability was incorporated using leave-site-out cross-validation (see Supplementary Methods) as recommended for multisite consortia.[26]

Next, to assess the relative contribution of clinical and proteomic data, we developed models using the same cross-validation and training framework but based on clinical (model 1b) and proteomic features (model 1c) separately.

Models 2a-b: parsimonious model

We sought to generate a parsimonious model based on the 10 highest-weighted proteomic predictors, and internally validate this model in unseen data (Supplementary Figure 3.1). As the largest site, London, United Kingdom was chosen as test site, and data for these participants were held out.

To derive the 10 highest-weighted proteins, we generated a model (model 2a) using proteomic data from all sites except London (CHR-T n=30 and CHR-NT n=50). A reduced model was then developed based solely on data for these 10 proteins in the non-London dataset (model 2b), then tested in the held-out London data (CHR-T n=19 and CHR-NT n=34). Both models used leave-site-out cross-validation.

Model 3: replication

Due to differences in protein identifications, it was not possible to apply models 1a-c and 2a-b to the replication dataset. We instead sought to replicate our initial findings by performing a second discovery analysis, generating a new model (with leave-site-out cross-validation) predicting transition based on clinical and proteomic data in the replication dataset.

Model 4: predicting PEs using proteomic data

We developed a model predicting PEs at age 18 in ALSPAC based on proteomic data at age 12. Repeated nested cross-validation with 5 inner and 5 outer folds was used.

3.2.5 Supplementary analyses

Several supplementary analyses were performed (see Supplementary Methods). These included: development of a model predicting transition in EU-GEI based on ELISA data (model S1); development of a model predicting functional outcome in EU-GEI (GAF disability subscale score Ò60 vs. and proteomic data (model S2); investigation of potential EU-GEI site effects; and development of multivariate-corrected versions of SVM models whereby the variance associated with multiple covariates was extracted using principal components analysis.

3.3 Results

3.3.1 Study 1: CHR sample

Sample characteristics

Of 344 participants at CHR who were recruited, 152 (44.2%) attended face-to-face interviews at 12 months and 105 (30.5%) at 24 months. Baseline characteristics of participants who did or did not attend at least one follow-up interview are compared in Supplementary Table 3.3. Following FDR correction, attenders had on average one more year of education and lower mean total SANS global score than non-attenders, but were otherwise comparable.

The subsample for the initial experiment comprised 133 (49 CHR-T and 84 CHR-NT) participants with baseline plasma available (Figure 3.1). The mean (SD) age of the participants was 22.6 (4.5) years; 68 participants (51.1%) were male. Following FDR

correction, participants included in the subsample had higher mean SANS total global, SANS total composite and BPRS total scores than non-included participants, but were otherwise comparable on baseline characteristics (Supplementary Table 3.4).

Subsample characteristics are presented in Table 3.1. Following FDR correction, there were no statistically significant group differences for CHR-T vs. CHR-NT on baseline characteristics. The median duration from baseline to transition was 219 days (interquartile range 424 days). The CHR-T participants had lower mean functional outcome scores at 2 years compared with CHR-NT participants (mean GAF symptoms score at 2 years, 42.3 in CHR-T vs 62.2 in CHR-NT; FDR-corrected $p_i < 0.007$; mean GAF disability vs 64.5 in CHR-NT; 2 year FDR-corrected $p_i < 0.007$).

Table 3.1: Sample characteristics for CHR-T and CHR-NT groups in the initial experiment

	Missing data, n (%)	CHR-T N=49	CHR-NT N=84	t / ^{2} g	*p*	Corrected *p* (FDR 5%)
Baseline age in years, mean (SD)	0	22.2 (5.0)	22.9 (4.2)	-0.824	0.412	0.783
Sex, n (%)	0	26 male (53.1%) 23 female (46.9%)	42 male (50.0%) 42 female (50.0%)	0.116	0.733	0.908
Baseline BMI in kg/m², mean (SD)	20 (15.0%)	24.5 (4.5)	24.4 (6.1)	0.116	0.908	0.908
Baseline years in education, mean (SD)	14 (10.5%)	14.1 (3.4)	14.4 (3.0)	-0.625	0.533	0.791
Ethnicity, n (%)	0	33 white (67.3%) 8 black (16.3%) 8 other (16.3%)	58 white (69.0%) 7 black (8.3%) 19 other (22.6%)	2.370	0.306	0.646
Ever used cannabis, n (%)	3 (2.3%)	36 yes (73.5%) 11 no (22.4%) 2 not known (4.1%)	65 yes (77.4%) 18 no (21.4%) 1 not known (1.2%)	0.051	0.821	0.908
Baseline cannabis use, n (%)	29 (21.8%)	15 yes (30.6%) 22 no (44.9%) 12 not known (24.5%)	26 yes (31.0%) 41 no (48.8%) 17 not known (20.2%)	0.030	0.862	0.908
Baseline tobacco use, n (%)	14 (10.5%)	21 yes (42.9%) 21 no (42.9%) 7 not known (14.2%)	43 yes (51.2%) 34 no (40.5%) 7 not known (8.3%)	0.373	0.541	0.791
Baseline alcohol use, n (%)	3 (2.3%)	35 yes (71.4%) 13 no (26.5%) 1 not known (2.0%)	58 yes (69.0%) 24 no (28.6%) 2 not known (2.4%)	0.071	0.790	0.908

Baseline medication use, n (%)	31 (23.3%)	19 yes (38.8%) Antidepressant 13 Antipsychotic 9 Hypnotic 2 Other 3 20 no (40.8%) 10 not known (20.4%)	32 yes (38.1%) Antidepressant 24 Antipsychotic 6 Hypnotic 6 Other 13 31 no (36.9%) 21 not known (25.0%)	0.042	0.839	0.908
Baseline GAF symptoms score, mean (SD)	12 (9.0%)	52.4 (10.3)	56.0 (10.0)	-1.906	0.059	0.190
Baseline GAF disability score, mean (SD)	5 (3.8%)	52.3 (12.4)	54.8 (11.3)	-1.148	0.253	0.601
Baseline SANS total composite score, mean (SD)	19 (14.3%)	20.9 (14.0)	16.2 (11.6)	1.903	0.060	0.190
Baseline SANS total global score, mean (SD)	11 (8.3%)	6.6 (4.1)	5.8 (3.7)	1.158	0.249	0.601
Baseline BPRS total score, mean (SD)	10 (7.5%)	49.1 (11.5)	44.2 (10.2)	2.452	0.016	0.076
Baseline MADRS total score, mean (SD)	7 (5.3%)	20.3 (10.4)	19.2 (9.2)	0.657	0.512	0.791
2-year GAF symptoms score, mean (SD) [a]	62 (46.7%)	42.3 (13.2)	62.2 (10.3)	-7.125	<0.001	<0.007
2-year GAF disability score, mean (SD) [b]	54 (40.6%)	44.7 (9.1)	64.5 (12.8)	-8.024	<0.001	<0.007

2-year GAF disability score, dichotomous outcome [c]	54 (40.6%)	29 poor functioning (59.2%) 1 good functioning (2.0%) 19 not known (38.8%)	18 poor functioning (21.4%) 31 good functioning (36.9%) 35 not known (41.7%)	27.734	<0.001	<0.007

[a] Data available for 71 of 133 participants (CHR-NT n=44, CHR-T n=27)

[b] Data available for 79 of 133 participants (CHR-NT n=49, CHR-T n=30)

[c] Poor functioning: GAF disability score Ò60; goo

Tobacco use was defined as daily use for at least 1 month over the previous 12 months. Alcohol use was defined as at least 12 or more alcoholic beverages over the previous 12 months.

Missing data excluded in hypothesis tests.

EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; CHR-T: clinical high risk, transitioned to psychosis; CHR-NT: clinical high risk, did not transition to psychosis; FDR: false discovery rate; BMI: body mass index; GAF: General Assessment of Functioning; SANS: Scale for the Assessment of Negative Symptoms; BPRS: Brief Psychiatric Rating Scale; MADRS: Montgomery Asberg Depression Rating Scale.

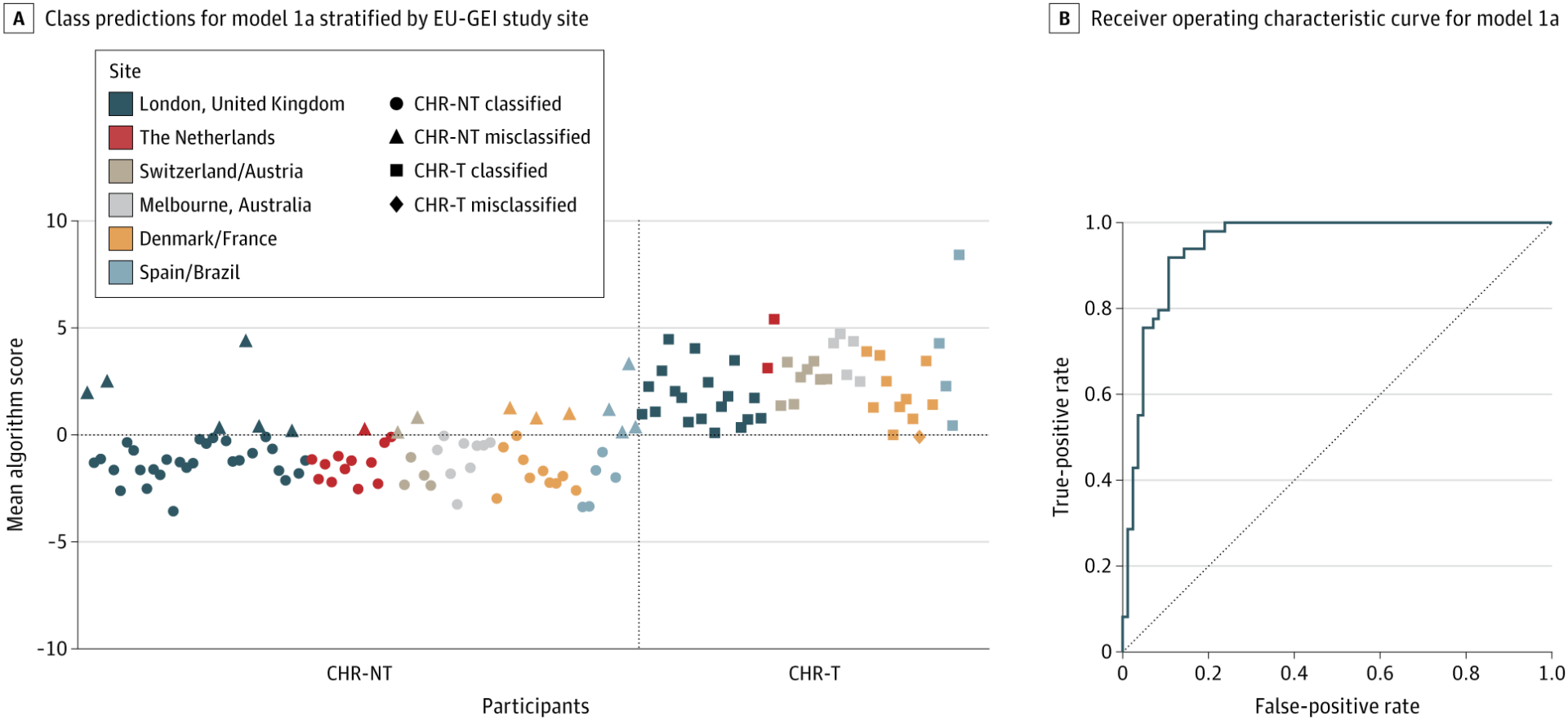
Differential expression

Of 345 proteins identified, 166 were quantified in more than 80% of plasma samples. There was nominally statistically significant ($p<0.05$) differential expression of 56 proteins in CHR-T vs. CHR-NT, of which 35 remained statistically significant following FDR correction (Supplementary Table 3.5, Supplementary Table 3.6). Supplementary Figure 3.2 shows a functional association network[27] for these proteins and Supplementary Table 3.7 details protein-protein interactions. On functional enrichment analysis, the topmost implicated pathway was the complement and coagulation cascade (Supplementary Table 3.8).

Model 1a: predicting transition using clinical and proteomic data

An SVM model predicted transition status based on clinical and proteomic features (model 1a) with excellent performance (AUC 0.95, $p<0.001$). Figure 3.2a shows mean algorithm scores and predicted outcomes, stratified by site. The receiver-operating characteristic curve is shown in Figure 3.2b. Performance metrics are presented in Table 3.2. Table 3.3 lists the 10% highest-weighted features according to mean feature weight. For example, the five highest-ranked predictive features were alpha-2-macroglobulin (A2M) (mean weight, ī0.3 constant mu (IGHM) (mean weight, ī0.256), C4 binding protein alpha chain (C4BPA) (mean weight, ī0.161), complement component 8 alpha chain (C8A) (mean weight, 0.158), and phospholipid transfer protein (PLTP) (mean weight, ī0.146).

Figure 3.2. Predicting transition to psychosis using clinical and proteomic data (model 1a)



A: The algorithm score is a decision score used to determine the predicted outcome class. Herein, a score greater than 0 is assigned as CHR-T, and a score less than 0 is assigned as CHR-NT. The dashed lines divide the graph into quadrants according to predicted vs actual outcome (ie, top right is true positive, bottom left is true negative, top left is false positive, and bottom right is false negative).
B: The dashed line is the line of no discrimination (area under the receiver operating characteristic curve, 0.5). CHR-NT indicates participants at clinical high risk who did not transition to psychosis; CHR-T, participants at clinical high risk who transitioned to first-episode psychosis; and EU-GEI, European Network of National Schizophrenia Networks Studying Gene-Environment Interactions.

Table 3.2: Performance metrics for unadjusted support vector machine models

Model description	True positives, n (%)	False negatives, n (%)	True negatives, n (%)	False positives, n (%)	Sens-itivity, %	Spec-ificity, %	Balanced accuracy, %	AUC (95% confidence interval)	Positive predictive value, %	Negative predictive value, %	LR+	LR-
Model 1a: clinical and proteomic *Dataset:* EU-GEI initial experiment, all sites *Features*: 69 clinical and 166 proteomic *Target*: transition status *N:* 49 transition, 84 non-transition	48 (98.0%)	1 (2.0%)	68 (81.0%)	16 (19.0%)	98.0	81.0	89.5	0.95 (0.91 ï 0.99)	75.0	98.6	5.1	<0.1
Model 1b: clinical *Dataset:* EU-GEI initial experiment, all sites *Features*: 69 clinical *Target*: transition status *N:* 49 transition, 84 non-transition	23 (46.9%)	26 (53.1%)	45 (53.6%)	39 (46.4%)	46.9	53.6	50.3	0.48 (0.38 ï 0.58)	37.1	63.4	1.0	1.0
Model 1c: proteomic *Dataset:* EU-GEI initial experiment, all sites *Features*: 166 proteomic *Target*: transition status *N:* 49 transition, 84 non-transition	49 (100.0%)	0 (0.0%)	71 (84.5%)	13 (15.5%)	100.0	84.5	92.3	0.96 (0.92 ï 1.00)	79.0	100.0	6.5	<0.1
Model 2a: proteomic (non-London) *Dataset:* EU-GEI initial experiment, all sites except London *Features*: 166 proteomic *Target*: transition status *N:* 30 transition, 50 non-transition	28 (93.3%)	2 (6.7%)	40 (80.0%)	10 (20.0%)	93.3	80.0	86.7	0.94 (0.88 ï 1.00)	73.7	95.2	4.7	0.1

Model 2b: top 10, training data *Dataset:* EU-GEI initial experiment, all sites except London *Features*: 10 proteomic *Target*: transition status *N:* 30 transition, 50 non-transition	30 (100.0%)	0 (0%)	41 (82.0%)	9 (18.0%)	100.0	82.0	91.0	0.99 (0.96 ï 1.00)	76.9	100.0	5.6	<0.1
Model 2b: top 10, test data *Dataset:* EU-GEI initial experiment, London site *Features*: 10 proteomic *Target*: transition status *N:* 19 transition, 34 non-transition	18 (94.7%)	1 (5.3%)	30 (88.2%)	4 (11.8%)	94.7	88.2	91.5	0.92 (0.83 ï 1.00)	81.8	96.8	8.1	0.1
Model 3: replication *Dataset:* EU-GEI replication experiment, all sites *Features*: 69 clinical and 119 proteomic *Target*: transition status *N:* 49 transition, 86 non-transition	48 (98.0%)	1 (2.0%)	77 (89.5%)	9 (10.5%)	98.0	89.5	93.7	0.98 (0.95 ï 1.00)	84.2	98.7	9.4	<0.1
Model 4: ALSPAC PEs *Dataset:* ALSPAC *Features*: 265 proteomic *Target*: PEs age 18 *N:* 55 PEs, 66 no PE	40 (72.7%)	15 (27.3%)	47 (71.2%)	19 (28.8%)	72.7	71.2	72.0	0.74 (0.65 ï 0.83)	67.8	75.8	2.5	0.4
Model S1: ELISA *Dataset:* EU-GEI initial experiment, all sites *Features:* 9 ELISA *Target*: transition status *N:* 44 transition, 82 non-transition	33 (75.0%)	11 (25.0%)	51 (62.2%)	31 (37.8%)	75.0	62.2	68.6	0.76 (0.67 ï 0.85)	51.6	82.3	2.0	0.4

Model S2: functional outcome *Dataset:* EU-GEI initial experiment, all sites *Features*: 69 clinical and 166 proteomic *Target*: functional outcome *N:* 47 poor functioning (GAFÒ60); 32 go (GAD>60)	27 (57.4%)	20 (42.6%)	22 (68.8%)	10 (31.2%)	57.4	68.8	63.1	0.74 (0.63 ï 0.85)	73.0	52.4	1.8	0.6

AUC: area under the receiver-operating characteristic curve; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; ALSPAC: Avon Longitudinal Study of Parents and Children; PEs: psychotic experiences; LR+: positive likelihood ratio; LR- negative likelihood ratio.

Models 1a-d, 2 and 3 are adjusted for age, sex, body mass index and years in education and model 4 is additionally adjusted for ethnicity and tobacco use. Model 5 is adjusted for sex, maternal social class at birth and body mass index at age 12.

Table 3.3: Ten percent highest-weighted features for model 1a, model 3 and model 4 (ranked according to mean feature weight for models selected in cross-validation inner loop)

Model 1a: EU-GEI clinical and proteomic data, initial experiment, all sites		**Model 3: EU-GEI clinical and proteomic data, replication experiment, all sites**		**Model 4: ALSPAC proteomic data**	
Feature	**Mean weight**	**Feature**	**Mean weight**	**Feature**	**Mean weight**
P01023 Alpha-2-macroglobulin	-0.330	P01023 Alpha-2-macroglobulin	-0.286	P04003 C4b-binding protein alpha chain	-0.227
P01871 Immunoglobulin heavy constant mu	-0.256	P22792 Carboxypeptidase N subunit 2	0.210	P27169 Serum paraoxonase/arylesterase 1	-0.180
P04003 C4b-binding protein alpha chain	-0.161	P01871 Immunoglobulin heavy constant mu	-0.193	Q03591 Complement factor H-related protein 1	-0.152
P07357 Complement component 8 alpha chain	0.158	P09871 Complement C1s subcomponent	-0.181	P07225 Vitamin K-dependent protein S	-0.145
P55058 Phospholipid transfer protein	-0.146	P01011 Alpha-1-antichymotrypsin	0.168	P61626 Lysozyme C	-0.142
O75636 Ficolin-3	-0.145	P00747 Plasminogen	0.163	P55103 Inhibin beta C chain	0.139
P02774 Vitamin D binding protein	0.135	P08571 Monocyte differentiation antigen CD14	0.161	Q08380 Galectin-3-binding protein	0.132
P07225 Vitamin K-dependent protein S	-0.132	P10909 Clusterin	0.158	P24593 Insulin-like growth factor-binding protein 5	0.122
P43320 Beta-crystallin B2	0.132	Q16610 Extracellular matrix protein 1	0.157	P00746 Complement factor D	0.120
P02766 Transthyretin	-0.130	G3XAM2 Complement factor I	0.140	P01019 Angiotensinogen	-0.118

P23142 Fibuln-1	0.125	P04003 C4b binding protein alpha chain	-0.140	P01871 Immunoglobulin heavy constant mu	-0.116
P10909 Clusterin	0.121	P13671 Complement component 6	0.132	O75636 Ficolin-3	0.115
P05155 Plasma protease C1 inhibitor	-0.114	P25311 Zinc alpha-2-glycoprotein	-0.131	Q9H4A9 Dipeptidase 2	-0.115
Sex	-0.111	P07359 Platelet glycoprotein Ib alpha chain	0.126	P01023 Alpha-2-macroglobulin	-0.113
P00747 Plasminogen	0.111	P01031 Complement C5	0.125	P04275 von Willebrand factor	-0.111
P13671 Complement component 6	0.111	O75882 Attractin	0.123	Q9NQ79 Cartilage acidic protein 1	0.107
P02747 Complement C1q subcomponent subunit C	0.109	P0DOY3 Immunoglobulin lambda constant 3	-0.120	P24592 Insulin-like growth factor-binding protein 6	0.106
P02753 Retinol-binding protein 4	0.109	P15169 Carboxypeptidase N catalytic chain (CPN)	0.115	P09871 Complement C1s subcomponent	-0.105
Q76LX8 A disintegrin and metalloproteinase with thrombospondin motifs 13	-0.108			P10909 Clusterin	-0.105
P08697 Alpha-2-antiplasmin	-0.106			O95497 Pantetheinase	0.105
P19827 Inter-alpha-trypsin inhibitor heavy chain H1	0.105			P02654 Apolipoprotein C-I	-0.099
MADRS: concentration difficulties	-0.104			P02679 Fibrinogen gamma chain	-0.099
P02489 Alpha-crystallin A chain	0.101			P07358 Complement component C8 beta chain	0.097

				Q5T7F0 Neuropilin	-0.097
				P04040 Catalase	0.094
				P43251 Biotinidase	0.094

Proteins are presented with their Uniprot accession number and corresponding protein name.

EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; ALSPAC: Avon Longitudinal Study of Parents and Children; BPRS: Brief Psychiatric Rating Scale; MADRS: Montgomery-Asberg Depression Rating Scale; SANS: Scale for the Assessment of Negative Symptoms.

Model 1b: clinical data

The clinical model (model 1b) demonstrated poor predictive performance (AUC 0.48, p=0.628; Table 3.2, Supplementary Figure 3.3). For example, sensitivity was 46.9%, specificity was 53.6%, PPV was 37.1%, and NPV was 63.4%.

Model 1c: proteomic data

The proteomic model (model 1c) demonstrated excellent predictive performance (AUC 0.96, p<0.001; Table 3.2, Supplementary Figure 3.4). For example, sensitivity was 100%, specificity was 84.5%, PPV was 79.0%, and NPV was 100%.

Models 2a-b: parsimonious model

The AUC for the model based on proteomic data from all sites except London (Model 2a) was 0.94, p<0.001 (Table 3.2, Supplementary Figure 3.5). The 10 highest-weighted features were: alpha-2-macroglobulin (A2M), immunoglobulin heavy constant mu (IGHM), C4b-binding protein alpha chain (C4BPA), vitamin K-dependent protein S, fibulin-1, transthyretin, N-acetylmuramoyl-L-alanine amidase, vitamin D-binding protein, clusterin and complement component 6 (C6).

A reduced model based solely on these 10 features was developed using data from all sites except London (model 2b) with an AUC of 0.99 (p<0.001), sensitivity of 100%, specificity of 82.0%, PPV of 76.9%, and NPV of 100% (Table 3.2, Supplementary Figure 3.6). This model predicted transition status in the held-out London data with an AUC of 0.92, sensitivity of 94.7%, specificity of 88.2%, PPV of 81.8%, and NPV of 96.8% (Table 3.2).

ELISA validation

Following FDR correction, two proteins assessed by ELISA showed statistically significant mean differences between CHR-T and CHR-NT: A2M and complement component 1r (C1r) (Supplementary Table 3.9, Supplementary Table 3.10). The A2M mean in CHR-T was 1173.1 εg/mL vs 11-T (FDR-corrected p = 0.02), CHR and the C
CHR-T was 65 008.9 εg/mL vs NT (FDR-corrected p = 0.04) in CI

Replication: sample characteristics and differential expression

Replication subsample characteristics are presented in Supplementary Table 3.11 and Supplementary Table 3.12. Of 485 proteins identified, 119 were quantified in more than

80% of plasma samples. There was nominally statistically significant (p<0.05) differential expression of 82 proteins, of which 78 remained significant following FDR correction (Supplementary Table 3.13).

Model 3: replication

This model demonstrated excellent performance for prediction of transition in the replication dataset (AUC 0.98, p<0.001; sensitivity 98.0%; specificity 89.5%; PPV 84.2%, NPV 98.7%; Table 3.2, Supplementary Figure 3.7). The highest-weighted 10% of features are presented in Table 3.3. Proteins among the highest-weighted 10% of features in both model 1a and model 3 (and weighted in similar directions) included: A2M, IGHM, C4BPA, plasminogen and complement component 6 (C6).

3.3.2 Study 2: general population sample

Sample characteristics

The initial subsample was comprised of 65 case and 67 control participants. Eleven plasma samples were excluded due to poor protein identification profiles, resulting in 55 cases and 66 controls from 121 participants (61 [50.4%] male). Case samples were more likely to be from female participants. There was no evidence for differences in BMI, ethnicity or maternal social class (Supplementary Table 3.14).

Differential expression

Of 506 proteins identified, 265 were quantified in more than 80% of samples. There was nominally statistically significant (p<0.05) differential expression of 40 proteins at age 12 (Supplementary Table 3.15) of which 5 remained significant following FDR correction (ratio of mean in cases vs. controls): C4BPA (0.77), serum paraoxonase/arylesterase 1 (0.80), IGHM (0.78), inhibin beta chain (1.31) and clusterin (0.92).

Model 4: predicting PEs using proteomic data

An SVM model using 265 proteomic features from age 12 plasma samples predicted PEs at age 18 with an AUC of 0.74 (p<0.001), sensitivity of 72.7%, specificity of 71.2%, PPV of 67.8% and NPV of 75.8% (Table 3.2, Supplementary Figure 3.8). The 10% highest-weighted features are listed in Table 3.3. For example, the 5 highest-ranked predictive features were C4BPA (mean weight, ī 0.227), ser weight, ī 0.180), ïcomplement protein 1 (mean weigh dependent protein S (mean weight, ī 0.145), and

3.3.3 Supplementary analyses

Model S1 used ELISA data to predict transition status in EU-GEI with an AUC of 0.76 (p<0.001). The results are summarised in Table 3.2 and Supplementary Figure 3.9.

Model S2 used clinical and proteomic data to predict poor (GAF disability subscale score Ò60) vs. good (>60) functioning in EU-GEI with an AUC of 0.74 at 2 years (p=0.003). The results are summarised in Table 3.2 and Supplementary Figure 3.10). The 10% highest-weighted features in this model are listed in Supplementary Table 3.16.

There was evidence of differences for the clinical data between the London and Netherlands sites compared with others (Supplementary Table 3.17, Supplementary Figure 3.11, Supplementary Figure 3.23), likely because of group differences in age, years in education and BPRS score (see Supplementary Methods and Supplementary Figures 13ï 22). There was no strong evidence of systematic site effects for the proteomic data (Supplementary Table 3.18, Supplementary Figure 3.12, Supplementary Figure 3.24).

Performance metrics of multivariate-corrected SVM models are presented in Supplementary Table 3.19. There were generally slight reductions in AUCs of the corrected models compared to their uncorrected counterparts (median change in AUC 0.04, range 0.01 ï 0.10) although in all cases the 95% confidence intervals overlapped.

3.4 Discussion

We report evidence of differential baseline plasma protein expression in CHR individuals who develop psychosis compared to those who do not. Machine learning algorithms incorporating clinical and proteomic data were used to predict transition outcome (AUC 0.95). Proteomic features were of greater predictive value than clinical features. A parsimonious model based on 10 highly predictive proteins showed excellent performance in training (AUC 0.99) and test data (AUC 0.92). Furthermore, we developed a predictive model using age 12 proteomic data for age 18 PEs in a general population sample (AUC 0.74).

Although only 16 to 35% of CHR individuals transition to FEP,[3] the CHR state remains a strong risk indicator.[28] Clinical data have previously shown value for prediction of transition[29-34] and the poor performance of the clinical features in our study does not infer that clinical data in general are of little prognostic use. Previous studies have attempted to

augment accuracy using neuroimaging[35-38] and neurocognitive[39] data, but blood-based tests have the advantage of greater accessibility. Perkins et al derived a panel of 15 proteins using immunoassays that distinguished between CHR-T and CHR-NT with AUC 0.88.[40] Chan et al used 22 blood-based biomarkers to predict schizophrenia onset with AUC 0.82, increasing to 0.90 with incorporation of CAARMS positive symptoms subscale.[41] Our parsimonious model used data for 10 proteins which, with further validation, may contribute to individualised prognosis and treatment stratification strategies.[42]

Supplementary Table 3.20 summarises our findings of differential expression in CHR-T vs. CHR-NT and the predicted functional implications (modelled in Supplementary Figure 3.25). We found particularly strong evidence for dysregulation of the complement and coagulation cascades, previously implicated in schizophrenia.[43-47] Similar processes have previously been implicated in proteomic studies of the development of PEs in the general population.[11,19] Changes in the current CHR study that overlapped with these previous PE studies include increases in plasminogen, C1r, clusterin and complement factor H, and decreases in A2M and IGHM. The primary causes of these changes remain unknown but are consistent with evidence of raised inflammatory tone preceding psychosis and other mental disorders[40,41,48-52] and schizophrenia risk associated with genetic variation of complement C4.[53]

Several complement proteins emerged as important predictors of transition including C4BPA, complement C1q subcomponent subunit C and C1r of the antibody-antigen complex mediated pathway, key regulatory protease complement factor I, ficolin-3 and terminal pathway components C6 and C8A. These arise from common pathways or functionally interact with coagulation proteins plasminogen and vitamin K-dependent protein S, supporting hypotheses of coagulation activation in psychosis.[54] In both the initial and replication experiments, the strongest predictor of transition was A2M (decreased in CHR-T vs. CHR-NT), a protease inhibitor with diverse functions including inhibition of pro-inflammatory cytokines such as interleukin-1 β[55] (consistently raised in FEP[56]). A2M is a key coagulation inhibitor[57] and thus links functionally to our observations of elevated plasminogen in CHR-T. This is intriguing in light of evidence that blood-derived plasminogen drives brain inflammation[58] and complement activation.[59] In models of multiple sclerosis, blood-brain barrier disruption facilitates transfer of fibrinogen into the brain, where it is deposited as fibrin, causing local inflammation.[60] Given evidence for blood-brain barrier

disruption in psychosis,[61] the effects of fibrin provide possible aetiopathogenic mechanisms and novel therapeutic avenues,[62] but this will require further investigation.

We validated differential expression of A2M and C1r using ELISA. The ELISA-based model (model S1) demonstrated fair, though reduced, predictive accuracy. This may reflect reduced sensitivity of ELISA and the inability to accurately quantify specific protein isoforms. Compellingly, several proteins in the highest-weighted 10% of features for transition in Study 1 were similarly highly weighted for PEs in Study 2, including C4BPA, vitamin K-dependent protein S, A2M and IGHM (Supplementary Table 3.21 summarises the directionality of effect of the 10% highest predictors in model 1a, model 3 and model 4). This could tentatively suggest a degree of similarity in proteomic changes between non-clinical youth who develop PEs and help-seeking CHR individuals who develop psychosis, but will require confirmation.

Outside of psychosis outcomes, several proteomic features contributed to prediction of functional outcome (model S2). A2M, IGHM, PLTP and clusterin were among the 10% highest-weighted predictors. The results of the present study are also in keeping with studies in bipolar disorder and depression reporting decreased A2M, IgM and C4BPA.[63] At least some of these proteomic changes may be common to multiple clinical phenotypes, including neurodegenerative disorders such as Alzheimer's disease.[64] Rather than considering such changes as biomarkers of individual disorders, phenotypic manifestations may be clinical manifestations of overlapping neuro-immune abnormalities that have their origin in combined genetic[53,65] and environmental influences.[66-69]

Limitations

This study has some limitations. Firstly, these models require validation in independent cohorts to assess generalisability and real-world applicability. Secondly, differences in protein identifications precluded application of models between studies. However, there are valid reasons not to do so, including differences in outcome (psychotic disorder vs. psychotic experiences) and age (post-pubertal vs. peri-pubertal). Thirdly, data on duration of follow-up and reasons for dropout were not systematically collected in EU-GEI and we were unable to fully assess the potential influence of these factors. Fourthly, the replication experiment was partial since only 2 CHR-T cases were different from the initial experiment. While our findings were generally replicated, no statement can be made regarding

generalisability of model sensitivity. Fifthly, participants were non-fasting and there were no restrictions on time of sample collection. Finally, factors such as childhood adversity may contribute to the proteomic changes we observe,[9,68] but this will require further study.

Conclusions

We developed models incorporating proteomic data for prediction of transition to psychotic disorder in CHR. In a general population sample, several of the same proteins contributed to prediction of PEs. Further studies are required to validate these findings, evaluate their causes and elucidate tractable targets for prediction and prevention of psychosis.

Article information

Author Contributions: Drs Cotter and McGuire are coï senior authors. Drs Mongan and Cotter had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Healy, Amminger, Krebs, Bressan, Borgwardt, Ruhrmann, de Haan, Pollak, Rutten, Cannon, Cagney, Cotter, McGuire.

Acquisition, analysis, or interpretation of data: Mongan, Föcking, Healy, Susai, Heurich, Wynne, Nelson, McGorry, Amminger, Nordentoft, Krebs, Riecher-Rössler, Bressan, Barrantes-Vidal, Borgwardt, Ruhrmann, Sachs, Pantelis, van der Gaag, de Haan, Valmaggia, Pollak, Kempton, Rutten, Whelan, Zammit, Cagney, Cotter, McGuire.

Drafting of the manuscript: Mongan, Föcking, Healy, Wynne, Cannon, Cotter, McGuire.

Critical revision of the manuscript for important intellectual content: Mongan, Föcking, Healy, Susai, Heurich, Nelson, McGorry, Amminger, Nordentoft, Krebs, Riecher-Rössler, Bressan, Barrantes-Vidal, Borgwardt, Ruhrmann, Sachs, Pantelis, van der Gaag, de Haan, Valmaggia, Pollak, Kempton, Rutten, Whelan, Zammit, Cagney, Cotter, McGuire.

Statistical analysis: Mongan, Healy, Whelan, Cagney, Cotter.

Obtained funding: Nelson, McGorry, Amminger, Nordentoft, Krebs, Riecher-Rössler, Bressan, Borgwardt, Pantelis, Valmaggia, Rutten, Cagney, Cotter, McGuire.

Administrative, technical, or material support: Susai, Wynne, Riecher-Rössler, Bressan, Barrantes-Vidal, Borgwardt, Pantelis, van der Gaag, Valmaggia, Pollak, Kempton, Cagney, Cotter, McGuire.

Supervision: Amminger, Nordentoft, Bressan, Borgwardt, Ruhrmann, Sachs, de Haan, Cannon, Cagney, Cotter, McGuire.

Conflict of Interest Disclosures: A United Kingdom (UK) patent application has been filed in relation to the development of a prognostic test derived from this work (UK patent application 1919155.0). Dr Mongan reported receiving grants from the Wellcome Trust and Health Research Board Ireland and having UK patent application 1919155.0 pending. Mr Healy reported receiving grants from the European Research Council. Dr Krebs reported receiving grants from the French Ministry Programme Hospitalier de Recherche Clinique AOM07-118 and Eisai, receiving grants and personal fees from Otsuka-Lundbeck and Janssen, and having a patent pending. Dr Borgwardt reported receiving grants from the European Communityôs Seventh Framework Program

F2-2010-241909 (Project EU-GEI). Dr Ruhrmann reported receiving grants from the European Commission and receiving nonfinancial support from Boehringer Ingelheim. Dr Sachs reported receiving honoraria for consulting and lectures on the topic of schizophrenia. Dr Pantelis reported receiving grants from the Australian National Health and Medical Research Council (NHMRC) and The Lundbeck Foundation and receiving personal fees from Lundbeck Australia Pty Ltd. Dr Kempton reported receiving grants from the European Commission and the Medical Research Council. Dr Cagney reported receiving grants from Health Research Board Ireland and having a patent for a biomarker panel pending. Dr Cotter reported receiving grants from Health Research Board Ireland and having UK patent 1919155.0 pending. No other disclosures were reported.

Funding/Support: EU-GEI was funded by a Framework 7 Grant (HEALTH-F2-2010-241909) for the European Network of National Schizophrenia Networks Studying Gene-Environment Interactions (EU-GEI) study and by Health Research Board Ireland through a Clinician Scientist Award to Dr Cotter. Additional support was provided by a Medical Research Council Fellowship to Dr Kempton (grant MR/J008915/1) and by the Ministerio de Ciencia, Innovación e Universidades (grant PSI2017-87512-C2-1-R) and Generalitat de Catalunya (grant 2017SGR1612 and Catalan Institution for Research and Advanced Studies [ICREA] Academia award) to Dr Barrantes-Vidal. The UK Medical Research Council and the Wellcome Trust (grant 102215/2/13/2) and the University of Bristol provide core support for the Avon Longitudinal Study of Parents and Children (ALSPAC). A comprehensive list of grant funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). The outcomes data collected in the ALSPAC study that were used in the present study were specifically funded by the Medical Research Council (grant G0701503/85179). Dr Zammit is supported by the Bristol National Institute for Health Research Biomedical Research Centre. Dr Mongan is a fellow of the Irish Clinical Academic Training (ICAT) Programme, which is supported by the Wellcome Trust and Health Research Board Ireland (grant 203930/B/16/Z), the Health Service Executive National Doctors Training and Planning, and the Health and Social Care Research and Development Division, Northern Ireland.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Group Information: EU-GEI High Risk Study Group members are Philip McGuire, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kingôs College London; Lucia Valmaggia, Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, Kingôs Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kingôs College London; Thomas A. Pollak, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kingôs Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kingôs College London; Tania Tognin, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kingôs Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kingôs College London; Academic Medical Centre (AMC), Academic Psychiatric Centre, Department Early Psychosis; Mark van der Gaag, VU University, Faculty of Behavioural and Movement Sciences, Department of Clinical Psychology and EMGO+ Institute for Health and Care Research, and Parnassia Psychiatric Institute, Department of Psychosis Research; Eva Velthorst, AMC, Academic Psychiatric Centre, Department Early Psychosis, and Icahn School of Medicine at Mount Sinai, Department of Psychiatry; Tamar C. Kraan, AMC, Academic Psychiatric Centre, Department Early Psychosis; Daniella S. van Dam, AMC, Academic Psychiatric Centre, Department Early Psychosis; Nadine Burger, Parnassia Psychiatric Institute, Department of Psychosis Research; Barnaby Nelson, Centre for Youth Mental Health, University of Melbourne; Patrick D. McGorry, Centre for Youth Mental Health, University of Melbourne; G. Paul Amminger, Centre for Youth Mental Health, University of Melbourne; Christos Pantelis, Center for Neuropsychiatric Schizophrenia Research (CNSR) and Center for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS), University of Copenhagen, and University of Copenhagen, Faculty of Health and Medical Sciences, Department of Clinical Medicine; Athena Politis, Centre for Youth Mental Health, University of Melbourne; Joanne Goodall, Centre for Youth Mental Health, University of Melbourne; Anita Riecher-Rössler, University of Basel Psychiatric Hospital; Stefan Borgwardt,

University of Basel Psychiatric Hospital; Charlotte Rapp, University of Basel Psychiatric Hospital; Sarah Ittig, University of Basel Psychiatric Hospital; Erich Studerus, University of Basel Psychiatric Hospital; Renata Smieskova, University of Basel Psychiatric Hospital; Rodrigo A. Bressan, LiNCï Lab Interdisciplinar Neurociências Clínicas, Depto Psiquiatria, Escola Paulista de Medicina, Universidade Federal de São Pauloï UNIFESP; Ary Gadelha, LiNC, Depto Psiquiatria, Escola Paulista de Medicina, UNIFESP; Elisa Brietzke, Depto Psiquiatria, Escola Paulista de Medicina, UNIFESP; Graccielle Asevedo, LiNC, Depto Psiquiatria, Escola Paulista de Medicina, UNIFESP; Elson Asevedo, LiNC, Depto Psiquiatria, Escola Paulista de Medicina, UNIFESP; Andre Zugman, LiNC, Depto Psiquiatria, Escola Paulista de Medicina, UNIFESP; Neus Barrantes-Vidal, Departament de Psicologia Clínica i de la Salut (Universitat Autònoma de Barcelona), Fundació Sanitària Sant Pere Claver, Spanish Mental Health Research Network (CIBERSAM); Tecelli Domínguez-Martínez, CONACYTï Dirección de Investigaciones Epidemiológicas y Psicosociales, Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz; Anna Racioppi, Departament de Psicologia Clínica i de la Salut (Universitat Autònoma de Barcelona); Thomas R. Kwapil, Department of Psychology, University of Illinois at Urbanaï Champaign; Manel Monsonet, Departament de Psicologia Clínica i de la Salut (Universitat Autònoma de Barcelona); Araceli Rosa, Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals (Universitat de Barcelona); Oussama Kebir, University Paris Descartes, Hôpital Sainte-Anne, CôJAAD, Université, INSERM U894, Institut de Psychiatrie (CNRS 3557); Claire Daban, University of Paris, GHU-Paris, Sainte-Anne, CôJAAD, Université, Department SHU; Julie Bourgin, University of Paris, GHU-Paris, Sainte-Anne, CôJAAD, Université, Department SHU; Boris Chaumette, University Paris Descartes, Hôpital Sainte-Anne, CôJAAD, Université, Hospital INSERM U894, Institut de Psychiatrie (CNRS 3557); Célia Mam-Lam-Fook, University of Paris, GHU-Paris, Sainte-Anne, CôJAAD, Université, Department SHU; Marie-Odile Krebs, University of Paris, GHU-Paris, Sainte-Anne, CôJAAD, Université, Hospital Department SHU; Dorte Nordholm, Mental Health Center Copenhagen and CINS, Mental Health Center Glostrup, Mental Health Services in the Capital Region of Copenhagen, University of Copenhagen; Lasse Randers, Mental Health Center Copenhagen and CINS, Mental Health Center Glostrup, Mental Health Services in the Capital Region of Copenhagen, University of Copenhagen; Kristine Krakauer, Mental Health Center Copenhagen and CINS, Mental Health Center Glostrup, Mental Health Services in the

Capital Region of Copenhagen, University of Copenhagen; Louise Birkedal Glenthøj, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kingôs College London, and CNSR and CINS, Univ CNSR and CINS, University of Copenhagen; Merete Nordentoft, Mental Health Center Copenhagen and CINS, Mental Health Center Glostrup, Mental Health Services in the Capital Region of Copenhagen, University of Copenhagen; Stephan Ruhrmann, Department of Psychiatry and Psychotherapy, University of Cologne; Dominika Gebhard, Department of Psychiatry and Psychotherapy, University of Cologne; Julia Arnhold, psyberlin, Berlin, Germany; Joachim Klosterkötter, Department of Psychiatry and Psychotherapy, University of Cologne; Gabriele Sachs, Department of Psychiatry and Psychotherapy, Medical University of Vienna; Iris Lasser, Department of Psychiatry and Psychotherapy, Medical University of Vienna; Bernadette Winklbaur, Department of Psychiatry and Psychotherapy, Medical University of Vienna; Philippe A. Delespaul, Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Centre, and Mondriaan Mental Health Trust; Bart P. F. Rutten, Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Centre; and Jim van Os, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kingôs College London, and Department of Psych University Medical Centre.

Additional Contributions: We thank Magda Hryniewiecka, MSc (Royal College of Surgeons in Ireland), for technical assistance and contributions in preparing samples for mass spectrometry. She was not compensated for her contributions. We also thank the Mass Spectrometry Core Facility at the Conway Institute, University College Dublin, for support in the development of our proteomics workflows. We are extremely grateful to all the participants, clinical teams, and research staff who contributed to the EU-GEI project. We are extremely grateful to all the families who took part in the ALSPAC, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses.

Chapter 3 References

1. Larsen TK, Melle I, Auestad B, et al. Early detection of psychosis: positive effects on 5-year outcome. *Psychological Medicine.* 2011;41(7):1461-1469.
2. Fusar-Poli P, Borgwardt S, Bechdolf A, et al. The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA Psychiatry.* 2013;70(1):107-120.
3. Fusar-Poli P, Bonoldi I, Yung AR, et al. Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. *Arch Gen Psych.* 2012;69(3):220-229.
4. Zammit S, Kounali D, Cannon M, et al. Psychotic experiences and psychotic disorders at age 18 in relation to psychotic experiences at age 12 in a longitudinal population-based cohort study. *Am J Psychiatry.* 2013;170(7):742-750.
5. Healy C, Brannigan R, Dooley N, et al. Childhood and adolescent psychotic experiences and risk of mental disorder: a systematic review and meta-analysis. *Psychological Medicine.* 2019;49(10):1589-1599.
6. Healy C, Campbell D, Coughlan H, Clarke M, Kelleher I, Cannon M. Childhood psychotic experiences are associated with poorer global functioning throughout adolescence and into early adulthood. *Acta Psychiatr Scand.* 2018;138(1):26-34.
7. McGorry P, Keshavan M, Goldstone S, et al. Biomarkers and clinical staging in psychiatry. *World Psychiatry.* 2014;13(3):211-223.
8. European Network of National Networks studying Gene-Environment Interactions in Schizophrenia, van Os J, Rutten BP, et al. Identifying gene-environment interactions in schizophrenia: contemporary challenges for integrated, large-scale investigations. *Schizophr Bull.* 2014;40(4):729-736.
9. Kraan TC, Velthorst E, Themmen M, et al. Child Maltreatment and Clinical Outcome in Individuals at Ultra-High Risk for Psychosis in the EU-GEI High Risk Study. *Schizophr Bull.* 2018;44(3):584-592.
10. Yung AR, Yuen HP, McGorry PD, et al. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. *The Australian and New Zealand Journal of Psychiatry.* 2005;39(11-12):964-971.
11. English JA, Lopez LM, O'Gorman A, et al. Blood-Based Protein Changes in Childhood Are Associated With Increased Risk for Later Psychotic Disorder:

Evidence From a Nested Case-Control Study of the ALSPAC Longitudinal Birth Cohort. *Schizophr Bull.* 2018;44(2):297-306.

12. Aas IH. Global Assessment of Functioning (GAF): properties and frontier of current knowledge. *Ann Gen Psychiatry.* 2010;9:20.
13. Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. *Am J Psychiatry.* 1992;149(9):1148-1156.
14. Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *Br J Psychiatry Suppl.* 1989(7):49-58.
15. Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. *Psychological Reports.* 1962;10(3):799-812.
16. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *The British journal of Psychiatry.* 1979;134:382-389.
17. English JA, Fan Y, Focking M, et al. Reduced protein synthesis in schizophrenia patient-derived olfactory cells. *Transl Psychiatry.* 2015;5:e663.
18. Focking M, Opstelten R, Prickaerts J, et al. Proteomic investigation of the hippocampus in prenatally stressed mice implicates changes in membrane trafficking, cytoskeletal, and metabolic function. *Dev Neurosci.* 2014;36(5):432-442.
19. Föcking M, Sabherwal S, Cates HM, et al. Complement pathway changes at age 12 are associated with psychotic experiences at age 18 in a longitudinal population-based study: evidence for a role of stress. *Molecular Psychiatry.* 2019.
20. Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s' - the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol.* 2013;42(1):111-127.
21. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol.* 2013;42(1):97-110.
22. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society.* 1995;Series B(57,1):289-300.
23. Cox J, Mann M. Quantitative, high-resolution proteomics for data-driven systems biology. *Annu Rev Biochem.* 2011;80:273-299.

24. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol.* 2008;26(12):1367-1372.
25. Lazar C. Impute LCMD. *R package version 20.* 2015.
26. Cearns M, Hahn T, Baune BT. Recommendations and future directions for supervised machine learning in psychiatry. *Translational Psychiatry.* 2019;9(1):271.
27. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015;43(Database issue):D447-452.
28. Radua J, Ramella-Cravaro V, Ioannidis JPA, et al. What causes psychosis? An umbrella review of risk and protective factors. *World Psychiatry.* 2018;17(1):49-66.
29. Studerus E, Ramyead A, Riecher-Rössler A. Prediction of transition to psychosis in patients with a clinical high risk for psychosis: a systematic review of methodology and reporting. *Psychological Medicine.* 2017;47(7):1163-1178.
30. Malda A, Boonstra N, Barf H, et al. Individualized Prediction of Transition to Psychosis in 1,676 Individuals at Clinical High Risk: Development and Validation of a Multivariable Prediction Model Based on Individual Patient Data Meta-Analysis. *Frontiers in Psychiatry.* 2019;10:345.
31. Mechelli A, Lin A, Wood S, et al. Using clinical information to make individualized prognostic predictions in people at ultra high risk for psychosis. *Schizophr Res.* 2017;184:32-38.
32. Schmidt A, Cappucciati M, Radua J, et al. Improving Prognostic Accuracy in Subjects at Clinical High Risk for Psychosis: Systematic Review of Predictive Models and Meta-analytical Sequential Testing Simulation. *Schizophrenia bulletin.* 2017;43(2):375-388.
33. Ruhrmann S, Schultze-Lutter F, Salokangas RK, et al. Prediction of psychosis in adolescents and young adults at high risk: results from the prospective European prediction of psychosis study. *Arch Gen Psychiatry.* 2010;67(3):241-251.
34. Cannon TD, Yu C, Addington J, et al. An Individualized Risk Calculator for Research in Prodromal Psychosis. *Am J Psychiatry.* 2016;173(10):980-988.
35. Koutsouleris N, Meisenzahl EM, Davatzikos C, et al. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Arch Gen Psychiatry.* 2009;66(7):700-712.

36. Koutsouleris N, Borgwardt S, Meisenzahl EM, Bottlender R, Möller H-J, Riecher-Rössler A. Disease Prediction in the At-Risk Mental State for Psychosis Using Neuroanatomical Biomarkers: Results From the FePsy Study. *Schizophrenia Bulletin.* 2011;38(6):1234-1246.
37. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, et al. Prediction Models of Functional Outcomes for Individuals in the Clinical High-Risk State for Psychosis or With Recent-Onset Depression: A Multimodal, Multisite Machine Learning Analysis. *JAMA Psychiatry.* 2018;75(11):1156-1172.
38. Das T, Borgwardt S, Hauke DJ, et al. Disorganized Gyrification Network Properties During the Transition to Psychosis. *JAMA Psychiatry.* 2018;75(6):613-622.
39. Koutsouleris N, Davatzikos C, Bottlender R, et al. Early Recognition and Disease Prediction in the At-Risk Mental States for Psychosis Using Neurocognitive Pattern Classification. *Schizophrenia Bull.* 2011;38(6):1200-1215.
40. Perkins DO, Jeffries CD, Addington J, 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):419-428.
41. Chan MK, Krebs MO, Cox D, et al. Development of a blood-based molecular biomarker test for identification of schizophrenia before disease onset. *Translational Psychiatry.* 2015;5:e601.
42. Ruhrmann S, Schultze-Lutter F, Schmidt SJ, Kaiser N, Klosterkotter J. Prediction and prevention of psychosis: current progress and future tasks. *European Archives of Psychiatry and Clinical Neuroscience.* 2014;264 Suppl 1:S9-16.
43. Woo JJ, Pouget JG, Zai CC, Kennedy JL. The complement system in schizophrenia: where are we now and what's next? *Molecular Psychiatry.* 2019.
44. Sabherwal S, English JA, Focking M, Cagney G, Cotter DR. Blood biomarker discovery in drug-free schizophrenia: the contribution of proteomics and multiplex immunoassays. *Expert Review of Proteomics.* 2016;13(12):1141-1155.
45. Li Y, Zhou K, Zhang Z, 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. *Molecular bioSystems.* 2012;8(10):2664-2671.
46. Kopczynska M, Zelek W, Touchard S, et al. Complement system biomarkers in first episode psychosis. *Schizophrenia Research.* 2017.

47. Levin Y, Wang L, Schwarz E, Koethe D, Leweke FM, Bahn S. Global proteomic profiling reveals altered proteomic signature in schizophrenia serum. *Molecular Psychiatry.* 2010;15(11):1088-1100.
48. Föcking M, Dicker P, Lopez LM, et al. Differential expression of the inflammation marker IL12p40 in the at-risk mental state for psychosis: a predictor of transition to psychotic disorder? *BMC Psychiatry.* 2016;16(1):326.
49. Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-analysis of cytokine alterations in schizophrenia: clinical status and antipsychotic effects. *Biol Psychiatry.* 2011;70(7):663-671.
50. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB. Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: a population-based longitudinal study. *JAMA Psychiatry.* 2014;71(10):1121-1128.
51. Laskaris L, Zalesky A, Weickert CS, et al. Investigation of peripheral complement factors across stages of psychosis. *Schizophrenia Research.* 2019;204:30-37.
52. Baumeister D, Russell A, Pariante CM, Mondelli V. Inflammatory biomarker profiles of mental disorders and their relation to clinical, social and lifestyle factors. *Soc Psychiatry Psychiatr Epidemiol.* 2014;49(6):841-849.
53. Sekar A, Bialas AR, de Rivera H, et al. Schizophrenia risk from complex variation of complement component 4. *Nature.* 2016;530(7589):177-183.
54. Hoirisch-Clapauch S, Amaral OB, Mezzasalma MA, Panizzutti R, Nardi AE. Dysfunction in the coagulation system and schizophrenia. *Translational Psychiatry.* 2016;6:e704.
55. Rehman AA, Ahsan H, Khan FH. alpha-2-Macroglobulin: a physiological guardian. *Journal of Cellular Physiology.* 2013;228(8):1665-1675.
56. Upthegrove R, Manzanares-Teson N, Barnes NM. Cytokine function in medication-naïve first episode psychosis: A systematic review and meta-analysis. *Schizophrenia Research.* 2014;155(1-3):101-108.
57. Borth W. Alpha 2-macroglobulin, a multifunctional binding protein with targeting characteristics. *Faseb J.* 1992;6(15):3345-3353.
58. Baker SK, Chen ZL, Norris EH, Revenko AS, MacLeod AR, Strickland S. Blood-derived plasminogen drives brain inflammation and plaque deposition in a mouse

model of Alzheimer's disease. *Proc Natl Acad Sci U S A.* 2018;115(41):E9687-E9696.

59. Amara U, Flierl MA, Rittirsch D, et al. Molecular intercommunication between the complement and coagulation systems. *J Immunol.* 2010;185(9):5628-5636.
60. Ryu JK, Petersen MA, Murray SG, et al. Blood coagulation protein fibrinogen promotes autoimmunity and demyelination via chemokine release and antigen presentation. *Nature Communications.* 2015;6:8164.
61. Pollak TA, Drndarski S, Stone JM, David AS, McGuire P, Abbott NJ. The blood-brain barrier in psychosis. *Lancet Psychiatry.* 2018;5(1):79-92.
62. Ryu JK, Rafalski VA, Meyer-Franke A, et al. Fibrin-targeting immunotherapy protects against neuroinflammation and neurodegeneration. *Nat Immunol.* 2018;19(11):1212-1223.
63. Comes AL, Papiol S, Mueller T, Geyer PE, Mann M, Schulze TG. Proteomics for blood biomarker exploration of severe mental illness: pitfalls of the past and potential for the future. *Translational Psychiatry.* 2018;8(1):160.
64. Westwood S, Leoni E, Hye A, et al. Blood-Based Biomarker Candidates of Cerebral Amyloid Using PiB PET in Non-Demented Elderly. *Journal of Alzheimer's disease.* 2016;52(2):561-572.
65. Ripke S, Neale BM, Corvin A, et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511(7510):421-427.
66. Fung TC, Olson CA, Hsiao EY. Interactions between the microbiota, immune and nervous systems in health and disease. *Nat Neurosci.* 2017;20(2):145-155.
67. Khandaker GM, Zimbron J, Lewis G, Jones PB. Prenatal maternal infection, neurodevelopment and adult schizophrenia: a systematic review of population-based studies. *Psychological Medicine.* 2013;43(2):239-257.
68. Rasmussen LJH, Moffitt TE, Arseneault L, et al. Association of Adverse Experiences and Exposure to Violence in Childhood and Adolescence With Inflammatory Burden in Young People. *JAMA Pediatrics.* 2019:1-11.
69. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature.* 2010;468(7321):203-212.

Chapter 3 Appendices

Appendix 3.1: Supplementary Methods é 1 3

Appendix 3.2: Supplementary Tables é . . .

Supplementary Table 3.1: Summary of support vector machine models

Supplementary Table 3.2: List of 69 baseline EU-GEI clinical variables included in model 1a, model 1b and model 3

Supplementary Table 3.3: Comparison of baseline characteristics for CHR participants who attended at least one follow-up interview vs. CHR participants who did not in EU-GEI

Supplementary Table 3.4: Comparison of characteristics for participants included in initial experiment (N=133) from total EU-GEI clinical high-risk cohort (N=344)

Supplementary Table 3.5: Results of ANCOVA (adjusted for age, sex, BMI and years in education) and fold changes (CHR-T vs. CHR-NT) for proteins identified in EU-GEI baseline plasma samples in the initial experiment

Supplementary Table 3.6: Coefficients of variation for proteins across quality control standards in EU-GEI initial experiment

Supplementary Table 3.7: Summary of protein-protein interactions identified from the BIOGRID database for significantly differentially expressed proteins between CHR-T and CHR-NT in EU-GEI initial experiment

Supplementary Table 3.8: Functional enrichment analysis of differentially expressed proteins (following false discovery rate correction) between CHR-T and CHR-NT in EU-GEI initial experiment: 6 KEGG pathways significantly enriched

Supplementary Table 3.9: Results of enzyme-linked immunosorbent assays in CHR-T and CHR-NT participants in EU-GEI initial experiment

Supplementary Table 3.10: Correlations between 5 proteins assessed by ELISA and by mass spectrometry in EU-GEI initial experiment

Supplementary Table 3.11: Comparison of characteristics for participants included in replication experiment (N=135) from total EU-GEI clinical high-risk cohort (N=344)

Supplementary Table 3.12: Sample characteristics for CHR-T and CHR-NT groups in the replication experiment

Supplementary Table 3.13: Results of ANCOVA (adjusted for age, sex, BMI, years in education, tobacco use and ethnicity) and fold changes (CHR-T vs. CHR-NT) for proteins identified in EU-GEI baseline plasma samples in the replication experiment

Supplementary Table 3.14: Sample characteristics for ALSPAC subsample cases and controls

Supplementary Table 3.15: Results of ANCOVA (adjusted for sex, BMI and maternal social class) and fold changes (definite PEs at 18 vs. no PEs at 18) for proteins identified in ALSPAC age 12 plasma samples

Supplementary Table 3.16: Ten percent highest-weighted features for Model S2 (support vector machine model predicting functional outcome at 24 months in EU-GEI)

Supplementary Table 3.17: Table comparing performance metrics for multi-class site prediction models based on 69 clinical features from model 1b

Supplementary Table 3.18: Table comparing performance metrics for multi-class site prediction models based on 166 proteomic features from model 1c

Supplementary Table 3.19: Performance metrics for uncorrected vs. corrected support vector machine models

Supplementary Table 3.20: Proteins differentially expressed in CHR-T vs. CHR-NT on ANCOVA (p<0.05) in EU-GEI baseline plasma samples in the initial and replication experiment and predicted systemic impact on coagulation and complement activation and regulation

Supplementary Table 3.21: Table summarising directionality of effect of ten percent highest weighted features in model 1a (EU-GEI initial experiment), model 3 (EU-GEI replication experiment) and model 4 (ALSPAC proteomic data)

Appendix 3.3: Supplementary Figures é 2 1

Supplementary Figure 3.1: Derivation and testing of model 2b: parsimonious (10-predictor) proteomic model

Supplementary Figure 3.2: STRING functional protein association network for proteins differentially expressed (following false discovery rate correction) between CHR-T and CHR-NT in EU-GEI initial experiment

Supplementary Figure 3.3: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 1b: clinical data

Supplementary Figure 3.4: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 1c: proteomic data

Supplementary Figure 3.5: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 2a: proteomic (non-London)

Supplementary Figure 3.6: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 2b: parsimonious (10-predictor) proteomic model, training data (non-London)

Supplementary Figure 3.7: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 3: replication

Supplementary Figure 3.8: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 4: ALSPAC proteomic data

Supplementary Figure 3.9: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for Model S1: ELISA

Supplementary Figure 3.10: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for Model S2: functional outcome

Supplementary Figure 3.11: Model 1b (clinical data) decision scores stratified by EU-GEI site

Supplementary Figure 3.12: Model 1c (proteomic data) decision scores stratified by EU-GEI site

Supplementary Figure 3.13: Age stratified by EU-GEI site

Supplementary Figure 3.14: Years in education stratified by EU-GEI site

Supplementary Figure 3.15: Body mass index stratified by EU-GEI site

Supplementary Figure 3.16: Sex stratified by EU-GEI site

Supplementary Figure 3.17: General Assessment of Functioning symptoms subscale stratified by EU-GEI site

Supplementary Figure 3.18: General Assessment of Functioning disability subscale stratified by EU-GEI site

Supplementary Figure 3.19: Scale for Assessment of Negative Symptoms (composite score) stratified by EU-GEI site

Supplementary Figure 3.20: Scale for Assessment of Negative Symptoms (global score) stratified by EU-GEI site

Supplementary Figure 3.21: Brief Psychiatric Rating Scale score stratified by EU-GEI site

Supplementary Figure 3.22: Montgomery- Åsberg Depression Rating Scale score stratified by EU-GEI site

Supplementary Figure 3.23: Multi-class receiver-operating curves for site prediction based on 69 clinical features from model 1b

Supplementary Figure 3.24: Multi-class receiver-operating curves for site prediction based on 166 proteomic features from model 1c

Supplementary Figure 3.25: Illustration of the complement and coagulation pathways depicting the impact model of differentially expressed complement and coagulation proteins in CHR-T vs. CHR-NT

Appendix 3.1: Supplementary Methods

Sample preparation: further details

Sample collection

Blood samples were collected according to a standardised protocol across the participating EU-GEI sites. BD Vacutainer P100 tubes (BD Biosciences, USA) were used for sample collection. Samples were stored on ice for a maximum of 90 minutes, then centrifuged at 2500g for 20 minutes. ALSPAC samples were collected into heparin Sarstedt S-Monovette tubes (Sarstedt, Germany) and stored on ice fo
In both studies, participants were non-fasting and there were no restrictions regarding time of sample collection. Aliquots of plasma were stored in a -80°C freezer until processing for proteomic analysis as described below. All samples were subject to one freeze-thaw cycle prior to analysis. The standard quality of the plasma samples was ensured by assessing the overall MS protein profile to facilitate the identification of outlier protein expression profiles.

Protein depletion of plasma samples

To improve the dynamic range for proteomic analysis, 40µl of plasma from each case in all samples was immunodepleted of the 14 most abundant proteins (Ŭ-1-antitrypsin, A1-acid glycoprotein, Ŭ-macroglobulin, Apolipoprotein A-I, Ŭ-2-Apolipoprotein A-II, Complement C3, Fibrinogen Ŭ/ƀ/ɔ, Haptoglobin, Serotransferrin), using the Agilent Hu14 Affinity Removal System (MARS) coupled to a High Performance Liquid Chromatography (HPLC) system.[1] Protein depletion was undertaken according to the manufacturerôs ins
performed with 50mM ammonium bicarbonate using spin columns with a 10kDA-molecular weight cut-off (Merck Millipore). Prior to sample preparation for mass spectrometry (MS), the protein concentration was determined using a Bradford Assay[2] according to the manufacturerôs (BioRad) instructions.

Sample preparation for mass spectrometry

Protein digestion and peptide purification was performed as previously described.[3] For quality control (QC), an equal aliquot from each protein digest in the experiment was pooled into one sample for use as an internal QC. This QC standard was injected three

times at the beginning of the MS study to condition the column, and after every 10 injections throughout the experiment to monitor the MS performance.

Discovery proteomic analysis using data dependent acquisition (DDA): further methodological details

All samples were injected on a Thermo Scientific Q Exactive mass spectrometer connected to a Dionex Ultimate 3000 (RSLCnano) chromatog digest) from each sample were loaded onto a fused silica e laser puller (Sutter Instruments P2000), packe media (nanoLCMS Solutions LCC) and separated by an increasing acetonitrile gradient over 90 minutes at a flow rate of 250 nL/min. The mass spectrometer was operated in data dependent TopN 8 mode, with the following settings: mass range 300-1600Th; resolution for MS1 scan 70000; AGC target 3e6; resolution for MS2 scan 17500; AGC target 2e4; charge exclusion unassigned, 1; dynamic exclusion 40 s. The peptide sequences used to match to parent proteins in the human FASTA database are provided in the eAppendix available via https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2769907. The co-efficients of variation (CV) across the injected pooled QC standards for the 35 differentially expressed proteins in the initial experiment are provided in Supplementary Table 3.6. This table also provides the CV for proteins not among these 35 that were either among the 10% highest-weighted proteins in model 1a, or among the 10 proteins included in model 2b. The majority of these (80%) had CV <20%.

Confirmatory analyses using enzyme-linked immunosorbent assays (ELISA)

To validate our proteomic differential expression findings using another method, we assessed several human complement and coagulation proteins and apolipoproteins in the plasma samples of the same CHR-T and CHR-NT subjects who contributed to the initial EU-GEI proteomic experiment using enzyme-linked immunoassays (ELISA). Candidate proteins were chosen on the basis of differential expression and machine learning results as well as previous studies from our group ([4-7]). Specifically, we tested macroglobulin (Abcam ab108888, 1:400), apolipoprotein E (ThermoFisher Scientific, EHAPOE, 1:2,000), Complement C1q (Abcam ab170246, 1:100,000), Complement C1r (Abcam ab170245, 1:40,000), Complement C4 binding protein (Abcam, ab222866, 1:40,000), Complement C8 (Abcam ab 137971, 1:10,000), complement factor H (Hykult Biotech HK342, 1:10,000), immunoglobulin M (Abcam, ab137982, 1:60,000), and

plasminogen (Abcam ab108893, 1:20,000). Each kit had an optimised and defined sample dilution recommendation. The assays were performed according to the manufacturer's instructions to ensure sample concentrations measured would fall in the linear portion of the standard curve in which analyte concentration can be determined accurately. Each plate contained at least 8 wells in duplicate used for a standard curve of known protein concentrations. The assays have been used reliably in other studies.[8-20] Results are provided in Supplementary Table 3.9. With respect to concordance with the mass spectrometry data, ELISA and mass spectrometry results were available for five proteins (alpha-2-macroglobulin, plasminogen, complement component 1r, complement factor H and apolipoprotein E) of which 3 were significantly positively correlated (alpha-2-macroglobulin, plasminogen and apolipoprotein E); see Supplementary Table 3.10.

Bioinformatics and statistical analysis: further details

Predictive models

We used Neurominer v.1.0 (https://www.pronia.eu/neurominer/) for MatLab 2018a (MathWorks Inc.) to develop support vector machine (SVM) models. For all models, hyperparameters were optimised in repeated nested cross-validation and area under the receiver-operating characteristic curve (AUC) was the performance evaluation criterion. We used the LIBLINEAR program with L2 regularisation to attenuate over-fitting[21] whereby weights of non-predictive features are minimised, but not reduced to zero (thus more closely modelling the biological effects of functionally inter-related proteins). Hyperplane weighting was enabled (increasing the misclassification penalty in the minority class) which reduces risk of bias in unbalanced group sizes.[22] Random-label permutation analysis[23-25] with 1000 permutations was used to verify models against a null distribution and derive *p*-values for statistical significance. Prior to model training, missing clinical data were replaced using the mean (for continuous) or modal value (for categorical variables). Continuous clinical variables were converted to z-scores and winsorised within ±3z.

EU-GEI clinical features

A full list of the baseline clinical features included in model 1a, model 1b and model 3 is provided in Supplementary Table 3.2.

Leave-site-out repeated nested cross-validation

The data were first split into a validation. To incorporate the ó o geographical generalisability, we split the data into groups by study site. Several of the smaller sites from the EU-GEI study were combined to ensure a large enough transition sample was present at each site. Thus, Amsterdam and The Hague were combined (Netherlands), Basel and Vienna were combined (Switzerland/ Austria), Copenhagen and Paris were combined (Denmark/ France), and Barcelona and Sao Paulo were combined (Spain/ Brazil). This resulted in six final groups (as shown below):

Group	CHR-NT, n (%)	CHR-T, n (%)
London	34 (64%)	19 (36%)
Denmark/ France	13 (54%)	11 (46%)
Netherlands	13 (87%)	2 (13%)
Melbourne	9 (64%)	5 (36%)
Switzerland/ Austria	6 (43%)	8 (57%)
Spain/ Brazil	9 (69%)	4 (31%)

These groups were folds in the outer cross-validation loop. For each cycle of cross-validation, data from each of the six groups were held out and the rest of the data moved into the óinner loopô for training.

Within the inner loop, we used five non-overlapping folds with iterative training-test cycles. Thus, training was applied to four-fifths of the data in the inner loop and then tested against the final one-fifth, with the five different inner loop folds as the test fold. Models were trained and tested within the inner loop using a range of regularisation parameter values (in 11 steps from 0.015625 to 16). The optimal models thus derived were tested against the held-out site in the outer loop. This process was then repeated, with each site in the outer loop as the test site, to determine the overall optimal model and final predictive accuracy. For a detailed description of repeated nested cross-validation, see the supplementary material of Koutsouleris et al[26] and the Neurominer manual (available from https://www.pronia.eu/neurominer/).

Confidence intervals for area under the receiver-operating characteristic curve

Ninety-five per cent confidence intervals for the area under the receiver-operating characteristic curve (AUC) for each model were calculated according to the method of Hanley & McNeil.[27]

ELISA results

Concentrations for the unknown samples were interpolated using 4 parameter logistic curve fit in GraphPad Prism 8 software. Means for each protein were compared using the t-test with unequal variances in Stata version 15.

Protein-protein interactions

To synthesise evidence for direct protein-protein interactions, searches for each of the 35 significantly differentially expressed proteins between CHR-T and CHR-NT in the initial experiment were performed using the BIOGRID curated database of protein-protein interactions (https://thebiogrid.org/). 34 of the 35 proteins were listed in the database (ADAMTS13 was not) and 11 unique interactions were identified (Supplementary Table 3.7). The same 35 proteins were entered into the STRING database (https://string-db.org/) to facilitate functional enrichment analysis (Supplementary Table 3.8) and generation of a functional protein association network (Supplementary Figure 3.2).

EU-GEI replication dataset

The replication dataset included 49 CHR-T participants (2 of whom were different from those in the original dataset used for models 1a-d) and 86 CHR-NT participants (all of whom were different from those in the original dataset). Baseline characteristics of participants included in the replication dataset are compared to those not included in Supplementary Table 3.11. Included participants were more likely to be male, although this difference was no longer statistically significant after false discovery rate (FDR) correction. Characteristics of the replication subsample are presented stratified by transition status in Supplementary Table 3.12. Regarding baseline characteristics, there was evidence of nominally significant differences between CHR-T and CHR-NT for ethnicity, tobacco use and mean total SANS composite, SANS global and BPRS scores, though not after FDR correction.

Supplementary analyses

Development of a model to predict transition outcome in EU-GEI based on ELISA results (Model S1)

Concentrations from the nine ELISAs performed (as described in Section 3 above) were z-transformed and winsorised within ±3z. The results were used as features in an L2-regularised SVM algorithm using leave-site-out repeated nested cross-validation. Hyperplane weighting was enabled and area under the receiver-operating characteristic curve (AUC) was the performance evaluation criterion.

The performance metrics for this model are presented in Table 3.2. The mean algorithm scores, predicted classes and receiver-operating characteristic (ROC) curve are shown in Supplementary Figure 3.9.

Development of a model to predict functional outcome at 24 months in EU-GEI based on clinical and proteomic data (Model S2)

To investigate the performance of the available clinical and proteomic data for prediction of a transdiagnostic outcome, we developed a model based on the 69 clinical and 166 proteomic features used in model 1a, but for prediction of functional outcome (model S2). We used the General Assessment of Functioning (GAF)[28,29] disability subscale, recorded at the follow-up assessment closest to two years from baseline, as a measure of general functioning. For use as a classification target variable (and in line with previous approaches[30]) the score was dichotomised into ópoor functio
or severe impairment) or ógood functioningô (>

An L2-regularised SVM algorithm was used to derive the model with AUC as the performance evaluation criterion. Hyperplane weighting was enabled. Compared to transition status, fewer participants had data available for functional outcome at 24 months (n=79). Therefore, this model used repeated nested cross-validation with five random inner and outer folds, irrespective of study site.

The performance metrics for this model are presented in Table 3.2 and the 10% highest-weighted features are shown in Supplementary Table 3.16. The mean algorithm scores, predicted classes and ROC curve are shown in Supplementary Figure 3.10.

Investigation of site associations in EU-GEI

Given the multisite nature of the EU-GEI study, we wished to investigate whether the clinical and proteomic data and predictions varied systematically by site.

First, we performed Kruskall-Wallis tests to determine whether the decision scores for model 1b (clinical model) and model 1c (proteomic model) differed according to the six site groupings used in the site-out cross validation schemes (i.e. London, Netherlands, Melbourne, Switzerland/Austria, Denmark/France and Spain/Brazil). The results are presented in Supplementary Figure 3.11 (for model 1b, clinical data) and Supplementary Figure 3.12 (for model 1c, proteomic data). There was evidence for differences in decision scores by site with respect to the clinical data (model 1b) and post-hoc Dunn's[31] tests with Bonferroni correction suggested that the pertinent sites were London and Netherlands. There was little evidence for differences in decision scores by site with respect to the proteomic data (model 1c).

On further investigation (Supplementary Figures 3.13 – 3.22), there were significant differences (p<0.05) between the six different site groups according to the Kruskall-Wallis test for six clinical variables: age, years in education, GAF symptoms subscale, GAF disability subscale, SANS total global score and BPRS total score. Post-hoc Dunn's tests with Bonferroni correction were performed for these six variables (Supplementary Figures 3.13, 3.14, 3.17, 3.18, 3.20, 3.21). The results indicated that site associations in the clinical data may have been driven primarily by differences in age between London vs. the other sites and Netherlands vs. the other sites (participants from London and Netherlands had slightly higher median age compared to the other sites), and possibly by differences in years in education between Netherlands vs. the other sites. Participants from the London site also had higher median total BPRS score compared to participants from most of the other sites.

Second, we trained SVM models to predict the site provenance of EU-GEI cases based firstly on the clinical data (69 clinical features as per model 1b) and secondly on the proteomic data (166 proteomic features as per model 1c). In both cases, in keeping with the original models, we used the LIBLINEAR program with area under the receiver-operating curve (AUC) as the performance evaluation criterion. Hyperplane weighting was enabled. Models were trained using five-fold pooled cross-validation with five inner and five

outer folds. We used a one-for multi-EU-GEI site predi
groupings that were used in cross-validation schemes, model
as shown below:

1. London vs. REST
2. Netherlands vs. REST
3. Melbourne vs REST
4. Switzerland/Austria vs. REST
5. Denmark/France vs REST
6. Spain/Brazil vs REST

The performance metrics and ROC curves for the multiclass site prediction models based on clinical data are shown in Supplementary Table 3.17 and Supplementary Figure 3.23, and for the proteomic data in Supplementary Table 3.18 and Supplementary Figure 3.24.

The results of this analysis indicated little evidence of systematic site associations for the proteomic data. The 95% confidence intervals (95% CI) for the AUC for prediction of each site included 0.5 (the value of no discrimination) with the exception of London vs. REST, although here the lower bound was close to 0.5 (AUC 0.61, 95% CI: 0.51 ï 0.71).

In keeping with the Kruskall-Wallis analyses above, there was greater evidence of site associations in relation to the clinical data for London vs. REST (AUC 0.76, 95% CI 0.67 ï 0.85) and Netherlands vs. REST (AUC 0.76, 95% CI 0.61 ï 0.91), likely driven by the differences in age, education and total BPRS score for these sites.

Multivariate correction

Since our models were predictive (rather than causal), the primary analyses did not include correction for possible covariates. However, to investigate this, we developed corrected versions of the SVM models using multivariate correction via principal components analysis (PCA). This was performed using the óExtract v processing function in Neurominer v1.0 (https://www.pronia.eu/neurominer/) whereby a PCA was performed on a source matrix of covari coefficient) between the eigenvariates from the PCA reduction were determined and those greater than 0.2 (indicating weak correlation) were identified. The target matrix was

projected to the source matrix PCA space (90% retained variance), then back-reconstructed to the original input space without the identified PCA components, thus removing the variance associated with those components.

The primary covariates of interest were age, sex, BMI and years in education (as a proxy socio-economic measure) in Study 1 and sex, BMI and maternal social class in Study 2. However, we additionally corrected for tobacco use and ethnicity in the replication dataset for Study 1, due to evidence of baseline differences for these variables between CHR-T and CHR-NT in this dataset (see Supplementary Table 3.12).

For each corrected model, the covariates for which correction was performed are listed below:

Model	**Covariates in PCA**
Model 1a (corrected): clinical and proteomic	Age, sex, BMI, years in education
Model 1b (corrected): clinical	Age, sex, BMI, years in education
Model 1c (corrected): proteomic	Age, sex, BMI, years in education
Model 2a (corrected): proteomic, non-London	Age, sex, BMI, years in education
Model 2b (corrected): top 10 proteomic, non-London	Age, sex, BMI, years in education
Model 3 (corrected): replication	Age, sex, BMI, years in education, tobacco use, ethnicity
Model 4 (corrected): ALSPAC PEs	Sex, BMI, maternal social class
Model S1 (corrected): ELISA	Age, sex, BMI, years in education
Model S2 (corrected): functional outcome	Age, sex, BMI, years in education

Performance metrics of uncorrected and corrected models are presented in Supplementary Table 3.19. There were generally slight reductions in the AUCs of the corrected models when compared to their uncorrected counterparts (median change in AUC 0.04, range 0.01 ï 0.10), although in all cases the 95% confidence intervals for the AUC of the uncorrected and corrected models overlapped.

Information on reporting of ethnicity

In the current study, data was collected for ethnicity on the basis of potential relationships to the outcomes of interest.[32,33] Options for ethnicity were investigator-defined according to the Medical Research Council Socio-Demographic Schedule (amended)[34] in EU-GEI and a self-reported questionnaire in ALSPAC.

Supplementary References

1. Levin Y, Wang L, Schwarz E, Koethe D, Leweke FM, Bahn S. Global proteomic profiling reveals altered proteomic signature in schizophrenia serum. *Mol Psychiatry.* 2010;15(11):1088-1100.
2. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976;72:248-254.
3. English JA, Fan Y, Focking M, et al. Reduced protein synthesis in schizophrenia patient-derived olfactory cells. *Translational Psychiatry.* 2015;5:e663.
4. English JA, Lopez LM, O'Gorman A, et al. Blood-Based Protein Changes in Childhood Are Associated With Increased Risk for Later Psychotic Disorder: Evidence From a Nested Case-Control Study of the ALSPAC Longitudinal Birth Cohort. *Schizophr Bull.* 2018;44(2):297-306.
5. Föcking M, Sabherwal S, Cates HM, et al. Complement pathway changes at age 12 are associated with psychotic experiences at age 18 in a longitudinal population-based study: evidence for a role of stress. *Molecular Psychiatry.* 2019.
6. Sabherwal S, English JA, Focking M, Cagney G, Cotter DR. Blood biomarker discovery in drug-free schizophrenia: the contributionof proteomics and multiplex immunoassays. *Expert Review of Proteomics.* 2016;13(12):1141-1155.
7. Sabherwal S, Focking M, English JA, et al. ApoE elevation is associated with the persistence of psychotic experiences from age 12 to age 18: Evidence from the ALSPAC birth cohort. *Schizophr Res.* 2019.
8. Ghale-Noie ZN, Hassani M, Kachooei AR, Kerachian MA. High Serum Alpha-2-Macroglobulin Level in Patients with Osteonecrosis of the Femoral Head. *The Archives of Bone and Joint Surgery.* 2018;6(3):219-224.

9. Bjelosevic S, Pascovici D, Ping H, et al. Quantitative Age-specific Variability of Plasma Proteins in Healthy Neonates, Children and Adults. *Molecular & Cellular Proteomics.* 2017;16(5):924-935.
10. Robinson MD, Mishra I, Deodhar S, et al. Water T2 as an early, global and practical biomarker for metabolic syndrome: an observational cross-sectional study. *Journal of Translational Medicine.* 2017;15(1):258.
11. Sunderic M, Mihailovic N, Nedic O. Protein molecular forms of insulin-like growth factor binding protein-2 change with aging. *Experimental Gerontology.* 2014;58:154-158.
12. Banville N, Burgess JK, Jaffar J, et al. A quantitative proteomic approach to identify significantly altered protein networks in the serum of patients with lymphangioleiomyomatosis (LAM). *PloS One.* 2014;9(8):e105365.
13. Zeng J, Deng P, Li J, Feng W, Chen J, Zeng Y. Increased serum protein levels by Yuanshi Shengmai Chenggu Tablet in treatment of avascular osteonecrosis of the femoral head. *Molecular Medicine Reports.* 2018;17(2):2121-2126.
14. Zhou ZQ, Tong DN, Guan J, et al. Follicular helper T cell exhaustion induced by PD-L1 expression in hepatocellular carcinoma results in impaired cytokine expression and B cell help, and is associated with advanced tumor stages. *American Journal of Translational Research.* 2016;8(7):2926-2936.
15. Guy JH, Pyne DB, Deakin GB, Miller CM, Edwards AM. Acclimation Training Improves Endurance Cycling Performance in the Heat without Inducing Endotoxemia. *Frontiers in Physiology.* 2016;7:318.
16. Ferrin G, Rodriguez-Peralvarez M, Aguilar-Melero P, et al. Plasma protein biomarkers of hepatocellular carcinoma in HCV-infected alcoholic patients with cirrhosis. *PloS One.* 2015;10(3):e0118527.
17. Liang Y, Xie SB, Wu CH, et al. Coagulation cascade and complement system in systemic lupus erythematosus. *Oncotarget.* 2018;9(19):14862-14881.
18. Ting KR, Henry M, Meiller J, et al. Novel panel of protein biomarkers to predict response to bortezomib-containing induction regimens in multiple myeloma patients. *BBA Clinical.* 2017;8:28-34.
19. McGuire JL, Gill AJ, Douglas SD, Kolson DL. The complement system, neuronal injury, and cognitive function in horizontally-acquired HIV-infected youth. *Journal of Neurovirology.* 2016;22(6):823-830.

20. Lui H, Zhang J, Makinson SR, et al. Progranulin Deficiency Promotes Circuit-Specific Synaptic Pruning by Microglia via Complement Activation. *Cell.* 2016;165(4):921-935.
21. Fan R-E, Chang K-W, Hsieh C-J, Wang X-R, Lin C-J. LIBLINEAR: A Library for Large Linear Classification. *J Mach Learn Res.* 2008;9:1871-1874.
22. Krawczyk B. Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence.* 2016;5(4):221-232.
23. Gaonkar B, Davatzikos C. Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification. *Neuroimage.* 2013;78:270-283.
24. Koutsouleris N, Kahn RS, Chekroud AM, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *Lancet Psychiatry.* 2016;3(10):935-946.
25. Golland P, Fischl B. Permutation Tests for Classification: Towards Statistical Significance in Image-Based Studies. Paper presented at: Information Processing in Medical Imaging; 2003; Berlin, Heidelberg.
26. Koutsouleris N, Kahn RS, Chekroud AM, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *Lancet Psychiatry.* 2016;3(10):935-946.
27. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143(1):29-36.
28. Aas IH. Global Assessment of Functioning (GAF): properties and frontier of current knowledge. *Ann Gen Psychiatry.* 2010;9:20.
29. Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. *Am J Psychiatry.* 1992;149(9):1148-1156.
30. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, et al. Prediction Models of Functional Outcomes for Individuals in the Clinical High-Risk State for Psychosis or With Recent-Onset Depression: A Multimodal, Multisite Machine Learning Analysis. *JAMA Psychiatry.* 2018;75(11):1156-1172.
31. Dinno A. Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. *Stata Journal.* 2015;15(1):292-300.
32. Morgan C, Charalambides M, Hutchinson G, Murray RM. Migration, ethnicity, and psychosis: toward a sociodevelopmental model. *Schizophr Bull.* 2010;36(4):655-664.

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

34. Mallett R. Sociodemographic Schedule. Section of Social Psychiatry, Institute of Psychiatry; London: 1997

Appendix 3.2: Supplementary Tables

Supplementary Table 3.1: Summary of support vector machine models

Model	Aim	Source cohort	Target outcome to be predicted	Data modalities	Repeated nested cross-validation method	Development data (*n* in each group)	Test data (*n* in each group)
Model 1a	Predict transition in CHR sample using baseline clinical and proteomic data	EU-GEI	Transition status	69 clinical and 166 proteomic features	Leave-site-out	Initial experiment, all sites (49 CHR-T, 84 CHR-NT)	N/A
Model 1b	Predict transition in CHR sample using baseline clinical data	EU-GEI	Transition status	69 clinical features	Leave-site-out	Initial experiment, all sites (49 CHR-T, 84 CHR-NT)	N/A
Model 1c	Predict transition in CHR sample using baseline proteomic data	EU-GEI	Transition status	166 proteomic features	Leave-site-out	Initial experiment, all sites (49 CHR-T, 84 CHR-NT)	N/A
Model 2a	Derivation of 10 highest-weighted proteomic features	EU-GEI	Transition status	166 proteomic features	Leave-site-out	Initial experiment, all sites except London (30 CHR-T, 50 CHR-NT)	N/A
Model 2b	Predict transition in CHR sample using 10 proteomic features, and test in held-out data	EU-GEI	Transition status	10 highest-weighted proteomic features from model 2a	Leave-site-out	Initial experiment, all sites except London (30 CHR-T, 50 CHR-NT)	Initial experiment, London site (19 CHR-T, 34 CHR-NT)

Model 3	Replicate findings from Model 1a	EU-GEI	Transition status	69 clinical and 119 proteomic features	Leave-site-out	Replication experiment, all sites (49 CHR-T, 86 CHR-NT)	N/A
Model 4	Predict age 18 psychotic experiences in general population sample using proteomic data at age 12	ALSPAC	Psychotic experiences (PEs)	265 proteomic features	Pooled (5 random inner and outer folds)	ALSPAC subsample (55 PE, 66 no PE)	N/A
Supplementary models							
Model	**Aim**	**Source cohort**	**Target outcome to be predicted**	**Data modalities**	**Repeated nested cross-validation method**	**Development data** **(*n* in each group)**	**Test data (*n* in each group)**
Model S1	Predict transition in CHR sample using baseline ELISA results	EU-GEI	Transition status	9 ELISA features	Leave-site-out	Initial experiment, all sites (44 CHR-T, 82 CHR-NT)	N/A
Model S2	Predict functional outcome in CHR sample using baseline clinical and proteomic data	EU-GEI	Functional outcome	69 clinical and 166 proteomic features	Pooled (5 random inner and outer folds)	Initial experiment, all sites (4 7 GAFÒ6 0 ,	N/A

CHR: clinical high risk; ELISA: enzyme-linked immunosorbent assay; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; ALSPAC: Avon Longitudinal Study of Parents and Children

Supplementary Table 3.2: List of 69 baseline EU-GEI clinical features included in Model 1a, Model 1b and Model 3

Age
Sex
Body mass index
Years in education
GAF symptoms
GAF disability
SANS: unchanging facial expression
SANS: decreased spontaneous movements
SANS: paucity of expressive gestures
SANS: poor eye contact
SANS: affective non-responsivity
SANS: inappropriate affect
SANS: lack of vocal inflections
SANS: global rating of affective flattening
SANS: poverty of speech
SANS: poverty of speech content
SANS: blocking
SANS: increased latency of response
SANS: global rating of alogia
SANS: grooming and hygiene
SANS: impersistence at work or school
SANS: physical anergia
SANS: global rating for avolition - apathy

SANS: recreational interests and activities
SANS: sexual activity
SANS: ability to feel intimacy and closeness
SANS: relationship with friends and peers
SANS: global rating of anhedonia - asociality
SANS: social inattentiveness
SANS: inattentiveness during mental status testing
SANS: global rating of attention
Total SANS composite score
Total SANS global score
BPRS: somatic concern
BPRS: anxiety
BPRS: depression
BPRS: suicidality
BPRS: guilt
BPRS: hostility
BPRS: elevated mood
BPRS: grandiosity
BPRS: suspiciousness
BPRS: hallucinations
BPRS: unusual thought content
BPRS: bizarre behaviour
BPRS: self-neglect
BPRS: disorientation
BPRS: conceptual disorganisation

BPRS: blunted affect
BPRS: emotional withdrawal
BPRS: motor retardation
BPRS: tension
BPRS: uncooperativeness
BPRS: excitement
BPRS: distractibility
BPRS: motor hyperactivity
BPRS: mannersims and posturing
Total BPRS score
MADRS: apparent sadness
MADRS: reported sadness
MADRS: inner tension
MADRS: reduced sleep
MADRS: reduced appetite
MADRS: concentration difficulties
MADRS: lassitude
MADRS: inability to feel
MADRS: pessimistic thoughts
MADRS: suicidal thoughts
Total MADRS score

GAF: General Assessment of Functioning; SANS: Scale for the Assessment of Negative Symptoms; BPRS: Brief Psychiatric Rating Scale; MADRS: Montgomery Asberg Depression Rating Scale; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions

Supplementary Table 3.3: Comparison of baseline characteristics for CHR participants who attended at least one follow-up interview vs. CHR participants who did not in EU-GEI

	Missing data, n (%)	Attended at least one follow-up interview, n=171	Did not attend at least one follow-up interview, n=173	t / ^{2} G	*p*	Corrected *p* (FDR 5%)
Baseline age in years, mean (SD)	0	23.0 (5.1)	21.8 (4.7)	2.240	0.026	0.104
Sex, n (%)	0	92 male (53.8%) 79 female (46.2%)	93 male (53.8%) 80 female (46.2%)	0.000	0.993	0.993
Baseline BMI in kg/m², mean (SD)	53 (15.4%)	23.9 (5.2)	24.0 (5.1)	-0.274	0.784	0.941
Baseline years in education, mean (SD)	41 (11.9%)	15.0 (2.8)	13.7 (3.2)	3.689	<0.001	<0.008
Ethnicity, n (%)	3 (0.9%)	124 white (72.1%) 18 black (10.5%) 30 other (17.4%)	122 white (72.2%) 15 black (8.9%) 32 other (18.9%)	0.327	0.849	0.941
Ever used cannabis, n (%)	13 (3.8%)	129 yes (77.2%) 38 no (22.8%)	112 yes (68.3%) 52 no (31.7%)	3.350	0.067	0.214
Baseline cannabis use, n (%)	98 (28.5%)	46 yes (35.4%) 84 no (64.6%)	40 yes (34.5%) 76 no (65.5%)	0.022	0.882	0.941
Baseline tobacco use, n (%)	41 (11.9%)	89 yes (55.6%) 71 no (44.4%)	71 yes (49.7%) 72 no (50.3%)	1.082	0.298	0.681
Baseline alcohol use, n (%)	15 (4.4%)	117 yes (69.6%) 51 no (30.4%)	116 yes (72.0%) 45 no (28.0%)	0.230	0.631	0.941

Baseline medication use, n (%)	89 (25.9%)	69 yes (50.4%) Antipsychotic 9 Antidepressant 54 Hypnotic 15 Other 21 68 no (49.6%)	57 yes (48.3%) Antipsychotic 22 Antidepressant 37 Hypnotic 8 Other 15 61 no (51.7%)	0.108	0.743	0.941
Baseline GAF symptoms score, mean (SD)	27 (7.8%)	55.3 (9.9)	54.9 (10.3)	0.280	0.779	0.941
Baseline GAF disability score, mean (SD)	12 (3.5%)	56.5 (12.0)	54.3 (12.5)	1.613	0.108	0.288
Baseline SANS total composite score, mean (SD)	48 (14.0%)	14.0 (10.5)	17.5 (12.7)	-2.600	0.010	0.053
Baseline SANS total global score, mean (SD)	32 (9.3%)	4.8 (3.4)	6.2 (3.7)	-3.357	0.001	0.008
Baseline BPRS total score, mean (SD)	28 (8.1%)	43.1 (10.7)	44.0 (9.7)	-0.794	0.428	0.856
Baseline MADRS total score, mean (SD)	19 (5.5%)	18.8 (9.1)	19.0 (9.2)	-0.261	0.794	0.941

Tobacco use was defined as daily use for at least 1 month over the previous 12 months. Alcohol use was defined as at least 12 or more alcoholic beverages over the previous 12 months. Missing data excluded in hypothesis tests. EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; CHR-T: clinical high risk, transitioned to psychosis; CHR-NT: clinical high risk, did not transition to psychosis; FDR: false discovery rate; BMI: body mass index; GAF: General Assessment of Functioning; SANS: Scale for the Assessment of Negative Symptoms; BPRS: Brief Psychiatric Rating Scale; MADRS: Montgomery Asberg Depression Rating Scale

Supplementary Table 3.4: Comparison of characteristics for participants included in initial experiment (N=133) from total EU-GEI clinical high-risk cohort (N=344)

	Missing data, n (%)	Included, n=133 (49 CHR-T, 84 CHR-NT)	Not included, n=211 (16 CHR-T, 195 CHR-NT)	t / ^{2} c	*p*	Corrected *p* (FDR 5%)
Baseline age in years, mean (SD)	0	22.6 (4.5)	22.3 (5.2)	0.686	0.493	0.681
Gender, n (%)	0	68 male (51.1%) 65 female (49.9%)	117 male (55.5%) 94 female (45.5%)	0.613	0.434	0.681
Baseline BMI in kg/m², mean (SD)	50 (14.5%)	24.4 (5.6)	23.7 (4.9)	1.190	0.235	0.439
Baseline years in education, mean (SD)	38 (11.0%)	14.3 (3.1)	14.4 (3.0)	-0.318	0.751	0.793
Ethnicity, n (%)	0	91 white (68.4%) 15 black (11.3%) 27 other (20.3%)	156 white (73.9%) 19 black (9.0%) 36 other (17.1%)	1.239	0.538	0.681
Ever used cannabis, n (%)	10 (2.9%)	101 yes (75.9%) 29 no (21.8%) 3 not known (2.3%)	143 yes (67.8%) 61 no (28.9%) 7 not known (3.3%)	2.326	0.127	0.302
Baseline cannabis use, n (%)	95 (27.6%)	41 yes (30.8%) 63 no (47.4%) 29 not known (21.8%)	47 yes (22.3%) 98 no (46.4%) 66 not known (31.3%)	1.302	0.254	0.439
Baseline tobacco use, n (%)	38 (11.0%)	64 yes (48.1%) 55 no (41.4%) 14 not known (10.5%)	97 yes (46.0%) 90 no (42.7%) 24 not known (11.4%)	0.106	0.744	0.793

Baseline alcohol use, n (%)	12 (3.5%)	93 yes (69.9%) 37 no (27.8%) 3 not known (2.3%)	141 yes (66.8%) 61 no (28.9%) 9 not known (4.3%)	0.115	0.735	0.793
Baseline medication use, n (%)	87 (25.3%)	51 yes (38.3%) Antidepressant 38 Antipsychotic 15 Hypnotic 9 Other 16 51 no (38.3%) 31 not known (23.3%)	77 yes (36.5%) Antidepressant 53 Antipsychotic 16 Hypnotic 14 Other 20 78 no (37.0%) 56 not known (26.5%)	0.003	0.960	0.960
Baseline GAF symptoms score, mean (SD)	27 (7.8%)	54.7 (10.2)	55.4 (10.0)	-0.634	0.526	0.681
Baseline GAF disability score, mean (SD)	12 (3.5%)	53.9 (11.7)	56.4 (12.5)	-1.827	0.069	0.187
Baseline SANS total composite score, mean (SD)	45 (13.1%)	18.0 (12.7)	14.2 (10.7)	2.782	0.006	0.029
Baseline SANS total global score, mean (SD)	29 (8.4%)	6.1 (3.9)	5.0 (3.4)	2.528	0.012	0.046
Baseline BPRS total score, mean (SD)	25 (7.3%)	46.0 (10.9)	42.2 (9.6)	3.174	0.002	0.013
Baseline MADRS total score, mean (SD)	16 (4.7%)	19.6 (9.7)	18.4 (8.8)	1.155	0.249	0.439

2-year GAF symptoms score, mean (SD)	142 (41.3%)	54.6 (15.0)	63.0 (11.6)	-4.083	<0.001	<0.010
2-year GAF disability score, mean (SD)	124 (36.0%)	56.9 (15.0)	63.6 (13.8)	-3.333	0.001	0.010
2-year GAF disability score, dichotomous outcome [a]	124 (36.0%)	32 good (24.1%) 47 poor (35.3%) 54 not known (40.6%)	80 good (37.9%) 61 poor (28.9%) 70 not known (33.2%)	5.337	0.021	0.067

[a] Poor functioning: GAF disability score Ò60; good func

Tobacco use was defined as daily use for at least 1 month over the previous 12 months. Alcohol use was defined as at least 12 or more alcoholic beverages over the previous 12 months. Missing data excluded in hypothesis tests.

EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; CHR-T: clinical high risk, transitioned to psychosis; CHR-NT: clinical high risk, did not transition to psychosis; FDR: false discovery rate; BMI: body mass index; GAF: General Assessment of Functioning; SANS: Scale for the Assessment of Negative Symptoms; BPRS: Brief Psychiatric Rating Scale; MADRS: Montgomery Asberg Depression Rating Scale

Supplementary Table 3.5: Results of ANCOVA (adjusted for age, sex, BMI and years in education) and fold changes (CHR-T vs. CHR-NT) for proteins identified in EU-GEI baseline plasma samples in the initial experiment

Uniprot No.	Protein name	F	*p*	Corrected *p* (5% FDR)	Direction of effect (T vs. NT)	Ratio of means (T vs. NT)
P01023	Alpha-2-macroglobulin	146	7.55E-23	1.25E-20	↓	0.33
P01871	Immunoglobulin heavy constant mu	72.16	4.53E-14	3.76E-12	↓	0.41
P07357	Complement component C8 alpha chain	44.25	7.76E-10	4.29E-08	↑	1.48
P02774	Vitamin D-binding protein	40.97	2.72E-09	1.13E-07	↑	1.43
P02747	Complement C1q subcomponent subunit C	36.52	1.56E-08	5.18E-07	↑	1.53
P00747	Plasminogen	31.39	1.25E-07	3.46E-06	↑	1.29
P10909	Clusterin	29.74	2.48E-07	5.88E-06	↑	1.29
P23142	Fibulin-1	28.56	4.06E-07	8.42E-06	↑	1.52
P55058	Phospholipid transfer protein	19.09	2.57E-05	0.0005	↓	0.67
P00736	Complement C1r subcomponent	18.07	4.09E-05	0.0007	↑	1.27
O75882	Attractin	16.85	0.0001	0.0011	↑	1.30
P03951	Coagulation factor XI	17.61	0.0001	0.0011	↑	1.36
P05156	Complement factor I	16.72	0.0001	0.0011	↑	1.23
P08603	Complement factor H	16.18	0.0001	0.0011	↑	1.16
P43320	Beta-crystallin B2	16.39	0.0001	0.0011	↑	1.80
P04003	C4b-binding protein alpha chain	14.71	0.0002	0.0020	↓	0.76
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	15.05	0.0002	0.0020	↑	1.19

O75636	Ficolin-3	14.09	0.0003	0.0028	↓	0.70
P01860	Immunoglobulin heavy constant gamma 3	13.19	0.0004	0.0033	↓	0.70
P15144	Aminopeptidase N	13.38	0.0004	0.0033	↓	0.72
P02489	Alpha-crystallin A chain	12.08	0.0007	0.0055	↑	1.63
P06396	Gelsolin	11.83	0.0008	0.0058	↑	1.17
Q14520	Hyaluronan-binding protein 2	11.87	0.0008	0.0058	↑	1.21
P05155	Plasma protease C1 inhibitor	11.01	0.0012	0.0083	↓	0.85
P02766	Transthyretin	10.79	0.0013	0.0086	↓	0.60
P02749	Beta-2-glycoprotein 1	9.93	0.0020	0.0123	↑	1.27
P04217	Alpha-1B-glycoprotein	9.97	0.0020	0.0123	↓	0.83
P22891	Vitamin K-dependent protein Z	9.86	0.0021	0.0125	↓	0.68
P00751	Complement factor B	9.41	0.0026	0.0149	↑	1.17
P05546	Heparin cofactor 2	8.91	0.0034	0.0188	↓	0.84
P06276	Cholinesterase	8.71	0.0038	0.0203	↑	1.18
P51884	Lumican	8.4	0.0044	0.0228	↑	1.25
P02649	Apolipoprotein E	8.25	0.0048	0.0241	↑	1.28
Q76LX8	A disintegrin and metalloproteinase with thrombospondin motifs 13	7.56	0.0068	0.0332	↓	0.84
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	7.15	0.0085	0.0403	↓	0.74
P02656	Apolipoprotein C-III	6.59	0.0114	0.0526	↑	1.37
P02751	Fibronectin	6.43	0.0124	0.0542	↑	1.30
P05543	Thyroxine-binding globulin	6.43	0.0124	0.0542	↑	1.18

P00450	Ceruloplasmin	6.32	0.0132	0.0562	↑	1.14
Q04756	Hepatocyte growth factor activator	5.69	0.0185	0.0768	↑	1.14
P05090	Apolipoprotein D	5.61	0.0193	0.0781	↑	1.17
Q08380	Galectin-3-binding protein	5.38	0.0219	0.0866	↓	0.77
P11226	Mannose-binding protein C	5.18	0.0246	0.0924	↑	1.36
P10643	Complement component C7	5.17	0.0247	0.0924	↑	1.17
P07225	Vitamin K-dependent protein S	5.12	0.0254	0.0924	↓	0.87
Q9BXR6	Complement factor H-related protein 5	5.1	0.0256	0.0924	↑	1.30
P49747	Cartilage oligomeric matrix protein	4.94	0.0280	0.0989	↑	1.16
P02675	Fibrinogen beta chain	4.79	0.0305	0.1033	↑	1.15
P02671	Fibrinogen alpha chain	4.77	0.0307	0.1033	↑	1.12
P04114	Apolipoprotein B-100	4.75	0.0311	0.1033	↓	0.89
P01042	Kininogen-1	4.69	0.0322	0.1048	↑	1.09
P43652	Afamin	4.49	0.0361	0.1152	↑	1.12
P05160	Coagulation factor XIII B chain	4.26	0.0410	0.1254	↑	1.14
Q9NZP8	Complement C1r subcomponent-like protein	4.23	0.0417	0.1254	↑	1.27
P02753	Retinol-binding protein 4	4.22	0.0421	0.1254	↑	1.36
P00742	Coagulation factor X	4.21	0.0423	0.1254	↑	1.11
P36980	Complement factor H-related protein 2	3.72	0.0559	0.1628	↑	1.31
P07195	L-lactate dehydrogenase B chain	3.68	0.0572	0.1637	↑	1.16

P02748	Complement component C9	3.64	0.0585	0.1646	↓	0.91
P01024	Complement C3	3.56	0.0615	0.1702	↑	1.13
P01876	Immunoglobulin heavy constant alpha 1	3.4	0.0675	0.1837	↑	1.14
P02743	Serum amyloid P-component	3.33	0.0705	0.1888	↑	1.12
P05452	Tetranectin	3.21	0.0756	0.1992	↑	1.11
P20742	Pregnancy zone protein	2.87	0.0930	0.2412	↓	0.76
P00740	Coagulation factor IX	2.83	0.0949	0.2424	↑	1.12
P01859	Immunoglobulin heavy constant gamma 2	2.55	0.1125	0.2830	↑	1.13
P00748	Coagulation factor XII	2.49	0.1170	0.2899	↑	1.14
P68871	Hemoglobin subunit beta	2.3	0.1315	0.3166	↑	1.33
P02787	Serotransferrin	2.3	0.1316	0.3166	↑	1.05
P01011	Alpha-1-antichymotrypsin	2.27	0.1347	0.3194	↑	1.06
P07737	Profilin-1	2.07	0.1526	0.3568	↓	0.75
Q92954	Proteoglycan 4	2.03	0.1565	0.3608	↓	0.88
P13671	Complement component C6	1.86	0.1745	0.3968	↑	1.06
P09871	Complement C1s subcomponent	1.84	0.1776	0.3984	↑	1.06
P07358	Complement component C8 beta chain	1.77	0.1856	0.4054	↓	0.91
Q03591	Complement factor H-related protein 1	1.77	0.1856	0.4054	↑	1.19
P02654	Apolipoprotein C-I	1.55	0.2147	0.4629	↓	0.69
P48740	Mannan-binding lectin serine protease 1	1.43	0.2336	0.4971	↑	1.10

P02760	Protein AMBP	1.41	0.2371	0.4982	↑	1.08
P02790	Hemopexin	1.37	0.2439	0.5033	↓	0.94
P29622	Kallistatin	1.34	0.2484	0.5033	↑	1.06
P04004	Vitronectin	1.34	0.2486	0.5033	↓	0.96
P14618	Pyruvate kinase PKM	1.32	0.2530	0.5060	↓	0.84
P22792	Carboxypeptidase N subunit 2	1.26	0.2639	0.5215	↑	1.06
P17936	Insulin-like growth factor-binding protein 3	1.21	0.2727	0.5275	↓	0.91
Q96PD5	N-acetylmuramoyl-L-alanine amidase	1.21	0.2734	0.5275	↓	0.93
P01019	Angiotensinogen	1.19	0.2767	0.5275	↓	0.89
Q96KN2	Beta-Ala-His dipeptidase	1.17	0.2808	0.5275	↓	0.92
P02652	Apolipoprotein A-II	1.16	0.2828	0.5275	↓	0.97
P27169	Serum paraoxonase/arylesterase 1	1.13	0.2894	0.5338	↑	1.11
P22105	Tenascin-X	1.08	0.3001	0.5474	↑	1.13
P09172	Dopamine beta-hydroxylase	0.95	0.3314	0.5980	↑	1.11
P60174	Triosephosphate isomerase	0.89	0.3481	0.6213	↓	0.86
P15169	Carboxypeptidase N catalytic chain	0.8	0.3724	0.6411	↑	1.07
P08185	Corticosteroid-binding globulin	0.8	0.3725	0.6411	↓	0.87
P23528	Cofilin-1	0.8	0.3737	0.6411	↑	1.11
P80108	Phosphatidylinositol-glycan-specific phospholipase D	0.79	0.3746	0.6411	↑	1.08

P02775	Platelet basic protein	0.78	0.3801	0.6438	↓	0.78
P06681	Complement C2	0.72	0.3974	0.6663	↓	0.94
P00734	Prothrombin	0.67	0.4130	0.6745	↑	1.03
P60709	Actin, cytoplasmic 1	0.67	0.4157	0.6745	↑	1.03
P35858	Insulin-like growth factor-binding protein complex acid labile subunit	0.66	0.4173	0.6745	↓	0.94
P02647	Apolipoprotein A-I	0.66	0.4185	0.6745	↓	0.89
P02763	Alpha-1-acid glycoprotein 1	0.63	0.4285	0.6769	↓	0.99
P04075	Fructose-bisphosphate aldolase A	0.62	0.4332	0.6769	↓	0.95
P18428	Lipopolysaccharide-binding protein	0.62	0.4335	0.6769	↓	0.91
P01834	Immunoglobulin kappa constant	0.61	0.4363	0.6769	↓	0.98
P04196	Histidine-rich glycoprotein	0.58	0.4474	0.6877	↑	1.08
Q9Y6R7	IgGFc-binding protein	0.56	0.4560	0.6906	↑	1.03
P08697	Alpha-2-antiplasmin	0.55	0.4594	0.6906	↓	0.97
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	0.54	0.4618	0.6906	↑	1.02
P00738	Haptoglobin	0.53	0.4680	0.6936	↓	0.60
O43866	CD5 antigen-like	0.5	0.4793	0.6981	↓	0.93
P0C0L5	Complement C4-B	0.5	0.4794	0.6981	↑	1.12
Q12913	Receptor-type tyrosine-protein phosphatase eta	0.49	0.4863	0.7020	↑	1.04
P04070	Vitamin K-dependent protein C	0.44	0.5103	0.7303	↓	0.88
P12814	Alpha-actinin-1	0.41	0.5251	0.7388	↓	0.82

P11021	Endoplasmic reticulum chaperone BiP	0.41	0.5252	0.7388	↑	1.02
P04275	von Willebrand factor	0.36	0.5505	0.7679	↑	1.08
P36955	Pigment epithelium-derived factor	0.33	0.5696	0.7839	↑	1.04
P03952	Plasma kallikrein	0.32	0.5714	0.7839	↑	1.03
P07360	Complement component C8 gamma chain	0.3	0.5837	0.7928	–	1.00
P02750	Leucine-rich alpha-2-glycoprotein	0.29	0.5908	0.7928	↑	1.03
Q15582	Transforming growth factor-beta-induced protein ig-h3	0.29	0.5922	0.7928	↓	0.95
P49908	Selenoprotein P	0.27	0.6054	0.8027	↑	1.06
Q16610	Extracellular matrix protein 1	0.26	0.6124	0.8027	↑	1.01
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	0.26	0.6141	0.8027	↓	0.97
P19320	Vascular cell adhesion protein 1	0.23	0.6290	0.8142	↓	0.93
P43251	Biotinidase	0.23	0.6327	0.8142	↓	0.91
P01031	Complement C5	0.19	0.6677	0.8491	↓	0.98
Q86UX7	Fermitin family homolog 3	0.18	0.6701	0.8491	↓	0.87
P21333	Filamin-A	0.18	0.6763	0.8505	↓	0.94
P00488	Coagulation factor XIII A chain	0.13	0.7166	0.8917	↓	0.96
Q9UGM5	Fetuin-B	0.13	0.7198	0.8917	↓	0.98
P04278	Sex hormone-binding globulin	0.11	0.7403	0.9103	↓	0.86
P30041	Peroxiredoxin-6	0.09	0.7667	0.9266	↓	0.99

P02746	Complement C1q subcomponent subunit B	0.09	0.7703	0.9266	↑	1.02
P02679	Fibrinogen gamma chain	0.08	0.7720	0.9266	↑	1.02
Q9NPH3	Interleukin-1 receptor accessory protein	0.08	0.7759	0.9266	↓	0.86
P00338	L-lactate dehydrogenase A chain	0.07	0.7934	0.9407	↑	1.02
P07996	Thrombospondin-1	0.06	0.8058	0.9439	↓	0.91
P25311	Zinc-alpha-2-glycoprotein	0.06	0.8074	0.9439	↑	1.01
Q9Y490	Talin-1	0.06	0.8149	0.9460	↓	0.89
P0CG06	Immunoglobulin lambda constant 3	0.05	0.8228	0.9485	↓	0.88
Q96IY4	Carboxypeptidase B2	0.03	0.8738	0.9920	↓	0.96
O00533	Neural cell adhesion molecule L1-like protein	0.02	0.8774	0.9920	↑	1.03
P08571	Monocyte differentiation antigen CD14	0.02	0.8899	0.9920	↑	1.03
P02765	Alpha-2-HS-glycoprotein	0.02	0.8933	0.9920	↑	1.01
P01857	Immunoglobulin heavy constant gamma 1	0.01	0.9034	0.9920	↓	0.98
P01008	Antithrombin-III	0.01	0.9108	0.9920	↑	1.01
P00915	Carbonic anhydrase 1	0.01	0.9113	0.9920	↑	1.05
Q6EMK4	Vasorin	0.01	0.9118	0.9920	↓	0.96
O00391	Sulfhydryl oxidase 1	0.01	0.9143	0.9920	↓	0.99
P12259	Coagulation factor V	0.01	0.9247	0.9928	–	1.00
P07359	Platelet glycoprotein Ib alpha chain	0.01	0.9291	0.9928	↓	0.98
P06727	Apolipoprotein A-IV	0.01	0.9391	0.9928	↑	1.01

P26927	Hepatocyte growth factor-like protein	0	0.9457	0.9928	↓	0.98
P13796	Plastin-2	0	0.9503	0.9928	↑	1.01
P63104	14-3-3 protein zeta/delta	0	0.9538	0.9928	↓	0.99
P01623	Immunoglobulin kappa variable 3-20	0	0.9579	0.9928	↓	0.92
P05154	Plasma serine protease inhibitor	0	0.9629	0.9928	↓	0.95
P0C0L4	Complement C4-A	0	0.9748	0.9978	↑	1.01
P22352	Glutathione peroxidase 3	0	0.9805	0.9978	↓	0.94
P18206	Vinculin	0	0.9858	0.9978	↓	0.91
Q9UK55	Protein Z-dependent protease inhibitor	0	0.9926	0.9986	↓	0.99
P01009	Alpha-1-antitrypsin	0	0.9993	0.9993	↓	0.86

CHR-T: clinical high-risk participants who transitioned to first episode psychosis; CHR-NT: clinical high-risk participants who did not transition; FDR: false discovery rate; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions

Supplementary Table 3.6: Coefficients of variation for proteins across quality control standards in EU-GEI initial experiment

Protein	Coefficient of variation (%)
35 significantly differentially expressed proteins on analysis of covariance:	
Alpha-2-macroglobulin	5.9
Immunoglobulin heavy constant mu	8.7
Complement component C8 alpha chain	20.2
Vitamin D-binding protein	20.3
Complement C1q subcomponent subunit C	15.0
Plasminogen	17.0
Clusterin	13.5
Fibulin-1	29.7
Phospholipid transfer protein	16.5
Complement C1r subcomponent	14.8
Attractin	12.0
Coagulation factor XI	10.5
Complement factor I	7.8
Complement factor H	9.1
Beta-crystallin B2	33.6
C4b-binding protein alpha chain	10.0
Inter-alpha-trypsin inhibitor heavy chain H1	13.5
Ficolin-3	12.0
Immunoglobulin heavy constant gamma 3	5.1
Aminopeptidase N	8.6
Alpha-crystallin A chain	25.7

Gelsolin	12.4
Hyaluronan-binding protein 2	11.1
Plasma protease C1 inhibitor	13.7
Transthyretin	29.8
Beta-2-glycoprotein 1	17.8
Alpha-1B-glycoprotein	14.0
Vitamin K-dependent protein Z	19.6
Complement factor B	5.6
Heparin cofactor 2	11.0
Cholinesterase	11.0
Lumican	8.9
Apolipoprotein E	13.2
A disintegrin and metalloproteinase with thrombospondin motifs 13	38.1
Inter-alpha-trypsin inhibitor heavy chain H3	14.1
Proteins not included in list above that were among 10% highest-weighted proteins in Model 1a:	
Vitamin K dependent protein S	12.3
Complement component 6	7.0
Retinol-binding protein 4	33.1
Alpha-2-antiplasmin	14.2
Proteins not included in both lists above that were among the 10 proteins included in Model 2b:	
N-acetylmuramoyl-L-alanine- amidase	9.5

Supplementary Table 3.7: Summary of protein-protein interactions identified from the BIOGRID database for significantly differentially expressed proteins between CHR-T and CHR-NT in EU-GEI initial experiment

Interactor A	Interactor B	Experimental system type	Notes	Source (Pubmed ID)
Alpha-2-macroglobulin	Apolipoprotein E	Affinity capture ï Western	Two-dimensional non-denaturing gradient gel electrophoresis of human plasma resulted in the separation of an apolipoprotein E-containing complex which co-migrated with plasma alpha-2-macroglobulin	Krimbou et al 1998 (9831625)
Alpha-2-macroglobulin	Apolipoprotein E	Reconstituted complex	Isolation of alpha-2-macroglobulin by gel filtration chromatography, electroelution, or immunoprecipitation resulted in co-isolation of apolipoprotein E	Krimbou et al 1998 (9831625)
Transthyretin	Alpha-2-macroglobulin	Two-hybrid	Yeast two-hybrid study. Directionality predicted by naïve Bayesian classifier (alpha-2-microglobulin upstream of transthyretin)	Vinayagam et al 2011 (21900206)
Alpha-2-macroglobulin	Apolipoprotein E	Two-hybrid	Yeast two-hybrid study	Soler-Lopez et al 2011 (21163940)
Transthyretin	Clusterin	Reconstituted complex	Interaction between clusterin and transthyretin (TTR) proteins including wild-type TTR and TTR variants V30M and L55P was assessed using a glutathione S-transferase pull-down assay. Clusterin was found to strongly interact with wild-type TTR and TTR variants V30M and L55P under acidic conditions. Clusterin was also found to inhibit the amyloid fibril formation of TTR variants	Lee et al 2009 (19664600)

Apolipoprotein E	Phospholipid transfer protein	Affinity capture ï Western	Co-elution on chromatography of HepG2 (human hepatoma cell line) cell culture medium. When medium subjected to chromatography on an anti-apolipoprotein E affinity column, a portion of phospholipid transfer protein applied was bound and could be eluted together with apolipoprotein E at low pH	Siggins et al 2003 (12810820)
Complement C1r	Plasma protease C1 inhibitor	Co-fractionation	Interactions analysed by sucrose-density-gradient ultracentrifugation and sodium dodecyl sulphate/polyacrylamide-gel electrophoresis. The interaction of C1 inhibitor with dimeric C1r in the presence of EDTA resulted into two bimolecular complexes accounting for a disruption of C1r	Chesne et al 1982 (6282262)
Alpha crystallin A chain	Beta-crystallin B2	Two-hybrid	Mammalian two-hybrid study	Fu et al 2002 (11700327)
Beta-crystallin B2	Alpha crystallin A chain	Two-hybrid	Mammalian two-hybrid study	Fu et al 2002 (11700327)
Beta-crystallin B2	Alpha crystallin A chain	Affinity capture ï Western	Complex formation on Western blot	Fu et al 2002 (11700327)
Alpha crystallin A chain	Beta-crystallin B2	Two-hybrid	Mammalian two-hybrid study	Fu et al 2003 (12601044)

BIOGRID database: https://thebiogrid.org/

Supplementary Table 3.8: Functional enrichment analysis of differentially expressed proteins (following false discovery rate correction) between CHR-T and CHR-NT in EU-GEI initial experiment: 6 KEGG pathways significantly enriched

KEGG pathway	Count in gene set	Fisher's *p* exact (corrected for false discovery rate)
Complement and coagulation cascades	13 of 78	2.23E-21
Staphylococcus aureus infection	6 of 51	5.29E-09
Pertussis	4 of 74	6.38E-05
Cholesterol metabolism	3 of 48	0.00047
Systemic lupus erythematosus	3 of 94	0.0025
Prion diseases	2 of 33	0.0058

CHR-T: clinical high-risk participants who transitioned to first episode psychosis; CHR-NT: clinical high-risk participants who did not transition; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; KEGG: Kyoto Encyclopedia of Genes and Genomes https://www.genome.jp/kegg/

Supplementary Table 3.9: Results of enzyme-linked immunosorbent assays in CHR-T and CHR-NT participants in EU-GEI initial experiment

	CHR-T mean (SD)	CHR-NT mean (SD)	t	*p*	Corrected *p* (5% FDR)
Alpha-2-macroglobulin (εg/ml)	1173.1 (459.1)	11501.7 (711.1)	3.2202	0.002	0.018
Apolipoprotein E (ng/ml)[a]	163751.3 (47433.8)	151740.6 (50903.5)	-1.3449	0.182	0.387
Complement C1q (ng/ml)	82811.7 (35347.1)	80204.6 (33535.8)	-0.4153	0.679	0.724
Complement C1r (εg/	65008.9 (27901.6)	52803.9 (18481.6)	-2.7099	0.008	0.036
Complement C4 binding protein (ng/ml)	495765.9 (222274.7)	482208.2 (192019.3)	-0.3538	0.724	0.724
Complement C8 (ng/ml)	58233.5 (22885.2)	55706.0 (21938.2)	-0.6196	0.537	0.690
Complement factor H (ng/ml)[b]	701713.1 (207717.4)	663292.1 (199397.1)	-1.0112	0.315	0.473
Immunoglobulin M (ng/ml)	1752941.0 (770671.5)	1941142.0 (936493.0)	1.2460	0.215	0.387
Plasminogen (ng/ml)	206880.2 (73232.4)	176929.9 (62709.8)	-2.3786	0.020	0.060

Data available for 48 CHR-T and 84 CHR-NT, except for [a]Apolipoprotein E (46 CHR-T, 84 CHR-NT) and [b]Complement factor H (45 CHR-T, 82 CHR-NT).

Means (and standard deviations) are presented and are compared using 2-sided t-test with unequal variances.

CHR-T: clinical high-risk participants who transitioned to first episode psychosis; CHR-NT: clinical high-risk participants who did not transition; SD: standard deviation; FDR: false discovery rate; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions.

Supplementary Table 3.10: Correlations between 5 proteins assessed by ELISA and by mass spectrometry in EU-GEI initial experiment

Protein	Spearman ρ	*p*	Corrected *p* (5% FDR)
Alpha-2-macroglobulin	0.25	0.0049	0.0080
Plasminogen	0.47	<0.0001	<0.0003
Complement component 1r	0.11	0.2089	0.2611
Complement factor H	0.07	0.4160	0.4160
Apolipoprotein E	0.41	<0.0001	<0.0003

ELISA: enzyme immunosorbent immunoassay; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; FDR: false discovery rate

Supplementary Table 3.11: Comparison of characteristics for participants included in replication experiment (N=135) from total EU-GEI clinical high-risk cohort (N=344)

	Missing data, n (%)	Included, n=135 (49 CHR-T, 86 CHR-NT)	Not included, n=209 (CHR-T 16, CHR-NT 193)	t / ^{2} G	*p*	Corrected *p* (5% FDR)
Baseline age in years, mean (SD)	0	22.2 (4.8)	22.5 (5.0)	-0.570	0.569	0.751
Sex, n (%)	0	82 male (60.7%) 53 female (39.3%)	103 male (49.3%) 106 female (50.7%)	4.322	0.037	0.320
Baseline BMI in kg/m^2, mean (SD)	50 (14.5%)	23.7 (4.1)	24.2 (5.8)	-0.936	0.350	0.700
Baseline years in education, mean (SD)	38 (11.0%)	14.3 (2.9)	14.5 (3.2)	-0.574	0.566	0.751
Ethnicity, n (%)	0	93 white (68.9%) 14 black (10.4%) 28 other (20.7%)	154 white (73.7%) 20 black (9.6%) 35 other (16.7%)	1.030	0.597	0.751
Ever used cannabis, n (%)	10 (2.9%)	101 yes (74.8%) 31 no (23.0%) 3 not known (2.2%)	143 yes (68.4%) 59 no (28.2%) 7 not known (3.3%)	1.328	0.249	0.619
Baseline cannabis use, n (%)	95 (27.6%)	43 yes (31.9%) 60 no (44.4%) 32 not known (23.7%)	45 yes (21.5%) 101 no (48.3%) 63 not known (30.1%)	3.155	0.076	0.320

Baseline tobacco use, n (%)	38 (11.0%)	72 yes (53.3%) 51 no (37.8%) 12 not known (8.9%)	89 yes (42.6%) 94 no (45.0%) 26 not known (12.4%)	2.893	0.089	0.320
Baseline alcohol use, n (%)	12 (3.5%)	36 yes (26.7%) 97 no (71.9%) 2 not known (1.5%)	137 yes (65.6%) 62 no (29.7%) 10 not known (4.8%)	0.640	0.424	0.751
Baseline medication use, n (%)	87 (25.3%)	51 yes (37.8%) Antidepressant 34 Antipsychotic 15 Hypnotic 4 Other 12 50 no (37.0%) 34 not known (25.2%)	77 yes (36.8%) Antidepressant 57 Antipsychotic 16 Hypnotic 19 Other 24 79 no (37.8%) 53 not known (25.4%)	0.032	0.859	0.960
Baseline GAF symptoms score, mean (SD)	27 (7.8%)	54.4 (10.2)	55.6 (10.0)	-1.103	0.271	0.619
Baseline GAF disability score, mean (SD)	12 (3.5%)	55.5 (13.7)	55.4 (11.3)	0.006	0.996	0.996
Baseline SANS total composite score, mean (SD)	45 (13.1%)	17.1 (12.7)	14.6 (10.8)	1.784	0.075	0.320
Baseline SANS total global score, mean (SD)	29 (8.4%)	5.6 (3.8)	5.4 (3.5)	0.510	0.610	0.751

Baseline BPRS total score, mean (SD)	25 (7.3%)	44.9 (11.2)	42.9 (9.6)	1.650	0.100	0.320
Baseline MADRS total score, mean (SD)	16 (4.7%)	18.8 (9.5)	18.9 (8.9)	-0.126	0.900	0.960

[a] Poor functioning: GAF disability score Ò60; good func

Tobacco use was defined as daily use for at least 1 month over the previous 12 months. Alcohol use was defined as at least 12 or more alcoholic beverages over the previous 12 months. Missing data excluded in hypothesis tests.

EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; CHR-T: clinical high risk, transitioned to psychosis; CHR-NT: clinical high risk, did not transition to psychosis; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; CHR-T: clinical high risk, transitioned to psychosis; CHR-NT: clinical high risk, did not transition to psychosis; FDR: false discovery rate; BMI: body mass index; GAF: General Assessment of Functioning; SANS: Scale for the Assessment of Negative Symptoms; BPRS: Brief Psychiatric Rating Scale; MADRS: Montgomery Asberg Depression Rating Scale; BMI: body mass index; GAF: General Assessment of Functioning; SANS: Scale for the Assessment of Negative Symptoms; BPRS: Brief Psychiatric Rating Scale; MADRS: Montgomery Asberg Depression Rating Scale

Supplementary Table 3.12: Sample characteristics for CHR-T and CHR-NT groups in the replication experiment

	Missing data, n (%)	CHR-T N = 49	CHR-NT N = 86	t / ^{2} G	*p*	Corrected *p* (5% FDR)
Baseline age in years, mean (SD)	0	22.0 (4.7)	22.3 (4.9)	t = -0.339	0.735	0.776
Sex, n (%)	0	26 male (53%) 23 female (47%)	56 male (65%) 30 female (35%)	G^2 = 1.902	0.168	0.290
Baseline BMI in kg/m², mean (SD)	21 (15.6%)	24.5 (4.5)	23.2 (3.8)	t = 1.722	0.088	0.186
Baseline years in education, mean (SD)	12 (8.8%)	14.0 (3.1)	14.3 (2.6)	t = -0.573	0.568	0.762
Ethnicity, n (%)	0	31 white (63.3%) 10 black (20.4%) 8 other (16.3%)	62 white (72.1%) 4 black (4.7%) 20 other (23.3%)	G^2 = 8.549	0.014	0.057
Ever used cannabis, n (%)	3 (2.2%)	35 yes (71.4%) 12 no (24.5%) 2 not known (4.1%)	66 yes (76.7%) 19 no (22.1%) 1 not known (1.2%)	G^2 = 0.170	0.680	0.762
Baseline cannabis use, n (%)	32 (23.7%)	14 yes (28.6%) 22 no (44.9%) 13 not known (26.5%)	29 yes (33.7%) 38 no (44.2%) 19 not known (22.1%)	G^2 = 0.186	0.666	0.762
Baseline tobacco use, n (%)	12 (8.9%)	19 yes (38.8%) 23 no (46.9%) 7 not known (14.3%)	53 yes (61.6%) 28 no (32.6%) 5 not known (5.8%)	G^2 = 4.647	0.031	0.074

Baseline alcohol use, n (%)	2 (1.5%)	34 yes (69.4%) 14 no (28.6%) 1 not known (2.0%)	63 yes (73.3%) 22 no (25.6%) 1 not known (1.2%)	$G^2 = 0.168$	0.682	0.762
Baseline medication use, n (%)	34 (25.2%)	20 yes (40.8%) Antidepressant 13 Antipsychotic 10 Hypnotic 2 Other 4 20 no (40.8%) 9 not known (18.4%)	31 yes (36.0%) Antidepressant 21 Antipsychotic 5 Hypnotic 2 Other 8 30 no (34.9%) 25 not known (29.1%)	$G^2 = 0.006$	0.936	0.936
Baseline GAF symptoms score, mean (SD)	4 (3.0%)	53.0 (10.1)	55.1 (10.2)	t = -1.089	0.278	0.440
Baseline GAF disability score, mean (SD)	4 (3.0%)	53.0 (12.6)	56.8 (14.1)	t = -1.531	0.128	0.243
Baseline SANS total composite score, mean (SD)	15 (11.1%)	20.9 (14.1)	14.9 (11.4)	t = 2.389	0.019	0.060
Baseline SANS total global score, mean (SD)	10 (7.4%)	6.6 (4.1)	5.0 (3.5)	t = 2.252	0.026	0.071
Baseline BPRS total score, mean (SD)	10 (7.4%)	48.1 (11.2)	43.1 (10.8)	t = 2.456	0.015	0.057

Baseline MADRS total score, mean (SD)	6 (4.4%)	19.9 (10.2)	18.1 (9.1)	t = 1.004	0.317	0.463
2-year GAF symptoms score, mean (SD) [a]	49 (36.3%)	43.6 (14.1)	63.5 (10.6)	t = -7.281	<0.001	<0.007
2-year GAF disability score, mean (SD) [b]	44 (32.6%)	45.3 (9.5)	65.1 (13.9)	t = -7.969	<0.001	<0.007
2-year GAF disability score, dichotomous outcome [c]	44 (32.6%)	28 poor functioning (57.1%) 2 good functioning (4.1%) 19 not known (38.8%)	23 poor functioning (26.7%) 38 good functioning (44.2%) 25 not known (29.1%)	G^2 = 25.261	<0.001	<0.007

[a] Data available for 86 of 135 participants (CHR-NT n=27, CHR-T n=59)

[b] Data available for 91 of 135 participants (CHR-NT n=30, CHR-T n=61)

[c] Poor functioning: GAF disability score Ò60; good func

Tobacco use was defined as daily use for at least 1 month over the previous 12 months.
Alcohol use was defined as at least 12 or more alcoholic beverages over the previous 12 months.

Missing data excluded in hypothesis tests.

EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions; CHR-T: clinical high risk, transitioned to psychosis; CHR-NT: clinical high risk, did not transition to psychosis; FDR: false discovery rate; BMI: body mass index; GAF: General Assessment of Functioning; SANS: Scale for the Assessment of Negative Symptoms; BPRS: Brief Psychiatric Rating Scale; MADRS: Montgomery Asberg Depression Rating Scale

Supplementary Table 3.13: Results of ANCOVA (adjusted for age, sex, BMI, years in education, tobacco use and ethnicity) and fold changes (CHR-T vs. CHR-NT) for proteins identified in EU-GEI baseline plasma samples in the replication experiment

Uniprot no.	Protein name	F	*p*	Corrected *p* (5% FDR)	Direction of effect	Ratio of means (T vs. NT)
P01023	Alpha-2-macroglobulin	264.41	9.81E-33	1.17E-30	↓	0.24
P01871	Immunoglobulin heavy constant mu	109.92	7.11E-19	4.23E-17	↓	0.33
P00747	Plasminogen	78.47	6.43E-15	2.55E-13	↑	1.30
P10909	Clusterin	73.41	3.17E-14	9.44E-13	↑	1.36
P01011	Alpha-1-antichymotrypsin	68.88	1.38E-13	3.28E-12	↑	1.38
G3XAM2	Complement factor I	65.48	4.24E-13	8.4E-12	↑	1.27
P0DOY3	Immunoglobulin lambda constant 3	58.29	4.9E-12	8.33E-11	↓	0.56
P01860	Immunoglobulin heavy constant gamma 3	56.09	1.06E-11	1.42E-10	↓	0.52
P08603	Complement factor H (H factor 1)	56.05	1.08E-11	1.42E-10	↑	1.18
P13671	Complement component C6	55.01	1.55E-11	1.7E-10	↑	1.39
P04004	Vitronectin	54.97	1.57E-11	1.7E-10	↑	1.33
P22792	Carboxypeptidase N subunit 2	52.23	4.18E-11	4.15E-10	↑	1.29
P07360	Complement component C8 gamma chain	49.11	1.3E-10	1.19E-09	↑	1.28
P08571	Monocyte differentiation antigen CD14	47.31	2.52E-10	2.14E-09	↑	1.39
P01031	Complement C5	44.69	6.71E-10	5.32E-09	↑	1.19
B7ZKJ8	ITIH4 protein	43.28	1.14E-09	8.51E-09	↑	1.21

P07359	Platelet glycoprotein Ib alpha chain	35.10	2.79E-08	1.96E-07	↑	1.40
P15169	Carboxypeptidase N catalytic chain (CPN)	33.90	4.53E-08	2.99E-07	↑	1.33
P36955	Pigment epithelium-derived factor	32.23	8.97E-08	5.62E-07	↑	1.22
P07357	Complement component C8 alpha chain	31.63	1.15E-07	6.82E-07	↑	1.25
P27169	Serum paraoxonase/arylesterase 1	30.66	1.71E-07	9.69E-07	↑	1.30
Q96IY4	Carboxypeptidase B2	29.85	2.39E-07	1.29E-06	↑	1.41
Q16610	Extracellular matrix protein 1	26.68	9.12E-07	4.72E-06	↑	1.38
P04003	C4b-binding protein alpha chain	26.26	1.09E-06	5.41E-06	↓	0.77
P12259	Coagulation factor V	24.43	2.4E-06	1.14E-05	↑	1.30
P06727	Apolipoprotein A-IV	23.96	2.95E-06	1.35E-05	↑	1.16
P80108	Phosphatidylinositol-glycan-specific phospholipase D (PI-G PLD)	22.81	4.87E-06	2.15E-05	↑	1.27
P05155	Plasma protease C1 inhibitor	22.67	5.19E-06	2.00E-04	↑	1.23
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	20.64	1.28E-05	0.0001	↑	1.15
P0C0L4	Complement C4-A (Acidic complement C4)	19.53	2.11E-05	0.0001	↓	0.57
B4E1Z4	cDNA FLJ55673	18.98	2.71E-05	0.0001	↑	1.13
P51884	Lumican	18.90	2.81E-05	0.0001	↑	1.26
P00736	Complement C1r subcomponent	18.64	3.17E-05	0.0001	↑	1.13

O43866	CD5 antigen-like	18.47	3.42E-05	0.0001	↓	0.65
P02765	Alpha-2-HS-glycoprotein	17.11	0.0001	0.0003	↑	1.16
P06396	Gelsolin	16.15	0.0001	0.0003	↑	1.15
Q04756	Hepatocyte growth factor activator	16.22	0.0001	0.0003	↑	1.19
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	15.64	0.0001	0.0003	↑	1.11
O00391	Sulfhydryl oxidase 1	14.82	0.0002	0.0006	↑	1.18
P02748	Complement component C9	13.25	0.0004	0.0012	↑	1.19
P00738	Haptoglobin	11.99	0.0007	0.0020	↓	0.49
P02790	Hemopexin	11.97	0.0007	0.0020	↑	1.10
P08697	Alpha-2-antiplasmin	11.36	0.0010	0.0027	↑	1.10
P23142	Fibulin-1	11.37	0.0010	0.0027	↑	1.37
P05452	Tetranectin	11.15	0.0011	0.0029	↑	1.15
P06681	Complement C2	10.96	0.0012	0.0031	↑	1.18
P02751	Fibronectin	10.62	0.0014	0.0035	↑	1.39
P02649	Apolipoprotein E	10.26	0.0017	0.0041	↑	1.21
P07358	Complement component C8 beta chain	10.24	0.0017	0.0041	↑	1.18
O75882	Attractin	10.07	0.0019	0.0044	↑	1.19
P01834	Immunoglobulin kappa constant	10.09	0.0019	0.0044	↓	0.74
P00734	Prothrombin	9.98	0.0020	0.0046	↑	1.09
P17936	Insulin-like growth factor-binding protein 3	9.46	0.0026	0.0058	↑	1.30
P00746	Complement factor D	9.15	0.0030	0.0066	↑	1.23
P02675	Fibrinogen beta chain	8.48	0.0043	0.0093	↓	0.87

P00488	Coagulation factor XIII A chain	8.32	0.0046	0.0098	↑	1.26
P43652	Afamin	8.04	0.0053	0.0111	↑	1.10
P35858	Insulin-like growth factor-binding protein complex acid labile subunit (ALS)	7.33	0.0077	0.0158	↑	1.11
P05543	T4-binding globulin	7.30	0.0079	0.0159	↑	1.27
P10643	Complement component C7	7.09	0.0088	0.0175	↑	1.29
P02787	Serotransferrin	6.92	0.0096	0.0187	↓	0.83
P26927	Hepatocyte growth factor-like protein (Macrophage stimulatory protein)	6.80	0.0102	0.0193	↑	1.24
P18428	Lipopolysaccharide-binding protein (LBP)	6.81	0.0102	0.0193	↑	1.27
P02760	Protein AMBP	6.63	0.0112	0.0208	↑	1.05
P00450	Ceruloplasmin	6.38	0.0128	0.0234	↓	0.93
P05154	Plasma serine protease inhibitor	6.25	0.0137	0.0247	↑	1.15
P02766	Transthyretin	6.14	0.0145	0.0258	↓	0.82
Q14520-2	Hyaluronan-binding protein 2	6.09	0.0149	0.0260	↑	1.20
P02679-2	Fibrinogen gamma chain	6.07	0.0151	0.0260	↓	0.90
P00740	Coagulation factor IX	5.92	0.0164	0.0279	↑	1.10
Q96KN2	Beta-Ala-His dipeptidase	5.82	0.0173	0.0290	↑	1.27
P01019	Angiotensinogen	5.42	0.0215	0.0355	↑	1.12
P04114	Apolipoprotein B-100	5.29	0.0231	0.0377	↑	1.07
P02647	Apolipoprotein A-I	5.03	0.0266	0.0428	↓	0.89
P05546	Heparin cofactor 2	4.81	0.0302	0.0479	↑	1.06

P01009	Alpha-1-antitrypsin	4.77	0.0307	0.0481	↓	0.79
P04275	von Willebrand factor	4.71	0.0319	0.0493	↑	1.37
P02743	Serum amyloid P-component (SAP)	4.68	0.0324	0.0494	↑	1.22
P01857	Immunoglobulin heavy constant gamma 1	4.63	0.0334	0.0503	↑	1.14
P01008	Antithrombin-III	4.27	0.0408	0.0607	↓	0.91
P25311	Zinc-alpha-2-glycoprotein	4.08	0.0456	0.0670	↑	1.01
P05160	Coagulation factor XIII B chain	3.40	0.0675	0.0980	↑	1.06
P22891	Vitamin K-dependent protein Z	3.13	0.0795	0.1140	↑	1.10
P02746	Complement C1q subcomponent subunit B	2.85	0.0936	0.1326	↑	1.17
P0C0L5	Complement C4-B (Basic complement C4)	2.72	0.1014	0.1420	↑	1.09
Q96PD5	N-acetylmuramoyl-L-alanine amidase	2.63	0.1075	0.1488	↑	1.05
P01876	Immunoglobulin heavy constant alpha 1	2.54	0.1135	0.1546	↓	0.72
P02763	Alpha-1-acid glycoprotein 1	2.51	0.1154	0.1546	↓	0.84
P01042	Kininogen-1	2.51	0.1156	0.1546	↑	1.03
K7ERI9	Apolipoprotein C-I	2.17	0.1433	0.1895	↓	0.84
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	2.07	0.1524	0.1993	↑	1.13
P02750	Leucine-rich alpha-2-glycoprotein	2.01	0.1589	0.2055	↑	1.08
D6RF35	Vitamin D-binding protein	1.89	0.1721	0.2202	↑	1.02
P00748	Coagulation factor XII	1.85	0.1762	0.2231	↓	0.93

P09871	Complement C1s subcomponent	1.70	0.1951	0.2444	↓	0.92
P07996	Thrombospondin-1	1.11	0.2932	0.3634	↑	1.10
Q9NZP8	Complement C1r subcomponent-like protein	1.06	0.3051	0.3743	↑	1.12
P02747	Complement C1q subcomponent subunit C	1.02	0.3148	0.3823	↑	1.05
P04278	Sex hormone-binding globulin	0.64	0.4244	0.5072	↑	1.17
P29622	Kallistatin (Kallikrein inhibitor)	0.64	0.4262	0.5072	↑	1.02
P03952	Plasma kallikrein	0.58	0.4476	0.5274	↑	1.08
P43251	Biotinidase (Biotinase)	0.55	0.4607	0.5375	↓	0.92
P01024	Complement C3	0.45	0.5025	0.5775	↑	1.01
P04196	Histidine-rich glycoprotein	0.45	0.5047	0.5775	↑	1.02
P02656	Apolipoprotein C-III	0.43	0.5145	0.5831	↑	1.09
O75636	Ficolin-3	0.36	0.5475	0.6146	↑	1.10
P27918	Properdin	0.30	0.5819	0.6472	↑	1.05
P00742	Coagulation factor X	0.27	0.6070	0.6688	↑	1.06
P02749	Beta-2-glycoprotein 1	0.23	0.6357	0.6940	↑	1.01
P06276	Acylcholine acylhydrolase	0.20	0.6588	0.7127	↑	1.01
P08185	Corticosteroid-binding globulin	0.13	0.7170	0.7687	–	1.00
P02671	Fibrinogen alpha chain	0.08	0.7845	0.8335	↓	0.97
P03951	Coagulation factor XI	0.05	0.8228	0.8628	↑	1.12
P19652	Alpha-1-acid glycoprotein 2 (AGP 2)	0.05	0.8265	0.8628	↓	0.97
Q9UK55	Protein Z-dependent protease inhibitor	0.04	0.8393	0.8685	↑	1.01

P07225	Vitamin K-dependent protein S	0.03	0.8521	0.8741	–	1.00
Q5VY30	Retinol-binding protein	0.02	0.8787	0.8937	↓	0.97
P68871	Hemoglobin subunit beta	0.02	0.9022	0.9079	↑	1.06
P22352	Glutathione peroxidase 3 (GPx-3)	0.01	0.9079	0.9079	↑	1.03

CHR-T: clinical high-risk participants who transitioned to first episode psychosis; CHR-NT: clinical high-risk participants who did not transition; FDR: false discovery rate; EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions

Supplementary Table 3.14: Sample characteristics for ALSPAC subsample cases and controls

	Cases **n=55**	**Controls** **n=66**	**t /^{2} G**	***p***
Sex, n (%)	22 male (40.0%) 33 female (60.0%)	39 male (59.1%) 27 female (40.9%)	4.374	0.036
BMI age 12 in kg/m^2, mean (SD)	18.1 (2.8)	17.7 (2.5)	0.749	0.455
Maternal social class, n (%)	40 non-manual (72.7%) 7 manual (12.7%) 8 not known (14.6%)	44 non-manual (66.7%) 8 manual (12.1%) 14 not known (21.2%)	0.005	0.946

Cases: participants with no PEs age 12 and definite PEs age 18; Controls: participants with no PEs age 12 and no PEs age 18

PEs: psychotic experiences; BMI: body mass index
Missing data excluded in hypothesis tests.

Data for ethnicity are not presented in the table because of potential risk of statistical disclosure due to small cell counts. There were no significant differences bet χ^2=0.729, p= 0.202).

ALSPAC: Avon Longitudinal Study of Parents and Children

Supplementary Table 3.15: Results of ANCOVA (adjusted for sex, BMI and maternal social class) and fold changes (definite PEs at 18 vs. no PEs at 18) for proteins identified in ALSPAC age 12 plasma samples

Uniprot No.	Protein name	F	*p*	Corrected *p* (5%FDR)	Direction of effect	Ratio of means (PE vs. no PE)
P04003	C4b-binding protein alpha chain	24.59	2.44E-06	0.0006	↓	0.77
P27169	Serum paraoxonase/arylesterase 1	17.78	4.94E-05	0.0065	↓	0.80
P01871	Immunoglobulin heavy constant mu	16.04	0.0001	0.0066	↓	0.78
P55103	Inhibin beta C chain	16.68	0.0001	0.0066	↑	1.31
P10909	Clusterin	12.71	0.0005	0.0265	↓	0.92
P01591	Immunoglobulin J chain	9.69	0.0023	0.1007	↓	0.70
P01860	Immunoglobulin heavy constant gamma 3	9.25	0.0029	0.1007	↓	0.80
P0DOY3	Immunoglobulin lambda constant 3	8.90	0.0035	0.1007	↓	0.81
P07225	Vitamin K-dependent protein S	8.83	0.0036	0.1007	↓	0.90
Q03591	Complement factor H-related protein 1	8.74	0.0038	0.1007	↓	0.79
P01023	Alpha-2-macroglobulin	7.92	0.0057	0.1373	↓	0.85
P01623	Immunoglobulin kappa variable 3-20	7.49	0.0072	0.1590	↓	0.81
P24593	Insulin-like growth factor-binding protein 5	7.34	0.0078	0.1590	↑	1.26
P01019	Angiotensinogen	6.74	0.0106	0.1943	↓	0.91
P26038	Moesin	6.68	0.0110	0.1943	↑	1.16
P04040	Catalase	6.53	0.0119	0.1971	↑	1.36

P12109	Collagen alpha-1 (VI) chain	6.36	0.0130	0.1988	↑	1.16
P08571	Monocyte differentiation antigen CD14	6.29	0.0135	0.1988	↓	0.91
B9A064	Immunoglobulin lambda-like polypeptide 5	6.03	0.0155	0.2107	↓	0.85
P09871	Complement C1s subcomponent	5.99	0.0159	0.2107	↓	0.93
P08697	Alpha-2-antiplasmin	5.50	0.0207	0.2495	↓	0.94
P15151	Poliovirus receptor	5.44	0.0214	0.2495	↑	1.16
P01717	Immunoglobulin lambda variable 3-25	5.34	0.0226	0.2495	↓	0.85
Q12884	Prolyl endopeptidase FAP	5.34	0.0226	0.2495	↑	1.24
P00746	Complement factor D	5.16	0.0249	0.2554	↑	1.12
P01615	Immunoglobulin kappa variable 2D-28	5.10	0.0258	0.2554	↓	0.84
P02671	Fibrinogen alpha chain	5.05	0.0266	0.2554	↓	0.89
P23142	Fibulin-1	4.99	0.0275	0.2554	↑	1.36
P01834	Immunoglobulin kappa constant	4.95	0.0280	0.2554	↓	0.90
P24592	Insulin-like growth factor-binding protein 6	4.85	0.0297	0.2554	↑	1.15
P02679	Fibrinogen gamma chain	4.77	0.0310	0.2554	↓	0.89
P02675	Fibrinogen beta chain	4.76	0.0311	0.2554	↓	0.90
P03951	Coagulation factor XI	4.72	0.0318	0.2554	↓	0.91
P80748	Immunoglobulin lambda variable 3-21	4.63	0.0334	0.2603	↓	0.85
P55290	Cadherin-13	4.43	0.0374	0.2832	↓	0.88
P02652	Apolipoprotein A-II	4.37	0.0388	0.2856	↓	0.85

P07358	Complement component C8 beta chain	4.12	0.0446	0.3061	↑	1.08
D6RAR4	Hepatocyte growth factor activator	4.11	0.0450	0.3061	↓	0.89
P15144	Aminopeptidase N	4.07	0.0460	0.3061	↑	1.06
Q99878	Histone H2A type 1-J	4.06	0.0462	0.3061	↑	1.37
P08294	Extracellular superoxide dismutase [Cu-Zn]	3.91	0.0505	0.3264	↑	1.13
P02750	Leucine-rich alpha-2-glycoprotein	3.83	0.0528	0.3331	↓	0.92
P61626	Lysozyme C	3.61	0.0600	0.3686	↓	0.91
P12111	Collagen alpha-3	3.55	0.0621	0.3686	↑	1.09
P00915	Carbonic anhydrase 1	3.53	0.0626	0.3686	↑	1.24
P14151	L-selectin	3.35	0.0700	0.4033	↑	1.08
O43866	CD5 antigen-like	3.26	0.0737	0.4155	↓	0.86
P02654	Apolipoprotein C-I	3.14	0.0791	0.4220	↓	0.91
P01024	Complement C3	3.09	0.0814	0.4220	↓	0.94
P29622	Kallistatin	3.08	0.0820	0.4220	↑	1.11
P02647	Apolipoprotein A-I	3.07	0.0823	0.4220	↓	0.86
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	3.06	0.0828	0.4220	↓	0.91
O00187	Mannan-binding lectin serine protease 2	2.92	0.0901	0.4404	↑	1.07
P16403	Histone H1.2	2.91	0.0907	0.4404	↑	1.19
P06276	Cholinesterase	2.90	0.0914	0.4404	↑	1.03
Q9NQ79	Cartilage acidic protein 1	2.84	0.0946	0.4435	↑	1.08
Q9H4A9	Dipeptidase 2	2.83	0.0954	0.4435	↓	0.84

H0Y755	Low affinity immunoglobulin gamma Fc region receptor III-A	2.74	0.1008	0.4606	↑	1.11
Q9UBQ6	Exostosin-like 2	2.50	0.1164	0.5142	↑	1.08
Q9UGM5	Fetuin-B	2.48	0.1178	0.5142	↓	0.94
O75636	Ficolin-3	2.47	0.1191	0.5142	↑	1.07
O95497	Pantetheinase	2.45	0.1203	0.5142	↑	1.15
P43251	Biotinidase	2.38	0.1254	0.5220	↑	1.06
Q08380	Galectin-3-binding protein	2.36	0.1276	0.5220	↑	1.08
P63261	Actin, cytoplasmic 2	2.34	0.1289	0.5220	↑	1.10
Q9Y5Y7	Lymphatic vessel endothelial hyaluronic acid receptor 1	2.33	0.1300	0.5220	↑	1.09
P00747	Plasminogen	2.26	0.1358	0.5361	↑	1.07
P22792	Carboxypeptidase N subunit 2	2.21	0.1395	0.5361	↓	0.95
P01033	Metalloproteinase inhibitor 1	2.21	0.1396	0.5361	↓	0.87
Q9ULI3	Protein HEG homolog 1	2.11	0.1488	0.5633	↑	1.13
Q96PD5	N-acetylmuramoyl-L-alanine amidase	2.06	0.1539	0.5638	↑	1.09
P05019	Insulin-like growth factor I	2.06	0.1540	0.5638	↑	1.24
P35858	Insulin-like growth factor-binding protein complex acid labile subunit	2.05	0.1553	0.5638	↑	1.10
P02753	Retinol-binding protein 4	2.01	0.1592	0.5701	↑	1.13
P02452	Collagen alpha-1 (I) chain	1.90	0.1706	0.5923	↑	1.18
P39060	Collagen alpha-1 (XVIII) chain	1.90	0.1707	0.5923	↑	1.06
P02747	Complement C1q subcomponent subunit C	1.89	0.1721	0.5923	↓	0.96

P04004	Vitronectin	1.78	0.1853	0.6295	↑	1.06
P07942	Laminin subunit beta-1	1.75	0.1887	0.6330	↑	1.01
P02745	Complement C1q subcomponent subunit A	1.70	0.1944	0.6440	↓	0.93
P07359	Platelet glycoprotein Ib alpha chain	1.65	0.2010	0.6535	↑	1.03
P07737	Profilin-1	1.65	0.2022	0.6535	↑	1.17
P00751	Complement factor B	1.58	0.2116	0.6756	↓	0.97
H0YD13	CD44 antigen	1.53	0.2185	0.6818	↑	1.09
P48740	Mannan-binding lectin serine protease 1	1.53	0.2192	0.6818	↓	0.91
P07357	Complement component C8 alpha chain	1.50	0.2227	0.6818	↑	1.05
P08603	Complement factor H	1.48	0.2263	0.6818	↑	1.04
Q16706	Alpha-mannosidase 2	1.48	0.2264	0.6818	↓	0.92
P20851	C4b-binding protein beta chain	1.46	0.2290	0.6819	↓	0.89
Q12913	Receptor-type tyrosine-protein phosphatase eta	1.43	0.2341	0.6870	↓	0.93
Q15582	Transforming growth factor-beta-induced protein ig-h3	1.42	0.2359	0.6870	↑	1.04
Q6UXB8	Peptidase inhibitor 16	1.39	0.2407	0.6933	↑	1.09
P69905	Hemoglobin subunit alpha	1.31	0.2553	0.7051	↑	3.63
F5GZZ9	Scavenger receptor cysteine-rich type 1 protein M130	1.29	0.2586	0.7051	↑	1.20
P13591	Neural cell adhesion molecule 1	1.27	0.2619	0.7051	↑	1.03
Q15113	Procollagen C-endopeptidase enhancer 1	1.26	0.2633	0.7051	↑	1.07

P26927	Hepatocyte growth factor-like protein	1.24	0.2671	0.7051	↑	1.03
Q12841	Follistatin-related protein 1	1.24	0.2675	0.7051	↑	1.04
Q15848	Adiponectin	1.24	0.2687	0.7051	↓	0.95
P27487	Dipeptidyl peptidase 4	1.23	0.2706	0.7051	↑	1.06
Q86U17	Serpin A11	1.22	0.2711	0.7051	↑	1.11
Q9UNW1	Multiple inositol polyphosphate phosphatase 1	1.22	0.2714	0.7051	↓	0.93
P43652	Afamin	1.20	0.2758	0.7096	↑	1.07
P51884	Lumican	1.18	0.2798	0.7130	↑	1.04
P13671	Complement component C6	1.15	0.2868	0.7195	↓	0.99
P04278	Sex hormone-binding globulin	1.14	0.2883	0.7195	↓	0.89
P03950	Angiogenin	1.13	0.2905	0.7195	↓	0.96
E9PBC5	Plasma kallikrein	1.10	0.2973	0.7295	↓	0.96
P04275	von Willebrand factor	1.07	0.3030	0.7309	↓	0.93
P02748	Complement component C9	1.07	0.3034	0.7309	↓	0.97
P08185	Corticosteroid-binding globulin	1.02	0.3139	0.7494	↑	1.02
P14625	Endoplasmin	1.00	0.3189	0.7545	↑	1.05
O75882	Attractin	0.97	0.3256	0.7636	↑	1.02
P68871	Hemoglobin subunit beta	0.96	0.3300	0.7671	↑	2.75
P11021	Endoplasmic reticulum chaperone BiP	0.94	0.3342	0.7701	↓	0.96
P02787	Serotransferrin	0.90	0.3459	0.7862	↓	0.97
P18206	Vinculin	0.89	0.3471	0.7862	↑	1.08
Q5T7F0	Neuropilin	0.87	0.3521	0.7907	↓	0.96

P07360	Complement component C8 gamma chain	0.83	0.3655	0.8139	↑	1.04
P55058	Phospholipid transfer protein	0.78	0.3783	0.8265	↑	1.05
P01765	Immunoglobulin heavy variable 3-23	0.77	0.3806	0.8265	↓	0.89
P13727	Bone marrow proteoglycan	0.74	0.3900	0.8265	↑	1.09
P12259	Coagulation factor V	0.73	0.3962	0.8265	↓	0.97
P00742	Coagulation factor X	0.72	0.3970	0.8265	↑	1.03
P18065	Insulin-like growth factor-binding protein 2	0.70	0.4043	0.8265	↓	0.92
P54289	Voltage-dependent calcium channel subunit alpha-2/delta-1	0.70	0.4060	0.8265	↑	1.03
P13473	Lysosome-associated membrane glycoprotein 2	0.70	0.4061	0.8265	↓	0.92
P32119	Peroxiredoxin-2	0.69	0.4084	0.8265	↑	1.02
P04196	Histidine-rich glycoprotein	0.69	0.4087	0.8265	↑	1.05
P02649	Apolipoprotein E	0.69	0.4092	0.8265	↓	0.96
P02749	Beta-2-glycoprotein 1	0.68	0.4107	0.8265	↓	0.98
P02655	Apolipoprotein C-II	0.68	0.4117	0.8265	↑	1.06
P00734	Prothrombin	0.65	0.4218	0.8404	↓	0.97
Q9UNN8	Endothelial protein C receptor	0.64	0.4267	0.8438	↑	1.10
P02743	Serum amyloid P-component	0.62	0.4319	0.8478	↓	0.98
P08709	Coagulation factor VII	0.59	0.4429	0.8591	↑	1.02
P02746	Complement C1q subcomponent subunit B	0.59	0.4448	0.8591	↑	1.05
P04180	Phosphatidylcholine-sterol acyltransferase	0.58	0.4474	0.8591	↓	0.95

P24821	Tenascin	0.55	0.4606	0.8756	↑	1.05
P01042	Kininogen-1	0.54	0.4626	0.8756	↓	0.98
P23470	Receptor-type tyrosine-protein phosphatase gamma	0.53	0.4680	0.8791	↓	1.00
P07195	L-lactate dehydrogenase B chain	0.52	0.4712	0.8791	↓	0.95
Q13822	Ectonucleotide pyrophosphatase/ phosphodiesterase family member 2	0.51	0.4744	0.8791	↑	1.02
P49908	Selenoprotein P	0.50	0.4810	0.8852	↓	1.00
P04114	Apolipoprotein B-100	0.49	0.4845	0.8855	↓	0.99
P02766	Transthyretin	0.47	0.4941	0.8909	↓	0.95
P27918	Properdin	0.47	0.4942	0.8909	↑	1.05
H0Y897	Target of Nesh-SH3	0.44	0.5101	0.9024	↑	1.06
P01344	Insulin-like growth factor II	0.44	0.5104	0.9024	↓	0.97
P01031	Complement C5	0.42	0.5163	0.9024	↑	1.01
P02776	Platelet factor 4	0.41	0.5214	0.9024	↓	0.94
P06727	Apolipoprotein A-IV	0.41	0.5222	0.9024	↑	1.01
P05155	Plasma protease C1 inhibitor	0.41	0.5226	0.9024	↓	0.96
P15169	Carboxypeptidase N catalytic chain	0.41	0.5251	0.9024	↓	0.97
Q9BXR6	Complement factor H-related protein 5	0.39	0.5319	0.9024	↓	1.00
P02656	Apolipoprotein C-III	0.39	0.5337	0.9024	↓	0.95
Q16270	Insulin-like growth factor-binding protein 7	0.39	0.5361	0.9024	↓	0.96
O95445	Apolipoprotein M	0.37	0.5426	0.9024	↓	0.98

P00488	Coagulation factor XIII A chain	0.37	0.5441	0.9024	↑	1.01
Q76LX8	A disintegrin and metalloproteinase with thrombospondin motifs 13	0.36	0.5492	0.9024	↓	0.97
Q8IUL8	Cartilage intermediate layer protein 2	0.35	0.5553	0.9024	↑	1.07
P61769	Beta-2-microglobulin	0.34	0.5628	0.9024	↑	1.05
Q99983	Osteomodulin	0.33	0.5661	0.9024	↑	1.04
P54802	Alpha-N-acetylglucosaminidase	0.32	0.5723	0.9024	↑	1.11
P11226	Mannose-binding protein C	0.32	0.5724	0.9024	↓	0.98
P05543	Thyroxine-binding globulin	0.32	0.5731	0.9024	↓	0.99
P10721	Mast/stem cell growth factor receptor Kit	0.32	0.5743	0.9024	↑	1.01
Q14520	Hyaluronan-binding protein 2	0.31	0.5759	0.9024	↓	0.98
P01861	Immunoglobulin heavy constant gamma 4	0.31	0.5783	0.9024	↓	0.82
P00748	Coagulation factor XII	0.31	0.5789	0.9024	↓	0.98
P09172	Dopamine beta-hydroxylase	0.29	0.5883	0.9081	↓	0.97
Q16610	Extracellular matrix protein 1	0.29	0.5894	0.9081	↓	0.98
P36980	Complement factor H-related protein 2	0.28	0.5952	0.9084	↓	0.91
Q9UHG3	Prenylcysteine oxidase 1	0.28	0.5982	0.9084	↑	1.02
P01008	Antithrombin-III	0.28	0.5999	0.9084	↓	0.98
P01009	Alpha-1-antitrypsin	0.24	0.6231	0.9357	↓	0.99
P33151	Cadherin-5	0.24	0.6250	0.9357	↓	0.96
P17936	Insulin-like growth factor-binding protein 3	0.23	0.6329	0.9416	↑	1.06

P01876	Immunoglobulin heavy constant alpha 1	0.23	0.6360	0.9416	↓	0.96
P22352	Glutathione peroxidase 3	0.22	0.6413	0.9439	↓	1.00
P01034	Cystatin-C	0.21	0.6447	0.9439	↑	1.03
P28827	Receptor-type tyrosine-protein phosphatase mu	0.20	0.6555	0.9504	↓	0.98
P36955	Pigment epithelium-derived factor	0.20	0.6563	0.9504	↓	1.00
P59666	Neutrophil defensin 3	0.19	0.6605	0.9513	↓	0.99
P18428	Lipopolysaccharide-binding protein	0.18	0.6698	0.9594	↑	1.05
P43121	Cell surface glycoprotein MUC18	0.17	0.6784	0.9665	↓	0.98
P19652	Alpha-1-acid glycoprotein 2	0.15	0.7034	0.9799	↑	1.13
P04075	Fructose-bisphosphate aldolase A	0.14	0.7057	0.9799	↓	0.97
Q6EMK4	Vasorin	0.14	0.7064	0.9799	↓	0.99
Q14515	SPARC-like protein 1	0.14	0.7085	0.9799	↑	1.05
P05109	Protein S100-A8	0.14	0.7099	0.9799	↑	1.01
P05546	Heparin cofactor 2	0.14	0.7100	0.9799	↑	1.03
P49747	Cartilage oligomeric matrix protein	0.13	0.7175	0.9812	↓	1.00
Q9NPY3	Complement component C1q receptor	0.13	0.7183	0.9812	↑	1.01
Q9Y4L1	Hypoxia up-regulated protein 1	0.12	0.7318	0.9927	↑	1.02
P02774	Vitamin D-binding protein	0.11	0.7460	0.9927	↓	0.97
P00736	Complement C1r subcomponent	0.10	0.7467	0.9927	↓	0.99

P19827	Inter-alpha-trypsin inhibitor heavy chain H1	0.10	0.7490	0.9927	↓	0.98
P06702	Protein S100-A9	0.10	0.7506	0.9927	↓	1.00
O00391	Sulfhydryl oxidase 1	0.10	0.7547	0.9927	↑	1.01
Q92820	Gamma-glutamyl hydrolase	0.09	0.7590	0.9927	↑	1.03
P80108	Phosphatidylinositol-glycan-specific phospholipase D	0.09	0.7649	0.9927	↑	1.01
Q9NPH3	Interleukin-1 receptor accessory protein	0.09	0.7649	0.9927	↑	1.06
P01011	Alpha-1-antichymotrypsin	0.08	0.7712	0.9927	↓	0.98
P35542	Serum amyloid A-4 protein	0.08	0.7743	0.9927	↑	1.03
P02763	Alpha-1-acid glycoprotein 1	0.08	0.7752	0.9927	↓	0.90
P20742	Pregnancy zone protein	0.08	0.7787	0.9927	↑	1.05
P02775	Platelet basic protein	0.08	0.7815	0.9927	↓	0.99
O95479	GDH/6PGL endoplasmic bifunctional protein	0.08	0.7835	0.9927	↓	0.99
P05362	Intercellular adhesion molecule 1	0.07	0.7875	0.9927	↓	0.95
O14791	Apolipoprotein L1	0.07	0.7904	0.9927	↑	1.06
P33908	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA	0.07	0.7947	0.9934	↓	0.98
P22891	Vitamin K-dependent protein Z	0.06	0.8010	0.9934	↓	0.91
Q13740	CD166 antigen	0.06	0.8022	0.9934	↓	0.97
P22105	Tenascin-X	0.05	0.8160	0.9976	↓	1.00
P05156	Complement factor I	0.05	0.8205	0.9976	↑	1.01
Q96KN2	Beta-Ala-His dipeptidase	0.05	0.8245	0.9976	↑	1.03
P10643	Complement component C7	0.04	0.8343	0.9976	↓	1.00

O00533	Neural cell adhesion molecule L1-like protein	0.04	0.8359	0.9976	↓	0.97
Q9BWP8	Collectin-11	0.04	0.8409	0.9976	↓	0.98
Q99969	Retinoic acid receptor responder protein 2	0.04	0.8425	0.9976	↓	0.99
P04070	Vitamin K-dependent protein C	0.04	0.8443	0.9976	↑	1.02
P04066	Tissue alpha-L-fucosidase	0.04	0.8513	0.9976	↓	0.94
P55056	Apolipoprotein C-IV	0.03	0.8585	0.9976	↑	1.12
P0C0L5	Complement C4-B	0.03	0.8646	0.9976	↓	0.99
P00450	Ceruloplasmin	0.03	0.8653	0.9976	↓	0.99
P02760	Protein AMBP	0.03	0.8721	0.9976	↓	1.00
P19320	Vascular cell adhesion protein 1	0.02	0.8840	0.9976	↓	0.98
P22692	Insulin-like growth factor-binding protein 4	0.02	0.8848	0.9976	↓	0.59
P02751	Fibronectin	0.02	0.8998	0.9976	↓	0.97
P08519	Apolipoprotein	0.01	0.9049	0.9976	↑	1.20
P12830	Cadherin-1	0.01	0.9146	0.9976	↓	0.94
Q96IY4	Carboxypeptidase B2	0.01	0.9190	0.9976	↑	1.01
P17813	Endoglin	0.01	0.9194	0.9976	↓	0.97
P35443	Thrombospondin-4	0.01	0.9285	0.9976	↑	1.11
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	0.01	0.9306	0.9976	↓	0.98
P00738	Haptoglobin	0.01	0.9311	0.9976	↓	0.98
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	0.01	0.9331	0.9976	↓	1.00
P02790	Hemopexin	0.01	0.9342	0.9976	↓	1.00

Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	0.01	0.9358	0.9976	↓	0.97
Q07954	Prolow-density lipoprotein receptor-related protein 1	0.01	0.9388	0.9976	↓	0.98
P0C0L4	Complement C4-A	0.01	0.9390	0.9976	↑	1.03
P00740	Coagulation factor IX	0.00	0.9441	0.9976	↑	1.01
P07333	Macrophage colony-stimulating factor 1 receptor	0.00	0.9442	0.9976	↓	0.98
Q92954	Proteoglycan 4	0.00	0.9448	0.9976	↑	1.01
Q6YHK3	CD109 antigen	0.00	0.9453	0.9976	↓	0.97
Q99784	Noelin	0.00	0.9490	0.9976	↓	1.00
Q01459	Di-N-acetylchitobiase	0.00	0.9502	0.9976	↓	0.99
Q9Y6R7	IgGFc-binding protein	0.00	0.9507	0.9976	↑	1.06
Q9NZP8	Complement C1r subcomponent-like protein	0.00	0.9535	0.9976	↑	1.01
P04217	Alpha-1B-glycoprotein	0.00	0.9550	0.9976	↓	1.00
K7EMN2	6-phosphogluconate dehydrogenase, decarboxylating	0.00	0.9590	0.9976	↓	0.95
G3V2W1	Protein Z-dependent protease inhibitor	0.00	0.9607	0.9976	↓	1.00
P12955	Xaa-Pro dipeptidase	0.00	0.9610	0.9976	↓	0.94
P06681	Complement C2	0.00	0.9688	0.9976	↑	1.02
B7ZKJ8	ITIH4 protein	0.00	0.9701	0.9976	↓	0.99
P05154	Plasma serine protease inhibitor	0.00	0.9707	0.9976	↓	1.00
P13796	Plastin-2	0.00	0.9799	0.9976	↓	0.98
P09486	SPARC	0.00	0.9809	0.9976	↓	0.95
P02765	Alpha-2-HS-glycoprotein	0.00	0.9812	0.9976	↑	1.01

P05160	Coagulation factor XIII B chain	0.00	0.9826	0.9976	↓	1.00
P54108	Cysteine-rich secretory protein 3	0.00	0.9883	0.9976	↓	0.98
Q16853	Membrane primary amine oxidase	0.00	0.9943	0.9976	↓	0.98
P05090	Apolipoprotein D	0.00	0.9949	0.9976	↓	0.92
Q10588	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 2	0.00	0.9976	0.9976	↓	0.97

PEs: psychotic experiences; FDR: false discovery rate; ALSPAC: Avon Longitudinal Study of Parents and Children

Supplementary Table 3.16: Ten percent highest-weighted features for model S2 (support vector machine model predicting functional outcome at 24 months in EU-GEI)

Feature	Mean weight
BPRS: suspiciousness	0.197
P01023 Alpha-2-macroglobulin	-0.191
P55058 Phospholipid transfer protein	-0.186
P01871 Immunoglobulin heavy constant mu	-0.182
Q9UGM5 Fetuin-B	0.148
O43866 CD5 antigen-like	-0.145
P14618 Pyruvate kinase	-0.133
P19827 Inter-alpha-trypsin inhibitor heavy chain H1	0.129
SANS: blocking	0.118
SANS: increased latency of response	0.111
P10909 Clusterin	0.107
P08603 Complement factor H	0.104
P36955 Pigment epithelium-derived factor	0.103
MADRS: suicidal thoughts	-0.103
Impersistence at work or school	0.102
P17936 Insulin-like growth factor-binding protein 3	0.102
Age	-0.101
P04196 Histidine-rich glycoprotein	0.099
Q08380 Galectin-3-binding protein	0.097
SANS: grooming and hygeine	0.097
SANS: ability to feel intimacy and closeness	0.095
SANS: sexual activity	-0.093
P11226 Mannose-binding protein C	0.090

Features are ranked according to mean feature weight for models selected in the cross-validation inner loop.
EU-GEI: European Network of National Schizophrenia Networks Studying Gene-Environment Interactions;

BPRS: Brief Psychiatric Rating Scale; SANS: Scale for Assessment of Negative Symptoms; MADRS: Montgomery-Asberg Depression Rating Scale.

Supplementary Table 3.17: Table comparing performance metrics for multi-class site prediction models based on 69 clinical features from model 1b

Multiclass prediction	TP, n	FN, n	TN, n	FP, n	Sens., %	Spec., %	Balanced accuracy, %	AUC (95% confidence interval)	PPV, %	NPV, %	LR+	LR-
London vs. REST	42	11	58	22	79.2	72.5	75.9	0.76 (0.67 ï 0.85)	65.6	84.1	2.9	0.3
Netherlands vs. REST	12	3	85	33	80.0	72.0	76.0	0.76 (0.61 ï 0.91)	26.7	96.6	2.9	0.3
Melbourne vs. REST	1	13	103	16	7.1	86.6	46.8	0.47 (0.31 ï 0.63)	5.9	88.8	0.5	1.1
Switzerland/ Austria vs. REST	1	13	116	3	7.1	97.5	52.3	0.52 (0.36 ï 0.68)	25.0	89.9	2.8	1.0
Denmark/ France vs. REST	3	21	109	0	12.5	100.0	56.3	0.56 (0.43 ï 0.69)	100.0	83.8	N/A	0.9
Spain/ Brazil vs. REST	0	13	120	0	0.0	100.0	50.0	0.50 (0.33 ï 0.67)	N/A	90.2	N/A	1.0

TP: true positives; FN: false negatives; TN: true negatives; FP: false positives; Sens.: sensitivity; Spec.: specificity; AUC: area under the receiver-operating curve; PPV: positive predictive value; NPV: negative predictive value; LR+: positive likelihood ratio; LR-: negative likelihood ratio; N/A: not applicable

Supplementary Table 3.18: Table comparing performance metrics for multi-class site prediction models based on 166 proteomic features from model 1c

Multiclass prediction	TP, n	FN, n	TN, n	FP, n	Sens., %	Spec., %	Balanced accuracy, %	AUC (95% confidence interval)	PPV, %	NPV, %	LR+	LR-
London vs. REST	37	16	42	38	69.8	52.5	61.2	0.61 (0.51 ï 0.71)	49.3	72.4	1.5	0.6
Netherlands vs. REST	8	7	77	41	53.3	65.3	59.3	0.59 (0.43 ï 0.75)	16.3	91.7	1.5	0.7
Melbourne vs. REST	0	14	112	7	0.0	94.1	47.1	0.47 (0.31 ï 0.63)	0.0	88.9	0.0	1.1
Switzerland/ Austria vs. REST	1	13	118	1	7.1	99.2	53.2	0.53 (0.37 ï 0.69)	50.0	90.1	8.5	0.9
Denmark/ France vs. REST	0	24	109	0	0.0	100.0	50.0	0.50 (0.37 ï 0.63)	N/A	82.0	N/A	1.0
Spain/ Brazil vs. REST	0	13	120	0	0.0	100.0	50.0	0.50 (0.33 ï 0.67)	N/A	90.2	N/A	1.0

TP: true positives; FN: false negatives; TN: true negatives; FP: false positives; Sens.: sensitivity; Spec.: specificity; AUC: area under the receiver-operating curve; PPV: positive predictive value; NPV: negative predictive value; LR+: positive likelihood ratio; LR-: negative likelihood ratio; N/A: not applicable

Supplementary Table 3.19: Performance metrics for uncorrected vs. corrected support vector machine models

	Model 1a: clinical and proteomic (EU-GEI initial, all sites)		Model 1b: clinical (EU-GEI initial, all sites)		Model 1c: proteomic (EU-GEI initial, all sites)		Model 2a: proteomic, non-London (EU-GEI initial, all sites except London)		Model 2b: top 10, training (EU-GEI initial, all sites except London)	
	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected
True positives, n (%)	48 (98%)	40 (82%)	23 (47%)	29 (59%)	49 (100%)	43 (88%)	28 (93%)	26 (87%)	30 (100%)	23 (77%)
False negatives, n (%)	1 (2%)	9 (18%)	26 (53%)	20 (41%)	0 (0%)	6 (12%)	2 (7%)	4 (13%)	0 (0%)	7 (23%)
True negatives, n (%)	68 (81%)	71 (85%)	45 (54%)	47 (56%)	71 (85%)	71 (85%)	40 (80%)	42 (84%)	41 (82%)	44 (88%)
False positives, n (%)	16 (19%)	13 (15%)	39 (46%)	37 (44%)	13 (15%)	13 (15%)	10 (20%)	8 (16%)	9 (18%)	6 (12%)
Sensitivity, %	98.0	81.6	46.9	59.2	100.0	87.8	93.3	86.7	100.0	76.7
Specificity, %	81.0	84.5	53.6	56.0	84.5	84.5	80.0	84.0	82.0	88.0
Balanced accuracy, %	89.5	83.1	50.3	57.6	92.3	86.1	86.7	85.3	91.0	82.3
Area under the curve (95% confidence interval)	0.95 (0.91 ï 0.99)	0.91 (0.85 ï 0.97)	0.48 (0.38 ï 0.58)	0.52 (0.42 ï 0.62)	0.96 (0.92 ï 1.00)	0.94 (0.89 ï 0.99)	0.94 (0.88 ï 1.00)	0.96 (0.91 ï 1.00)	0.99 (0.96 ï 1.00)	0.91 (0.84 ï 0.98)
Positive predictive value, %	75.0	75.5	37.1	43.9	79.0	76.8	73.7	76.5	76.9	79.3

Negative predictive value, %	98.6	88.8	63.4	70.1	100.0	92.2	95.2	91.3	100.0	86.3
Positive likelihood ratio	5.1	5.3	1.0	1.3	6.5	5.7	4.7	5.4	5.6	6.4
Negative likelihood ratio	<0.1	0.2	1.0	0.7	<0.1	0.1	0.1	0.2	<0.1	0.3

Supplementary Table 3.19, Continued:

	Model 3: replication (EU-GEI replication, all sites)		Model 4: Psychotic experiences (ALSPAC)		Model S1: ELISA (EU-GEI initial, all sites)		Model S2: functional outcome (EU-GEI initial, all sites)	
	Uncorrected	**Corrected**	**Uncorrected**	**Corrected**	**Uncorrected**	**Corrected**	**Uncorrected**	**Corrected**
True positives, n (%)	48 (98%)	46 (94%)	40 (73%)	29 (53%)	33 (75%)	29 (66%)	27 (57%)	32 (68%)
False negatives, n (%)	1 (2%)	3 (6%)	15 (27%)	26 (47%)	11 (25%)	15 (34%)	20 (43%)	15 (32%)
True negatives, n (%)	77 (90%)	79 (92%)	47 (71%)	39 (59%)	51 (62%)	54 (66%)	22 (69%)	19 (59%)
False positives, n (%)	9 (10%)	7 (8%)	19 (29%)	27 (41%)	31 (38%)	28 (34%)	10 (31%)	13 (41%)
Sensitivity, %	98.0	93.9	72.7	52.7	75.0	65.9	57.4	68.1
Specificity, %	89.5	91.9	71.2	59.1	62.2	65.9	68.8	59.4
Balanced accuracy, %	93.7	92.9	72.0	55.9	68.6	65.9	63.1	63.7
Area under the curve (95% confidence interval)	0.98 (0.95 ï 1.00)	0.97 (0.94 ï 1.00)	0.74 (0.65 ï 0.83)	0.64 (0.54 ï 0.74)	0.76 (0.67 ï 0.85)	0.68 (0.58 ï 0.78)	0.74 (0.63 ï 0.85)	0.71 (0.60 ï 0.82)
Positive predictive value, %	84.2	86.8	67.8	51.8	51.6	50.9	73.0	71.1
Negative predictive value, %	98.7	96.3	75.8	60.0	82.3	78.3	52.4	55.9
Positive likelihood ratio	9.4	11.5	2.5	1.3	2.0	1.9	1.8	1.7
Negative likelihood ratio	<0.1	0.1	0.4	0.8	0.4	0.5	0.6	0.5

EU-GEI: European Network of National Schizophrenia Networks studying Gene-Environment Interactions; ALSPAC: Avon Longitudinal Study of Parents and Children; ELISA: enzyme-linked immunosorbent assay.

Supplementary Table 3.20: Proteins differentially expressed in CHR-T vs. CHR-NT on ANCOVA (*p*<0.05*)* in EU-GEI baseline plasma samples in the initial and replication experiment and predicted systemic impact on coagulation and complement activation and regulation

Pathway	CHR-T vs. CHR-NT, initial experiment (p<0.05)	CHR-T vs. CHR-NT, replication experiment (p<0.05)	Predicted impact
Complement	C1Q (ŷ), C1R SERPING1 (Ź) FIC3 (Ź) CFB (ŷ) CFI (ŷ), CFH C8 (ŷ), CLU (	C1Q (ŷ), C1R -A(ŷ() C4BP (Ź) SERPING1 (Ź) FIC3 (Ź) CFD (ŷ) CFI (ŷ() ŷ) CFH C5 (ŷ), C6 (ŷ), (ŷ), CLU (ŷ), VTN CPB2 (ŷ)	Classical pathway: activation ŷ i C3b generati Terminal pathway: activation ŷ
Coagulation	FXI (ŷ) SERPING1 (Ź) SERPIND1 (Ź) ADAMTS13 (Ź) PLG (ŷ)	IX (ŷ), CPN (ŷ) VWF (ŷ) FII (ŷ), FV (ŷ), SERPIND1 (-) CPB2 (ŷ) XIII (ŷ), FIB (Ź PLG (ŷ)	Intrinsic pathway activation ŷ regulation Thrombin gene Thrombin regu Fibrin genera Plasmin gener
Associated components	IGM, IGG (Ź) A2M (Ź)	IGM, IGG (Ź) A2M (Ź)	Complement classical pathway ŷ Plasmin/thrombin regulation Ź

Notes

The directions of the differentially expressed proteins indicate a general upregulation of activation components and downregulation of regulatory proteins.

With regards to the complement system, we observe upregulation of complement classical pathway components (C1Q, C1R, C2, C4 and IGG, IGM), while key regulators (SERPING1, C4BP) are downregulated. Classical pathway activation would lead to alternative pathway activation and amplification leading to opsonin C3b generation.[1] Key alternative pathway regulatory components (CFH, CFI) are upregulated resulting in increased iC3b, both C3b and iC3b serve as ligands for selective complement receptors on leukocytes.[2] Terminal pathway components are upregulated (C5, C6, C7, C8, C9) leading to an increase in the terminal complement complex C5b-9, while terminal pathway regulators are also upregulated (CLU, VTN), which bind to the nascent amphiphilic C5b-9 complex, rendering it soluble and lytically inactive.[3] Overall, we expect an enhanced output of complement activation products, as a result of enhanced classical pathway activation and dysregulation, amplified through the alternative pathway, while increased terminal pathway components drive the generation of soluble terminal complement component (sTCC).

In the coagulation system, the primary effect is a downregulation of intrinsic pathway control (SERPING1), thrombin (SERPIND1) and plasmin (A2M) regulation while increased plasminogen (PLG) drives plasmin generation. Overall, we expect an increase in thrombin generation and dysregulation, with increased plasmin generation.

1. Merle NS, Church SE, Fremeaux-Bacchi V, Roumenina LT. Complement System Part I - Molecular Mechanisms of Activation and Regulation. Front Immunol. 2015;6:262.

2. van Lookeren Campagne M, Wiesmann C, Brown EJ. Macrophage complement receptors and pathogen clearance. Cellular microbiology. 2007;9(9):2095-102.

3. Choi NH, Nakano Y, Tobe T, Mazda T, Tomita M. Incorporation of SP-40,40 into the soluble membrane attack complex (SMAC, SC5b-9) of complement. International immunology. 1990;2(5):413-7.

Supplementary Table 3.21: Table summarising directionality of effect of ten percent highest weighted features in model 1a (EU-GEI initial experiment), model 3 (EU-GEI replication) and model 4 (ALSPAC proteomic data)

Model 1a: EU-GEI clinical and proteomic data, initial experiment		**Model 3: EU-GEI clinical and proteomic data, replication experiment**		**Model 4: ALSPAC proteomic data**	
Feature	**Directionality of effect (CHR-T vs. CHR-NT)**	**Feature**	**Directionality of effect (CHR-T vs. CHR-NT)**	**Feature**	**Directionality of effect (PE at 18 vs. no PE at 18)**
P01023 Alpha-2-macroglobulin	↓	P01023 Alpha-2-macroglobulin	↓	P04003 C4b-binding protein alpha chain	↓
P01871 Immunoglobulin heavy constant mu	↓	P22792 Carboxypeptidase N subunit 2	↑	P27169 Serum paraoxonase/arylesterase 1	↓
P04003 C4b-binding protein alpha chain	↓	P01871 Immunoglobulin heavy constant mu	↓	Q03591 Complement factor H-related protein 1	↓
P07357 Complement component 8 alpha chain	↑	P09871 Complement C1s subcomponent	↓	P07225 Vitamin K-dependent protein S	↓
P55058 Phospholipid transfer protein	↓	P01011 Alpha-1-antichymotrypsin	↑	P61626 Lysozyme C	↓
O75636 Ficolin-3	↓	P00747 Plasminogen	↑	P55103 Inhibin beta C chain	↑
P02774 Vitamin D binding protein	↑	P08571 Monocyte differentiation antigen CD14	↑	Q08380 Galectin-3-binding protein	↑
P07225 Vitamin K-dependent protein S	↓	P10909 Clusterin	↑	P24593 Insulin-like growth factor-binding protein 5	↑

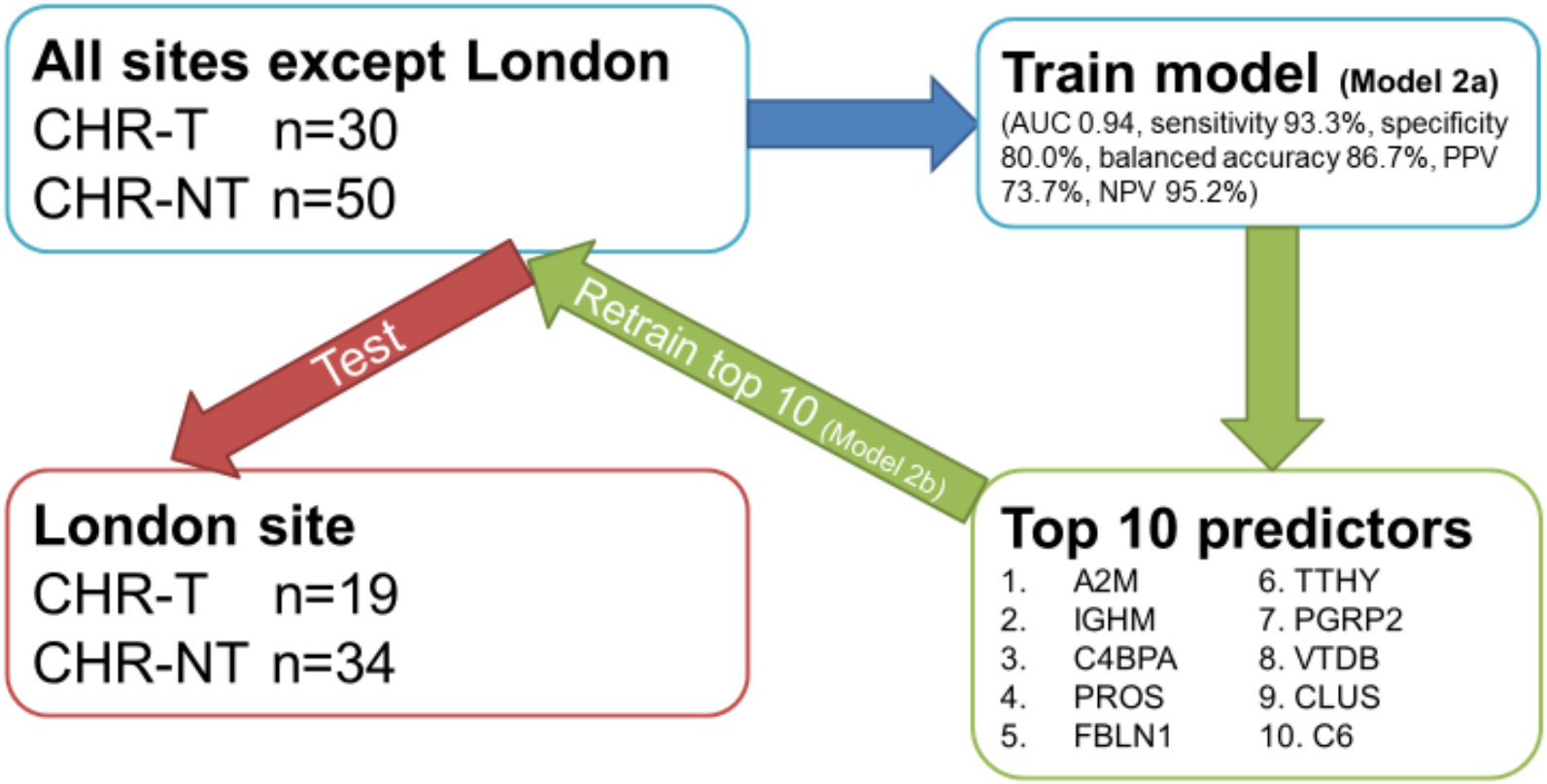
P43320 Beta-crystallin B2	↑	Q16610 Extracellular matrix protein 1	↑	P00746 Complement factor D	↑
P02766 Transthyretin	↓	G3XAM2 Complement factor I	↑	P01019 Angiotensinogen	↓
P23142 Fibuln-1	↑	P04003 C4b binding protein alpha chain	↓	P01871 Immunoglobulin heavy constant mu	↓
P10909 Clusterin	↑	P13671 Complement component 6	↑	O75636 Ficolin-3	↑
P05155 Plasma protease C1 inhibitor	↓	P25311 Zinc alpha-2-glycoprotein	↑	Q9H4A9 Dipeptidase 2	↓
Sex (female vs. male)	↓	P07359 Platelet glycoprotein Ib alpha chain	↑	P01023 Alpha-2-macroglobulin	↓
P00747 Plasminogen	↑	P01031 Complement C5	↑	P04275 von Willebrand factor	↓
P13671 Complement component 6	↑	O75882 Attractin	↑	Q9NQ79 Cartilage acidic protein 1	↑
P02747 Complement C1q subcomponent subunit C	↑	P0DOY3 Immunoglobulin lambda constant 3	↓	P24592 Insulin-like growth factor-binding protein 6	↑
P02753 Retinol-binding protein 4	↑	P15169 Carboxypeptidase N catalytic chain (CPN)	↑	P09871 Complement C1s subcomponent	↓
Q76LX8 A disintegrin and metalloproteinase with thrombospondin motifs 13	↓			P10909 Clusterin	↓
P08697 Alpha-2-antiplasmin	↓			O95497 Pantetheinase	↑

P19827 Inter-alpha-trypsin inhibitor heavy chain H1	↑			P02654 Apolipoprotein C-I	↓
MADRS: concentration difficulties	↓			P02679 Fibrinogen gamma chain	↓
P02489 Alpha-crystallin A chain	↑			P07358 Complement component C8 beta chain	↑
				Q5T7F0 Neuropilin	↓
				P04040 Catalase	↑
				P43251 Biotinidase	↑

EU-GEI: European Network of National Schizophrenia Networks studying Gene-Environment Interactions; ALSPAC: Avon Longitudinal Study of Parents and Children

Appendix 3.3: Supplementary Figures

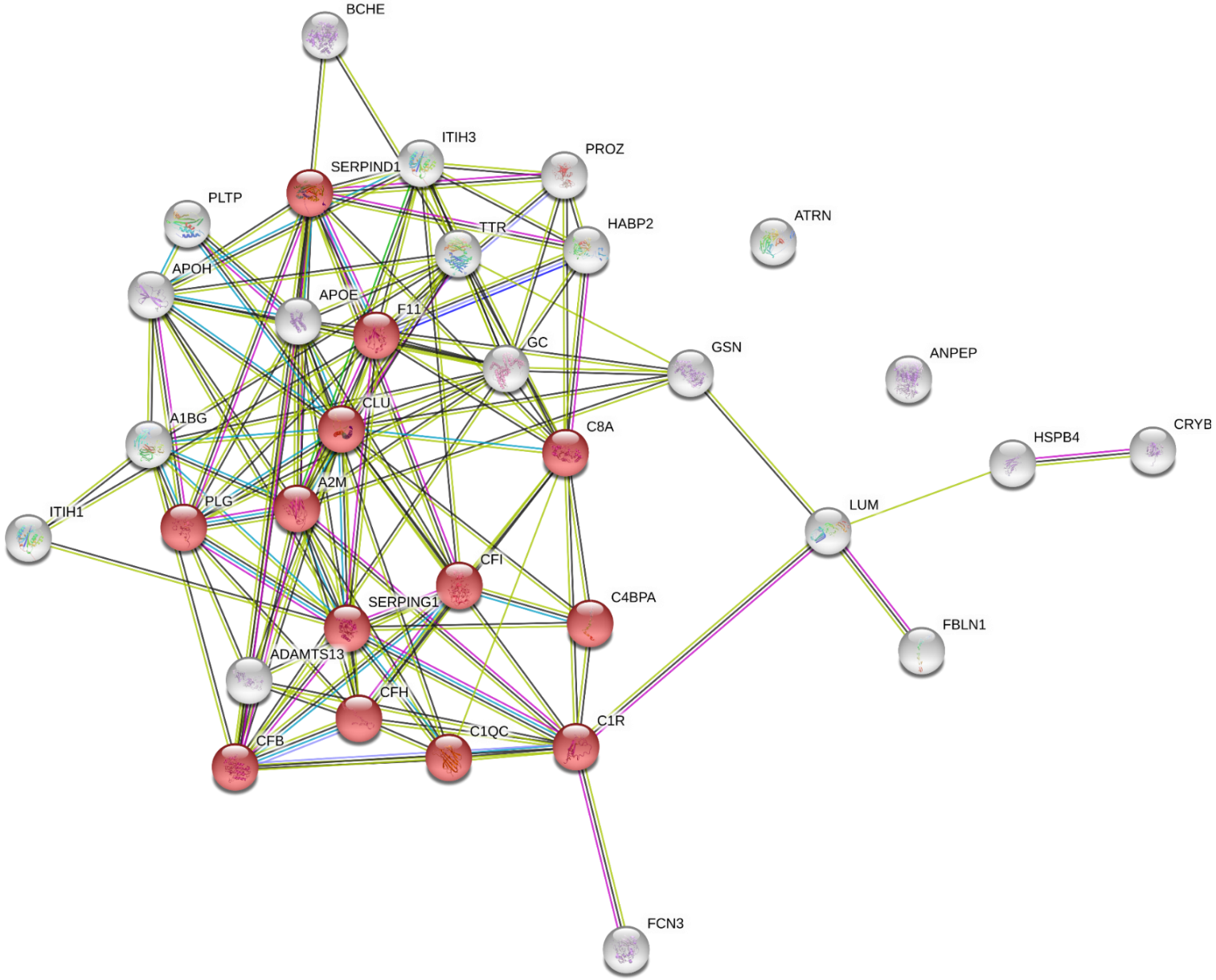
Supplementary Figure 3.1: Derivation and testing of model 2b: parsimonious (10-predictor) proteomic model



Notes. To derive the 10 highest-weighted proteins for prediction of transition outcome, an L2-regularised SVM model was trained using the proteomic data from all sites except London (CHR-T n=30, CHR-NT n=50), with leave-site-out cross-validation (Model 2a). The top 10 predictors, ranked by mean feature weight across models selected in the inner loop, are presented in the figure. Next, a reduced model based solely on these 10 features (Model 2b) was trained using data from all sites except London (Model 1d). This reduced model was then tested in the held -out London sample (CHR-T n=19, CHR-NT n=34).

AUC: area under the receiver-operating curve; PPV: positive predictive value; NPV: negative predictive value; A2M: alpha-2-macroglobulin; IGHM: immunoglobulin heavy constant mu; C4BPA: C4b-binding protein alpha chain; PROS: vitamin K-dependent protein S; FBLN1: Fibulin-1; TTHY: transthyretin; PGRP2: N-acetylmuramoyl-L-alanine amidase; VTDB: vitamin D binding protein; CLUS: clusterin; C6: complement component 6.

Supplementary Figure 3.2: STRING functional protein association network for proteins significantly differentially expressed (following false discovery rate correction) between CHR-T and CHR-NT in EU-GEI initial experiment



The network nodes are proteins. Proteins implicated in the complement and coagulation cascades are highlighted in red.

The edges represent functional associations between proteins, and the colour of each edge represents the source of evidence for that association:

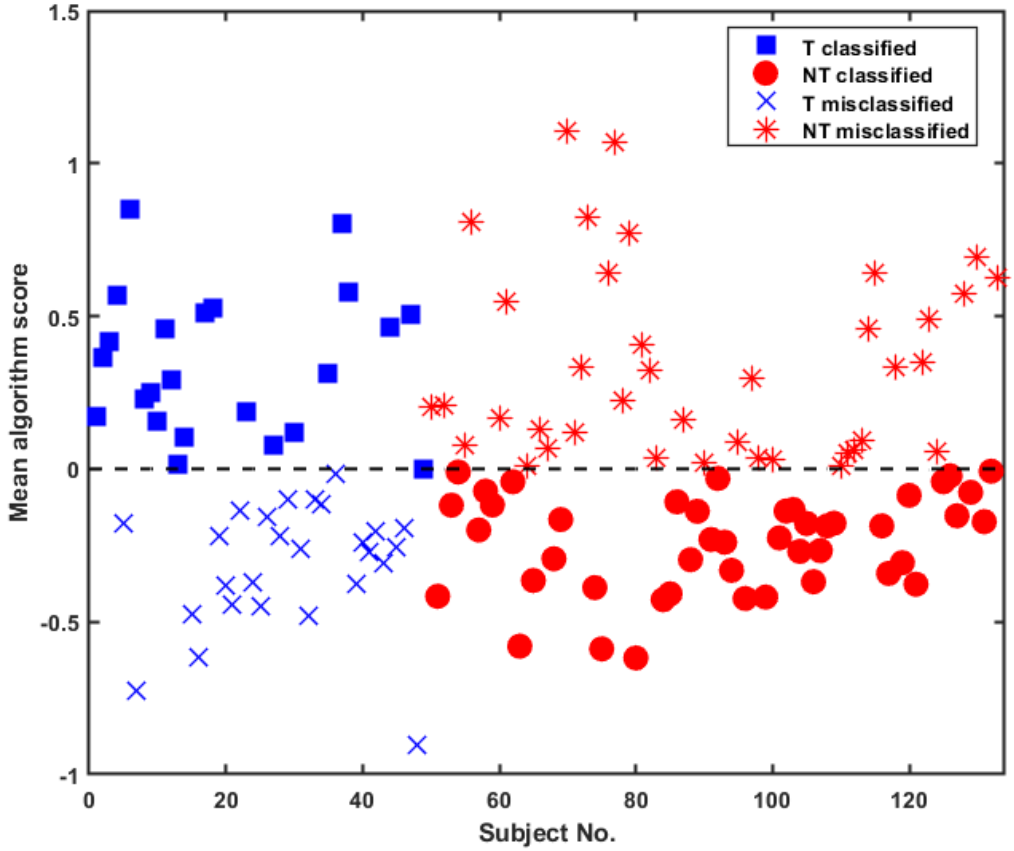
- Red line - fusion evidence
- Green line - neighbourhood evidence
- Blue line ï co-occurrence evidence
- Purple line - experimental evidence
- Yellow line ï text-mining evidence
- Light blue line - database evidence
- Black line ï co-expression evidence.

Note: IGHM is not listed in the STRING database and hence is not shown.

STRING database: https://string-db.org/

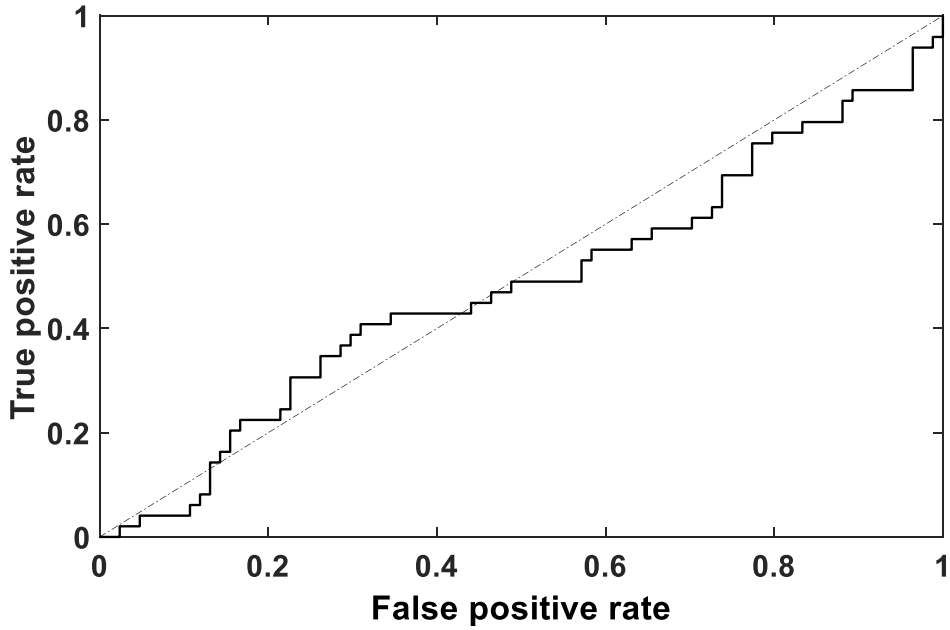
Supplementary Figure 3.3: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 1b: clinical data

A

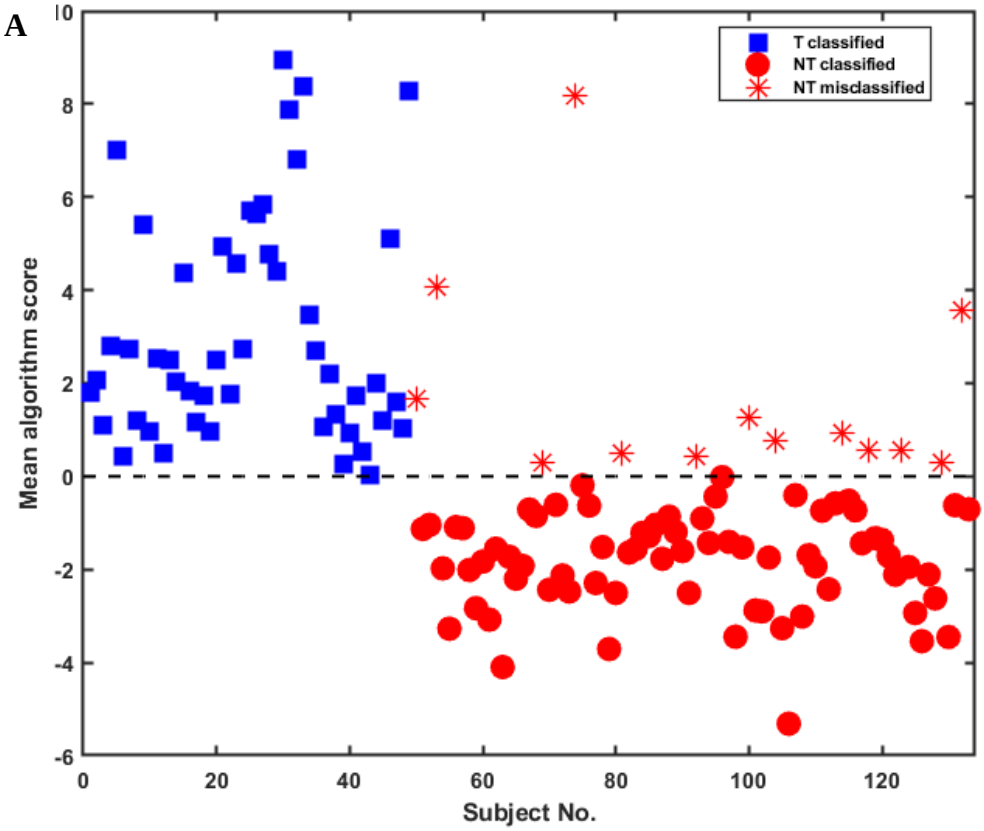


T: clinical high-risk participants who transitioned to first episode psychosis; NT: clinical high-risk participants who did not transition.

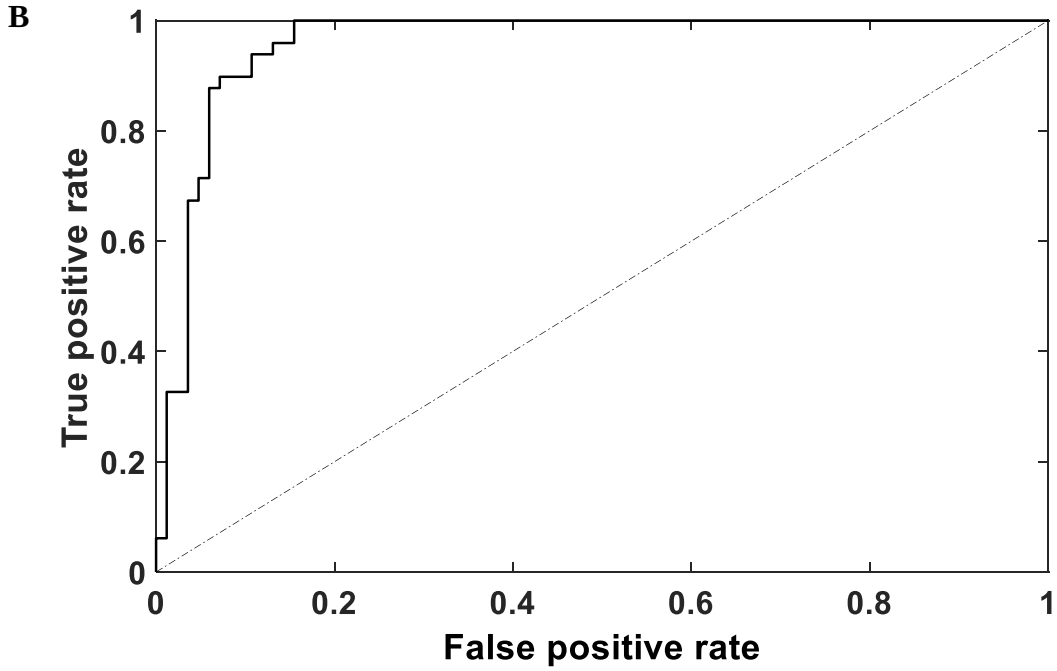
B



Supplementary Figure 3.4: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 1c: proteomic data

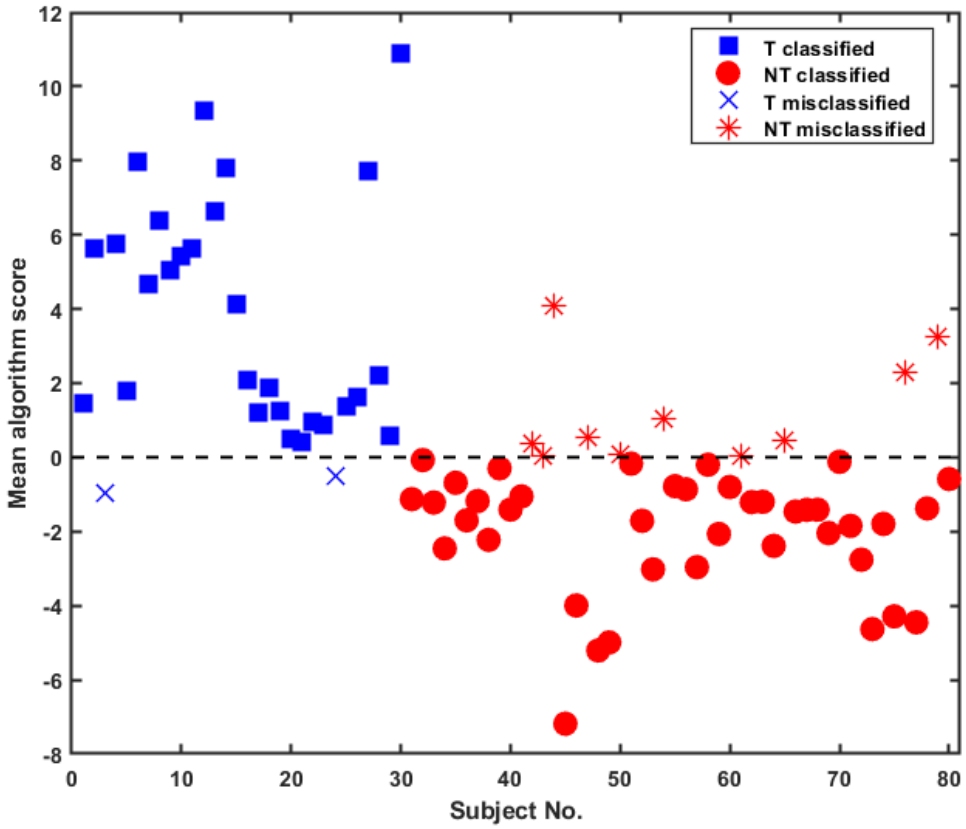


T: clinical high-risk participants who transitioned to first episode psychosis; NT: clinical high-risk participants who did not transition.



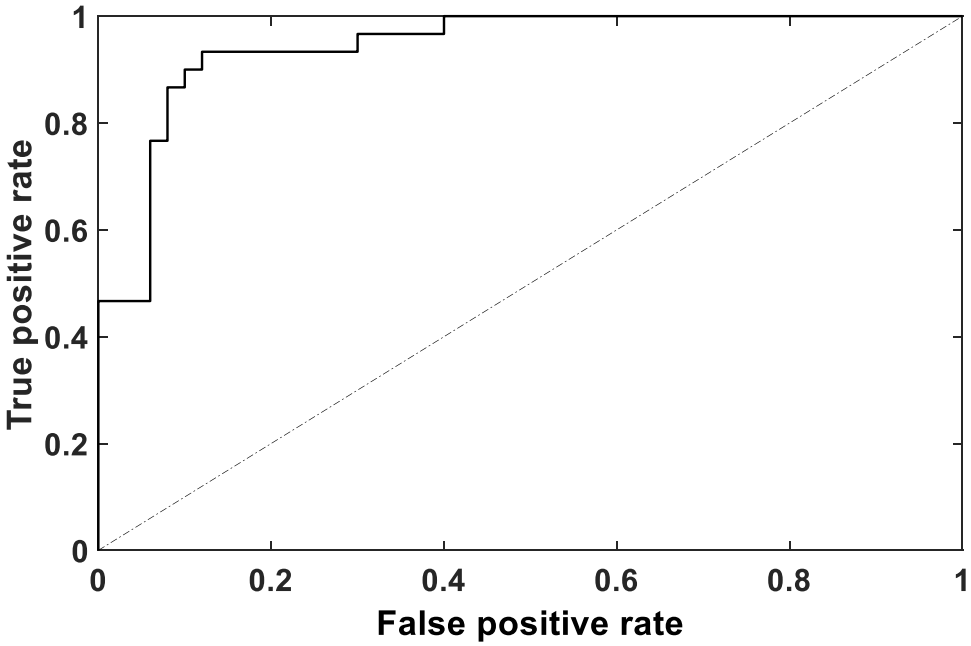
Supplementary Figure 3.5: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 2a: proteomic (non-London)

A

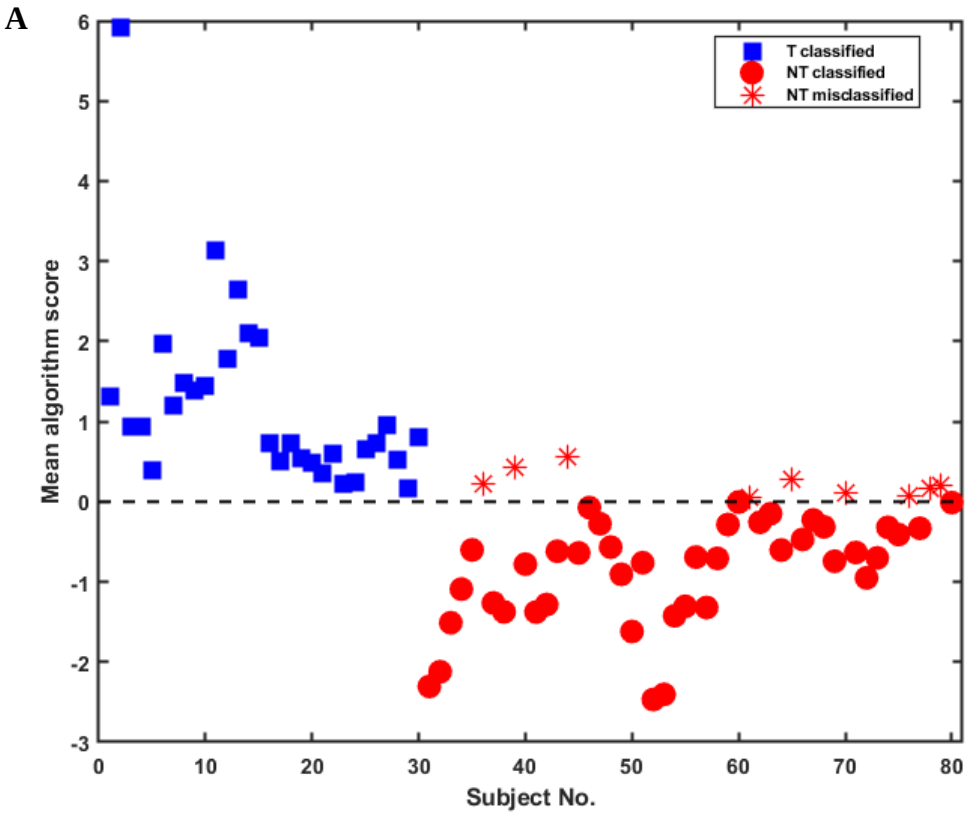


T: clinical high-risk participants who transitioned to first episode psychosis; NT: clinical high-risk participants who did not transition.

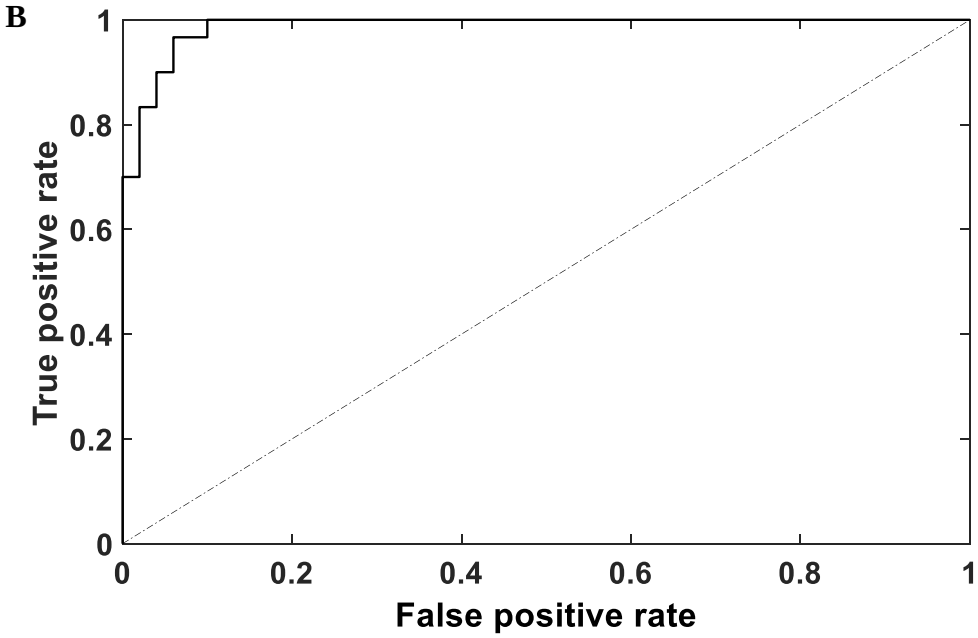
B



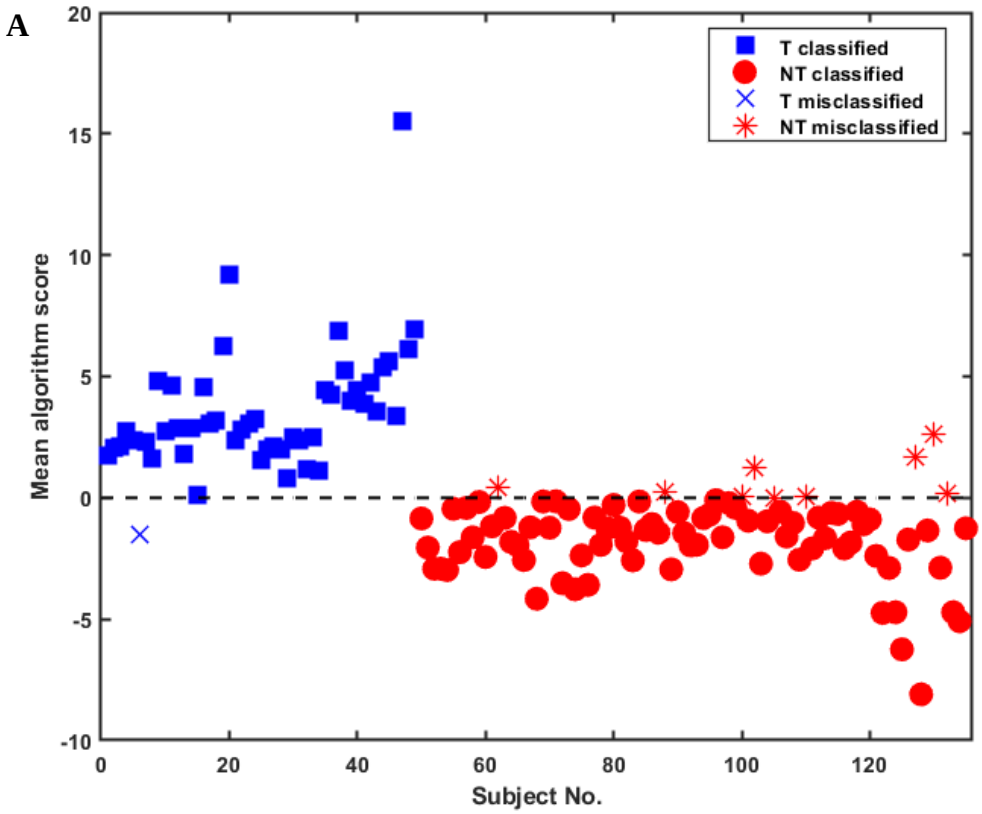
Supplementary Figure 3.6: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 2b: parsimonious (10-predictor) proteomic model, training data (non-London)



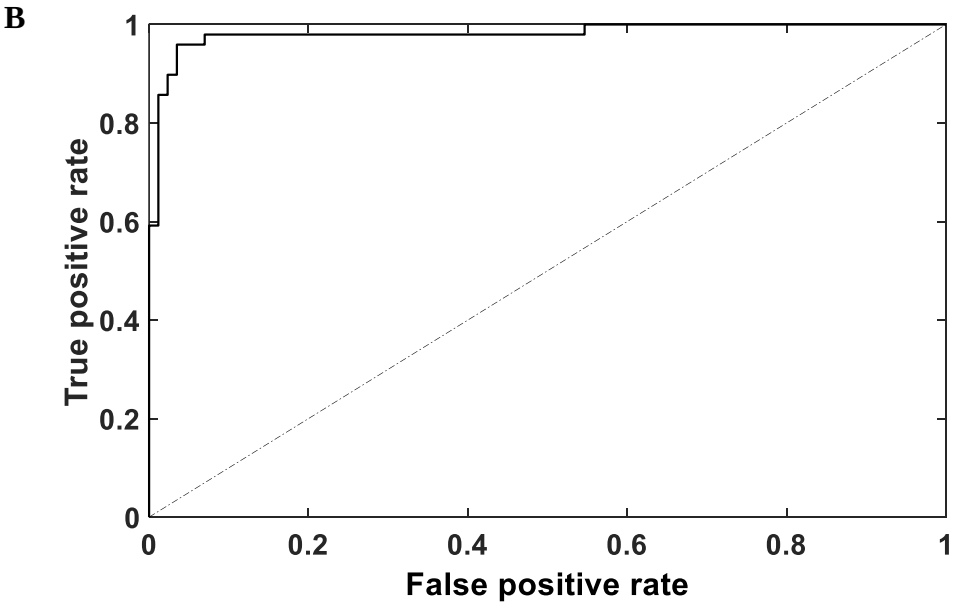
T: clinical high-risk participants who transitioned to first episode psychosis; NT: clinical high-risk participants who did not transition.



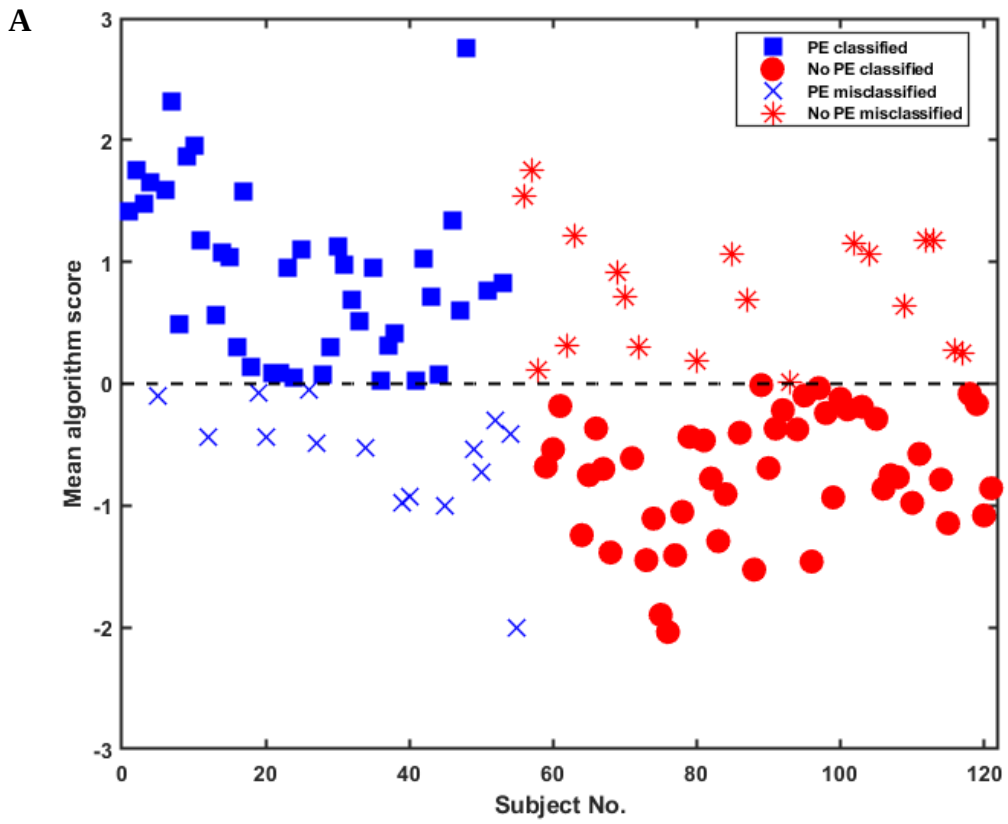
Supplementary Figure 3.7: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 3: replication



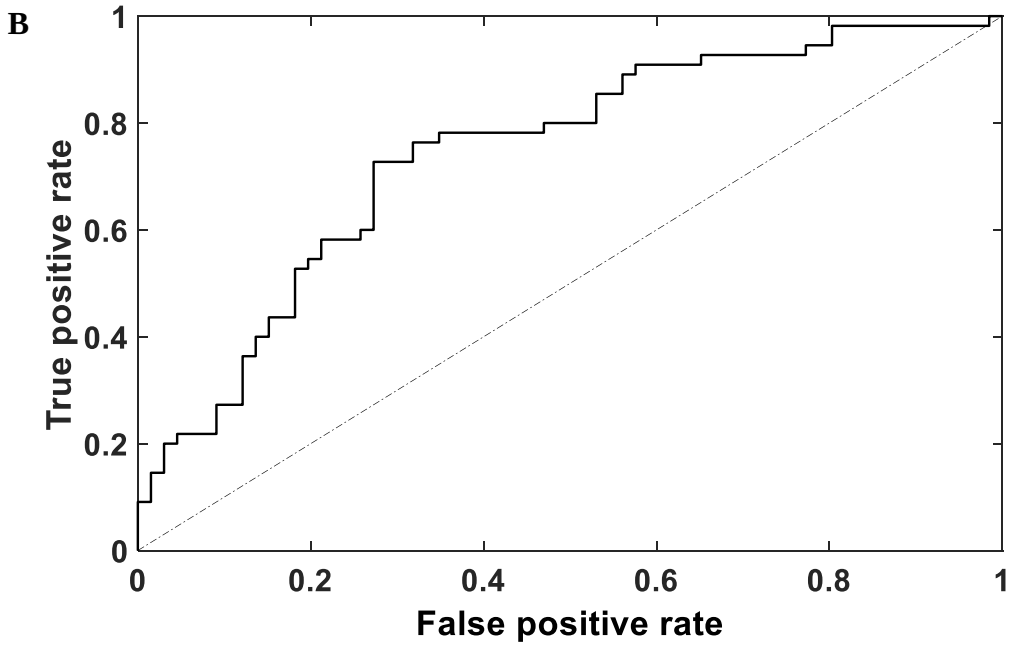
T: clinical high-risk participants who transitioned to first episode psychosis; NT: clinical high-risk participants who did not transition



Supplementary Figure 3.8: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model 4: ALSPAC proteomic data

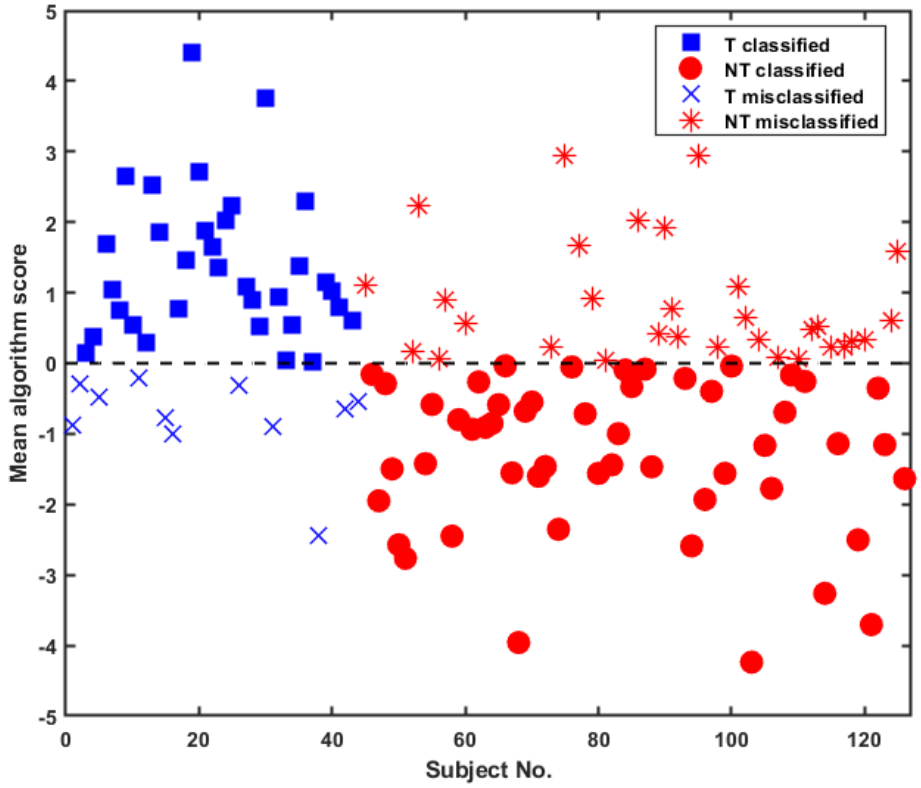


PE: psychotic experiences at 18; No PE: No psychotic experiences at 18.



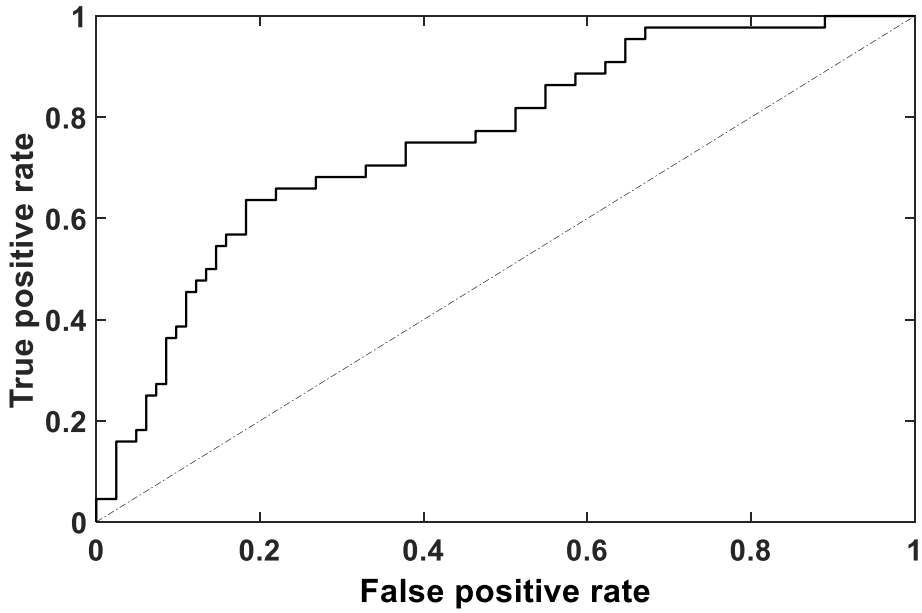
Supplementary Figure 3.9: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model S1: ELISA

A

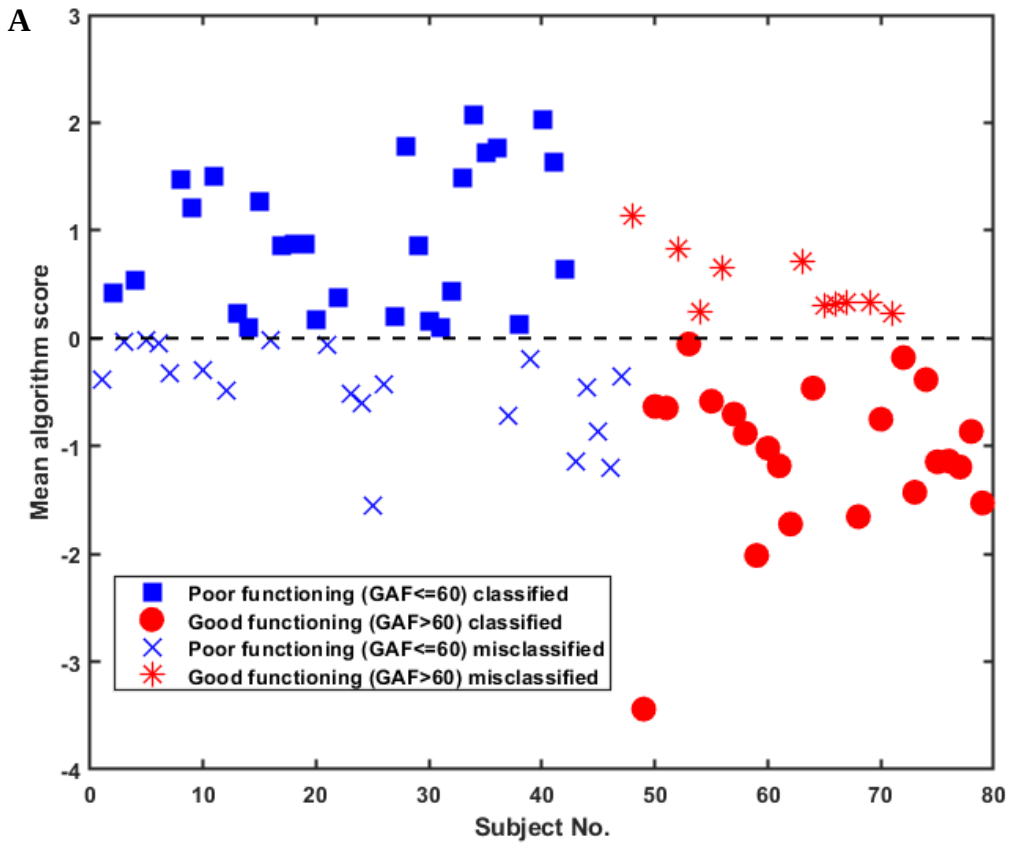


T: clinical high-risk participants who transitioned to first episode psychosis; NT: clinical high-risk participants who did not transition

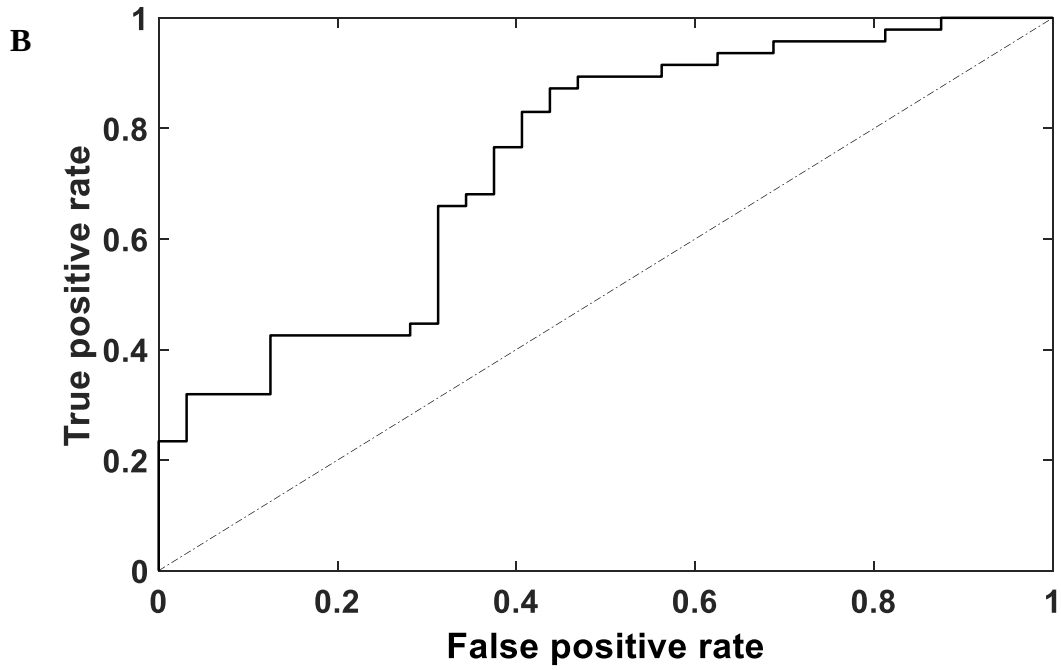
B



Supplementary Figure 3.10: Mean algorithm scores and class predictions (A) and receiver-operating characteristic curve (B) for model S2: functional outcome

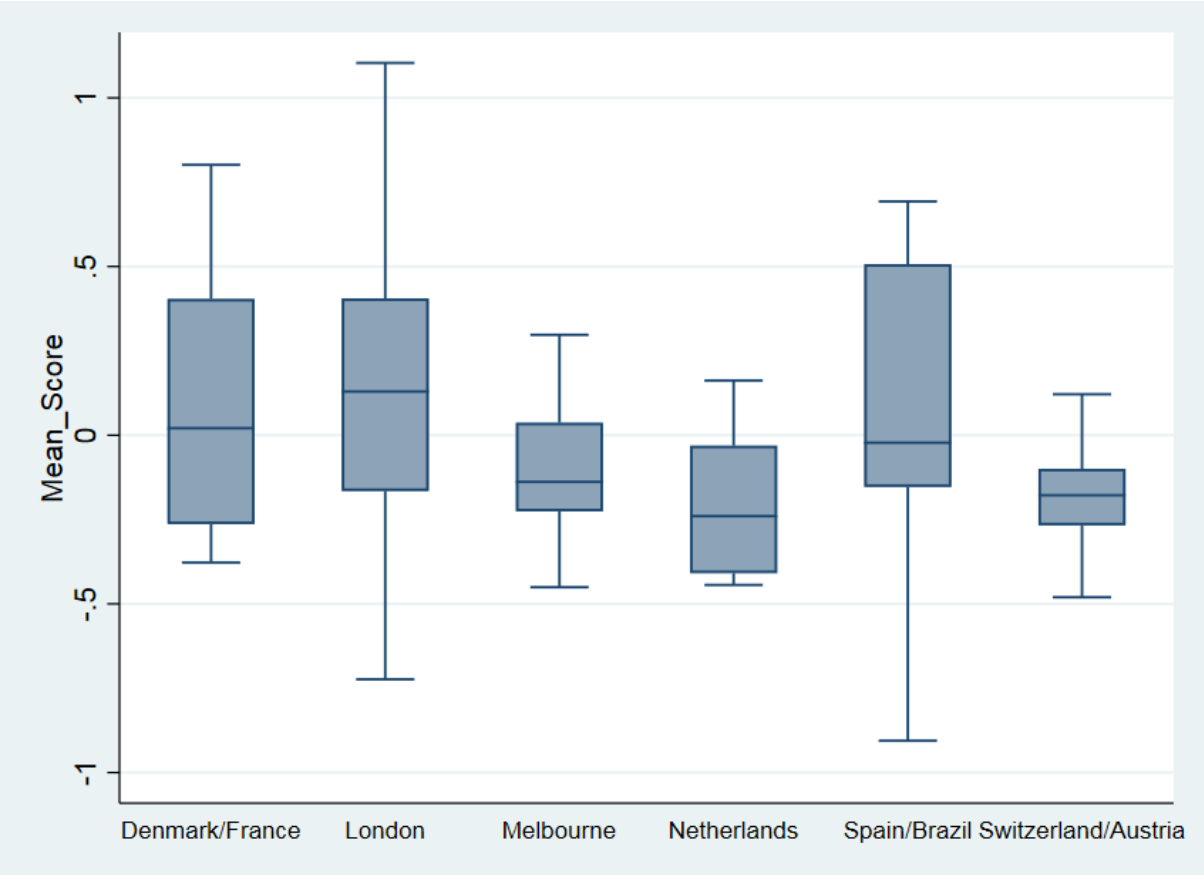


GAF: General Assessment of Functioning (disability su
Good: GAF score >60 at 24 months



Supplementary Figure 3.11: Model 1b (clinical data) decision scores stratified by EU-GEI site

a. Box plots for model 1b (clinical data) decision scores, stratified by EU-GEI site



b. Kruskall-Wallis test

Chi-squared (5 d.f.): 16.878

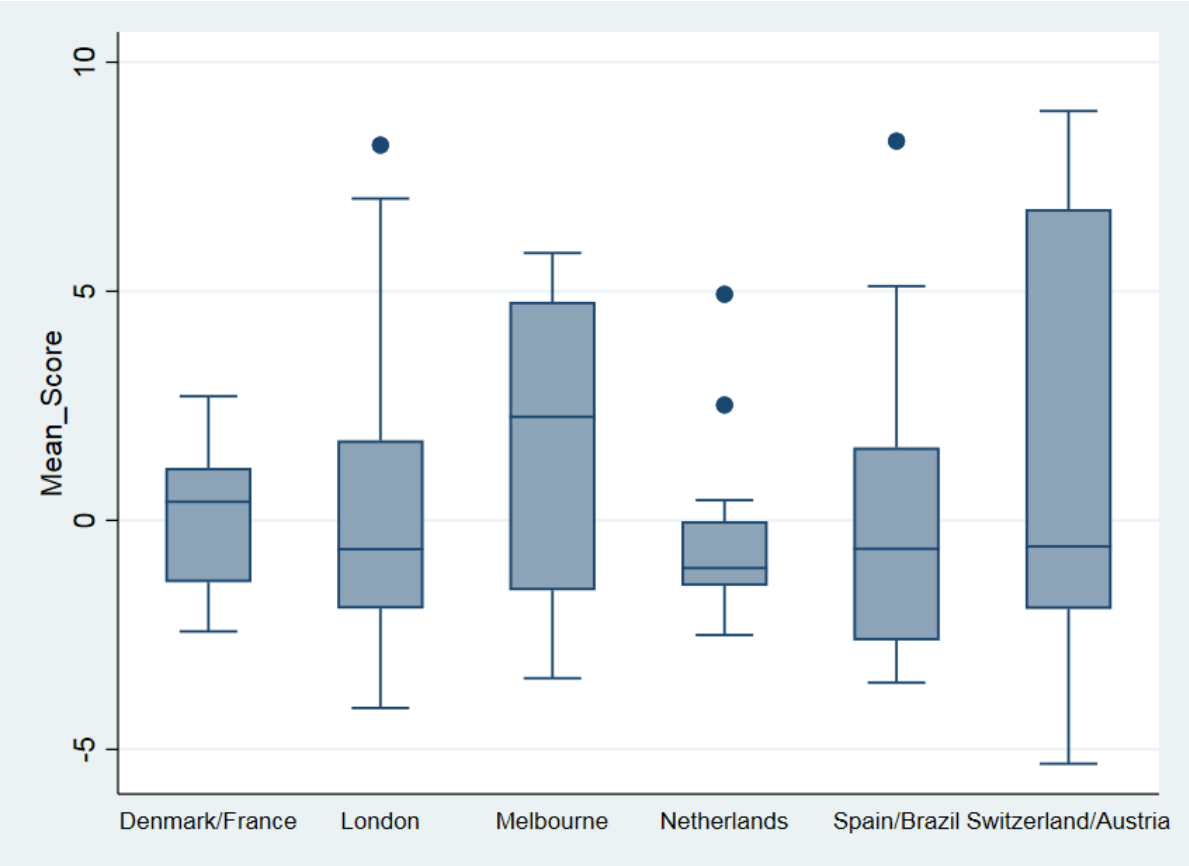
p = 0.0047

c. Post-hoc Dunnôs test pairwise comparisons with corrected p values

	London	Netherlands	Melbourne	Switzerland/Austria	Denmark/France
Netherlands	3.308				
p	**0.007**				
Melbourne	2.021	-0.969			
p	0.324	1.000			
Switzerland/Austria	2.626	-0.480	0.481		
p	0.065	1.000	1.000		
Denmark/France	0.786	-2.352	-1.231	-1.771	
p	1.000	0.140	1.000	0.574	
Spain/Brazil	0.349	-2.269	-1.297	-1.768	-0.248
p	1.000	0.175	1.000	0.578	1.000

Supplementary Figure 3.12: Model 1c (proteomic data) decision scores stratified by EU-GEI site

a. Box plots for model 1c (proteomic data) decision scores stratified by EU-GEI site



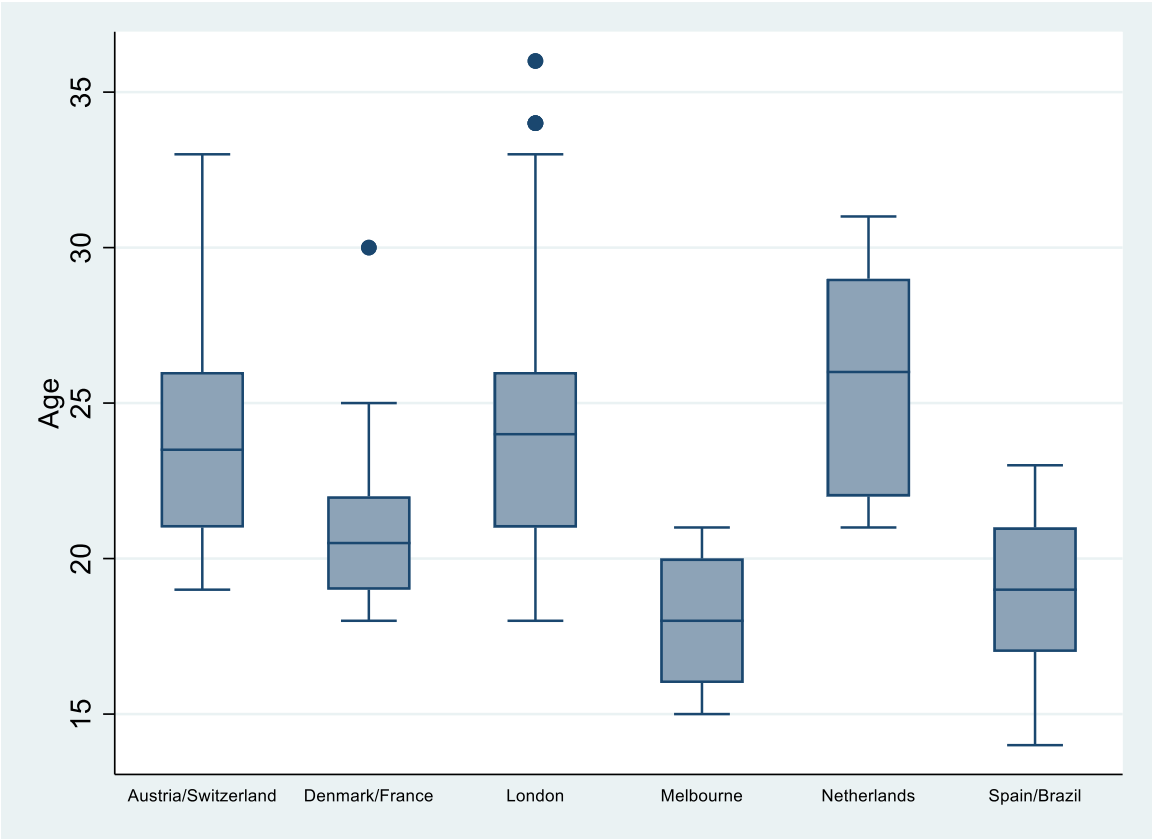
b. Kruskall-Wallis test

Chi-squared (5 d.f.): 3.512

p = 0.6216

Supplementary Figure 3.13: Age stratified by EU-GEI site

a. Box plots for age in years, stratified by EU-GEI site



b. Kruskall-Wallis test

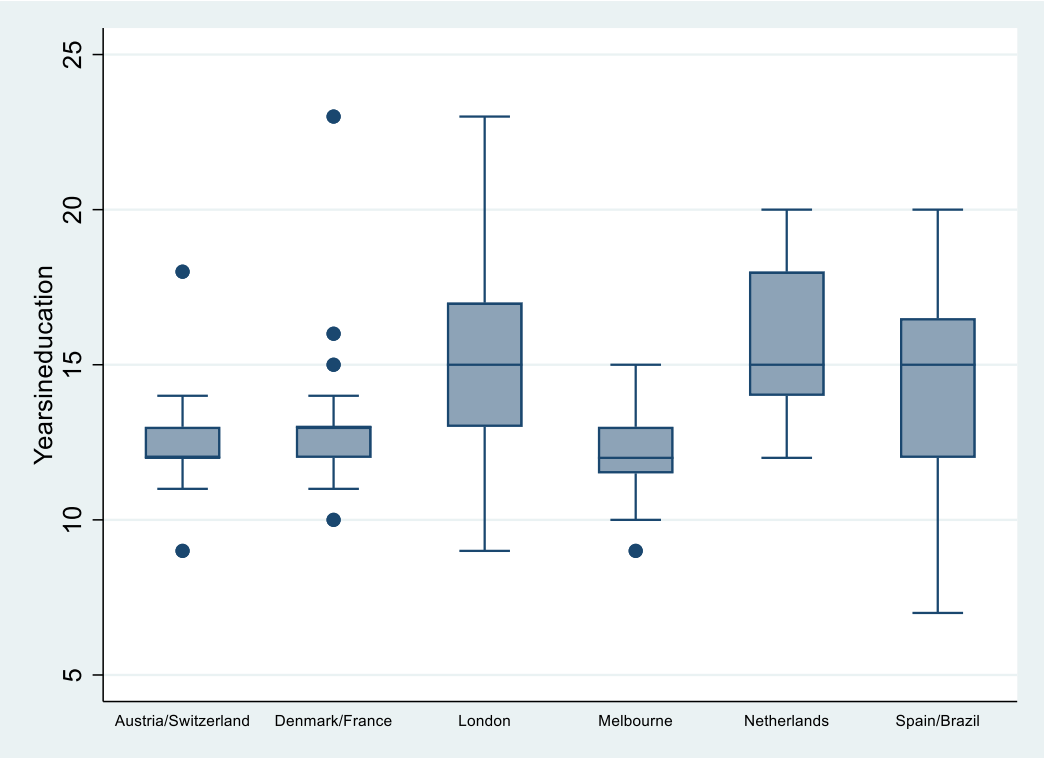
Chi-sq (5 d.f.): 49.157
p = 0.0001

c. Post-hoc Dunnôs test pairwise comparisons with Bonferroni-corrected p values

	Austria/Switzerland	Denmark/France	London	Melbourne	Netherlands
Denmark/France	2.283				
p	0.168				
London	0.077	-3.027			
p	1.000	**0.019**			
Melbourne	4.034	2.250	4.996		
p	**<0.001**	0.183	**<0.001**		
Netherlands	-1.012	-3.476	-1.366	-5.115	
p	1.000	**0.004**	1.000	**<0.001**	
Spain/Brazil	3.297	1.458	4.029	-0.661	4.344
p	**0.007**	1.000	**<0.001**	1.000	**<0.001**

Supplementary Figure 3.14: Years in education stratified by EU-GEI site

a. Box plots for years in education, stratified by EU-GEI site



b. Kruskall-Wallis test

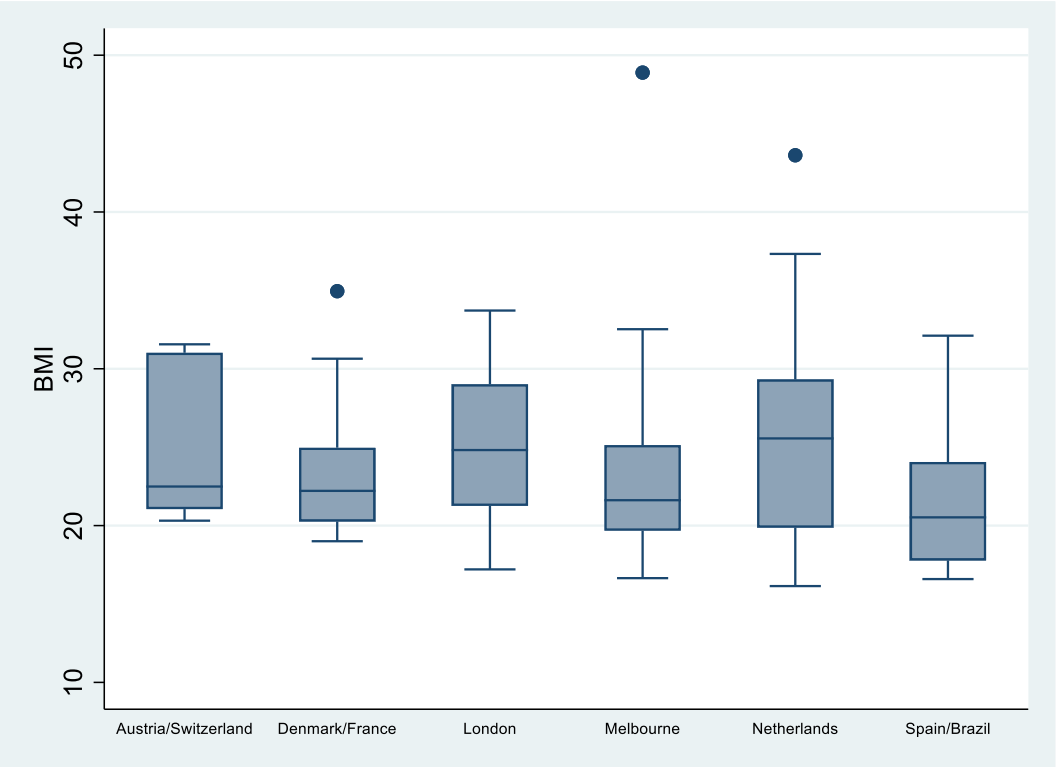
Chi-sq (5 d.f.): 23.921
p = 0.0002

c. Post-hoc Dunnôs test pairwise comparisons with corrected p values

	Austria/Switzerland	Denmark/France	London	Melbourne	Netherlands
Denmark/France	-0.350				
p	1.000				
London	-2.528	-2.688			
p	0.086	0.054			
Melbourne	0.291	0.704	3.115		
p	1.000	1.000	**0.014**		
Netherlands	-3.064	-3.184	-1.290	-3.551	
p	**0.016**	**0.011**	1.000	**0.003**	
Spain/Brazil	-1.773	-1.666	0.356	-2.164	1.271
p	0.572	0.718	1.000	0.229	1.000

Supplementary Figure 3.15: Body mass index stratified by EU-GEI site

a. Box plots for body mass index (BMI) in kg/m^2, stratified by EU-GEI site

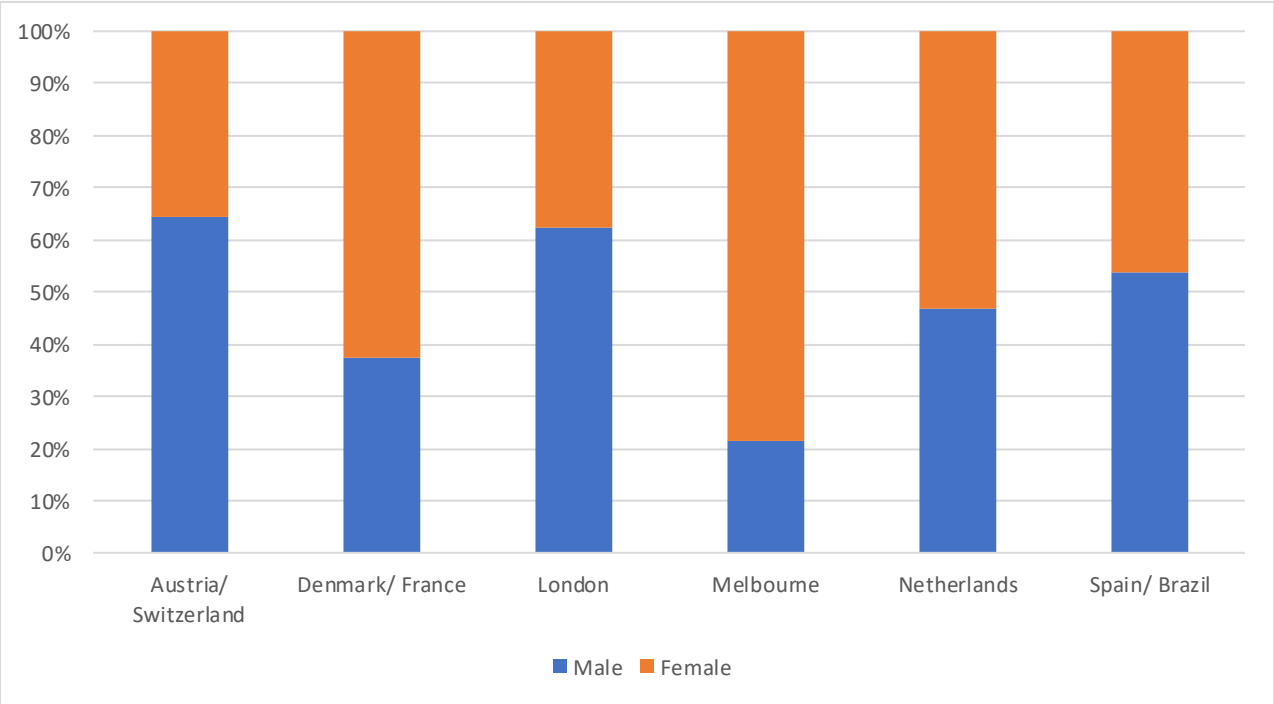


b. Kruskall-Wallis test

Chi-sq (5 d.f.): 8.985
p = 0.1097

Supplementary Figure 3.16: Sex stratified by EU-GEI site

a. Sex distribution, stratified by EU-GEI site



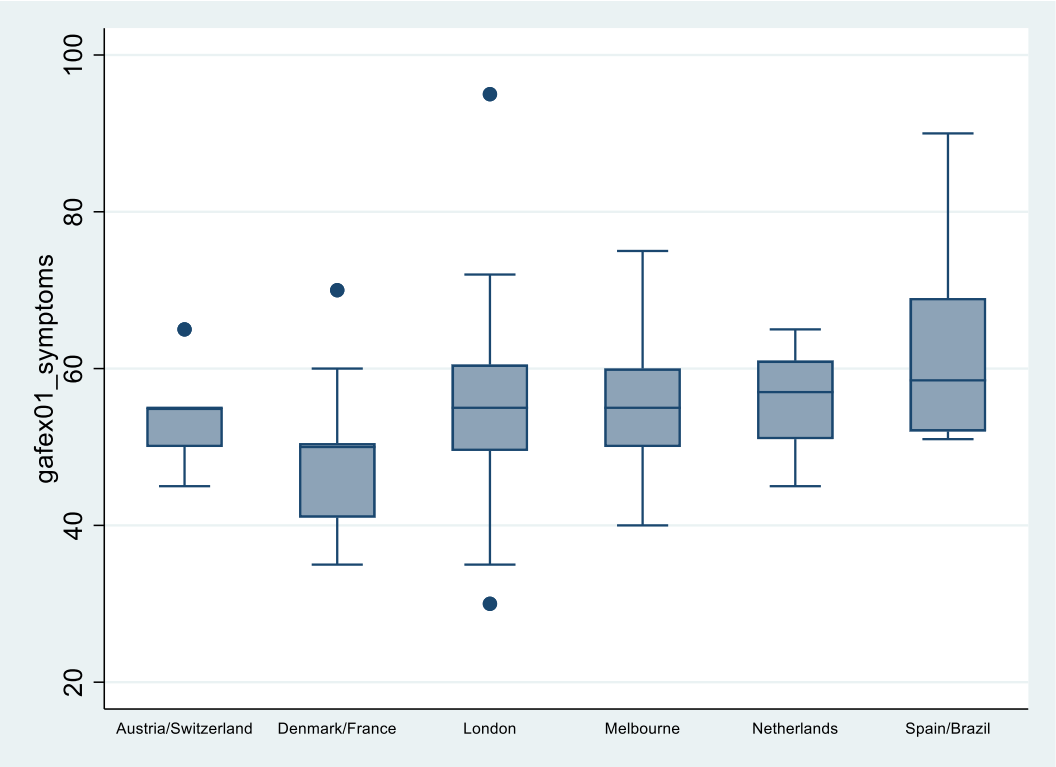
b. Chi-squared test

	Male, n (%)	Female, n (%)
Austria/ Switzerland	9 (64%)	5 (36%)
Denmark/ France	9 (28%)	15 (62%)
London	33 (62%)	20 (38%)
Melbourne	3 (21%)	11 (79%)
Netherlands	7 (47%)	8 (53%)
Spain/ Brazil	7 (54%)	6 (46%)

Chi-squared (5 d.f.): 10.4852, p=0.063

Supplementary Figure 3.17: General Assessment of Functioning symptoms subscale stratified by EU-GEI site

a. Box plots for General Assessment of Functioning (GAF) symptoms subscale score, stratified by EU-GEI site



b. Kruskall-Wallis test

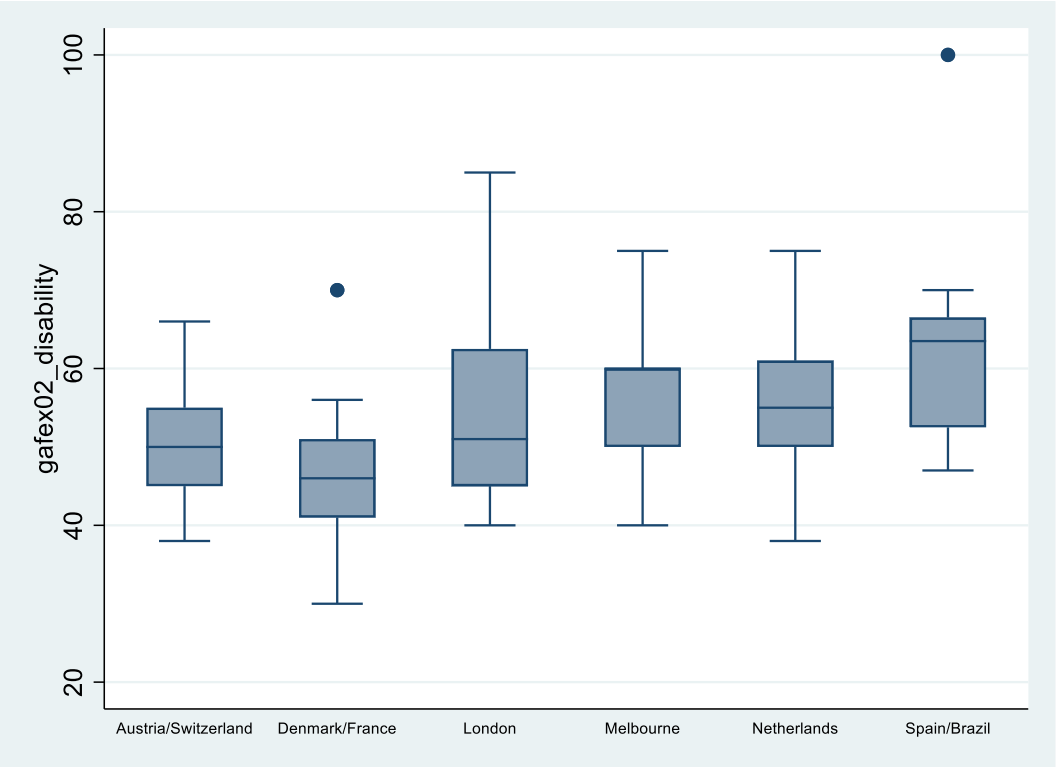
Chi-sq (5 d.f.): 15.575
p = 0.0082

c. Post-hoc Dunnôs test pairwise comparisons with Bonferro corrected p values

	Austria/Switzerland	Denmark/France	London	Melbourne	Netherlands
Denmark/France	1.218				
p	1.000				
London	-0.952	-2.624			
p	1.000	0.065			
Melbourne	-0.688	-1.941	0.082		
p	1.000	0.392	1.000		
Netherlands	-1.527	-2.876	-0.967	-0.815	
p	0.950	**0.030**	1.000	1.000	
Spain/Brazil	-2.275	-3.576	-1.922	-1.601	-0.858
p	0.172	**0.003**	0.410	0.820	1.000

Supplementary Figure 3.18: General Assessment of Functioning disability subscale stratified by EU-GEI site

a. Box plots for General Assessment of Functioning (GAF) disability subscale score, stratified by EU-GEI site



b. Kruskall-Wallis test

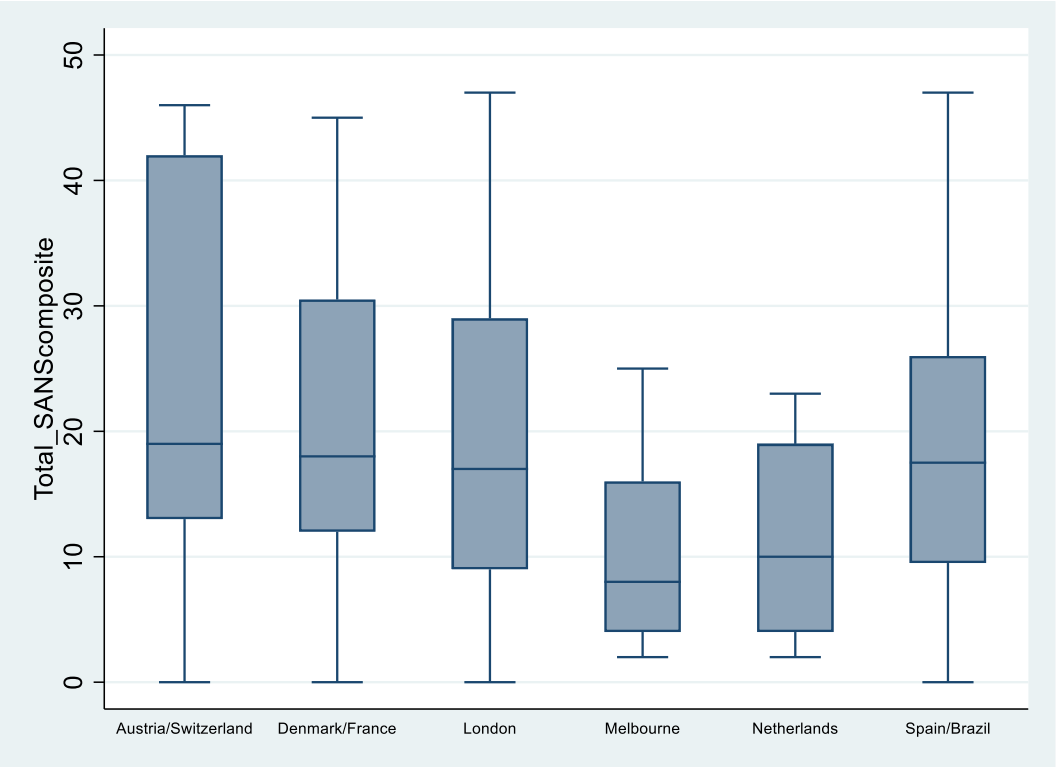
Chi-sq (5 d.f.): 15.838
p = 0.0073

c. Post-hoc Dunnôs test pairwise comparisons with Bonferroni-corrected p values

	Austria/Switzerland	Denmark/France	London	Melbourne	Netherlands
Denmark/France	1.088				
p	1.000				
London	-0.974	-2.714			
p	1.000	0.050			
Melbourne	-0.929	-2.138	-0.201		
p	1.000	0.244	1.000		
Netherlands	-0.853	-2.111	-0.072	0.108	
p	1.000	0.261	1.000	1.000	
Spain/Brazil	-2.366	-3.720	-2.014	-1.456	-1.611
p	0.135	**0.002**	0.330	1.000	0.804

Supplementary Figure 3.19: Scale for Assessment of Negative Symptoms (composite score) stratified by EU-GEI site

a. Box plots for Scale for Assessment of Negative Symptoms (SANS) total composite score, stratified by EU-GEI site

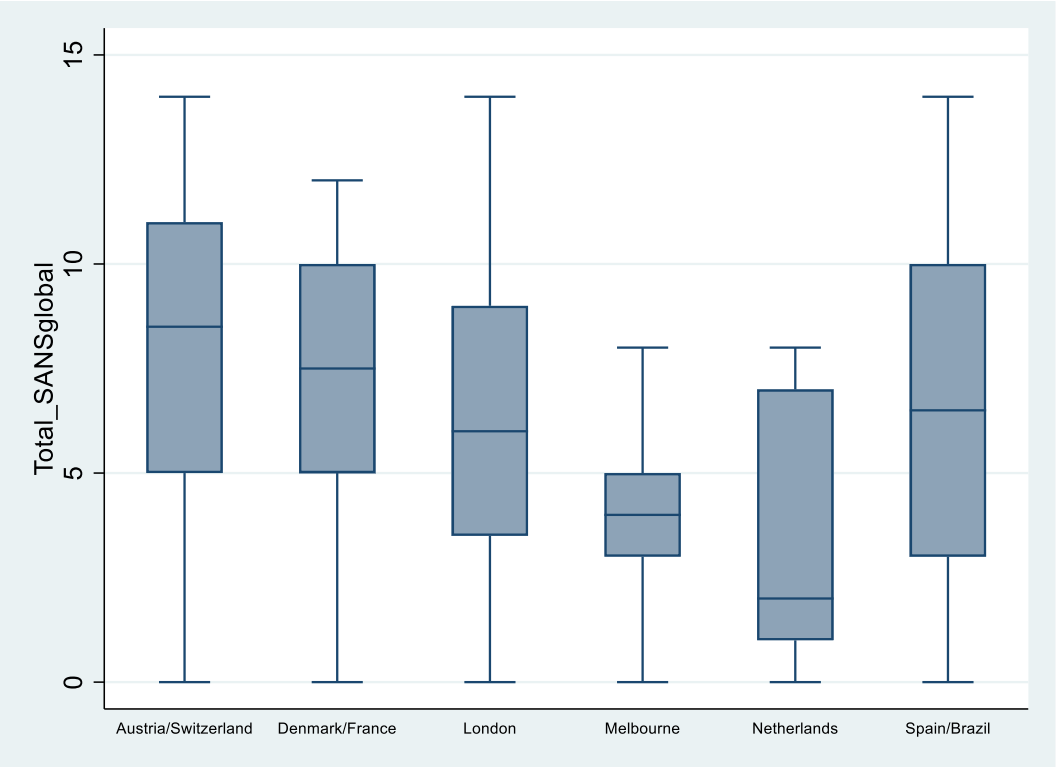


b. Kruskall-Wallis test

Chi-sq (5 d.f.): 9.030
p = 0.1079

Supplementary Figure 3. 20: Scale for Assessment of Negative Symptoms (global score) stratified by EU-GEI site

a. Box plots for Scale for Assessment of Negative Symptoms (SANS) total global score, stratified by EU-GEI site



b. Kruskall-Wallis test

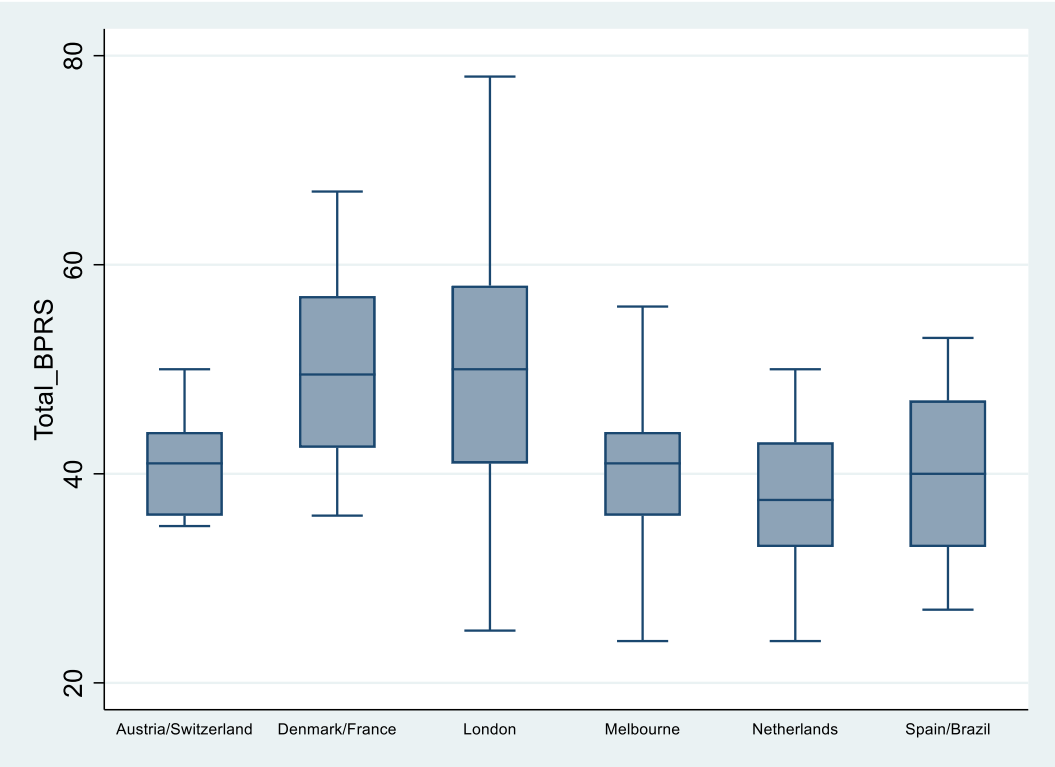
Chi-sq (5 d.f.): 11.823
p = 0.0373

c. Post-hoc Dunnôs test pairwise comparisons with Bonferroni-corrected p values

	Austria/Switzerland	Denmark/France	London	Melbourne	Netherlands
Denmark/France	0.159				
p	1.000				
London	0.709	0.747			
p	1.000	1.000			
Melbourne	2.066	2.350	1.991		
p	0.291	0.141	0.349		
Netherlands	2.194	2.540	2.195	0.071	
p	0.212	0.083	0.211	1.000	
Spain/Brazil	0.519	0.460	-0.075	-1.615	-1.739
p	1.000	1.000	1.000	0.797	0.615

Supplementary Figure 3.21: Brief Psychiatric Rating Scale score stratified by EU-GEI site

a. Box plots for Brief Psychiatric Rating Scale (BPRS) total score, stratified by EU-GEI site



b. Kruskall-Wallis test

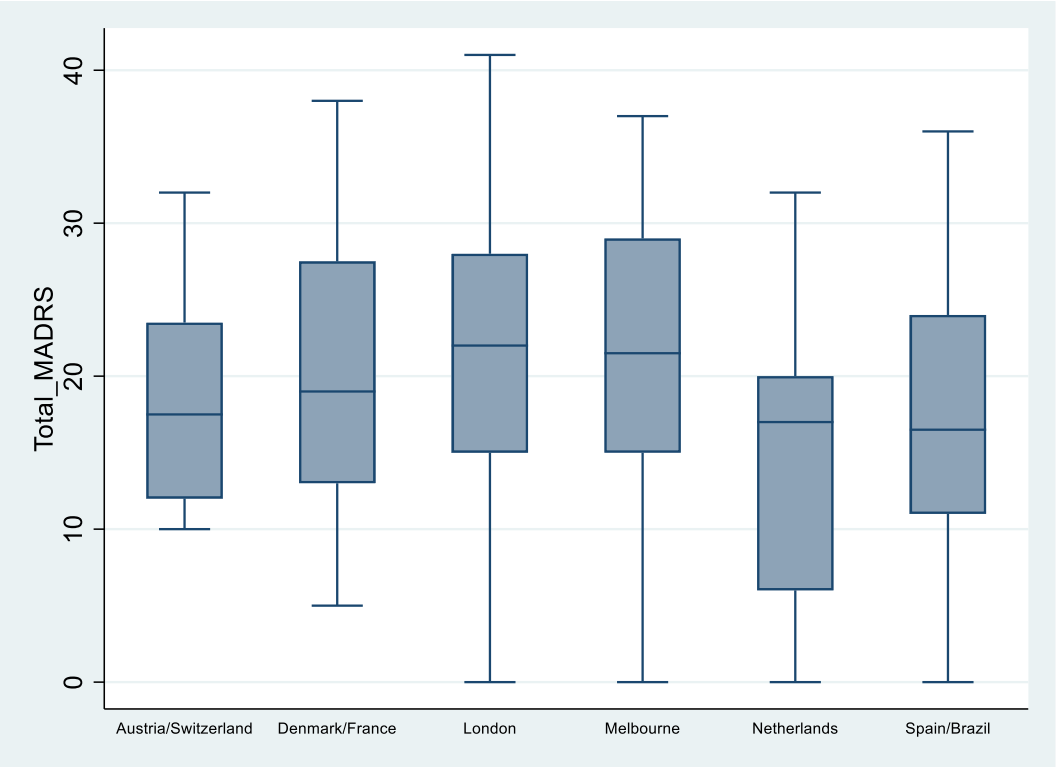
Chi-sq (5 d.f.): 28.083
p = 0.0001

c. Post-hoc Dunnôs test pairwise comparisons with Bonferroni-corrected p-values

	Austria/Switzerland	Denmark/France	London	Melbourne	Netherlands
Denmark/France	-2.426				
p	0.114				
London	-2.603	0.060			
p	0.069	1.000			
Melbourne	0.040	2.613	2.836		
p	1.000	0.067	**0.034**		
Netherlands	0.674	3.434	3.762	0.662	
p	1.000	**0.005**	**0.001**	1.000	
Spain/Brazil	0.083	2.597	2.804	0.045	-0.602
p	1.000	0.071	**0.038**	1.000	1.000

Supplementary Figure 3.22: Montgomery- Åsberg Depression Rating Scale score stratified by EU-GEI site

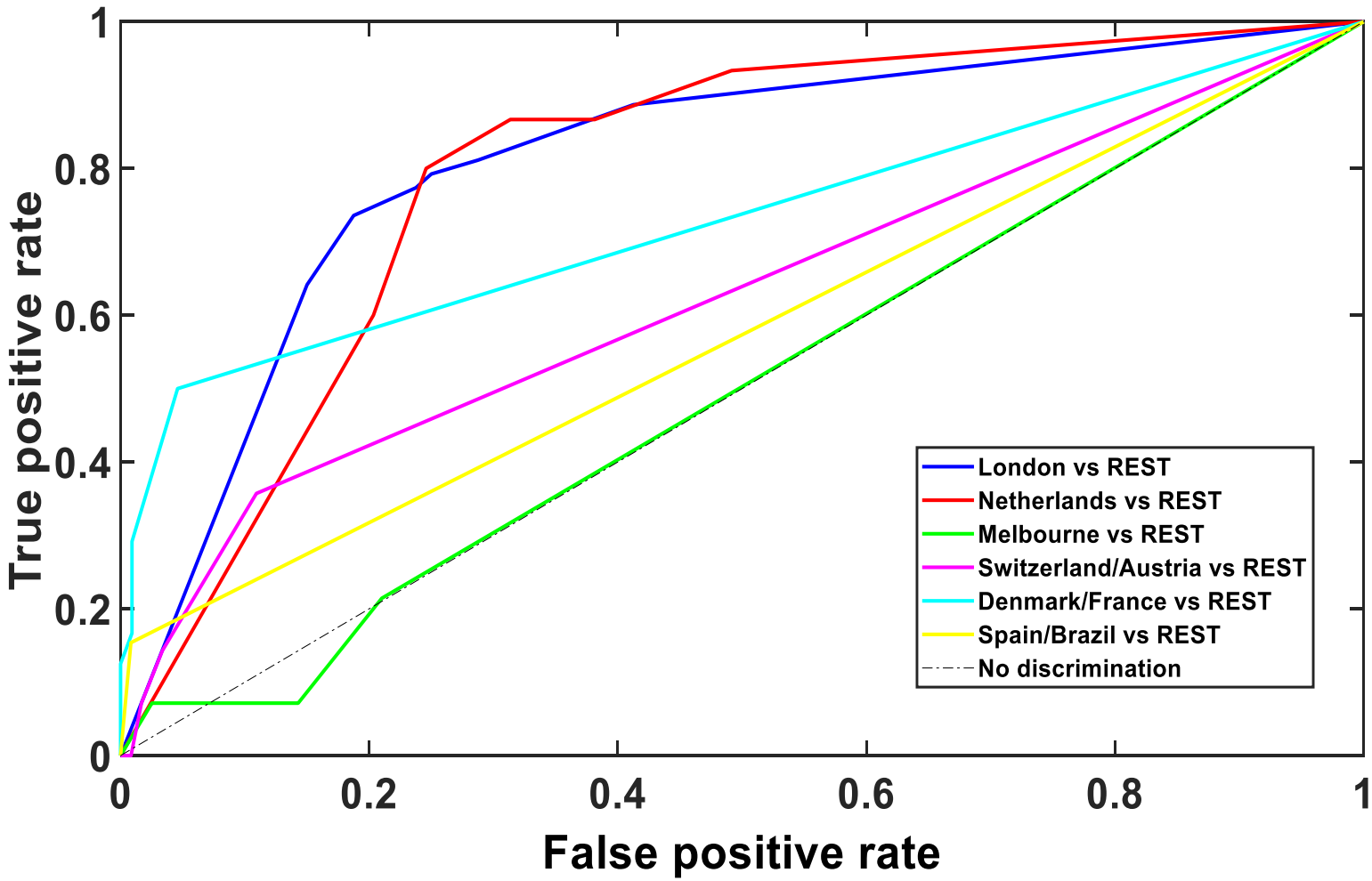
a. Box plots for Montgomery- Åsberg Depression Rating Scale (MADRS) total score, stratified by EU-GEI site



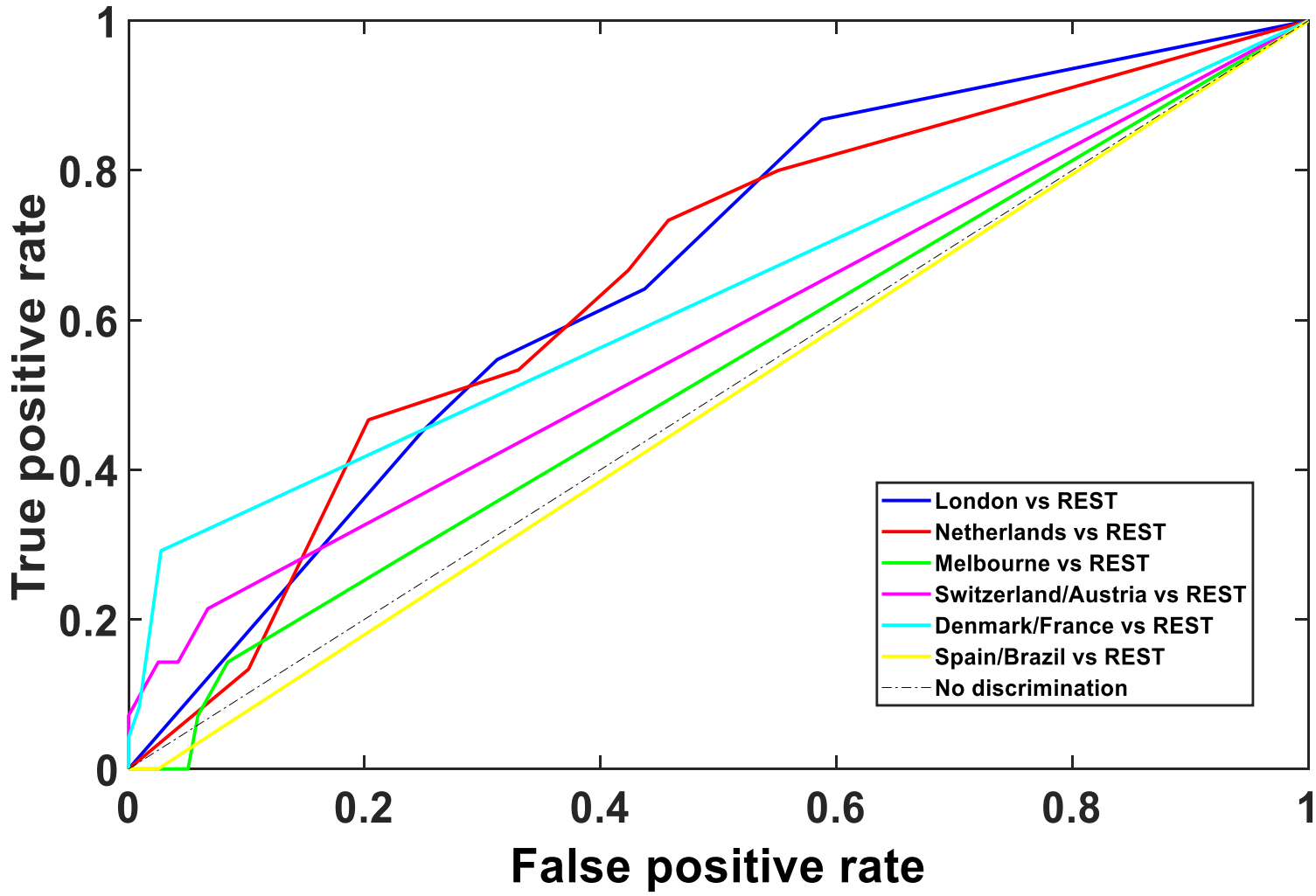
b. Kruskall-Wallis test

Chi-sq (5 d.f.): 5.398
p = 0.3693

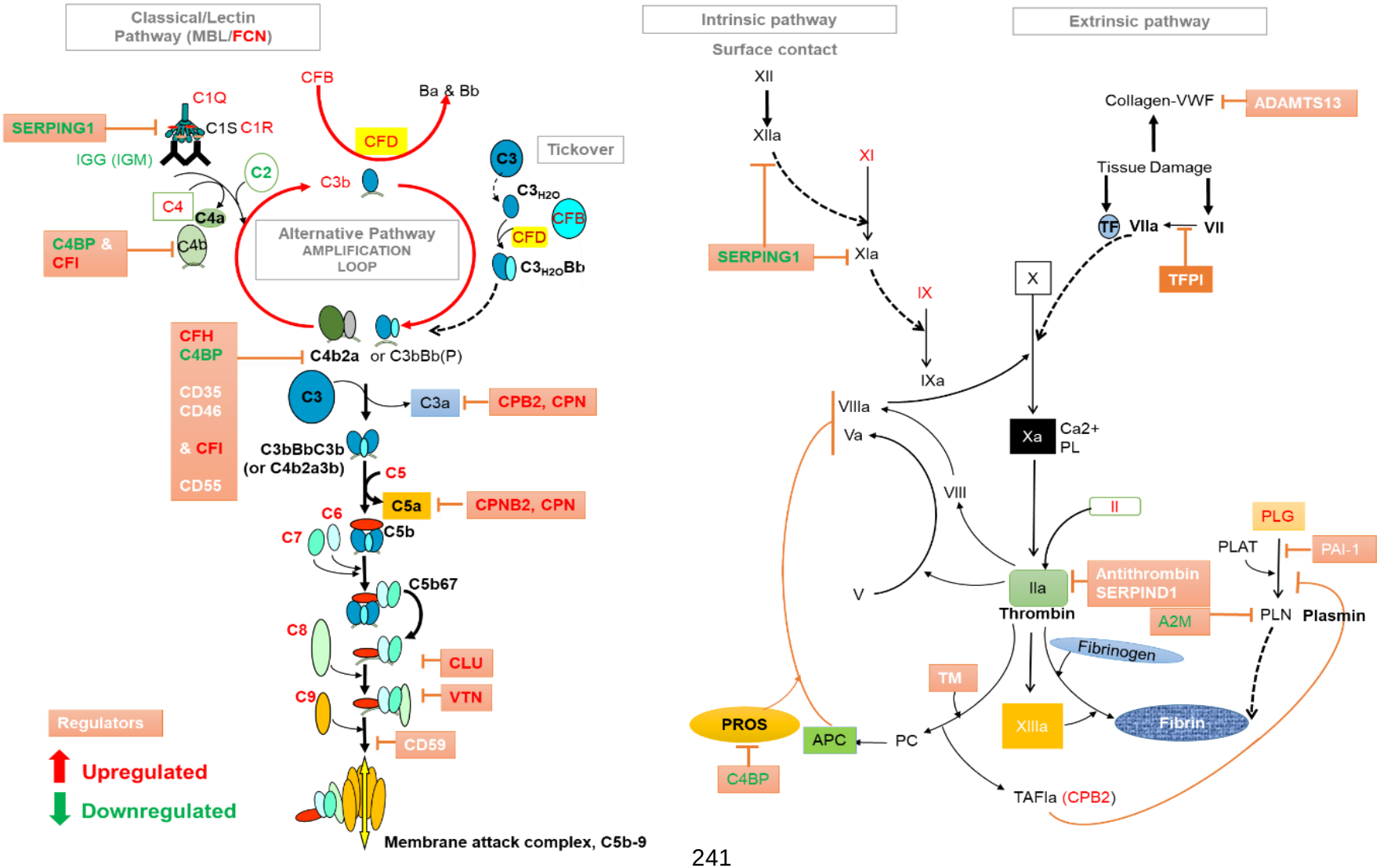
Supplementary Figure 3.23: Multi-class receiver-operating curves for site prediction based on 69 clinical features from model 1b



Supplementary Figure 3.24: Multi-class receiver-operating curves for site prediction based on 166 proteomic features from model 1c



Supplementary Figure 3.25: Illustration of the complement and coagulation pathways depicting the impact model of differentially expressed complement and coagulation proteins in CHR-T vs. CHR-NT



Chapter 4. Plasma polyunsaturated fatty acids and mental disorders in adolescence and early adulthood: cross-sectional and longitudinal associations in a general population cohort

Chapter overview

This chapter comprises a manuscript published in *Translational Psychiatry* examining associations between plasma polyunsaturated fatty acids and mental disorders in adolescence and early adulthood.

Role of the thesis author

Dr Mongan contributed to the conception and design of the study, formulated the research question, performed statistical analyses and interpretation of data, and drafted the initial manuscript and final version of the text following peer review. He was the first and corresponding author on the published manuscript.

Reference: Mongan D, Healy C, Jones HJ, Zammit S, Cannon M, Cotter DR. Plasma polyunsaturated fatty acids and mental disorders in adolescence and early adulthood: cross-sectional and longitudinal associations in a general population cohort. *Translational Psychiatry* 2021;11:321. doi: 10.1038/s41398-021-01425-4

An animation explaining the findings of this study for a general audience is available to view online at https://www.youtube.com/watch?v=vNWNWmIDXuI.

Plasma polyunsaturated fatty acids and mental disorders in adolescence and early adulthood: cross-sectional and longitudinal associations in a general population cohort

Authors: David Mongan[1]*, MB BCh BAO; Colm Healy[1], PhD; Hannah J. Jones[2,3,4], PhD; Stan Zammit[3,4,5], PhD; Mary Cannon[1], PhD; David R. Cotter[1], PhD.

1. Department of Psychiatry, Royal College of Surgeons in Ireland, Dublin, Ireland
2. MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom
3. Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom
4. National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust, University of Bristol, Bristol, United Kingdom
5. Division of Psychological Medicine and Clinical Neurosciences, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, United Kingdom

***Corresponding author:** Dr David Mongan, Department of Psychiatry, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin 9, Ireland.
E-mail: davidmongan@rcsi.ie

Keywords: polyunsaturated fatty acids, omega-6, omega-3, n-6, n-3, psychotic disorder, depressive disorder, generalised anxiety disorder, biomarker, plasma, ALSPAC

Word count (excluding abstract and references): 4998
Number of tables: 4
Number of figures: 3

Appendices:

Appendix 4.1: Supplementary Methods

Appendix 4.2: Supplementary Tables 4.1 ï 4.10

Appendix 4.3: Supplementary Figure 4.1

Abstract

Polyunsaturated fatty acids (PUFAs) may be pertinent to the development of mental disorders, for example via modulation of inflammation and synaptogenesis. We wished to examine cross-sectional and longitudinal associations between PUFAs and mental disorders in a large cohort of young people. Participants in the Avon Longitudinal Study of Parents and Children were interviewed and provided blood samples at two sampling periods when approximately 17 and 24 years old. Plasma PUFA measures (total omega-6 [n-6], total omega-3 [n-3], n-6:n-3 ratio and docosahexaenoic acid [DHA] percentage of total fatty acids) were assessed using nuclear magnetic resonance spectroscopy. Cross-sectional and longitudinal associations between standardised PUFA measures and three mental disorders (psychotic disorder, moderate/severe depressive disorder and generalised anxiety disorder [GAD]) were measured by logistic regression, adjusting for age, sex, body mass index and cigarette smoking. There was little evidence of cross-sectional associations between PUFA measures and mental disorders at age 17. At age 24, the n-6:n-3 ratio was positively associated with psychotic disorder, depressive disorder and GAD, while DHA was inversely associated with psychotic disorder. In longitudinal analyses, there was evidence of an inverse association between DHA at age 17 and incident psychotic disorder at age 24 (adjusted odds ratio 0.44, 95% confidence interval 0.22 ï 0.87) with little such evidence for depressive disorder or GAD. There was little evidence for associations between change in PUFA measures from 17 to 24 years and incident mental disorders at 24 years. These findings provide support for associations between PUFAs and mental disorders in early adulthood, and in particular for DHA in adolescence in relation to prevention of psychosis.

4.1 Introduction

Fatty acids are the major metabolic products of lipid metabolism. They are categorised by number of carbon-carbon double bonds: saturated fatty acids have none, monounsaturated fatty acids have one and polyunsaturated fatty acids (PUFAs) have two or more. PUFAs are further subdivided by the position of the first carbon-carbon double bond from the methyl end of the molecule. This occurs at the sixth carbon atom in omega-6 PUFAs and the third in omega-3 PUFAs. Omega-6 (n-6) PUFAs, such as linoleic acid and arachidonic acid, produce n-6 eicosanoids which have broadly pro-inflammatory effects. Omega-3 (n-3) PUFAs, including alpha-linolenic acid (ALA), eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), produce n-3 eicosanoids which are generally anti-inflammatory (1). Alpha-linolenic acid cannot be synthesised by humans and thus must be obtained directly from dietary intake. Although humans can convert ALA into EPA and DHA, the efficiency of this conversion is limited and thus EPA and DHA must also be obtained from the diet to maintain adequate levels (2). Dietary sources of ALA include flaxseed and soybean oils, while EPA and DHA are found in oily fish or supplements.

Humans may have evolved on a diet with a relatively balanced ratio of n-6:n-3 PUFAs, but in the average modern Western diet this ratio may be 10:1 or higher (3). Given the broadly opposing pro- and anti-inflammatory effects of their lipid mediators (1), the balance of n-6 and n-3 fatty acids is thought to be relevant to disease states characterised by inflammation such as cardiovascular disease (4). There is a body of evidence implicating low-grade inflammation in association with presence or risk of mental disorders such as schizophrenia and depression, at least in a subset of affected patients (5, 6). The precise biological mechanisms by which inflammation may contribute to development of mental disorders are complex but may include modulation of hypothalamic-pituitary-adrenal axis activity, neurotransmission, neurodegeneration and microglial activation (7, 8). A lower ratio of n-6 to n-3 fatty acids may be associated with reduction in inflammation (4) and thus presents one possible mechanism by which dietary factors can modulate inflammatory processes.

Aside from their effects on inflammation, PUFAs have several other biological actions which may be of relevance to mental disorders. For example, through reorganisation of lipid rafts (clusters of membrane proteins and receptors that contribute to signal transduction), n-3 PUFAs can modulate intracellular and intercellular signalling,

cytokine secretion and immune activation (9, 10). n-3 PUFAs also contribute to the structural integrity and fluidity of neuronal membranes (11). Of the n-3 PUFAs, DHA is particularly important in the brain, where it comprises approximately 90% of n-3 PUFA content and 10ï20% of the brain (12). In addition to its neuronal membrane effects, DHA may protect against oxidative damage (13) and promote neurite growth and synaptogenesis (14).

Observational studies provide evidence of decreased n-3 PUFA levels in patients with schizophrenia (15), depressive (16) and anxiety disorders (17) relative to controls. PUFA deficiency has also been implicated in conditions such as attention deficit hyperactivity disorder (18), autism spectrum disorder (19) and Alzheimerôs disease (20). It is however unclear whether PUFA abnormalities precede the onset of mental disorders. In addition, factors such as body mass index (BMI), smoking and socio-economic status (which may be associated with both dietary quality and, independently, mental disorders) may confound observed associations (21-23). In clinical trials, n-3 supplementation for treatment or prevention of mental disorders has produced mixed results (24). For example, in individuals at clinical high-risk of psychosis, n-3 supplementation initially showed beneficial effects in reducing transition to psychotic disorder (25). This finding was not replicated in a subsequent trial (26), although on further analysis several PUFA measures were associated with severity of psychopathology (27, 28). Associations between PUFAs and mental disorders in adolescence and early adulthood have not been extensively studied, even though most mental disorders have their onset during this critical period of development (29).

We sought to address three main questions regarding temporal associations between plasma PUFA measures and three mental disorders (psychotic disorder, depressive disorder and generalised anxiety disorder) in adolescence and early adulthood: 1) Are PUFAs associated with these disorders cross-sectionally at age 17 years? 2) Are PUFAs associated with these disorders cross-sectionally at age 24 years? 3) Are PUFAs at age 17 years associated with these disorders longitudinally at age 24 years?

4.2 Methods

4.2.1 Participants and study design

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a prospective birth cohort study (30-32). The study website contains details of all data available through a fully searchable data dictionary and variable search tool (http://www.bristol.ac.uk/alspac/researchers/access). Pregnant women in Avon, UK with expected delivery dates between 1st April 1991 to 31st December 1992 were invited to participate. 14,541 pregnancies were enrolled with 13,988 children alive at 1 year of age. When the oldest children were approximately 7 years old, an attempt was made to bolster the initial sample with eligible cases who did not join originally. The total sample size for data from age 7 years is 15,454 pregnancies with 14,901 children alive at 1 year of age. Data were collected and managed using REDCap electronic data capture tools (33, 34).

Participants were invited to attend clinics at multiple timepoints where anthropometric measurements, questionnaires, interviews and blood sample collection were completed. The current study was based primarily on data from two such clinics: first when the participants were approximately 17 years old, and subsequently when they were approximately 24 years old.

4.2.2 Exposures

Participants were requested to fast overnight or for at least 6 hours prior to attendance at both the age 17 and age 24 clinics. Blood samples were collected according to a standardised protocol. Samples were obtained between 8am and 2pm in >99% of cases at both clinics. Following collection, samples were centrifuged and stored at ī 80ÁC until analysis. For the age between plasma collection and analysis was 2.4 ï 5.1 years. For the age 24 clinic, this range was 0.3 ï 2.7 years.

Summary and specific fatty acid plasma levels were measured using a high-throughput nuclear magnetic resonance spectroscopy platform (35). We considered four PUFA measures as exposures in the current study: total n-6 fatty acids (mmol/l); total n-3 fatty acids (mmol/l); n-6:n-3 fatty acid ratio; and DHA expressed as percentage of total fatty acids.

4.2.3 Outcomes

We examined three non-mutually exclusive outcomes.

Psychotic disorder: At the age 17 and age 24 clinics, participants completed the semi-structured Psychosis-Like Symptoms Interview (PLIKSi) to assess for psychotic experiences (36). The PLIKSi asks 12 core questions regarding psychotic experiences comprising hallucinations, delusions and experiences of thought interference. Participants who answered óyes cross-questioned to establish whether the experiences were psychotic, and these were coded according to the Schedules for Clinical Assessment in Neuropsychiatry (37). Trained interviewers rated symptoms as ó and coded whether the experience was attributable to sleep or fever. In line with previous studies from this sample (38, 39), psychotic disorder was defined as having at least one definite psychotic experience not attributable to sleep or fever which recurred at least once per month over the previous six months, and was associated with severe distress, marked impairment of t functioning, or led them to seek help from a professional source.

Depressive disorder: At the age 17 and age 24 clinics, participants completed the self-administered computerised version of the Clinical Interview Schedule Revised (CIS-R) (40). The CIS-R includes questions about the recent occurrence, severity, duration and onset of depressive symptoms providing diagnoses of mild, moderate or severe depressive disorder according to the International Classification of Diseases version 10 (ICD-10) (41). In the current study, this outcome was defined as meeting ICD-10 criteria for moderate or severe depressive disorder.

Generalised anxiety disorder (GAD): The CIS-R also contains questions pertaining to anxiety symptoms to determine diagnosis of GAD according to ICD-10 criteria (41). In the current study, this outcome was defined as meeting ICD-10 criteria for GAD.

4.2.4 Confounders

Based on a systematic review of non-dietary factors associated with n-3 PUFA levels (42) we considered the following as a minimal set of confounders: precise age, sex, BMI and cigarette smoking (average number of cigarettes smoked per day in preceding 30 days).

With respect to other candidate confounders, we examined the following additional characteristics contemporaneous with exposures: alcohol intake as measured by Alcohol Use Disorders Identification Test (AUDIT) score (43) at 17 years and AUDIT-Concise (AUDIT-C) score (44) at 24 years; and regular cannabis use (defined as using cannabis more than once monthly at age 17 years, and weekly or daily use at age 24 years).

We examined the following prenatal socio-economic factors: family home ownership status during pregnancy (own home or mortgage/ rented or other) as measured by a questionnaire completed by mothers at 8 weeks gestation; highest maternal educational qualification (less than O-level/ O-level or higher) as measured by a questionnaire completed by mothers at 32 weeks gestation; and parental social class based on occupation (manual/non-manual) of the participantôs (whichever was highest) as measured by a questionnaire completed by mothers at 32 weeks gestation.

Finally, we also examined the following characteristics in childhood: average take-home family income as measured by a questionnaire completed by mothers when the participant was approximately 8 years ol and IQ as measured by the Wechsler Intelligence Scale for Children (Third Edition) (45) at age 8 years.

Further information regarding selection of confounding variables is provided in Supplementary Methods.

4.2.5 Statistical analyses

Statistical analyses were performed in Stata 16 (StataCorp).

Descriptive analyses

We compared sociodemographic characteristics in participants who attended either the age 17 or age 24 clinics versus the rest of the ALSPAC cohort. We also compared characteristics at both 17 and 24 years according to whether PUFA data were available. Finally, we compared characteristics at both ages according to whether participants met criteria for at least one mental disorder, and further stratified by type of disorder.

Categorical variables were summarised using proportions and compared using chi-squared tests. Continuous variables were summarised using the mean and standard deviation or median and interquartile range and compared using two-tailed *t*-tests (with unequal variances specified) or Mann-Whitney U tests as appropriate. $p<0.05$ were considered

Multiple imputation

To address potential bias arising from missing data, we imputed missing data for exposures and confounder variables for participants with outcomes data available. Fifty imputed datasets were generated using multiple imputation with chained equations. In addition to the outcomes, exposures and confounders above, several auxiliary variables were included in the imputation model to make the missing-at-random assumption more plausible. Further details are provided in Supplementary Methods.

Logistic regression analyses

Logistic regression was used to determine associations between exposures (PUFA measures) and outcomes (mental disorders) cross-sectionally at 17 years; cross-sectionally at 24 years; and longitudinally from 17 to 24 years. Two sets of models were used to evaluate longitudinal associations: firstly with PUFA measures at 17 years as the exposure variable, and secondly with change in PUFA measures (PUFA measure at 24 years minus PUFA measure at 17 years) as the exposure variable. Participants who met criteria for the relevant outcome at 17 years were excluded from longitudinal models to enable analysis of incident disorders in early adulthood. Estimates were combined across imputed datasets. PUFA measures were transformed to z-scores such that odds ratios (ORs) derived from logistic regression models may be interpreted per standard deviation increase in PUFA measure.

Confounders

We first assessed crude associations in models that were unadjusted for confounders. The minimum set of confounders (age, sex, BMI and average number of cigarettes smoked per day measured contemporaneously with exposures) were then included as covariates in adjusted models, which were considered as the base models. Longitudinal models examining change in PUFA measures were adjusted for

sex, change in BMI (BMI at 24 years minus BMI at 17 years) and change in average number of cigarettes smoked per day (number of cigarettes at 24 years minus number of cigarettes at 17 years). Finally, in sensitivity analyses, the remaining candidate confounders (AUDIT score at 17 years, regular cannabis use at 17 years, home ownership at 8 weeks gestation, parental social class at 32 weeks gestation, family income at 8 years and IQ at 8 years) were included as covariates one at a time and the effect estimates compared to those for the base models to assess for evidence of further potential confounding.

4.2.6 Ethical approval and consent

Ethical approval for the ALSPAC study was obtained from ALSPAC Ethics and Law Committee and local research ethics committees. Consent for biological samples was collected in accordance with the Human Tissue Act (2004). Informed consent for use of questionnaire and clinic data was obtained following recommendations of the ALSPAC Ethics and Law Committee at the time.

4.3 Results

4.3.1 Sample characteristics and descriptive analyses

Sample derivation is summarised in Supplementary Figure 4.1.

Characteristics of participants according to clinic attendance

9919 participants were invited to the age 17 clinic of whom 5215 (52.6%) attended. 9958 participants were invited to the age 24 clinic of whom 4019 (40.4%) attended. Participants who attended one or both clinics (*n*=6087) differed compared to the rest of the initially-recruited ALSPAC sample on multiple socio-demographic characteristics (Supplementary Table 4.1).

Characteristics of participants according to PUFA data availability

Of 5215 participants who attended the age 17 clinic, 3167 (60.7%) had PUFA data available. One participant with clearly outlying PUFA data was excluded, resulting in 3166 participants. Of 4019 participants who attended the age 24 clinic, 3257 (81.0%) had PUFA data available.

There was evidence of differences on several characteristics between participants with and without PUFA data available at 17 years (Supplementary Table 4.2) and at 24 years (Supplementary Table 4.3) including precise age, sex, BMI, alcohol intake,

smoking status, socio-economic factors, childhood IQ and presence of moderate/severe depressive disorder.

Characteristics of participants according to mental disorder

At age 17 years, 79 of 4718 participants assessed (1.7%) met criteria for psychotic disorder; 227 of 4563 (5.0%) met criteria for moderate/severe depressive disorder; and 263 of 4563 (5.8%) met criteria for GAD.

At age 24 years, 47 of 3889 participants assessed (1.2%) met criteria for psychotic disorder; 304 of 3966 (7.7%) met criteria for moderate/severe depressive disorder; and 386 of 3957 (9.8%) met criteria for GAD.

Among participants assessed at both 17 and 24 years, 2774 participants did not meet criteria for psychotic disorder at age 17 of which 20 (0.7%) subsequently met criteria for psychotic disorder at age 24. 2662 participants did not meet criteria for moderate/severe depressive disorder at age 17, of which 157 (5.9%) subsequently met criteria for moderate/severe depressive disorder at age 24. 2628 participants did not meet criteria for GAD at age 17, of which 205 (7.8%) subsequently met criteria for GAD at age 24.

There was evidence of differences between participants who met criteria for one or more of these disorders compared to those who did not for several characteristics at both age 17 and age 24 (Table 4.1) including precise age, sex, BMI, smoking status, alcohol intake, regular cannabis use, and prenatal socio-economic factors. Sample characteristics are further stratified according to specific disorder at age 17 years in Supplementary Table 4.4 and at age 24 years in Supplementary Table 4.5.

Table 4.1: Characteristics of age 17 and age 24 samples according to mental disorder status

		Met criteria for psychotic disorder, moderate/severe depressive disorder or generalised anxiety disorder at age 17				Met criteria for psychotic disorder, moderate/severe depressive disorder or generalised anxiety disorder at age 24			
		Yes *n*=444	**No *n*=4490**	***p***	**Missing data, n (%)**	**Yes *n*=549**	**No *n*=3459**	***p***	**Missing data, n (%)**
Age in years, mean (SD)		17.8 (0.4)	17.8 (0.4)	0.303	0 (0%)	24.1 (0.8)	24.0 (0.8)	0.015	0 (0%)
Sex*, n (%)	Female	Ó 7 4 . 9 %	Ó 5 4 . 7 %	<0.001	Ò 0 . 2	Ó 7 5 . 1	Ó 6 0 . 6 %	<0.001	Ò 0 . 2
	Male	Ò 2 5 . 1 %	Ò 4 5 . 3 %			Ò 2 5 . 0	Ò 3 9 . 4 %		
Ethnicity, n (%)	White	374 (94.7%)	3846 (95.7%)	0.349	520 (10.5%)	460 (95.6%)	2954 (96.0%)	0.678	451 (11.3%)
	Non-white	21 (5.3%)	173 (4.3%)			21 (4.4%)	122 (4.0%)		
BMI in kg/m², mean (SD)		23.1 (4.8)	22.8 (4.1)	0.244	100 (2.0%)	25.7 (6.3)	24.8 (4.8)	0.002	42 (1.0%)
Daily smoker, n (%)	No	278 (76.2%)	3385 (88.7%)	<0.001	753 (15.3%)	428 (78.4%)	3039 (89.1%)	<0.001	51 (1.3%)
	Yes	87 (23.8%)	431 yes (11.3%)			118 (21.6%)	372 (10.9%)		

Average number of cigarettes smoked per day if daily smoker, median (IQR)		10 (10)	10 (8)	0.288	753 (15.3%)	10 (7)	9 (7)	0.447	51 (1.3%)
AUDIT score, median (IQR)		8 (7)	7 (6)	<0.001	1083 (21.9%)	N/A	N/A	N/A	N/A
AUDIT-C score, median (IQR)		N/A	N/A	N/A	N/A	5 (4)	5 (3)	0.068	79 (2.0%)
Regular cannabis use[1], n (%)	No	307 (85.5%)	3529 (93.5%)	<0.001	801 (16.2%)	491 (89.6%)	3238 (94.4%)	<0.001	31 (0.8%)
	Yes	52 (14.5%)	245 (6.5%)			57 (10.4%)	191 (5.6%)		
Home ownership status at 8 weeks gestation, n (%)	Own or mortgage	315 (80.4%)	3496 (87.7%)	<0.001	554 (11.2%)	385 (81.2%)	2688 (88.7%)	<0.001	504 (12.6%)
	Rent or other	77 (19.6%)	492 (12.3%)			89 (18.8%)	342 (11.3%)		
Highest parental social class at 32 weeks gestation, n (%)	Non-manual	324 (85.9%)	3420 (87.3%)	0.283	639 (13.0%)	395 (85.0%)	2662 (88.9%)	0.013	549 (13.7%)
	Manual	53 (14.1%)	498 (12.7%)			70 (15.0%)	332 (11.1%)		

Highest maternal educational qualification at 32 weeks gestation, n (%)	O-level or higher	323 (80.0%)	3318 (81.4%)	0.488	452 (9.2%)	403 (83.0%)	2585 (83.1%)	0.578	406 (10.1%)
	Less than O-level	81 (20.0%)	760 (18.6%)			88 (17.0%)	526 (16.9%)		
Take-home family income at age 8 years, n (%)	£400+ per week	158 (52.2%)	1874 (57.0%)	0.102	1344 (27.2%)	200 (53.8%)	1466 (59.0%)	0.056	1151 (28.7%)
	<£400 per week	145 (47.8%)	1413 (43.0%)			172 (46.2%)	1019 (41.0%)		
IQ score at age 8 years, mean (SD)		106.2 (16.1)	107.2 (16.3)	0.255	852 (17.3%)	107.4 (15.9)	108.0 (15.9)	0.461	797 (19.9%)

AUDIT: Alcohol Use Disorders Identification Test; BMI: body mass index; IQR: interquartile range; PUFA: polyunsaturated fatty acid; SD: standard deviation; N/A: not applicable.

[1] Regular cannabis use was defined as using cannabis more than once monthly at age 17 and weekly or daily use at age 24.

* Exact numbers suppressed due to small cell counts for missing data.

4.3.2 Logistic regression analyses

Cross-sectional associations between PUFAs and mental disorders at 17 years

Results from logistic regression analyses examining associations between PUFA measures and mental disorders at 17 years are presented in Table 4.2 and Figure 4.1. There was little evidence of associations between PUFA measures and psychotic disorder, moderate/severe depressive disorder or GAD at 17 years in crude or adjusted models.

Cross-sectional associations between PUFAs and mental disorders at 24 years

Results from logistic regression analyses examining associations between PUFA measures and mental disorders at 24 years are presented in Table 4.3 and Figure 4.2. There was little evidence of associations between total n-6 or n-3 PUFAs and mental disorders at age 24. However, the n-6:n-3 ratio was positively associated with all three outcomes, with the strongest evidence for psychotic disorder (adjusted OR 1.54, 95% CI 1.16 ï 2.03). There was evidence of an inverse association between DHA % of total fatty acids and psychotic disorder (adjusted OR 0.58, 95% CI 0.36 ï 0.93) with weaker evidence for moderate/severe depressive disorder (adjusted OR 0.86, 95% CI 0.73 ï 1.01) and GAD (adjusted OR 0.89, 95% CI 0.78 ï 1.02).

Longitudinal associations between PUFAs at 17 years and incident mental disorders at 24 years

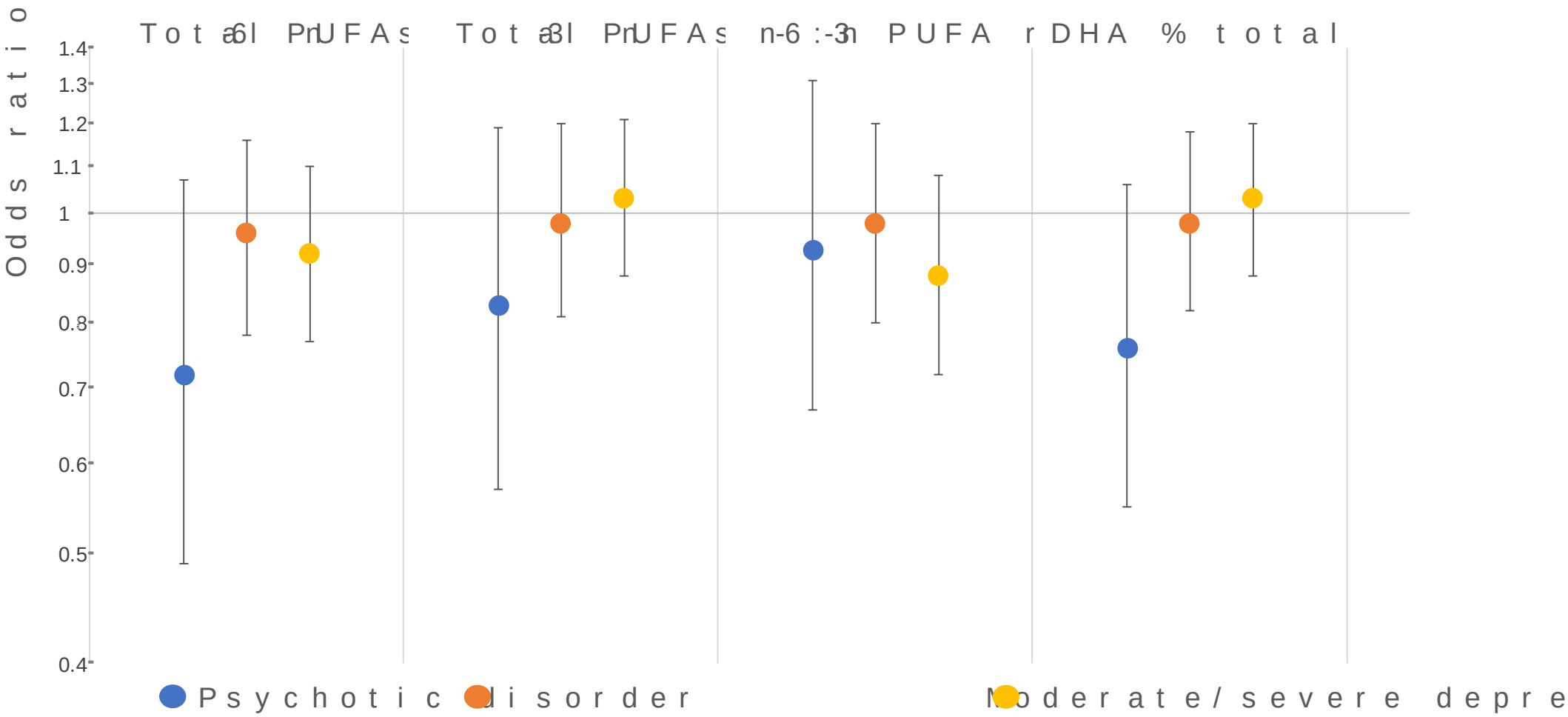
Results from logistic regression analyses examining associations between PUFA measures at 17 years and incident mental disorders at 24 years are presented in Table 4.4 and Figure 4.3. There was little evidence of associations between total n-6 fatty acids, total n-3 fatty acids or the n-6:n-3 ratio at age 17 and incident mental disorders at age 24. There was evidence of an inverse association between DHA % total fatty acids at age 17 and psychotic disorder at age 24 (adjusted OR 0.44, 95% CI 0.22 ï 0.87) but comparatively little evidence for depressive disorder (adjusted OR 0.89, 95% CI 0.70 ï 1.13) or GAD (adjusted OR 0.89, 95% CI 0.74 ï 1.06).

Table 4.2: Cross-sectional associations between plasma polyunsaturated fatty acid (PUFA) measures and odds of mental disorders at 17 years

Outcome (*n* with vs. without outcome)	**Exposure**	**Crude models**			**Adjusted models**		
		Odds ratio	**95% CI**	***p***	**Odds ratio**	**95% CI**	***p***
Psychotic disorder (79 vs. 4639)	Total n-6 PUFAs	0.90	0.65 ï 1.26	0.544	0.72	0.49 ï 1.07	0.107
	Total n-3 PUFAs	0.86	0.60 ï 1.23	0.396	0.83	0.57 ï 1.19	0.305
	n-6:n-3 ratio	1.04	0.76 ï 1.44	0.798	0.93	0.67 ï 1.31	0.688
	DHA % total fatty acids	0.78	0.57 ï 1.07	0.125	0.76	0.55 ï 1.06	0.112
Moderate/ severe depressive disorder (227 vs. 4336)	Total n-6 PUFAs	1.13	0.95 ï 1.34	0.161	0.96	0.78 ï 1.16	0.648
	Total n-3 PUFAs	1.03	0.85 ï 1.24	0.785	0.98	0.81 ï 1.20	0.861
	n-6:n-3 ratio	1.07	0.88 ï 1.30	0.497	0.98	0.80 ï 1.20	0.815
	DHA % total fatty acids	0.99	0.83 ï 1.18	0.918	0.98	0.82 ï 1.18	0.845
Generalised anxiety disorder (263 vs. 4300)	Total n-6 PUFAs	1.07	0.92 ï 1.25	0.373	0.92	0.77 ï 1.10	0.347
	Total n-3 PUFAs	1.08	0.93 ï 1.26	0.316	1.03	0.88 ï 1.21	0.691
	n-6:n-3 ratio	0.95	0.77 ï 1.16	0.582	0.88	0.72 ï 1.08	0.227
	DHA % total fatty acids	1.08	0.93 ï 1.25	0.299	1.03	0.88 ï 1.20	0.731

BMI: body mass index; CI: confidence interval; DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid. PUFA measures were converted to z-scores such that the odds ratios may be interpreted as per standard deviation increase in the PUFA measure at 17 years. Covariates in adjusted models included age at attendance at age 17 clinic, sex, BMI at 17 years and average number of cigarettes smoked per day at 17 years.

Figure 4.1 Odds ratios for cross-sectional associations between plasma polyunsaturated fatty acid measures and mental disorders at 17 years



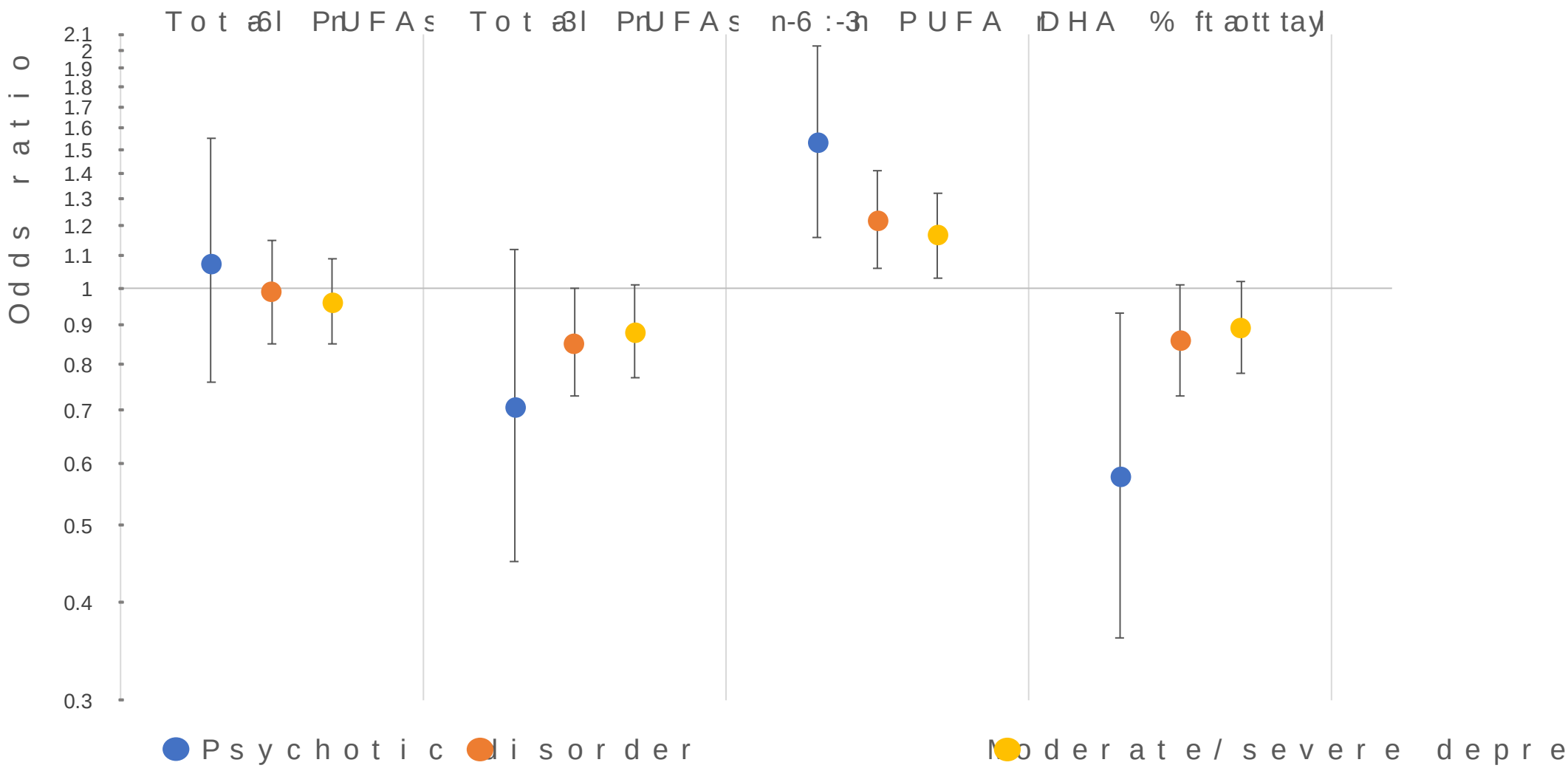
Data are presented on logarithmic scale. Odds ratios are per standard deviation increase in PUFA measure and adjusted for precise age, sex, body mass index and average number of cigarettes smoked per day. Error bars are 95% confidence intervals. DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid.

Table 4.3: Cross-sectional associations between plasma polyunsaturated fatty acid measures and odds of mental disorders at 24 years

Outcome (*n* with vs. without outcome)	**Exposure**	**Crude models**			**Adjusted models**		
		Odds ratio	**95% CI**	***p***	**Odds ratio**	**95% CI**	***p***
Psychotic disorder (47 vs. 3842)	Total n-6 PUFAs	1.15	0.82 ï 1.62	0.420	1.08	0.76 ï 1.55	0.661
	Total n-3 PUFAs	0.66	0.41 ï 1.06	0.083	0.71	0.45 ï 1.12	0.141
	n-6:n-3 ratio	1.72	1.36 ï 2.17	<0.001	1.54	1.16 ï 2.03	0.003
	DHA % total fatty acids	0.55	0.36 ï 0.85	0.007	0.58	0.36 ï 0.93	0.025
Moderate/ severe depressive disorder (304 vs. 3662)	Total n-6 PUFAs	1.07	0.93 ï 1.23	0.366	0.99	0.85 ï 1.15	0.898
	Total n-3 PUFAs	0.83	0.71 ï 0.98	0.026	0.85	0.73 ï 1.00	0.050
	n-6:n-3 ratio	1.37	1.20 ï 1.56	<0.001	1.22	1.06 ï 1.41	0.005
	DHA % total fatty acids	0.85	0.73 ï 0.98	0.027	0.86	0.73 ï 1.01	0.071
Generalised anxiety disorder (386 vs. 3571)	Total n-6 PUFAs	1.03	0.91 ï 1.16	0.692	0.96	0.85 ï 1.09	0.559
	Total n-3 PUFAs	0.88	0.77 ï 1.00	0.053	0.88	0.77 ï 1.01	0.064
	n-6:n-3 ratio	1.27	1.13 ï 1.42	<0.001	1.17	1.03 ï 1.32	0.017
	DHA % total fatty acids	0.90	0.79 ï 1.02	0.085	0.89	0.78 ï 1.02	0.094

BMI: body mass index; CI: confidence interval; DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid. PUFA measures were converted to z-scores such that the odds ratios may be interpreted as per standard deviation increase in the PUFA measure at 24 years. Covariates in adjusted models included precise age at attendance at age 24 clinic, sex, BMI at 24 years and average number of cigarettes smoked per day at 24 years.

Figure 4.2 Odds ratios for cross-sectional associations between plasma polyunsaturated fatty acid measures and mental disorders at 24 years



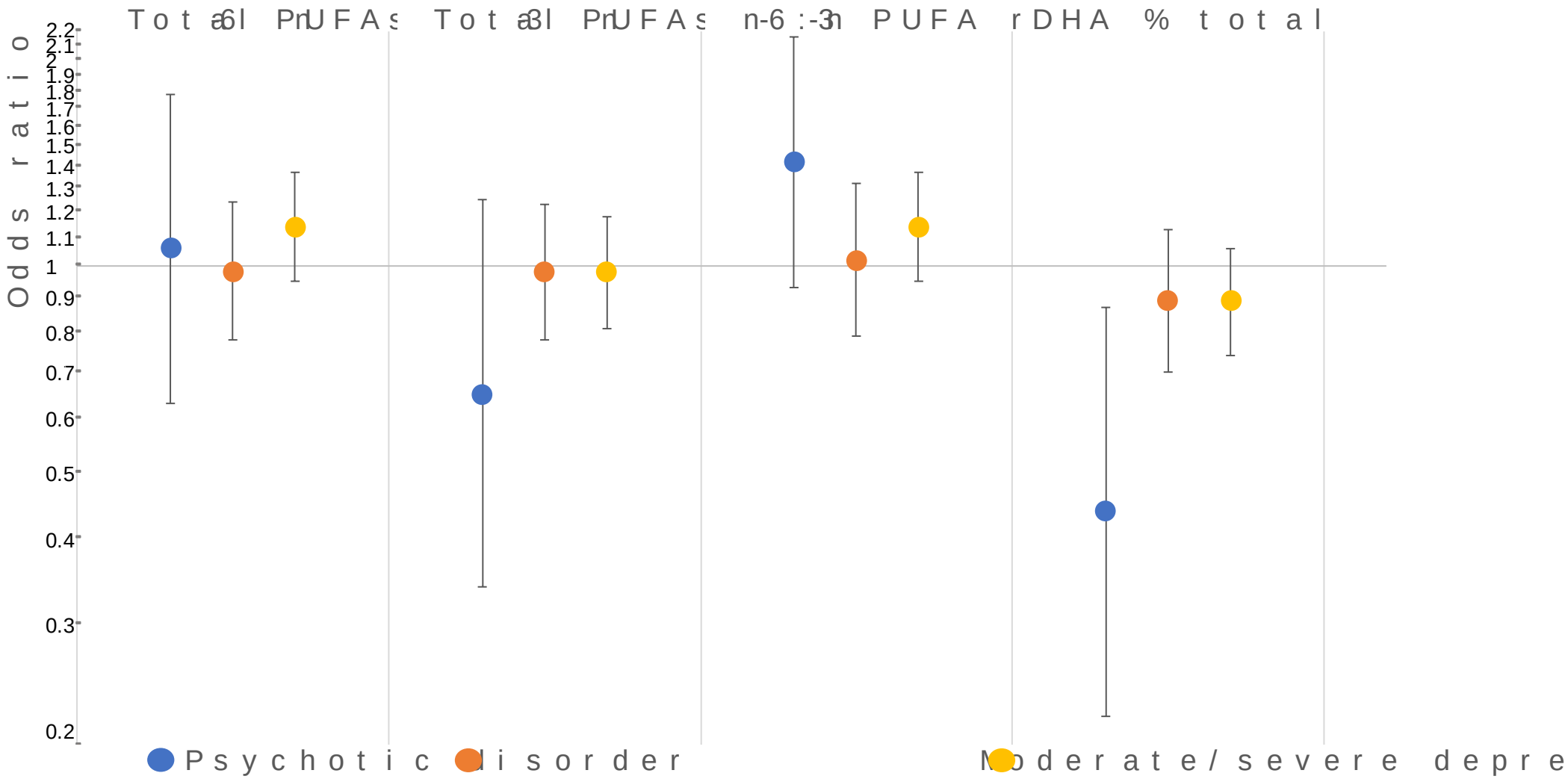
Data are presented on logarithmic scale. Odds ratios are per standard deviation increase in PUFA measure and adjusted for precise age, sex, body mass index and average number of cigarettes smoked per day. Error bars are 95% confidence intervals. DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid.

Table 4.4: Longitudinal associations between plasma polyunsaturated fatty acid measures at 17 years and odds of incident mental disorders at 24 years

Outcome (*n* with vs. without outcome)	**Exposure**	**Crude models**			**Adjusted models**		
		Odds ratio	**95% CI**	***p***	**Odds ratio**	**95% CI**	***p***
Psychotic disorder (20 vs. 2754)	Total n-6 PUFAs	1.14	0.71 ï 1.82	0.590	1.06	0.63 ï 1.78	0.824
	Total n-3 PUFAs	0.69	0.36 ï 1.31	0.254	0.65	0.34 ï 1.25	0.196
	n-6:n-3 ratio	1.43	0.94 ï 2.16	0.092	1.42	0.93 ï 2.16	0.105
	DHA % total fatty acids	0.44	0.23 ï 0.86	0.016	0.44	0.22 ï 0.87	0.019
Moderate/ severe depressive disorder (157 vs. 2505)	Total n-6 PUFAs	1.11	0.90 ï 1.36	0.332	0.98	0.78 ï 1.24	0.876
	Total n-3 PUFAs	1.04	0.84 ï 1.30	0.705	0.98	0.78 ï 1.23	0.876
	n-6:n-3 ratio	1.05	0.81 ï 1.35	0.703	1.02	0.79 ï 1.32	0.852
	DHA % total fatty acids	0.94	0.75 ï 1.18	0.595	0.89	0.70 ï 1.13	0.334
Generalised anxiety disorder (205 vs. 2423)	Total n-6 PUFAs	1.26	1.07 ï 1.48	0.006	1.14	0.95 ï 1.37	0.156
	Total n-3 PUFAs	1.04	0.87 ï 1.24	0.672	0.98	0.81 ï 1.18	0.815
	n-6:n-3 ratio	1.17	0.97 ï 1.41	0.096	1.14	0.95 ï 1.37	0.169
	DHA % total fatty acids	0.91	0.77 ï 1.09	0.311	0.89	0.74 ï 1.06	0.193

BMI: body mass index; CI: confidence interval; DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid. PUFA measures were converted to z-scores such that the odds ratios may be interpreted as per standard deviation increase in the PUFA measure at 17 years. Covariates in adjusted models included age at attendance at age 17 clinic, sex, BMI at 17 years and average number of cigarettes smoked per day at 17 years.

Figure 4.3 Odds ratios for longitudinal associations between plasma polyunsaturated fatty acid measures at 17 years and incident mental disorders at 24 years



Data are presented on logarithmic scale. Odds ratios are per standard deviation increase in PUFA measure and adjusted for precise age, sex, body mass index and average number of cigarettes smoked per day. Error bars are 95% confidence intervals. DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid.

Longitudinal associations between change in PUFAs from 17 to 24 years and incident mental disorders at 24 years

Results from logistic regression analyses examining associations between change in PUFA measures from 17 to 24 years and incident mental disorders at 24 years are presented in Supplementary Table 4.6. There was little evidence of associations between change in PUFA measures and incident mental disorders.

4.3.3 Sensitivity analyses

There was little evidence of substantial confounding for the candidate variables examined based on change in effect estimates compared to the base models (Supplementary Tables 4.7 ï 4.10).

4.4 Discussion

We examined associations between plasma PUFAs and psychotic disorder, moderate/severe depressive disorder and GAD in adolescence and early adulthood in a general population sample. In general, there was little evidence of cross-sectional or longitudinal associations between total n-6 or total n-3 plasma PUFAs and mental disorders. There was also little evidence of cross-sectional associations between any PUFA measure and mental disorders at age 17. Cross-sectionally at age 24, the n-6:n-3 fatty acid ratio was positively associated with psychotic disorder, depressive disorder and GAD. There was evidence for an inverse cross-sectional association between DHA % total fatty acids and psychotic disorder at age 24, but little evidence for depression or GAD. In longitudinal analyses, total n-6 fatty acids, total n-3 fatty acids and the n-6:n-3 ratio at 17 years were not associated with incident mental disorders at 24 years. However, DHA % total fatty acids at age 17 was inversely associated with odds of incident psychotic disorder (but not depression or GAD) at age 24. Finally, there was little evidence of associations between change in PUFA measures from 17 to 24 years and incident mental disorders at 24 years.

Total n-6 and total n-3 levels were not associated with mental disorders cross-sectionally or longitudinally. This contrasts with evidence for cross-sectional associations between the ratio of n-6:n-3 fatty acids and mental disorders at age 24. This might suggest that it is the balance of n-6 to n-3 fatty acids that is of greater importance in relation to these disorders in early adulthood rather than their absolute levels. One possible explanation is that a higher n-6:n-3 ratio predisposes to

increased inflammation (4). Higher levels of circulating inflammatory biomarkers have been detected in association with and preceding the onset of several mental disorders compared to controls (46-50). Increasing the proportion of n-3 PUFAs while decreasing n-6 PUFAs could potentially reduce risk of mental illness (at least in part) by reduction of neuroinflammation, for example via increased synthesis of specialised pro-resolving mediators by n-3 fatty acids (51). However, it remains unclear whether supplementation with n-3 PUFAs, either alone or in combination with dietary changes to reduce the n-6:n-3 ratio, is sufficient to induce putative protective effects on risk of mental illness as mediated via these inflammatory mechanisms (52, 53).

We observed cross-sectional associations between DHA and psychotic disorder at age 24, but not age 17. These findings could suggest that adult-onset psychotic disorder is more associated with PUFA abnormalities than early-onset psychosis. Early-onset psychosis may differ from adult-onset illness in terms of genetic liability and neurodevelopmental trajectory (54, 55), but whether or how these differences relate to PUFA status is unclear and will require further evaluation in future studies. Longitudinally, higher levels of DHA at age 17 were associated with lower odds of psychotic disorder at age 24. Given the low incidence of psychotic disorder in the longitudinal sample (0.7% of participants who were assessed at both timepoints) these results should be interpreted with caution. We can however speculate that any beneficial effects of DHA on reducing psychosis risk are dependent on developmental stage. Indeed, there was little evidence that changes in PUFAs from 17 to 24 years were associated with incident psychosis (Supplementary Table 4.6). This could indicate that adolescent PUFA status may be of greater importance in relation to psychosis risk in early adulthood than change in PUFA status after adolescence. Such sensitivity to developmental stage could also be one reason why clinical trials of n-3 fatty acids in schizophrenia generally suggest greater efficacy in reducing psychotic symptoms in early stages of the disorder compared to patients with chronic illness (56) and is also in keeping with the typical age of onset of schizophrenia in early adulthood (57).

Aside from modulating neuroinflammatory processes as discussed above, DHA has further neurobiological effects that may be of particular relevance to psychosis. The process of synaptic pruning extends normatively through adolescence and early

adulthood (58). Complement system proteins tag synapses for elimination by microglia (59). Microglial activation, leading to excessive synaptic pruning, is thought to contribute to the pathogenesis of schizophrenia (60). DHA is known to modulate activity of glial cells (61). It could be speculated that high against excessive synaptic pruning through modulation of complement system and microglial activity (14, 62). DHA is also thought to play a role in dopaminergic neural development and survival (63). It has been hypothesised that n-3 PUFA deficiency during critical phases of brain development may give rise to dopaminergic dysfunction as observed in schizophrenia (63). It is also possible that DHA depletion occurs as a consequence of underlying biological abnormalities that have been observed in association with schizophrenia, including genetic variation in fatty acid desaturase enzymes, abnormalities in fatty acid binding proteins, pathologically enhanced phospholipase A2 activity, increased oxidative stress and gut microbial dysbiosis (see (64) for a review of these factors).

It should be acknowledged that given the observational nature of this study, it is not possible to ascribe causality, especially for the cross-sectional analyses. It may be the case, for example, that participants with mental disorders in early adulthood ate a poorer diet (lower in n-3 PUFAs or with a higher ratio of n-6 to n-3 fatty acids) as a result of the functional impact of their symptoms. Nonetheless, the same associations were not observed cross-sectionally at age 17 years using the same outcome definitions, which together with the longitudinal findings potentially suggests that reverse causality is not the sole explanation, at least in the case of psychotic disorder. On the other hand, in cross-sectional analyses at age 17 years, participants may have been more likely to follow the diet of their family or carers rather than their own individual dietary choices (which might emerge to a greater extent in early adulthood).

It is unlikely that PUFA status measured at a single point in time is sufficient to predict risk of future psychotic illness, but rather that longer-term deficiency in DHA, particularly in early development and adolescence, leads to enhanced risk in concert with other genetic and environmental factors. In interpreting our findings, we make an underlying presumption that the ósnapshotô p partially reflects PUFA status over a longer duration of exposure. There is evidence that overall adolescent dietary patterns remain moderately stable over time (65-69)

but, without access to shorter-interval repeated measures of plasma PUFAs and dietary information, we cannot definitively measure the accuracy of this presumption in the present data.

Previous general population studies have reported on associations between PUFAs and psychosis phenotypes. For example, a longitudinal study of Swedish women found that those who ate fish 3 ï 4 times per week had reduced risk of subsequent high-frequency psychotic experiences compared to women who never ate fish (70). However, previous investigations in the ALSPAC cohort have found limited evidence of associations between plasma PUFA measures in childhood or mid-adolescence in relation to psychotic experiences or disorder in late adolescence (71, 72). Studies have reported inverse associations between dietary PUFA intake and depression (73, 74) and anxiety disorders (75, 76) in the general population, although conflicting results have also been reported in studies of samples of predominantly older adults (77, 78). A study in adolescents reported an inverse association between self-reported n-3 PUFA intake at age 14 years and depressive symptoms at age 17 years, although this did not persist when adjusted for confounding factors (79). In a longitudinal study of young adults at clinical high risk for psychosis, higher baseline n-6:n-3 ratio and lower DHA levels were predictive of onset of mood disorder on follow-up, but not anxiety or psychosis (80). Mixed results have been reported in trials of PUFAs as treatments for mental disorders, although meta-analyses suggest that EPA (rather than DHA) supplementation may have some efficacy in treatment of depression (81) and that high-dose n-3 preparations may reduce anxiety symptoms (82). It may be the case that only a subset of patients with already-identifiable PUFA abnormalities benefit optimally from these interventions (64).

Regarding implications for prevention, the present data provide tentative evidence to support a role for DHA in adolescence in relation to prevention of psychosis in early adulthood. The results provide less support for prevention of depression or anxiety, in keeping with a systematic review of clinical trials which found that n-3 PUFA supplementation had little effect in prevention of depressive or anxiety symptoms (83). Independent validation of the current findings in samples with larger incidence of psychosis (such as in clinical high-risk for psychosis populations (28)) is required. While we examined multiple potential confounders, the possibility of residual confounding cannot be excluded. Studies implementing causal methods such as

Mendelian randomisation may be helpful in determining whether PUFAs (and particularly DHA) have a causal role in development of psychosis. Further studies examining PUFA exposure earlier in development, such as in early childhood, may be helpful in evaluating longer-term associations with risk of mental disorders in later life. Considering our findings of a longitudinal inverse association between DHA in adolescence and odds of incident psychotic disorder, large-scale adequately powered trials of DHA supplementation or dietary intervention in adolescence with long-term follow-up may facilitate investigation of DHA as a modifiable preventative factor in relation to risk of psychosis and other mental disorders in early adulthood.

4.4.1 Strengths and limitations

Strengths of this study include its focus on a well-characterised large general population cohort of young adults; the use of biomarkers of PUFA status rather than self-reported intake; and the availability of repeated measures enabling longitudinal analyses. However, several limitations should be noted. In common with most longitudinal studies, the ALSPAC sample is subject to attrition. This occurred along a socio-economic gradient and may limit generalisability. Exposure and confounder variables had varying amounts of missing data, and participants differed on several characteristics depending on whether PUFA data were available. We attempted to address this using multiple imputation. Unlike the CIS-R, the PLIKSi does not provide ICD-10-defined diagnoses. However, individuals who fulfilled the definition of psychotic disorder in our study would also likely meet ICD criteria since they were experiencing regular psychotic symptoms associated with severe distress or impairment. The low number of incident cases of psychosis in the longitudinal sample (n=20) may increase the chance of type I error. We did not have access to data for other PUFAs (for example EPA and nervonic acid) and were unable to assess their influence. It is possible that unmeasured PUFAs have stronger or weaker associations with certain mental disorders that cannot be evaluated in this study (14, 81). Unobserved dietary components (such as fruit and vegetable intake or supplement use) and non-dietary factors (such as physical activity) may residually confound the observed associations. While nuclear magnetic resonance spectroscopy facilitates high-throughput reproducible analyses, other methods (such as mass spectrometry) may have higher sensitivity (84) and it is possible that the observed associations may thus be under-estimated due to measurement error.

Samples were stored for variable ranges of time prior to analysis during which degradation could have occurred. There is evidence that PUFAs are reasonably stable for up to seven years of storage (85, 86), however without access to individual-level data for storage time we were unable to investigate this. Finally, in contrast to much of the existing literature, plasma PUFA levels were measured as opposed to erythrocyte membrane PUFA composition. While plasma measures are considered robust biomarkers of PUFA status (87), there are two potential associated limitations. Firstly, plasma levels may correlate less strongly with brain PUFA content in comparison to erythrocyte membrane levels, although this is uncertain in humans (88, 89). Secondly, plasma levels reflect PUFA status over a shorter timeframe (approximately 1 ï 2 weeks) than membrane composition (approximately 1 ï 2 months) because the latter has slower turnover (90). Thus, membrane levels would likely provide a more accurate reflection of long-term PUFA status.

4.4.2 Conclusions

We report evidence of cross-sectional associations between the n-6:n-3 fatty acid ratio and mental disorders in early adulthood, but not late adolescence, in a general population sample. There was evidence of inverse associations between DHA and odds of psychotic disorder cross-sectionally in early adulthood and longitudinally in late adolescence. Long-term trials specifically in adolescent and clinical high-risk for psychosis samples may further clarify the potential role of PUFAs in prevention of mental disorders.

Acknowledgements

We are extremely grateful to all the families who took part in ALSPAC, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. DM is a Fellow on the Irish Clinical Academic Training (ICAT) Programme which is supported by the Wellcome Trust and the Health Research Board (Grant Number 203930/B/16/Z), the Health Service Executive National Doctors Training and Planning and the Health and Social Care, Research and Development Division, Northern Ireland. MC is supported by a European Research Council Consolidator Award (iHEAR 724809). CH is supported by the Health Research Board (HRB ILP-PHR-2019-009). DC is supported by the Health Research Board (HRB POR 2019-0005 and HRB ILP-2017-039). This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number 16/RC/3948 and co-funded under the European Regional Development Fund and by FutureNeuro industry partners. HJ and SZ are supported by the NIHR Biomedical Research Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. The UK Medical Research Council and Wellcome Trust (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and David Mongan will serve as guarantor for the contents of this paper. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). Data collection for primary outcomes data used in this research was funded by the Medical Research Council (MR/M006727/1 and MR/L022206/1) and Wellcome Trust (08426812/Z/07/Z). Data collection for primary exposures data used in this research was joint-funded by the Medical Research Council and Wellcome Trust (102215/Z/13/Z).

Conflicts of interest

DM, MC and DRC report a patent pending (UK Patent Application No. 1919155.0, ñBiomarkers to predict psychosisò). CH, HJH relationships with commercial interests.

Chapter 4 References

1. Schmitz G, Ecker J. The opposing effects in Lipid Research. 2008;47(2):147-55.

2. Burdge G. Alpha-linolenic acid metabolism in men and women: nutritional and biological implications. Current Opinion in Clinical Nutrition and Metabolic Care. 2004;7(2):137-44.

3. Simopoulos AP. Evolutionary aspects of diet, the omega-6/omega-3 ratio and genetic variation: nutritional implications for chronic diseases. Biomedicine & Pharmacotherapy. 2006;60(9):502-7.

4. DiNicolantonio JJ, O'Keefe JH. Importance of maintaining a low omega-6/omega-3 ratio for reducing inflammation. Open Heart. 2018;5(2):e000946-e.

5. Khandaker GM, Dantzer R, Jones PB. Immunopsychiatry: important facts. Psychological Medicine. 2017;47(13):2229-37.

6. Buckley PF. Neuroinflammation and Schizophrenia. Current Psychiatry Reports. 2019;21(8):72.

7. Baumeister D, Russell A, Pariante CM, Mondelli V. Inflammatory biomarker profiles of mental disorders and their relation to clinical, social and lifestyle factors. Social Psychiatry and Psychiatric Epidemiology. 2014;49(6):841-9.

8. Mongan D, Ramesar M, Föcking M, Cannon M, Cotter D. Role of inflammation in the pathogenesis of schizophrenia: A review of the evidence, proposed mechanisms and implications for treatment. Early Intervention in Psychiatry. 2020;14(4):385-97.

9. Reimers A, Ljung H. The emerging role of omega-3 fatty acids as a therapeutic option in neuropsychiatric disorders. Ther Adv Psychopharmacol. 2019;9:2045125319858901-.

10. Turk HF, Chapkin RS. Membrane lipid raft organization is uniquely modified by n-3 polyunsaturated fatty acids. Prostaglandins Leukot Essent Fatty Acids. 2013;88(1):43-7.

11. Cutuli D. Functional and Structural Benefits Induced by Omega-3 Polyunsaturated Fatty Acids During Aging. Curr Neuropharmacol. 2017;15(4):534-42.

12. Weiser MJ, Butt CM, Mohajeri MH. Docosahexaenoic Acid and Cognition throughout the Lifespan. Nutrients. 2016;8(2):99-.

13. Che H et al. Neuroprotective Effects of n-3 Polyunsaturated Fatty Acid-Enriched Phosphatidylserine Against Oxidative Damage in PC12 Cells. Cell Mol Neurobiol. 2018;38(3):657-68.

14. Dyall SC. Long-chain omega-3 fatty acids and the brain: a review of the independent and shared effects of EPA, DPA and DHA. Front Aging Neurosci. 2015;7:52-.

15. Hoen WP et al. Red blood cell polyunsaturated fatty acids measured in red blood cells and schizophrenia: a meta-analysis. Psychiatry Res. 2013;207(1-2):1-12.

16. Lin PY, Huang SY, Su KP. A meta-analytic review of polyunsaturated fatty acid compositions in patients with depression. Biol Psychiatry. 2010;68(2):140-7.

17. Ross BM. Omega-3 polyunsaturated fatty acids and anxiety disorders. Prostaglandins, Leukotrienes and Essential Fatty Acids. 2009;81(5):309-12.

18. Chang JP-C et al. High-dose eicosapentaenoic acid (EPA) improves attention and vigilance in children and adolescents with attention deficit hyperactivity disorder (ADHD) and low endogenous EPA levels. Translational Psychiatry. 2019;9(1):303.

19. Mazahery H et al. Relationship between Long Chain n-3 Polyunsaturated Fatty Acids and Autism Spectrum Disorder: Systematic Review and Meta-Analysis of Case-Control and Randomised Controlled Trials. Nutrients. 2017;9(2):155.

20. Cole GM, Ma Q-L, Frautschy SA. Omega-3 fatty acids and dementia. Prostaglandins Leukot Essent Fatty Acids. 2009;81(2-3):213-21.

21. Saraceno B, Levav I, Kohn R. The public mental health significance of research on socio-economic factors in schizophrenia and major depression. World Psychiatry. 2005;4(3):181-5.

22. Livingstone KM et al. Socioeconomic Inequities in Diet Quality and Nutrient Intakes among Australian Adults: Findings from a Nationally Representative Cross-Sectional Study. Nutrients. 2017;9(10).

23. Darmon N, Drewnowski A. Does social class predict diet quality? Am J Clin Nutr. 2008;87(5):1107-17.

24. Firth J et al. The efficacy and safety of nutrient supplements in the treatment of mental disorders: a meta-review of meta-analyses of randomized controlled trials. World psychiatry. 2019;18(3):308-24.

25. Amminger GP et al. Long-chain omega-3 fatty acids for indicated prevention of psychotic disorders: a randomized, placebo-controlled trial. Archives of General Psychiatry. 2010;67(2):146-54.

26. McGorry PD et al. Omega-3 Polyunsaturated Fatty Acids in Young People at Ultrahigh Risk for Psychotic Disorders: the NEURAPRO Randomized Clinical Trial. JAMA Psychiatry. 2017;74(1):19-27.

27. Berger M et al. Relationship Between Polyunsaturated Fatty Acids and Psychopathology in the NEURAPRO Clinical Trial. Frontiers in Psychiatry. 2019;10(393).

28. Amminger GP et al. The NEURAPRO Biomarker Analysis: Long-Chain Omega-3 Fatty Acids Improve 6-Month and 12-Month Outcomes in Youths at Ultra-High Risk for Psychosis. Biological Psychiatry. 2020;87(3):243-52.

29. de Girolamo G, Dagani J, Purcell R, Cocchi A, McGorry PD. Age of onset of mental disorders and use of mental health services: needs, opportunities and obstacles. Epidemiology and Psychiatric Sciences. 2012;21(1):47-57.

30. Boyd A et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol. 2013;42(1):111-27.

31. Fraser A et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. Int J Epidemiol. 2013;42(1):97-110.

32. Northstone K et al. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019 [version 1; peer review: 2 approved]. Wellcome Open Research. 2019;4(51).

33. Harris PA et al. Research electronic data capture (REDCap)ð A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of Biomedical Informatics. 2009;42(2):377-81.

34. Harris PA et al. The REDCap consortium: Building an international community of software platform partners. Journal of Biomedical Informatics. 2019;95:103208.

35. Soininen P et al. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. The Analyst. 2009;134(9):1781-5.

36. Horwood J et al. IQ and non-clinical psychotic symptoms in 12-year-olds: results from the ALSPAC birth cohort. The British Journal of Psychiatry. 2008;193(3):185-91.

37. World Health Organisation. Division of Mental Health. Schedules for clinical assessment in neuropsychiatry : version 2. World Health Organisation; 1994.

38. Zammit S et al. Psychotic Experiences and Psychotic Disorders at Age 18 in Relation to Psychotic Experiences at Age 12 in a Longitudinal Population-Based Cohort Study. American Journal of Psychiatry. 2013;170(7):742-50.

39. Sullivan SA et al. A Population-Based Cohort Study Examining the Incidence and Impact of Psychotic Experiences From Childhood to Adulthood, and Prediction of Psychotic Disorder. American Journal of Psychiatry. 2020;177(4):308-17.

40. Lewis G. Assessing psychiatric disorder with a human interviewer or a computer. Journal of Epidemiology and Community Health. 1994;48(2):207-10.

41. World Health Organisation. ICD-10 : international statistical classification of diseases and related health problems : tenth revision. 2nd ed ed. Geneva: World Health Organisation; 2004.

42. de Groot RHM, Emmett R, Meyer BJ. Non-dietary factors associated with n-3 long-chain PUFA levels in humans - a systematic literature review. Br J Nutr. 2019;121(7):793-808.

43. World Health Organisation. AUDIT: the Alcohol Use Disorders Identification Test : guidelines for use in primary health care / Thomas F. Babor ... [et al.]. 2nd ed. Geneva: World Health Organization; 2001.

44. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Archives of internal medicine. 1998;158(16):1789-95.

45. Wechsler D. Wechsler Intelligence Scale for Children (3rd ed.). San Antonio, Texas: The Psychological Corporation; 1991.

46. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB. Association of Serum Interleukin 6 and C-Reactive Protein in Childhood With Depression and Psychosis in Young Adult Life: A Population-Based Longitudinal Study. JAMA Psychiatry. 2014;71(10):1121-8.

47. Miller BJ, Goldsmith DR. Inflammatory biomarkers in schizophrenia: Implications for heterogeneity and neurobiology. Biomarkers in Neuropsychiatry. 2019;1:100006.

48. Costello H, Gould RL, Abrol E, Howard R. Systematic review and meta-analysis of the association between peripheral inflammatory cytokines and generalised anxiety disorder. BMJ Open. 2019;9(7):e027925.

49. Goldsmith DR, Rapaport MH, Miller BJ. A meta-analysis of blood cytokine network alterations in psychiatric patients: comparisons between schizophrenia, bipolar disorder and depression. Molecular Psychiatry. 2016;21(12):1696-709.

50. Köhler CA et al. Peripheral cytokine and chemokine alterations in depression: a meta-analysis of 82 studies. Acta Psychiatrica Scandinavica. 2017;135(5):373-87.

51. Joffre C, Rey C, Layé S. N-3 Polyunsaturated Fatty Acids and the Resolution of Neuroinflammation. Frontiers in Pharmacology. 2019;10:1022-.

52. Aucoin M, LaChance L, Cooley K, Kidd S. Diet and Psychosis: A Scoping Review. Neuropsychobiology. 2020;79(1-2):20-42.

53. Marx W, Moseley G, Berk M, Jacka F. Nutritional psychiatry: the present state of the evidence. Proceedings of the Nutrition Society. 2017;76(4):427-36.

54. Fernandez A et al. Childhood-Onset Schizophrenia: A Systematic Overview of Its Genetic Heterogeneity From Classical Studies to the Genomic Era. Front Genet. 2019;10:1137-.

55. Ahn K et al. High rate of disease-related copy number variations in childhood onset schizophrenia. Molecular Psychiatry. 2014;19(5):568-72.

56. Chen AT, Chibnall JT, Nasrallah HA. A meta-analysis of placebo-controlled trials of omega-3 fatty acid augmentation in schizophrenia: Possible stage-specific effects. Ann Clin Psychiatry. 2015;27(4):289-96.

57. Gogtay N, Vyas NS, Testa R, Wood SJ, Pantelis C. Age of onset of schizophrenia: perspectives from structural neuroimaging studies. Schizophrenia Bull. 2011;37(3):504-13.

58. Petanjek Z et al. Extraordinary neoteny of synaptic spines in the human prefrontal cortex. Proceedings of the National Academy of Sciences. 2011;108(32):13281.

59. Presumey J, Bialas AR, Carroll MC. Complement System in Neural Synapse Elimination in Development and Disease. Adv Immunol. 2017;135:53-79.

60. Sellgren CM et al. Increased synapse elimination by microglia in schizophrenia patient-derived models of synaptic pruning. Nature neuroscience. 2019;22(3):374-85.

61. Heras-Sandoval D, Pedraza-Chaverri J, Pérez-Rojas JM. Role of docosahexaenoic acid in the modulation of glial cells in Alzheimer's disease. J Neuroinflammation. 2016;13(1):61-.

62. Madore C et al. Essential omega-3 fatty acids tune microglial phagocytosis of synaptic elements in the mouse developing brain. Nature Communications. 2020;11(1):6133.

63. Healy-Stoffel M, Levant B. N-3 (Omega-3) Fatty Acids: Effects on Brain Dopamine Systems and Potential Role in the Etiology and Treatment of Neuropsychiatric Disorders. CNS Neurol Disord Drug Targets. 2018;17(3):216-32.

64. Hsu M-C, Huang Y-S, Ouyang W-C. Beneficial effects of omega-3 fatty acid supplementation in schizophrenia: possible mechanisms. Lipids in Health and Disease. 2020;19(1):159.

65. Cutler GJ, Flood A, Hannan P, Neumark-Sztainer D. Major patterns of dietary intake in adolescents and their stability over time. J Nutr. 2009;139(2):323-8.

66. Movassagh EZ, Baxter-Jones ADG, Kontulainen S, Whiting SJ, Vatanparast H. Tracking Dietary Patterns over 20 Years from Childhood through Adolescence into Young Adulthood: The Saskatchewan Pediatric Bone Mineral Accrual Study. Nutrients. 2017;9(9):990.

67. Ambrosini GL, Emmett PM, Northstone K, Jebb SA. Tracking a dietary pattern associated with increased adiposity in childhood and adolescence. Obesity (Silver Spring). 2014;22(2):458-65.

68. Biazzi Leal D et al. Changes in Dietary Patterns from Childhood to Adolescence and Associated Body Adiposity Status. Nutrients. 2017;9(10):1098.

69. Harris C et al. Changes in dietary intake during puberty and their determinants: results from the GINIplus birth cohort study. BMC Public Health. 2015;15(1):841.

70. Hedelin M et al. Dietary intake of fish, omega-3, omega-6 polyunsaturated fatty acids and vitamin D and the prevalence of psychotic-like symptoms in a cohort of 33 000 women from the general population. BMC Psychiatry. 2010;10(1):38.

71. Ramsay H et al. Cognition, psychosis risk and metabolic measures in two adolescent birth cohorts. Psychological Medicine. 2018;48(15):2609-23.

72. Thompson AD, Jones HJ, Heron J, Hibbeln J, Sullivan S, Zammit S. Omega-3 and Omega-6 fatty acids and risk of psychotic outcomes in the ALSPAC birth cohort. Schizophrenia Research. 2020;224:108-15.

73. Lai JS et al. A systematic review and meta-analysis of dietary patterns and depression in community-dwelling adults. Am J Clin Nutr. 2014;99(1):181-97.

74. Grosso G et al. Dietary n-3 PUFA, fish consumption and depression: A systematic review and meta-analysis of observational studies. J Affect Disord. 2016;205:269-81.

75. Jacka FN et al. Dietary intake of fish and PUFA, and clinical depressive and anxiety disorders in women. British Journal of Nutrition. 2013;109(11):2059-66.

76. Natacci L et al. Omega 3 Consumption and Anxiety Disorders: A Cross-Sectional Analysis of the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). Nutrients. 2018;10(6):663.

77. Astorg P et al. Long-chain n-3 fatty acid levels in baseline serum phospholipids do not predict later occurrence of depressive episodes: a nested case-control study within a cohort of middle-aged French men and women. Prostaglandins Leukot Essent Fatty Acids. 2009;81(4):265-71.

78. Persons JE et al. Omega-3 fatty acid biomarkers and subsequent depressive symptoms. Int J Geriatr Psychiatry. 2014;29(7):747-57.

79. Oddy WH et al. Dietary intake of omega-3 fatty acids and risk of depressive symptoms in adolescents. Depression and Anxiety. 2011;28(7):582-8.

80. Berger ME et al. Omega-6 to omega-3 polyunsaturated fatty acid ratio and subsequent mood disorders in young people with at-risk mental states: a 7-year longitudinal study. Translational Psychiatry. 2017;7(8):e1220-e.

81. Liao Y et al. Efficacy of omega-3 PUFAs in depression: A meta-analysis. Translational Psychiatry. 2019;9(1):190.

82. Su K-P et al. Association of Use of Omega-3 Polyunsaturated Fatty Acids With Changes in Severity of Anxiety Symptoms: A Systematic Review and Meta-analysis. JAMA Netw Open. 2018;1(5):e182327-e.

83. Deane KHO et al. Omega-3 and polyunsaturated fat for prevention of depression and anxiety symptoms: systematic review and meta-analysis of randomised trials. Br J Psychiatry. 2021;218(3):135-42.

84. Emwas A-H, Salek R, Griffin J, Merzaban J. NMR-based metabolomics in human disease diagnosis: Applications, limitations, and recommendations. Metabolomics. 2013;9.

85. Jonasdottir HS, Brouwers H, Toes REM, Ioan-Facsinay A, Giera M. Effects of anticoagulants and storage conditions on clinical oxylipid levels in human plasma. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids. 2018;1863(12):1511-22.

86. Wagner-Golbs A et al. Effects of Long-Term Storage at -80 °C on the Human Plasma Metabolome. Metabolites. 2019;9(5).

87. Serra-Majem L, Nissensohn M, Øverby NC, Fekete K. Dietary methods and biomarkers of omega 3 fatty acids: a systematic review. British Journal of Nutrition. 2012;107(S2):S64-S76.

88. Guest J, Garg M, Bilgin A, Grant R. Relationship between central and peripheral fatty acids in humans. Lipids in Health and Disease. 2013;12(1):79.

89. Kuratko CN, Salem N. Biomarkers of DHA status. Prostaglandins, Leukotrienes and Essential Fatty Acids. 2009;81(2):111-8.

90. Sun Q, Ma J, Campos H, Hankinson SE, Hu FB. Comparison between plasma and erythrocyte fatty acid content as biomarkers of fatty acid intake in US women. Am J Clin Nutr. 2007;86(1):74-81.

Chapter 4 Appendices

Appendix 4.1: Supplementary Methods é 2 7

Appendix 4.2: Supplementary Tables é . . .

Supplementary Table 4.1: Comparison of sociodemographic characteristics for participants who did versus did not attend age 17 and/or age 24 clinics

Supplementary Table 4.2: Comparison of characteristics for participants with and without polyunsaturated fatty acid data available at age 17 years (total age 17 sample *n*=5215)

Supplementary Table 4.3: Comparison of characteristics for participants with and without polyunsaturated fatty acid data available at age 24 years (total age 24 sample *n*=4019)

Supplementary Table 4.4: Characteristics of age 17 sample according to specific mental disorder at age 17 years

Supplementary Table 4.5: Characteristics of age 24 sample according to specific mental disorder at age 24 years

Supplementary Table 4.6: Effect estimates for associations between change in plasma polyunsaturated fatty acid measures from 17 to 24 years and odds of incident mental disorders at 24 years

Supplementary Table 4.7: Odds ratios for cross-sectional associations between PUFA measures and mental disorders at age 17 years in models with additional adjustment for further potential confounding factors

Supplementary Table 4.8: Odds ratios for cross-sectional associations between PUFA measures and mental disorders at age 24 years in models with additional adjustment for further potential confounding factors

Supplementary Table 4.9: Odds ratios for longitudinal associations between PUFA measures at age 17 years and incident mental disorders at age 24 years in models with additional adjustment for further potential confounding factors

Supplementary Table 4.10: Odds ratios for longitudinal associations between change in PUFA measures from age 17 to age 24 years and incident mental disorders at age 24 years in models with additional adjustment for further potential confounding factors

Appendix 4.3: Supplementary Figure é 2 9

Supplementary Figure 4.1: Flow diagram depicting sample derivation

Appendix 4.1: Supplementary Methods

Further details regarding confounder variables

Age, sex, BMI, tobacco smoking

Age, sex, BMI and tobacco smoking were included as confounders on the basis of a recent systematic review of non-dietary factors associated with n-3 PUFA levels in humans (1). At both age 17 and age 24 clinics, weight was measured using electronic scales to the nearest 50g and height with a Harpenden stadiometer to the nearest cm. The tobacco smoking variable was the average number of cigarettes smoked per day in the preceding 30 days, as derived from a questionnaire regarding tobacco smoking history administered at both age 17 and age 24 clinics.

Alcohol intake

There is little evidence for a general association between alcohol consumption and polyunsaturated fatty acids (PUFAs), although there is evidence for an association with wine drinking specifically (1). It has therefore been recommended to consider adjusting for alcohol, but specified according to type and amount (1). We did not have access to data on the specific types of alcohol consumed. However, at age 17 participants completed the Alcohol Use Disorders Identification Test (AUDIT) (2). The AUDIT contains questions on alcohol consumption, dependence behaviour and alcohol-related problems with a maximum score of 40 points. At age 24 participants completed the Alcohol Use Disorders Identification Test ï Concise (AUDIT-C) (3), a shorter screening form of the AUDIT, with a maximum score of 12 points.

Regular cannabis use

PUFAs may interact biologically with endocannabinoid metabolism (4, 5). Cannabis use may be associated with psychotic, depressive and anxiety symptoms (6-10). We hence examined regular use of cannabis as a potential confounding factor.

At age 17, participants were asked about frequency of cannabis use in the past year: monthly or less; 2 ï 4 times per month; 2 ï 3 times per week; or 4+ times per week. We collapsed the latter three categories and defined regular cannabis use at age 17 years as using cannabis more than once monthly (yes/ no).

At age 24, participants were again asked about frequency of cannabis use in the past year: not in the past year; once or twice; less than monthly; weekly; daily or almost daily. We collapsed the latter two categories and defined regular cannabis use at age 24 years as using cannabis on a weekly or daily basis (yes/ no).

Prenatal and childhood socio-economic factors

Given evidence of associations between socio-economic status and PUFA intake (11, 12) and between socio-economic status and mental disorders (13) we considered several socio-economic variables as potential confounders.

We examined family home ownership status in early pregnancy (own home or mortgage/ rent or other) as measured by questionnaire completed by mothers at 8 weeks gestation. We examined social class (manual/ non-manual) of the participant ôs mother or father (whichever wa questionnaire completed by mothers at 32 weeks gestation. We also considered highest maternal educational qualification as measured by questionnaire at 32 weeks gestation (less than O-level/ O-level or higher). O-levels are the standard educational qualification taken by UK students at approximately age 16 years. With respect to childhood, we examined average take-home family income as measured by a questionnaire completed by mothers when the participant was aged approximately 8 years (<£400 per week/ £400 or more per week).

Childhood IQ

PUFA measures may be associated with cognitive function (14). Higher IQ in childhood may be associated with healthier dietary choices in later life (15) while lower IQ may be associated with increased risk of mental disorders (16, 17). ALSPAC participants were invited to attend a clinic when aged approximately 8 years old at which IQ was measured using the Wechsler Intelligence Scale for Children (Third Edition) (18). We examined the total IQ score derived from this measure as a potential confounder.

Further details regarding multiple imputation

We used multiple imputation with chained equations to impute missing data for exposures and covariates. In addition to the outcomes, exposures and confounders included above and in the main analyses, we included several auxiliary variables in the imputation model (based on their relationship with the outcomes of interest and/or missingness) to make the missing-at-random assumption more plausible. These included: the part-white); BMI at age 11; ethnicity psychotic experiences at age 12 according to the Psychosis-Like Symptom Interview (none/ suspected/ definite); Short Mood and Feelings Questionnaire (19) score at age 12; whether or not the participant had smoked tobacco at age 12; whether or not

the participant had drank alcohol at age 12; whether or not the participant had tried cannabis by age 12; whether or not the participant was in education or employment at age 17; Cognitive Style Questionnaire ï Short Form (20) score at age 17; whether or not the participant reported having thoughts of self-harm in the past week at age 17; whether or not the participant had experienced trauma (physical abuse, emotional abuse, domestic violence, sexual abuse or bullying) from birth to age 17 (see (21) for further details); whether or not the participant had reported being in trouble with police in the past year at age 17; whether or not the participant was in employment, education or training at age 24; whether or not the participant reported a history of previous self-harm at age 24; the fatherôs highest qualification prior to birth of the child; whether or not the pregnancy was intentional; whether or not the mother smoked tobacco during the first trimester of pregnancy; maternal Edinburgh Postnatal Depression Scale (22) score measured at 18 weeks gestation; maternal birth age; and the Family Adversity Index (FAI) score (23). The FAI is a cumulative measure consisting of 18 items assessed by postal questionnaires completed by mothers of ALSPAC participants during pregnancy. The specific items included related to multiple socio-economic factors including maternal age and education, housing, financial difficulties, partner relationship, family size, social support, maternal affective disorder, substance abuse and crime. Each individual adversity item was rated one if present and zero if not present, then summed to derive a total score.

Binary variables were imputed in the imputation model using logistic regression, categorical variables using multinomial regression and ordinal variables using ordinal regression. Continuous variables were imputed using predictive mean matching with *k*=10 nearest neighbours. Fifty imputed datasets were created and estimates were combined according to Rubinôs rules (24).

Supplementary References

1. de Groot RHM, Emmett R, Meyer BJ. Non-dietary factors associated with n-3 long-chain PUFA levels in humans - a systematic literature review. Br J Nutr. 2019;121(7):793-808.
2. World Health Organisation. AUDIT: the Alcohol Use Disorders Identification Test: guidelines for use in primary health care / Thomas F. Babor ... [et al.]. 2nd ed ed. Geneva: World Health Organization; 2001.
3. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Archives of Internal Medicine. 1998;158(16):1789-95.
4. Watson JE, Kim JS, Das A. Emerging class of omega-3 fatty acid endocannabinoids & their derivatives. Prostaglandins Other Lipid Mediat. 2019;143:106337.
5. Dyall SC. Interplay Between n-3 and n-6 Long-Chain Polyunsaturated Fatty Acids and the Endocannabinoid System in Brain Protection and Repair. Lipids. 2017;52(11):885-900.
6. Shakoor S, Zavos HMS, McGuire P, Cardno AG, Freeman D, Ronald A. Psychotic experiences are linked to cannabis use in adolescents in the community because of common underlying environmental risk factors. Psychiatry Research. 2015;227(2-3):144-51.
7. Fergusson DM, Horwood LJ, Swain-Campbell NR. Cannabis dependence and psychotic symptoms in young people. Psychological Medicine. 2003;33(1):15-21.
8. Lev-Ran S, Roerecke M, Le Foll B, George TP, McKenzie K, Rehm J. The association between cannabis use and depression: a systematic review and meta-analysis of longitudinal studies. Psychological Medicine. 2014;44(4):797-810.
9. Hayatbakhsh MR, Najman JM, Jamrozik K, Mamun AA, Alati R, Bor W. Cannabis and anxiety and depression in young adults: a large prospective study. Journal of the American Academy of Child and Adolescent Psychiatry. 2007;46(3):408-17.
10. Horwood LJ, Fergusson DM, Coffey C, Patton GC, Tait R, Smart D, et al. Cannabis and depression: An integrative data analysis of four Australasian cohorts. Drug and Alcohol Dependence. 2012;126(3):369-78.
11. Livingstone KM, Olstad DL, Leech RM, Ball K, Meertens B, Potter J, et al. Socioeconomic Inequities in Diet Quality and Nutrient Intakes among Australian Adults: Findings from a Nationally Representative Cross-Sectional Study. Nutrients. 2017;9(10):1092.
12. Darmon N, Drewnowski A. Does social class predict diet quality? The American Journal of Clinical Nutrition. 2008;87(5):1107-17.
13. Saraceno B, Levav I, Kohn R. The public mental health significance of research on socio-economic factors in schizophrenia and major depression. World psychiatry. 2005;4(3):181-5.
14. Ramsay H, Barnett JH, Murray GK, Miettunen J, Mäki P, Järvelin MR, et al. Cognition, psychosis risk and metabolic measures in two adolescent birth cohorts. Psychological Medicine. 2018;48(15):2609-23.
15. Wraw C, Der G, Gale CR, Deary IJ. Intelligence in youth and health behaviours in middle age. Intelligence. 2018;69:71-86.
16. Zammit S, Allebeck P, David AS, Dalman C, Hemmingsson T, Lundberg I, et al. A longitudinal study of premorbid IQ Score and risk of developing schizophrenia,

bipolar disorder, severe depression, and other nonaffective psychoses. Archives of General Psychiatry. 2004;61(4):354-60.
17. Koenen KC, Moffitt TE, Roberts AL, Martin LT, Kubzansky L, Harrington H, et al. Childhood IQ and adult mental disorders: a test of the cognitive reserve hypothesis. American Journal of Psychiatry. 2009;166(1):50-7.
18. Wechsler D. Wechsler Intelligence Scale for Children (3rd ed.). San Antonio, Texas: The Psychological Corporation; 1991.
19. Ancold A, Stephen C. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. International Journal of Methods in Psychiatric Research, 5(4), 237ï 249.
20. Meins E, McCarthy-Jones S, Fernyhough C, Lewis G, Bentall RP, Alloy LB. Assessing negative cognitive style: Development and validation of a Short-Form version of the Cognitive Style Questionnaire. Personality and Individual Differences. 2012;52(5):581-5.
21. Croft J, Heron J, Teufel C, Cannon M, Wolke D, Thompson A, et al. Association of Trauma Type, Age of Exposure, and Frequency in Childhood and Adolescence With Psychotic Experiences in Early Adulthood. JAMA Psychiatry. 2018.
22. Cox JL, Holden JM, Sagovsky R. Detection of Postnatal Depression: Development of the 10-item Edinburgh Postnatal Depression Scale. British Journal of Psychiatry. 1987;150(6):782-6.
23. Bowen E, Heron J, Waylen A, Wolke D. Domestic violence risk during and after pregnancy: findings from a British longitudinal study. BJOG. 2005;112(8):1083-9.
24. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.

Appendix 4.2: Supplementary Tables

Supplementary Table 4.1: Comparison of characteristics for participants who did versus did not attend age 17 and/or age 24 clinics

	Attended at age 17 and/or age 24, *n*=6087	Missing data, *n* (%)	Did not attend at age 17 and/or age 24, *n*=9559	Missing data, *n* (%)	Test statistic	*p*
Sex, n (%)	3487 female (57.3%) 2595 male (42.7%)	5 (0.1%)	3861 female (43.1%) 5100 male (56.9%)	598 (6.3%)	G^2 = 294.3	<0.001
Ethnicity, n (%)	5095 white (95.6%) 232 non-white (4.4%)	760 (12.5%)	6432 white (94.4%) 381 non-white (5.6%)	2746 (28.7%)	G^2 = 9.5	0.002
Home ownership status at 8 weeks gestation, n (%)	4557 own home or mortgage (86.0%) 743 rent or other (14.0%)	787 (12.9%)	5329 own home or mortgage (69.0%) 2400 rent or other (31.0%)	1830 (19.1%)	G^2 = 498.3	<0.001
Highest parental social class at 32 weeks gestation, n (%)	4478 non-manual (86.7%) 689 manual (13.3%)	920 (15.1%)	4836 non-manual (75.6%) 1559 manual (24.4%)	3164 (33.1%)	G^2 = 222.5	<0.001
Highest maternal educational qualification at 32 weeks gestation, n (%)	4339 O-level or higher (80.3%) 1068 less than O-level (19.7%)	680 (11.2%)	4398 O-level or higher (62.1%) 2688 less than O-level (37.9%)	2473 (25.9%)	G^2 = 482.2	<0.001
Take-home family income at age 8 years, n (%)	2326 £400+ per week (55.7%) 1852 <£400 per week (44.3%)	1909 (31.4%)	1289 £400+ per week (44.1%) 1632 <£400 per week (55.9%)	6638 (69.4%)	G^2 = 91.7	<0.001
IQ score at age 8 years, mean (SD)	106.6 (16.3)	1369 (22.5%)	99.3 (15.9)	7329 (76.7%)	t = -18.7	<0.001

IQR: interquartile range; SD: standard deviation.

Data for highest parental social class, take-home family income, home ownership status and highest maternal qualification were collected by questionnaires completed by the mother prior to birth of the participant.

O-levels (Ordinary levels) are the standard examinations taken by students in the United Kingdom at approximately age 16.

Supplementary Table 4.2: Comparison of characteristics for participants with and without polyunsaturated fatty acid data available at age 17 years (total age 17 sample *n*=5215)

	PUFA data available, *n*=3166	Missing data, *n* (%)	PUFA data not available, *n*=2049	Missing data, *n* (%)	Test stat-istic	*p*
Age in years at attendance of age 17 clinic, mean (SD)	17.8 (0.4)	0 (0%)	17.9 (0.5)	0 (0%)	t = 7.8	<0.001
Sex, n (%)	Ó52.0% fe Ò48.0% mal	Ò0.2%	Ó63.1% fe Ò36.9% mal	Ò0.2%	G^2 = 61.8	<0.001
Ethnicity, n (%)	2718 white (95.6%) 125 non-white (4.4%)	323 (10.2%)	1707 white (95.6%) 79 non-white (4.4%)	263 (12.8%)	G^2 <0.01	0.966
BMI in kg/m² at age 17 years, mean (SD)	22.7 (3.9)	82 (2.6%)	23.1 (4.6)	71 (3.5%)	t = 2.9	0.004
Daily smoker at age 17 years, n (%)	2316 no (87.6%) 328 yes (12.4%)	522 (16.5%)	1359 no (87.4%) 196 yes (12.6%)	494 (24.1%)	G^2 = 0.03	0.851
Average number of cigarettes smoked per day if daily smoker, median (IQR)	9.5 (10)	522 (16.5%)	10 (6)	494 (24.1%)	z = 0.337	0.736
AUDIT score at age 17 years, median (IQR)	7 (6)	718 (22.7%)	7 (6)	646 (31.5%)	z = 0.3	0.743
Regular cannabis use[a] at age 17 years, n (%)	2416 no (92.4%) 200 yes (7.6%)	550 (17.4%)	1435 no (93.5%) 100 yes (6.5%)	514 (25.1%)	G^2 = 1.8	0.174
Psychotic disorder at age 17 years, n (%)	2943 no (98.5%) 44 yes (1.5%)	179 (5.7%)	1696 no (98.0%) 35 yes (2.0%)	318 (15.5%)	G^2 = 2.0	0.157
Moderate/ severe depressive disorder at age 17 years, n (%)	2774 no (95.8%) 123 yes (4.3%)	269 (8.5%)	1562 no (93.8%) 104 yes (6.2%)	383 (18.7%)	G^2 = 8.9	0.003
Generalised anxiety disorder at age 17 years, n (%)	2741 no (94.6%) 156 yes (5.4%)	269 (8.5%)	1559 no (93.6%) 107 yes (6.4%)	383 (18.7%)	G^2 = 2.1	0.148

Home ownership status at 8 weeks gestation, n (%)	2433 own home or mortgage (86.8%) 369 rent or other (13.2%)	364 (11.5%)	1548 own home or mortgage (85.9%) 255 rent or other (14.1%)	246 (12.0%)	G^2 = 0.9	0.346
Highest parental social class at 32 weeks gestation, n (%)	2416 non-manual (87.4%) 349 manual (12.6%)	401 (12.7%)	1483 non-manual (85.8%) 245 manual (14.2%)	321 (15.7%)	G^2 = 2.2	0.134
Highest maternal educational qualification at 32 weeks gestation, n (%)	2366 O-level or higher (82.2%) 514 less than O-level (17.8%)	286 (9.0%)	1428 O-level or higher (78.3%) 395 less than O-level (21.7%)	226 (11.0%)	G^2 = 10.4	0.001
Take-home family income at age 8 years, n (%)	1337 £400+ per week (57.3%) 995 <£400 per week (42.7%)	834 (26.3%)	752 £400+ per week (46.2%) 752 <£400 per week (53.8%)	652 (31.8%)	G^2 = 4.4	0.037
IQ score at age 8 years, mean (SD)	108.0 (16.5)	546 (17.2%)	105.2 (16.1)	430 (21.0%)	t = -5.4	<0.001

[a] Regular cannabis use at age 17 was defined as using cannabis more than once monthly.
* Data suppressed due to small cell counts
AUDIT: Alcohol Use Disorders Identification Test; IQR: interquartile range; PUFA: polyunsaturated fatty acid; SD: standard deviation.

Supplementary Table 4.3: Comparison of characteristics for participants with and without polyunsaturated fatty acid data available at age 24 years (total age 24 sample *n*=4019)

	PUFA data available, *n*=3257	Missing data, *n* (%)	PUFA data not available, *n*=762	Missing data, *n* (%)	Test stat-istic	*p*
Age in years at attendance of age 17 clinic, mean (SD)	24.0 (0.9)	0 (0%)	24.0 (0.8)	0 (0%)	t = 0.5	0.596
Sex, n (%)	Ó60.2% fe Ò39.8% mal	Ò0.2%'	553 female (72.6%) 209 male (27.4%)	0 (0%)	G^2 = 40.4	<0.001
Ethnicity, n (%)	2781 white (96.3%) 107 non-white (3.7%)	369 (11.3%)	642 white (94.7%) 36 non-white (5.3%)	84 (11.0%)	G^2 = 3.7	0.055
BMI in kg/m² at age 24 years, mean (SD)	24.7 (4.9)	33 (1.0%)	25.8 (5.8)	11 (1.4%)	t = 4.5	<0.001
Daily smoker at age 24 years, n (%)	2827 no (88.2%) 379 yes (11.8%)	51 (1.6%)	641 no (85.2%) 111 yes (14.8%)	10 (1.3%)	G^2 = 4.9	0.028
Average number of cigarettes smoked per day if daily smoker, median (IQR)	10 (7)	51 (1.6%)	10 (7)	10 (1.3%)	z <0.1	0.987
AUDIT-C score at age 24 years, median (IQR)	5 (3)	71 (2.2%)	5 (4)	19 (2.5%)	z = -4.0	<0.001
Regular cannabis use[a] at age 24 years, n (%)	3016 no (93.5%) 209 yes (6.5%)	32 (0.1%)	Ó94.8% no Ò5.2% yes	Ò0.7%	G^2 = 1.8	0.184
Psychotic disorder at age 24 years, n (%)	3127 no (98.8%) 37 yes (1.4%)	93 (2.9%)	715 no (98.6%) 10 yes (1.4%)	37 (4.9%)	G^2 = 0.2	0.641
Moderate/ severe depressive disorder at age 24 years, n (%)	2986 no (92.8%) 230 yes (7.2%)	41 (1.3%)	676 no (90.1%) 74 yes (9.9%)	12 (1.6%)	G^2 = 6.3	0.012

Generalised anxiety disorder at age 24 years, n (%)	2907 no (90.6%) 301 yes (9.4%)	49 (1.5%)	664 no (88.7%) 85 yes (11.4%)	13 (1.7%)	G^2 = 2.7	0.103
Home ownership status at 8 weeks gestation, n (%)	2525 own home or mortgage (88.4%) 331 rent or other (11.6%)	401 (12.3%)	554 own home or mortgage (84.5%) 102 rent or other (15.5%)	106 (13.9%)	G^2 = 7.7	0.005
Highest parental social class at 32 weeks gestation, n (%)	2502 non-manual (88.9%) 314 manual (11.1%)	441 (13.5%)	563 non-manual (86.4%) 89 manual (13.6%)	110 (14.4%)	G^2 = 3.2	0.073
Highest maternal educational qualification at 32 weeks gestation, n (%)	2442 O-level or higher (83.6%) 480 less than O-level (16.4%)	335 (10.3%)	554 O-level or higher (80.4%) 135 less than O-level (19.6%)	73 (9.6%)	G^2 = 4.0	0.047
Take-home family income at age 8 years, n (%)	1372 £400+ per week (58.9%) 958 <£400 per week (41.1%)	927 (28.5%)	301 £400+ per week (56.3%) 234 <£400 per week (43.7%)	227 (29.8%)	G^2 = 1.2	0.267
IQ score at age 8, mean (SD)	108.7 (15.8)	653 (20.0%)	104.4 (16.0)	146 (19.2%)	t = -6.0	<0.001

[a] Regular cannabis use at age 24 was defined as using cannabis weekly or daily.
* Data suppressed due to small cell counts.
AUDIT-C: Alcohol Use Disorders Identification Test ï Concise; IQR: interquartile range; PUFA: polyunsaturated fatty acid; SD: standard deviation.

Supplementary Table 4.4: Characteristics of age 17 sample according to specific mental disorder at age 17 years

		Psychotic disorder				Moderate/severe depressive disorder				Generalised anxiety disorder			
		Yes *n*=79	**No *n*=4639**	***p***	**Missing data, n (%)**	**Yes *n*=227**	**No *n*=4336**	***p***	**Missing data, n (%)**	**Yes *n*=263**	**No *n*=4300**	***p***	**Missing data, n (%)**
Age in years, mean (SD)		17.8 (0.5)	17.8 (0.4)	0.262	0 (0%)	17.8 (0.4)	17.8 (0.4)	0.850	0 (0%)	17.8 (0.4)	17.8 (0.4)	0.293	0 (0%)
Sex*, n (%)	Female	Ó 7 8 . 5	Ó 5 6 . 2	<0.001	Ò 0 . 2	Ó 7 7 . 1 %	Ó 5 5 . 2 %	<0.001	Ò 0 . 2	Ó 7 4 . 8	Ó 5 5 . 2	<0.001	Ò 0 . 2
	Male	Ò 2 1% 5	Ò 4 3 . 8			Ò 2 2 . 9 %	Ò 4 4 . 8 %			Ò 2 5 . 2	Ò 4 4 . 8		
Ethnicity, n (%)	White	63 (92.6%)	3982 (95.8%)	0.203	493 (10.4%)	186 (93.5%)	3729 (95.8%)	0.112	472 (10.3%)	227 (95.8%)	3688 (95.7%)	0.948	472 (10.3%)
	Non-white	5 (7.4%)	175 (4.2%)			13 (6.5%)	163 (4.2%)			10 (4.2%)	166 (4.3%)		
BMI in kg/m² at age 17 years, mean (SD)		23.2 (4.5)	22.8 (4.2)	0.474	73 (1.5%)	23.4 (5.3)	22.8 (4.1)	0.079	85 (1.9%)	22.7 (4.4)	22.8 (4.2)	0.631	85 (1.9%)
Daily smoker at age 17 years, n (%)	No	42 (66.7%)	3499 (88.0%)	<0.001	680 (14.4%)	126 (69.2%)	3390 (88.5%)	<0.001	552 (12.1%)	178 (80.5%)	3338 (88.1%)	0.001	552 (12.1%)
	Yes	21 (33.3%)	476 (12.0%)			56 (30.8%)	439 (11.5%)			43 (19.5%)	452 (11.9%)		

Average number of cigarettes smoked per day if daily smoker, median (IQR)		10 (10)	10 (8)	0.544	680 (14.4%)	10 (8.5)	8 (8)	0.008	552 (12.1%)	8 (10)	10 (9.5)	0.811	552 (12.1%)
AUDIT score at age 17 years, median (IQR)		8.5 (7.5)	7 (6)	0.081	1012 (21.4%)	9 (7)	7 (6)	<0.001	884 (19.4%)	8 (6)	7 (6)	0.003	884 (19.4%)
Regular cannabis use[a] at age 17 years, n (%)	No	51 (82.3%)	3655 (93.0%)	0.001	724 (15.3%)	149 (83.2%)	3533 (93.3%)	<0.001	598 (13.1%)	185 (85.2%)	3497 (93.3%)	<0.001	598 (13.1%)
	Yes	11 (17.7%)	277 (7.0%)			30 (16.8%)	253 (6.7%)			32 (14.8%)	251 (6.7%)		
Home ownership status at 8 weeks gestation, n (%)	Own home or mortgage	47 (68.1%)	3600 (87.4%)	<0.001	531 (11.3%)	160 (81.6%)	3378 (87.6%)	0.015	509 (11.2%)	194 (82.5%)	3344 (87.6%)	0.025	509 (11.2%)
	Rent or other	22 (31.9%)	518 (12.6%)			36 (18.4%)	480 (12.4%)			41 (17.5%)	475 (12.4%)		
Highest parental social class at 32 weeks gestation, n (%)	Non-manual	47 (77.0%)	3543 (87.4%)	0.016	603 (12.8%)	164 (86.3%)	3311 (87.3%)	0.680	582 (12.8%)	201 (87.8%)	3274 (87.3%)	0.821	582 (12.8%)
	Manual	14 (23.0%)	511 (12.6%)			26 (13.7%)	480 (12.7%)			28 (12.2%)	478 (12.7%)		
Highest maternal educational qualification at 32 weeks gestation, n (%)	O-level or higher	51 (71.8%)	3432 (81.3%)	0.042	427 (9.1%)	170 (82.9%)	3204 (81.3%)	0.550	415 (9.1%)	201 (83.4%)	3173 (81.2%)	0.397	415 (9.1%)
	Less than O-level	20 (28.2%)	788 (18.7%)			35 (17.1%)	739 (18.7%)			40 (16.6%)	734 (18.8%)		

Take-home family income at age 8 years, n (%)	£400+ per week	22 (47.8%)	1926 (56.8%)	0.222	1282 (27.2%)	77 (47.5%)	1806 (56.9%)	0.019	1229 (26.9%)	100 (56.5%)	1783 (56.5%)	0.996	1229 (26.9%)
	<£400 per week	24 (52.3%)	1464 (43.2%)			85 (52.5%)	1366 (43.1%)			77 (43.5%)	1374 (43.5%)		
IQ score at age 8, mean (SD)		104.3 (15.8)	107.4 (16.3)	0.150	804 (17.0%)	107.1 (15.6)	107.4 (16.0)	0.766	764 (16.7%)	105.9 (15.8)	10.7.5 (16.0)	0.167	764 (16.7%)

AUDIT: Alcohol Use Disorders Identification Test; IQR: interquartile range; DHA: decosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid; SD: standard deviation; N/A: not applicable. Exact numbers are not given for cell counts with less than 5 participants to avoid statistical disclosure.

[a] Regular cannabis use at age 17 was defined as using cannabis more than once monthly.

* Exact numbers not shown due to small cell counts for missing data.

Supplementary Table 4.5: Characteristics of age 24 sample according to specific mental disorder at age 24 years

		Psychotic disorder				Moderate/severe depressive disorder				Generalised anxiety disorder			
		Yes *n*=47	**No *n*=3842**	***p***	**Missing data, n (%)**	**Yes *n*=304**	**No *n*=3662**	***p***	**Missing data, n (%)**	**Yes *n*=386**	**No *n*=3571**	***p***	**Missing data, n (%)**
Age in years, mean (SD)		24.1 (0.9)	24.0 (0.8)	0.571	0 (0%)	24.1 (0.8)	24.0 (0.8)	0.182	0 (0%)	24.1 (0.8)	24.0 (0.8)	0.909	0 (0%)
Sex*, n (%)	Female	Ó 7 4 . 5	Ó 6 2 . 3	0.088	Ò 0 . 2	Ó 7 7 . 3	Ó 6 1 . 4	<0.001	Ò 0 . 2	Ó 7 5 . 1	Ó 6 1 . 3	<0.001	Ò 0 . 2
	Male	Ò 2 5 . 5	Ò 3 7 .			Ò 2 2 . 7	Ò 3 8 . 6			Ò 2 4 . 9	Ò 3 8 . 7		
Ethnicity*, n (%)	White	Ó 8 7 . 5	Ó 9 6 . 1	0.251	441 (11.3%)	253 (95.1%)	3127 (96.1%)	0.430	446 (11.2%)	324 (96.1%)	3047 (96.0%)	0.898	446 (11.3%)
	Non-white	Ò 1 2 . 5	Ò 3 . 9 9			13 (4.9%)	127 (3.9%)			13 (3.9%)	127 (4.0%)		
BMI in kg/m² at age 24 years, mean (SD)		26.7 (7.5)	24.9 (5.0)	0.098	37 (1.0%)	26.4 (6.9)	24.8 (4.9)	<0.001	40 (1.0%)	25.6 (6.5)	24.8 (4.9)	0.019	40 (1.0%)
Daily smoker at age 24 years, n (%)	No	30 (65.2%)	3347 (88.3%)	<0.001	51 (1.3%)	237 (78.5%)	3219 (88.3%)	<0.001	51 (1.3%)	302 (78.7%)	3146 (88.6%)	<0.001	21 (0.5%)
	Yes	16 (34.8%)	445 (11.7%)			65 (21.5%)	425 (11.7%)			82 (21.3%)	406 (11.4%)		

Average number of cigarettes smoked per day if daily smoker, median (IQR)		10 (8.5)	9 (7)	0.105	51 (1.3%)	9 (7)	10 (7)	0.920	51 (1.3%)	10 (7)	9.5 (7)	0.910	21 (0.5%)
AUDIT-C score at age 24 years, median (IQR)		4 (4)	5 (3)	0.004	77 (2.0%)	5 (4)	5 (3)	0.057	50 (1.3%)	5 (4)	5 (3)	0.060	49 (1.2%)
Regular cannabis use[a]* at age 24 years, n (%)	No	41 (89.1%)	3577 (93.8%)	0.189	31 (0.8%)	Ó 8 8 . 5	Ó 9 4 . 2	<0.001	Ò 0 . 2	Ó 8 8 . 3	Ó 9 4 . 3	<0.001	Ò 0 . 2
	Yes	5 (10.9%)	235 (6.2%)			Ò 1 1 . 5	Ò 5 . 8 %			Ò 1 1 . 7	Ò 5 . 7 9		
Home ownership status at 8 weeks gestation, n (%)	Own home or mortgage	30 (75.0%)	2957 (88.2%)	0.011	495 (12.7%)	209 (79.5%)	2836 (88.5%)	<0.001	420 (10.6%)	274 (82.5%)	2762 (88.3%)	0.002	496 (12.5%)
	Rent or other	10 (25.0%)	397 (11.8%)			54 (20.5%)	370 (11.5%)			58 (17.5%)	367 (11.7%)		
Highest parental social class at 32 weeks gestation*, n (%)	Non-manual	Ó 8 9 . 2	Ó 8 8 . 5	0.900	532 (13.7%)	218 (84.8%)	2811 (88.7%)	0.060	541 (13.6%)	278 (85.3%)	2740 (88.7%)	0.066	542 (13.7%)
	Manual	Ò 1 0 . 8	Ò 1 1 . 5			39 (15.2%)	357 (11.3%)			48 (14.7%)	349 (11.3%)		
Highest maternal educational qualification at 32 weeks gestation, n (%)	O-level or higher	33 (80.5%)	2878 (83.5%)	0.610	400 (10.3%)	226 (83.1%)	2734 (83.0%)	0.970	323 (8.1%)	280 (81.4%)	2670 (83.1%)	0.417	401 (10.1%)
	Less than O-level	8 (19.5%)	570 (16.5%)			46 (16.9%)	560 (17.0%)			64 (18.6%)	542 (16.9%)		

Take-home family income at age 8 years, n (%)	£400+ per week	13 (48.2%)	1607 (58.6%)	0.273	1119 (28.8%)	105 (52.2%)	1549 (58.9%)	0.064	1059 (26.7%)	139 (53.7%)	1512 (58.9%)	0.102	1132 (28.6%)
	<£400 per week	14 (51.8%)	1136 (41.4%)			96 (47.8%)	1080 (41.1%)			120 (46.3%)	1054 (41.1%)		
IQ score at age 8, mean (SD)		99.9 (13.8)	108.2 (15.8)	0.003	765 (19.7%)	107.4 (16.8)	108.1 (15.8)	0.518	787 (19.8%)	107.8 (15.8)	108.1 (15.9)	0.798	787 (19.9%)

AUDIT-C: Alcohol Use Disorders Identification Test ï Concise; IQR: interquartile range; DHA: decosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid; SD: standard deviation; N/A: not applicable. Exact numbers are not given for cell counts with less than 5 participants to avoid statistical disclosure.

[a] Regular cannabis use at age 24 was defined as using cannabis weekly or daily.

* Exact numbers not shown due to small cell counts.

Supplementary Table 4.6: Effect estimates for associations between change in plasma polyunsaturated fatty acid measures from 17 to 24 years and odds of incident mental disorders at 24 years

Outcome (*n* with vs. without outcome)	**Exposure**	**Crude**			**Adjusted**		
		Odds ratio	**95% CI**	***p***	**Odds ratio**	**95% CI**	***p***
Psychotic disorder (20 vs. 2754)	Change in total n-6 PUFAs	1.28	0.75 ï 2.16	0.362	1.23	0.72 ï 2.09	0.444
	Change in total n-3 PUFAs	1.18	0.70 ï 1.97	0.531	1.18	0.71 ï 1.98	0.520
	Change in n-6:n-3 ratio	1.16	0.65 ï 2.04	0.618	1.09	0.64 ï 1.86	0.750
	Change in DHA % total fatty acids	0.97	0.59 ï 1.57	0.888	1.03	0.63 ï 1.69	0.900
Moderate/ severe depressive disorder (157 vs. 2505)	Change in total n-6 PUFAs	0.97	0.76 ï 1.23	0.792	0.98	0.77 ï 1.24	0.876
	Change in total n-3 PUFAs	0.85	0.68 ï 1.06	0.152	0.87	0.69 ï 1.08	0.208
	Change in n-6:n-3 ratio	1.21	0.92 ï 1.60	0.166	1.18	0.90 ï 1.53	0.227
	Change in DHA % total fatty acids	0.94	0.76 ï 1.16	0.553	0.96	0.78 ï 1.19	0.720
Generalised anxiety disorder (205 vs. 2423)	Change in total n-6 PUFAs	0.79	0.66 ï 0.95	0.011	0.83	0.69 ï 1.00	0.045
	Change in total n-3 PUFAs	0.85	0.71 ï 1.00	0.057	0.88	0.74 ï 1.04	0.133
	Change in n-6:n-3 ratio	1.02	0.84 ï 1.25	0.81-	1.02	0.84 ï 1.24	0.837
	Change in DHA % total fatty acids	0.98	0.83 ï 1.16	0.794	0.97	0.83 ï 1.15	0.759

BMI: body mass index; CI: confidence interval; DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid.

Change in PUFA measures were converted to z-scores such that the odds ratios may be interpreted as per standard deviation increase in change in PUFA measure between 17 and 24 years. Covariates in adjusted models included sex; change in BMI between 17 and 24 years; and change in average number of cigarettes smoked per day between 17 and 24 years.

Supplementary Table 4.7: Odds ratios for cross-sectional associations between PUFA measures and mental disorders at age 17 years in models with additional adjustment for further potential confounding factors

Outcome	Exposure	Base model: adjusted for age, sex, BMI, cigarettes at age 17 years	Base + AUDIT score at age 17 years	Base + regular cannabis use[a] at age 17 years	Base + home ownership status at 8 weeks gestation	Base + parental social class at 32 weeks gestation	Base + maternal education at 32 weeks gestation	Base + family income at age 8 years	Base + IQ at age 8 years
Psychotic disorder	Total n-6 PUFAs	0.83	0.83	0.82	0.84	0.83	0.83	0.83	0.83
	Total n-3 PUFAs	0.89	0.89	0.89	0.90	0.90	0.90	0.90	0.90
	n-6:n-3 ratio	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.95
	DHA % total fatty acids	0.77	0.77	0.77	0.80	0.79	0.78	0.79	0.78
Moderate/ severe depressive disorder	Total n-6 PUFAs	0.97	0.97	0.96	0.97	0.97	0.97	0.97	0.96
	Total n-3 PUFAs	0.98	0.98	0.97	0.98	0.97	0.97	0.99	0.97
	n-6:n-3 ratio	0.99	0.99	0.98	0.99	0.99	1.00	0.98	0.99
	DHA % total fatty acids	0.95	0.95	0.95	0.95	0.95	0.93	0.97	0.93
Generalised anxiety disorder	Total n-6 PUFAs	0.91	0.92	0.90	0.91	0.91	0.91	0.91	0.91
	Total n-3 PUFAs	1.03	1.03	1.02	1.03	1.03	1.02	1.03	1.04
	n-6:n-3 ratio	0.89	0.89	0.89	0.89	0.89	0.90	0.89	0.88
	DHA % total fatty acids	1.02	1.02	1.02	1.03	1.02	1.01	1.02	1.03

AUDIT: Alcohol Use Disorders Identification Test; CI: confidence interval; DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid.

[a] Regular cannabis use at age 17 was defined as using cannabis more than once monthly.

PUFA measures were converted to z-scores such that the odds ratios may be interpreted as per standard deviation increase in the PUFA measure.

Supplementary Table 4.8: Odds ratios for cross-sectional associations between PUFA measures and mental disorders at age 24 years in models with additional adjustment for further potential confounding factors

Outcome	Exposure	Base model: adjusted for age, sex, BMI, cigarettes at age 24 years	Base + AUDIT-C score at age 24 years	Base + regular cannabis use[a] at age 24 years	Base + home ownership status at 8 weeks gestation	Base + parental social class at 32 weeks gestation	Base + maternal education at 32 weeks gestation	Base + family income at age 8 years	Base + IQ at age 8 years
Psychotic disorder	Total n-6 PUFAs	1.07	1.08	1.07	1.07	1.07	1.07	1.07	1.07
	Total n-3 PUFAs	0.71	0.75	0.71	0.73	0.71	0.71	0.72	0.73
	n-6:n-3 ratio	1.43	1.41	1.43	1.41	1.43	1.43	1.43	1.40
	DHA % total fatty acids	0.60	0.63	0.60	0.61	0.60	0.60	0.61	0.62
Moderate/ severe depressive disorder	Total n-6 PUFAs	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	Total n-3 PUFAs	0.87	0.87	0.87	0.88	0.87	0.87	0.87	0.87
	n-6:n-3 ratio	1.18	1.18	1.18	1.17	1.18	1.18	1.18	1.18
	DHA % total fatty acids	0.87	0.87	0.87	0.88	0.87	0.87	0.87	0.87
Generalised anxiety disorder	Total n-6 PUFAs	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
	Total n-3 PUFAs	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
	n-6:n-3 ratio	1.14	1.13	1.14	1.13	1.14	1.14	1.13	1.14
	DHA % total fatty acids	0.89	0.90	0.90	0.90	0.89	0.89	0.90	0.89

AUDIT-C: Alcohol Use Disorders Identification Test ï Concise ; CI: confidence interval; DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid. [a] Regular cannabis use at age 24 was defined as using cannabis weekly or daily.

PUFA measures were converted to z-scores such that the odds ratios may be interpreted as per standard deviation increase in the PUFA measure.

Supplementary Table 4.9: Odds ratios for longitudinal associations between PUFA measures at age 17 years and incident mental disorders at age 24 years in models with additional adjustment for further potential confounding factors

Outcome	Exposure	Base model: adjusted for age, sex, BMI, cigarettes at age 17 years	Base + AUDIT score at age 17 years	Base + regular cannabis use[a] at age 17 years	Base + home ownership status at 8 weeks gestation	Base + parental social class at 32 weeks gestation	Base + maternal education at 32 weeks gestation	Base + family income at age 8 years	Base + IQ at age 8 years
Psychotic disorder	Total n-6 PUFAs	1.06	1.06	1.06	1.07	1.06	1.05	1.08	1.06
	Total n-3 PUFAs	0.65	0.66	0.66	0.66	0.65	0.67	0.68	0.68
	n-6:n-3 ratio	1.42	1.42	1.40	1.42	1.42	1.39	1.38	1.36
	DHA % total fatty acids	0.44	0.44	0.45	0.46	0.44	0.45	0.47	0.48
Moderate/ severe depressive disorder	Total n-6 PUFAs	0.98	0.98	0.98	0.99	0.98	0.98	0.99	0.98
	Total n-3 PUFAs	0.98	0.98	0.98	0.99	0.99	0.97	0.99	0.98
	n-6:n-3 ratio	1.02	1.03	1.02	1.02	1.02	1.04	1.02	1.02
	DHA % total fatty acids	0.89	0.89	0.89	0.90	0.89	0.88	0.90	0.89
Generalised anxiety disorder	Total n-6 PUFAs	1.14	1.15	1.14	1.14	1.14	1.14	1.14	1.14
	Total n-3 PUFAs	0.98	0.98	0.98	0.98	0.98	0.97	0.98	0.97
	n-6:n-3 ratio	1.14	1.14	1.13	1.14	1.14	1.15	1.14	1.15
	DHA % total fatty acids	0.89	0.88	0.88	0.89	0.89	0.88	0.89	0.88

AUDIT: Alcohol Use Disorders Identification Test; CI: confidence interval; DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid.

[a] Regular cannabis use at age 17 was defined as using cannabis more than once monthly.

PUFA measures were converted to z-scores such that the odds ratios may be interpreted as per standard deviation increase in the PUFA measure.

Supplementary Table 4.10: Odds ratios for longitudinal associations between change in PUFA measures from age 17 to age 24 years and incident mental disorders at age 24 years in models with additional adjustment for further potential confounding factors

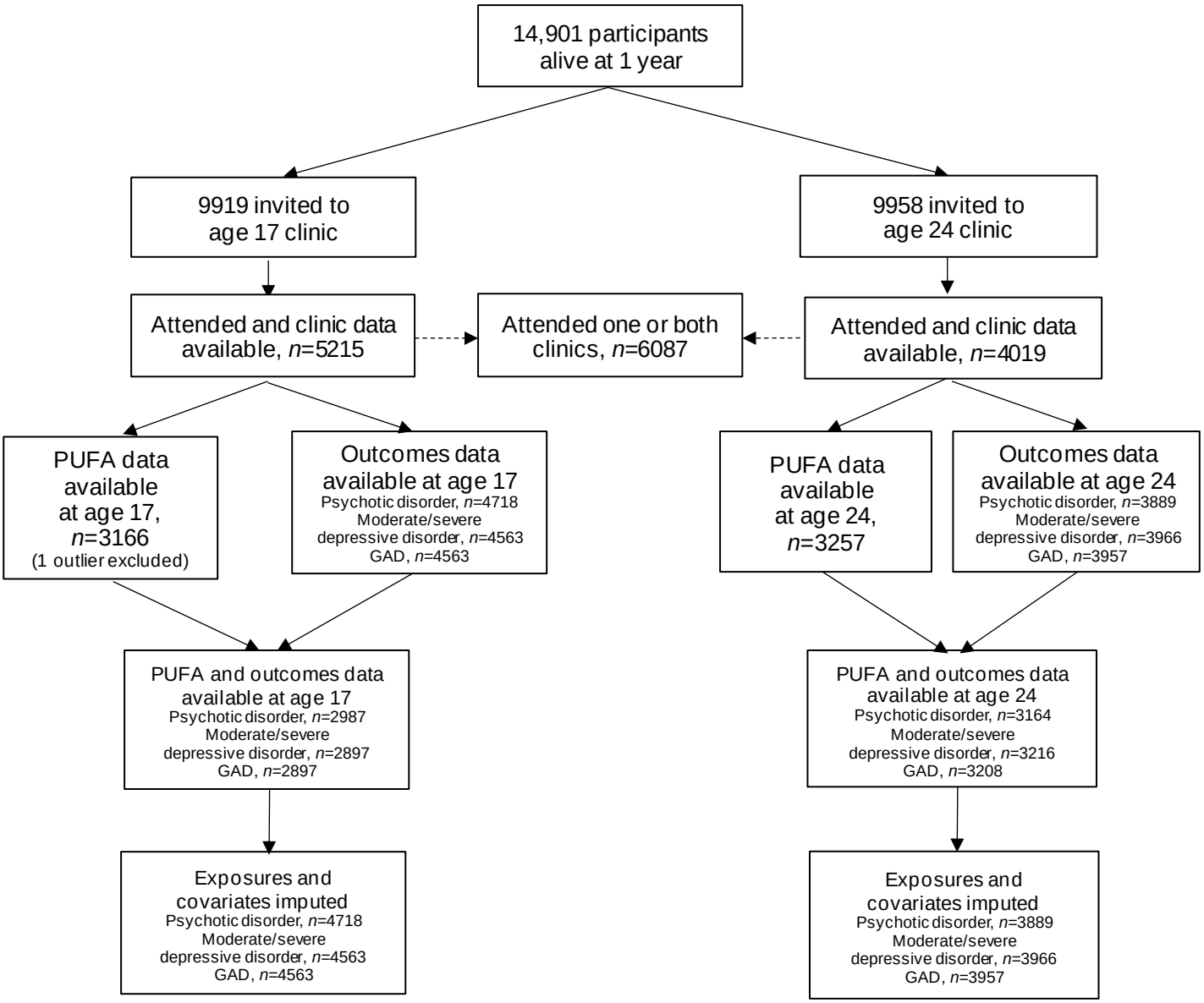
Outcome	Exposure	Base model: adjusted for sex, change in BMI, change in cigarettes	Base + AUDIT score at age 17 years	Base + regular cannabis use[a] at age 17 years	Base + home ownership status at 8 weeks gestation	Base + parental social class at 32 weeks gestation	Base + maternal education at 32 weeks gestation	Base + family income at age 8 years	Base + IQ at age 8 years
Psychotic disorder	Change in total n-6 PUFAs	1.23	1.21	1.24	1.23	1.23	1.23	1.25	1.22
	Change in total n-3 PUFAs	1.18	1.16	1.18	1.21	1.18	1.16	1.18	1.16
	Change in n-6:n-3 ratio	1.09	1.10	1.11	1.07	1.09	1.11	1.10	1.11
	Change in DHA % total fatty acids	1.03	1.02	1.03	1.04	1.03	1.02	1.01	0.99
Moderate/ severe depressive disorder	Change in total n-6 PUFAs	0.98	0.97	0.99	0.98	0.98	0.98	0.98	0.99
	Change in total n-3 PUFAs	0.87	0.86	0.87	0.87	0.86	0.87	0.87	0.87
	Change in n-6:n-3 ratio	1.18	1.19	1.18	1.17	1.18	1.17	1.18	1.18
	Change in DHA % total fatty acids	0.96	0.95	0.96	0.96	0.96	0.97	0.96	0.96
Generalised anxiety disorder	Change in total n-6 PUFAs	0.83	0.82	0.83	0.82	0.83	0.83	0.83	0.83
	Change in total n-3 PUFAs	0.88	0.87	0.87	0.88	0.88	0.88	0.88	0.88
	Change in n-6:n-3 ratio	1.02	1.03	1.03	1.01	1.02	1.02	1.02	1.02
	Change in DHA % total fatty acids	0.97	0.97	0.97	0.98	0.97	0.97	0.97	0.97

AUDIT: Alcohol Use Disorders Identification Test; CI: confidence interval; DHA: docosahexaenoic acid; n-6: omega-6; n-3: omega-3; PUFA: polyunsaturated fatty acid.
[a] Regular cannabis use at age 17 was defined as using cannabis more than once monthly.

PUFA measures were converted to z-scores such that the odds ratios may be interpreted as per standard deviation increase in the PUFA measure.

Appendix 4.3: Supplementary Figure

Supplementary Figure 4.1: Flow diagram depicting sample derivation



GAD: generalised anxiety disorder; PUFA: polyunsaturated fatty acid.

Chapter 5. General Discussion

Chapter overview

This chapter reviews and integrates the main findings of the studies contributing to this thesis. Implications of the presented research are discussed with consideration of future directions and overall methodological strengths and limitations.

Role of the thesis author

Dr Mongan authored this chapter.

5.1 Review of main findings

This thesis presents three investigations focusing on specific aspects of the molecular pathophysiology of psychotic disorders with implications for prediction and prevention. There follows a brief review of the main findings from each chapter.

5.1.1 Peripheral complement proteins in schizophrenia: a systematic review and meta-analysis of serological studies

This study aimed to systematically review the available evidence from published studies comparing peripheral complement protein concentrations or activity in patients with a diagnosis of schizophrenia compared to controls without the disorder. In general, there was little clearly consistent evidence of raised complement protein levels or activity in people with compared to without schizophrenia across the included studies. A meta-analysis of studies measuring peripheral concentrations of complement components 3 (C3) and 4 (C4) provided little quantitative evidence for consistent differences in levels of these proteins in patients relative to controls. Notably, there was significant methodological and clinical heterogeneity between the included studies.

5.1.2 Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence

The main objectives of this study were to use machine learning algorithms to develop models incorporating plasma proteomic data for prediction of, firstly, transition to first episode psychosis (FEP) in a clinical high-risk (CHR) for psychosis sample from the European Network of Schizophrenia Networks studying Gene Environment Interactions (EU-GEI) study and, secondly, development of adolescent psychotic experiences (PEs) in a general population sample from the Avon Longitudinal Study of Parents and Children (ALSPAC). For prediction of transition from CHR to FEP in EU-GEI, proteomic data were of greater predictive value than clinical information alone. Functional enrichment analysis of plasma proteins differentially expressed at baseline in CHR who developed FEP compared to those who did not provided evidence for early dysregulation of the complement and coagulation cascades. A model based on the ten most predictive proteins accurately predicted transition status in training (area under the receiver-operating

characteristic curve [AUC] 0.99, positive predictive value [PPV] 76.9%, negative predictive value [NPV] 100.0%) and test data (AUC 0.92, PPV 81.8%, NPV 96.8%). In the general population sample from ALSPAC, a model using proteomic data at age 12 had fair, though comparatively reduced, predictive accuracy for development of PEs at age 18 (AUC 0.74, PPV 67.8%, NPV 75.8%).

5.1.3 Plasma polyunsaturated fatty acids and mental disorders in adolescence and early adulthood: cross-sectional and longitudinal associations in a general population cohort

This study sought to evaluate cross-sectional and longitudinal associations between plasma polyunsaturated fatty acids (PUFAs) and psychotic disorder, depressive disorder and generalised anxiety disorder (GAD) in late adolescence and early adulthood in ALSPAC. Cross-sectionally, the ratio of omega-6 (n-6) to omega-3 (n-3) PUFAs was associated with all three disorders in early adulthood (age 24 years) but not in late adolescence (age 17 years). Plasma levels of docosahexaenoic acid at age 17 years were inversely associated with odds of incident psychotic disorder at age 24 years (adjusted odds ratio [OR] 0.44, 95% confidence interval [CI] 0.22 ï 0.87), with little such longitudinal evidence in relation to depressive disorder or GAD. There was also little evidence for associations between the change in plasma PUFA measures from 17 to 24 years and incident psychotic disorder, depressive disorder or GAD at 24 years. These results were consistent when adjusted for several potential confounding factors.

5.2 Evidence and implications

In the following sections, three major themes of evidence and associated implications arising from the studies in this thesis are discussed (Figure 5.1).

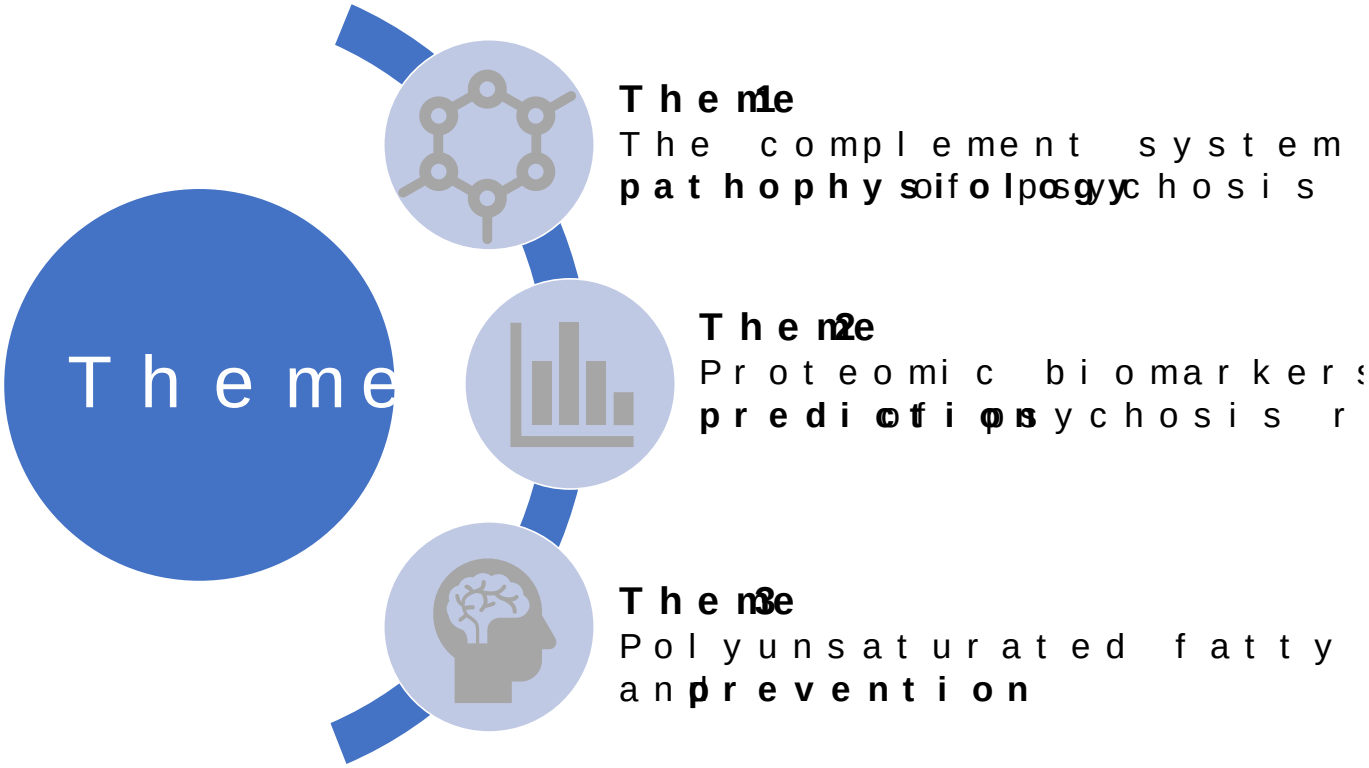


Figure 5.1 Major themes of evidence and associated implications arising from the studies included in this thesis

5.2.1 Theme 1: The complement system in the pathophysiology of psychosis

The systematic review of serological studies in Chapter 2 presents a mixed picture regarding complement deregulation in psychosis when the evidence from these studies is considered in totality. Qualitatively, there was little evidence of a consistent pattern of complement dysregulation across the included studies. There was also little quantitative evidence that C3 or C4 were increased in patients with schizophrenia relative to controls. Methodological differences in assay type, case definition, ascertainment of controls, sample preparation, control of confounding variables, medication use and other environmental factors likely contribute to this picture. It is also possible that peripheral measures of complement function do not adequately reflect complement activity in the brain (160). However, there are at least two further explanations that require further exploration.

Biological heterogeneity

It is possible that schizophrenia and other psychotic disorders have distinct and heterogeneous clinico-pathological subtypes. Previous studies have attempted to specify subgroups of schizophrenia patients based on clinical and behavioural criteria (for example, based on symptom profiles) ó distre types (161). Biomarkers may be helpful in delineating subgroups based on immunophenotypes (63). For example, a previous study comparing antipsychotic-naïve schizophrenia patients and matched controls identified two subgroups of patients with distinct molecular profiles in either immune molecules or growth factors and hormones (162). This study did not specifically examine markers of the complement or coagulation systems. However, in principle, there may be subgroups of patients with schizophrenia with varying degrees of associated complement dysregulation. This could partially explain the resulting inconsistent findings when the evidence from serological studies is synthesised, and is in keeping with the fact that measurable immune dysfunction occurs in a subset of, but not all, affected patients (163). However, the complement and coagulation cascades were strongly and consistently implicated in the proteomic studies of early psychosis phenotypes presented in Chapter 3, suggesting the presence of dysregulation early in the course of psychotic disorders which might subsequently vary over time.

Temporal changes

The degree of complement and coagulation dysfunction (and/or immune disturbance more generally) may be time-dependent, varying by age and stage of illness. It is possible that the underlying hyper-activation leads to óburnoutô these systems in the later stages of psychotic disorders, with the result that it is difficult to detect significant abnormalities in peripheral levels relative to controls in cross-sectional studies. Notably, all the studies included in the systematic review in Chapter 2 were cross-sectional comparisons of patients diagnosed with schizophrenia versus controls. This hypothesis of temporal change may explain why, in general, there was little clear evidence of consistently raised peripheral complement protein concentrations on synthesis of the evidence, since most of the patients in the included studies were chronically unwell.

Looking to earlier stages of the psychosis phenotypic continuum, the proteomic studies in Chapter 3 provide evidence that dysregulation of the complement and coagulation systems precedes the onset of FEP in CHR individuals. There was also evidence for similar dysregulation (though to a lesser degree) six years prior to the onset of PEs at age 18 in a general population sample. The overall pattern of proteomic changes in these studies is in keeping with a net activation of the complement and coagulation systems. For example, among the top predictors in both studies were alpha-2-macroglobulin (A2M) and complement C4b binding protein alpha chain (C4BPA), both decreased at baseline in cases compared to controls. These proteins ordinarily serve regulatory functions through inhibition of the complement and coagulation cascades (164, 165). That downregulation of key regulatory proteins already exists prior to development of FEP in CHR individuals, and six years before the onset of PEs in the general population sample, suggests that deficiencies of these factors occur early in the psychosis continuum and potentially years before the onset of even subthreshold psychotic symptoms. It remains to be determined whether these regulatory protein deficiencies are primary phenomena or a result of depletion due to over-consumption in the face of chronic complement and coagulation over-activation.

These proteomic findings are broadly in keeping with the small number of published studies that have specifically examined complement biomarkers in early psychosis phenotypes. One previous study reported a multivariable model including several complement protein biomarkers that distinguished between FEP patients and controls with fair accuracy (AUC 0.79), although on univariate analysis only levels of one complement analyte (the terminal complement complex) differed significantly in patients compared to controls (166). A further study compared peripheral C1q, C3 and C4 levels in groups of patients that varied by stage of psychosis. Relative to controls, there were no differences in levels of these proteins in FEP patients when adjusting for body mass index (BMI); C4 was increased in chronic schizophrenia; and both C3 and C4 were significantly increased in a small sample of CHR individuals (167). While not a longitudinal study, these preliminary findings suggest that the extent of complement dysfunction varies by stage of illness and may in fact be maximally detectable in early stages of psychosis.

Longitudinal investigations of complement biomarkers for psychosis are sparse in the existing literature. However, a previous proteomic study in the ALSPAC cohort reported evidence for similar proteomic dysregulation at age 12 years, 6 years prior to the onset of psychotic disorder at age 18 (168). The pattern of differentially expressed proteins in this previous study showed remarkably similar findings to the CHR and PE studies in Chapter 3, including decreases in A2M, immunoglobulin heavy constant mu (IGHM) and C4 binding protein alpha chain (C4BPA) and increases in complement factor H, complement factor I, plasminogen, attractin, gelsolin and cholinesterase. Comparing between the two studies in Chapter 3, several proteins were highly weighted for prediction of both PEs in the general population sample and FEP in the CHR sample, including A2M, IGHM, C4BPA and vitamin Kï dependent protein S.

Taken together, these overlapping findings suggest a degree of proteomic, in addition to phenotypic, continuity across the early temporal phases of the psychosis continuum. It is worth noting that, compared to the CHR study, the proteomic dysregulation preceding PEs occurred to a lesser extent in terms of number, effect size and predictive accuracy of differentially expressed proteins. This could be in keeping with the earlier age of exposure (12 years) examined in the PE study, as well as the ómilderô phenotypic outcome (PEs FEP in a CHR sample). The proteomic signature of complement and coagulation dysfunction may take time to develop, beginning in early life or peripubertally in children with no or few clinical signs of psychosis, before increasing to a maximum in early adulthood when attenuated psychotic symptoms emerge (as in the CHR state prior to FEP). This is in keeping with the typical age of onset of psychotic disorders in early adulthood (169). Given that complement proteins ar for elimination by microglia (170), this hypothesis may also be in keeping with results of imaging studies showing microglial activation in CHR (171), although other studies report inconsistent findings in this regard (172). One possible underlying mechanism is that complement hyper-activation (whatever the contributing cause or causes) represents napaćhowamô leading to excess synapt adolescence and hence the characteristic findings of synaptic loss observed in schizophrenia (173), at least in a subset of vulnerable individuals. However, mechanistic investigations are required to confirm whether this is the case.

5.2.2 Theme 2: Proteomic biomarkers for individualised prediction of psychosis risk

The findings of Chapter 3 showed that a model using proteomic data alone for prediction of psychosis had higher predictive accuracy (AUC 0.96, PPV 79%, NPV 100%) compared to a model using clinical (symptom and functioning) data alone (AUC 0.48, PPV 37%, NPV 63%). Indeed, the model based on proteomic data alone compared favourably to a combined clinical-and-protein model (AUC 0.95, PPV 75%, NPV 99%). In comparison to the existing literature, these proteomic models are among the most accurate described for blood biomarker-based prediction of transition to psychosis in CHR. For example, a study from the NAPLS cohort reported development of an algorithm based on 15 plasma analytes reflecting multiple biological processes (including inflammation, oxidative stress, hormones and metabolism) with an AUC of 0.88 for prediction of transition (130). A further study applied a model of 26 blood analytes developed for discrimination of schizophrenia patients versus controls, achieving an AUC of 0.82 for prediction of psychosis in help-seeking prodromal individuals (131). The greater predictive performance of biological compared to clinical data is also in keeping with the literature more generally, and may in part reflect a greater degree of heterogeneity in clinical and environmental data compared to biological data in this context (45).

While the proteomic model achieved a high predictive accuracy, this model drew on data for 166 proteins. This highlights a further germane issue in prediction studies, in that each additional predictive parameter adds a degree of complexity to the model, potentially risking over-fitting, obscuring interpretability and increasing the pragmatic challenges of implementation in practice. A balance between model complexity, predictive performance and clinical feasibility is required. Some studies have advocated for a stepped sequential testing approach. A systematic review investigated whether sequential application of predictive models improved overall prediction, finding that an optimal sequential testing approach consisted of a combined (clinical and EEG) model followed by a neuroimaging model and then a blood marker model (achieving a PPV of 98%) (174). However, from a pragmatic perspective, a sequential approach may be more difficult and costly to implement in clinical practice than a parsimonious model drawing on data from a single timepoint that does not adversely compromise predictive accuracy.

Considering these concerns, we developed a more parsimonious model consisting of solely the ten most predictive proteins. This reduced model maintained a high predictive performance in training data (AUC 0.99, PPV 77%, NPV 100%). Importantly, the model retained high performance when internally validated in test data that did not contribute to the model-building process (AUC 0.92, PPV 82%, NPV 97%). The implication is that transition outcome in CHR was predicted with a high degree of accuracy using a relatively simpler model based on data for a small pool of highly predictive proteins. This reduces the potential complexity of the model and the number of parameters that may need to be measured. Critically, these models require external validation in other CHR cohorts (both retrospectively and prospectively) before further implementation. All models were cross-validated by study site, and the parsimonious model was internally validated using held-out data from the largest study site. While these procedures suggest a degree of geographical and clinical generalisability, this needs to be confirmed in future studies to meet stringent clinical biomarker requirements. Formal integration of expert clinical opinion may augment prediction further, as has been shown with respect to clinical and neuropsychological models (47).

The proteomic model for prediction of PEs in the ALSPAC general population sample had lower performance (AUC 0.74, PPV 67.8%, NPV 75.8%) than that for prediction of FEP in the CHR study. As previously discussed, the younger sampling age and milder phenotypic outcome may contribute to this. These results are in keeping with performance of previous machine learning models based on functional neuroimaging for prediction of adolescent psychotic symptoms (175). However, an additional challenge is that the incidence of adolescent PEs in the general population is relatively low (approximately 2.5% per year according to one meta-analysis (176)) compared to the incidence of FEP in the CHR subpopulation, which is enriched for psychosis risk. As a result, direct application of these models for prediction of PEs in young people in the wider unselected general population may be less likely to yield accurate individualised prediction of future adolescent PEs. Furthermore, in contrast to the EU-GEI study, all the participants in the ALSPAC sample lived in the same geographical area, which may limit generalisability to other populations.

5.2.3 Theme 3: Polyunsaturated fatty acids and prevention

The findings from Chapter 4 add to the existing literature regarding PUFAs and mental disorders in several ways. There was little evidence for cross-sectional associations between plasma PUFA biomarkers and mental disorders in late adolescence, but the n-6:n-3 ratio was positively associated with psychotic disorder, depressive disorder and GAD in early adulthood. However, as discussed in Chapter 4, cross-sectional associations are difficult to interpret with confidence given the potential for reverse causality.

Longitudinally, there was evidence of an inverse relationship between plasma DHA in late adolescence and odds of incident psychotic disorder in early adulthood: each standard deviation increase in DHA % of total fatty acids at age 17 years was associated with a 56% reduction in odds of psychotic disorder at age 24 years (adjusted odds ratio 0.44, 95% confidence interval 0.22 ï 0.87). In contrast, there was little evidence for similar associations in relation to incident depressive disorder or GAD. These latter findings are in keeping with a systematic review and meta-analysis which found that increasing long-chain omega-3 fatty acids had little or no effect for prevention of depressive or anxiety symptoms in clinical trials (177).

Causality cannot be assumed in observational studies. However, the longitudinal relationship between DHA and psychosis demonstrated in Chapter 4 is consistent with some of the traditional causal criteria as originally defined by Bradford Hill (178): the strength of the observed relationship is clinically important; the relationship appears to be more specific to psychotic disorder than to depression or anxiety; the temporal sequence is consistent; and the relationship is biologically plausible. In common with other n-3 PUFAs, DHA has anti-inflammatory properties which may explain the associations with psychosis through reductions in (neuro)inflammatory tone (179). However, a range of other biological actions may be of particular relevance to the pathophysiology of psychosis (137, 180). For example, n-3 PUFAs contribute to neuronal membrane integrity (181), influence neurotransmission (182), reduce oxidative stress (183), modulate glial cell activity (184) and promote synaptogenesis (185), all of which are processes which have previously been implicated in early psychosis pathophysiology (129).

Considering the proteomic findings in Chapter 3, the role of n-3 PUFAs in modulating complement-dependent synaptic pruning requires elucidation as a further potential preventative mechanism. Excess synaptic pruning was originally proposed in a hypothesis by Feinberg to underlie the pathogenesis of schizophrenia (186). As previously discussed, dysregulation of the complement and coagulation cascades may precede the onset of PEs and FEP and may be maximal in early, as compared to chronic, phases of psychotic disorders. This upregulation could result in the normal process of adolescent synaptic pruning being driven to excess (170), giving rise to increased brain dysconnectivity in late adolescence or early adulthood and the onset of psychosis. In a compelling series of mechanistic studies in mice, maternal exposure to an n-3 PUFA deficient diet resulted in increased microglial activity and excessive synaptic pruning in the offspring in a complement-dependent manner (187). Compared to n-3 sufficient mice, n-3 deficient mice showed increased hippocampal C1q immunostaining and higher synaptosomal C3 expression (187). This research identifies the complement system as a driver of synaptic elimination in n-3 PUFA deficiency. While further confirmatory evidence is required in humans, it is possible that PUFA exposure may be relevant to psychosis from a neurodevelopmental perspective via effects on synaptic pruning.

Chapter 4 examined adolescent rather than prenatal PUFA exposure, but the neuroprotective effects of DHA in terms of promotion of synaptogenesis and neurite growth remain relevant in later development (185, 188). While adolescent plasma DHA levels were inversely associated with psychosis, there was little evidence that changes in plasma PUFA markers from adolescence to early adulthood were similarly associated with psychosis or other mental disorders. It is thus possible that the actions of PUFAs on psychosis risk are long-term effects, and that the adolescent and/or pre-adolescent period is especially sensitive in this regard. It is not possible to extrapolate mechanistic understanding from these observations. However, the findings are at least temporally consistent with a speculative biological mechanism by which DHA could ó b u f f e r ô a g a i n s t t h e e x c e s s i v associated with excessive complement activity in adolescents who may go on to develop psychosis. Whether this is in fact the protective mechanism requires confirmation in future studies.

The primary implication from Chapter 4 is that adolescent DHA levels may be a potentially modifiable protective factor in relation to risk of psychotic disorders in early adulthood. Psychotic disorders are relatively rare in the general population (5). However, the direct and indirect socio-economic benefits of preventing even a small number of psychotic disorders would likely outweigh the costs of primary prevention aimed at the population level, especially if such interventions are relatively inexpensive, safe and associated with other health benefits (as has been claimed of n-3 PUFAs) (189). For example, if the present findings are replicated in future studies, one recommendation might be for public health authorities to issue guidelines regarding DHA intake for adolescents (potentially with advice regarding oily fish consumption and/or n-3 supplementation). Large-scale trials of adolescents with long-term follow-up periods would be helpful in assessing whether n-3 supplementation or dietary interventions aid in prevention of psychotic disorders.

5.3 An integrated view of findings and implications from this thesis

Figure 5.2 provides an integrated overview of the major findings and implications arising from the studies in this thesis in the form of one possible longitudinal model of psychosis pathophysiology focused on activity of the complement system, beginning prenatally and spanning childhood, puberty, early adulthood and later stages of disorder.

The prenatal and early childhood periods were not explicitly examined in this work. However, in line with existing evidence implicating prenatal inflammation in people who go on to develop psychosis (190, 191), global activity of the complement and coagulation is hypothesised to be increased early in development in this model. Several prenatal and postnatal biopsychosocial factors (including genetic factors, maternal infection and childhood adversity) may cause, contribute to or maintain this increased activity in early and later life. Several examples of other established risk factors for psychosis are provided in the Figure (192, 193), although it is important to note that the relationship of these risk factors to complement dysfunction in the context of psychosis pathophysiology requires confirmation.

The hypothesised model implies that there are detectable proteomic changes in puberty (prior to the beginning of phenotypic expression) as supported by the findings in the proteomic study for prediction of PEs in **Chapter 3**. During

adolescence, when synaptic pruning normatively escalates, this process may be driven to excess because of the unchecked complement activity. However, as per the findings in **Chapter 4**, DHA represents a potential preventative factor, perhaps by ó b u fhis effect (and/or acting through other mechanisms such as reduction of neuroinflammation) to decrease the risk of psychosis in early adulthood. Persistently dysregulated complement activity leads to increasing synaptic loss, dysconnectivity and eventually greater phenotypic expression of early psychotic symptoms, consistent with the strong proteomic prediction findings in the CHR state preceding FEP in **Chapter 3**. It is unknown how complement activity tracks longitudinally following FEP; nor is it known how or if complement activity responds to available treatments (hence the dashed line in the Figure). However, the model implies that by the chronic stages of disorder as examined in **Chapter 2**, complement activation becomes quiescent (for example following multiple episodes of psychosis in patients with chronic schizophrenia).

This model is not intended to represent a definitive or exhaustive explanation of psychosis vulnerability, but rather serves as an integrated overview of the major findings of this thesis and how they may relate to one another. Further studies will be required to confirm these findings and hypotheses, to which our attention now turns.

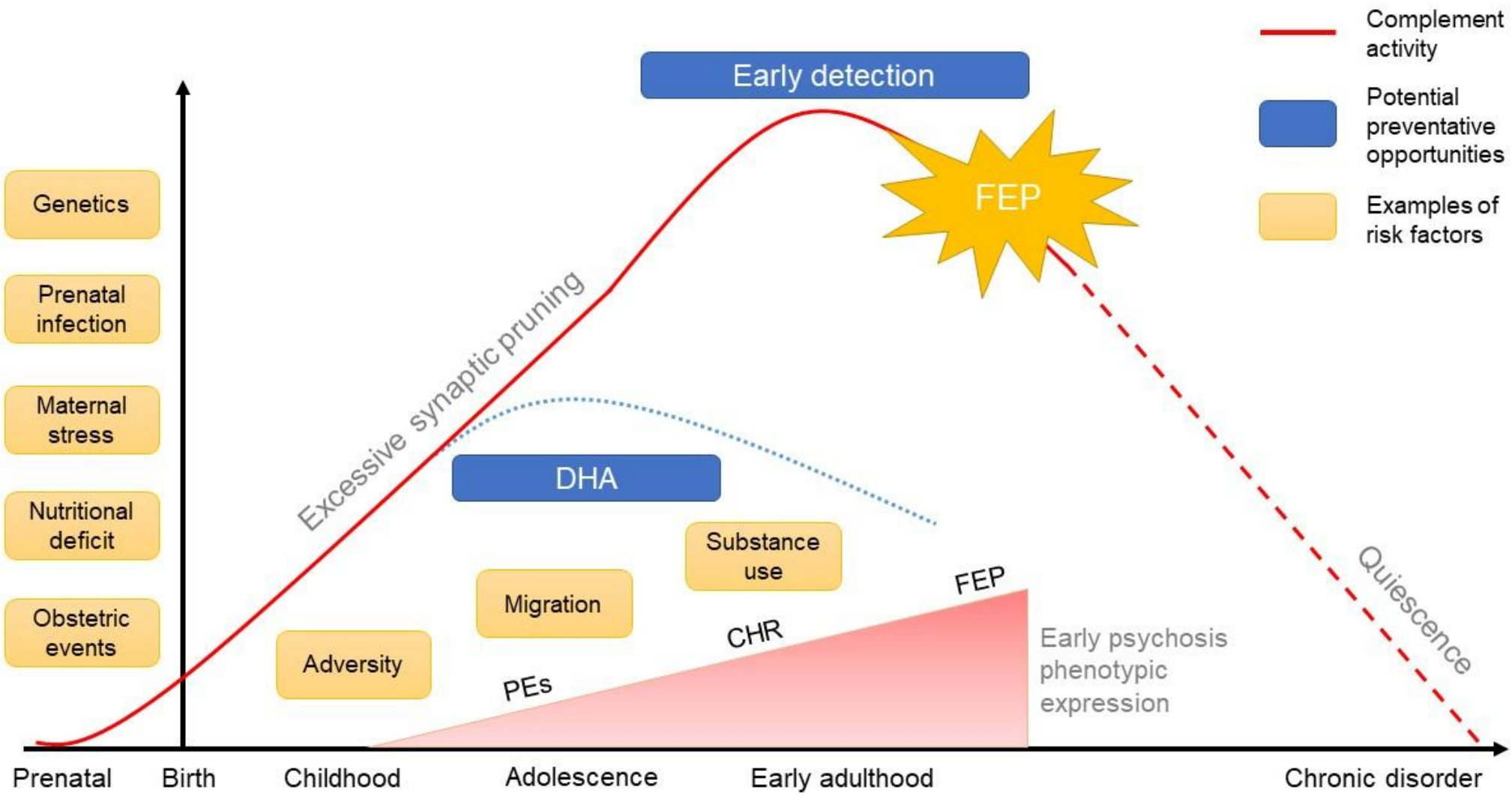


Figure 5.2 An hypothesised longitudinal trajectory of complement activity in psychosis development with examples of risk factors and preventative opportunities. CHR: clinical high-risk; DHA: docosahexaenoic acid; FEP: first episode psychosis; PEs: psychotic experiences.

5.4 Future directions

Several potential questions arise from the findings presented in this thesis which could form the basis of future studies.

1) What are the causes of complement and coagulation activation in psychosis?

The proteomic studies in Chapter 3 suggest that dysregulation of the complement and coagulation cascades occurs early in psychosis, potentially several years in advance of subthreshold PEs. While this early evidence may be consistent with complement and coagulation activation being a primary cause of psychosis, there may be further underlying causes for this dysregulation which could be important in terms of intervention and prevention. The present findings are in keeping with the risk of psychosis associated with genetic variation of C4 (170) but, given the relatively small effect sizes of individual genetic abnormalities, these alone are probably insufficient to cause psychosis. Complex gene-environment interactions may contribute to biological dysregulation and psychosis risk (194). Prenatal and perinatal factors associated with schizophrenia risk include maternal and paternal history of psychopathology, maternal infections, maternal hypertension, famine, obstetric complications, congenital malformations, low birthweight and premature birth (193). Several studies have associated adverse pregnancy outcomes with excessive complement activation (195) and complement is thought to play a role in embryonic development (196). However, the question of whether prenatal or perinatal factors influence risk of psychosis via complement activation requires investigations in further studies and biosampling in large longitudinal cohorts.

Postnatally, multiple factors are associated with psychosis risk including cannabis and tobacco use, traumatic brain injury, physical and mental illnesses and urbanicity (197-202). In particular, childhood adversity has been extensively studied as an established risk factor for psychotic experiences and psychotic disorders (203). Several psychosocial factors may mediate these relationships including attachment, affective symptoms and substance use (204). Neurobiological explanations have also been advanced such as effects of trauma on the hypothalamic-pituitary axis, dopaminergic systems and neurocognition (203, 205). With regard to inflammation, markers of chronic inflammation have been found to be increased in association with

childhood adverse experiences and exposure to violence (206). Hyperactivation of the complement pathway has been noted in patients with post-traumatic stress disorder relative to controls (207). In an intriguing proteomic study, mice subjected to social defeat stress showed increased complement activity, and several upregulated proteins overlapped with findings of complement activation in human plasma samples prior to development of PEs (including increased C1R, complement factor H and C5) (208). This could suggest a role for severe childhood psychological stress or adversity in potentially increasing complement activation, eventually leading to elevated risk of psychosis. Studies investigating the influence of childhood adversity on the proteomic markers implicated in Chapter 3 would be helpful to evaluate this hypothesis.

2) Is complement and coagulation activation distinct from or secondary to more general inflammatory processes?

As discussed in Chapter 1, there is a wealth of evidence showing that pro-inflammatory cytokines are increased prior to and in association with psychotic disorders (209-211). It remains unclear whether complement and coagulation dysregulation in psychosis is a primary phenomenon, or secondary to a more general immune or inflammatory activation. The complement system requires a trigger for activation, and it is possible that increased circulation of pro-inflammatory cytokines underlie this. Longitudinal studies with repeated biosampling of inflammatory cytokines and complement biomarkers may help in delineating the temporal sequence of immune activation, although it is also possible that the precise sequence of events varies substantially between individuals.

3) Does peripheral complement activation reflect activity in the brain? What is the role of the blood-brain barrier?

The proteomic findings in Chapter 3 were derived from peripheral (plasma) samples. From a pathophysiological perspective, further studies are required to confirm the degree to which complement activity in the blood accurately reflects complement activity in the central nervous system in early psychosis. This may help to clarify the relationship between complement dysfunction in the periphery and in the brain in this context. There is evidence from post-mortem studies that complement expression in the brain is increased in schizophrenia (170). Virtually all complement components

can be produced locally in the brain (212) and most complement proteins do not cross an intact blood-brain barrier (213), so findings of significant peripheral complement upregulation may be a manifestation of systemic dysfunction that is relatively independent of activity in the brain. However, there is growing evidence of blood-brain barrier dysfunction in psychosis (214). Increased synthesis of complement components locally in the brain could directly contribute to neuroinflammation and blood-brain barrier injury (215, 216). This could potentially permit movement of complement proteins from the periphery to the brain and vice versa, but the degree to which this speculative communication is relevant to the pathophysiology of psychosis is unknown.

4) When is complement dysfunction first detectable?

Chapter 3 showed that complement dysfunction may be detectable up to six years prior to the development of PEs in apparently well 12-year-old children. However, as implied by the model in Figure 5.2, this dysfunction could begin to occur earlier than this, and even prenatally. It is also possible that dysfunction occurs first in the brain, followed by abnormalities in the periphery. Repeated sampling early in development could be helpful in establishing when the earliest signs of detectable dysfunction begin to occur. This would also enable temporal sequencing of dysfunction in the same set of vulnerable individuals, potentially shedding light on whether the observed downregulation of key regulatory proteins (such as A2M) results from a primary deficiency or from over-consumption as part of a compensatory response to chronic excessive complement and coagulation activity.

5) What are the precise biological mechanisms by which dysfunction of the complement and coagulation systems may give rise to psychosis?

The action of complement causing excessive synaptic pruning is one major hypothesised mechanism by which complement dysfunction contributes to psychosis risk. However, the exact biological pathways that explain how this eventually leads to the diversity of positive, negative, cognitive, affective and other symptoms of psychotic disorders requires further mechanistic elucidation. Dysconnectivity in the brain arising from excessive synaptic pruning may account for some, but potentially not all, of these symptoms. It is likely that heterogeneous levels of dysconnectivity in different brain regions exert differential effects (217).

6) What are the effects of hypercoagulation?

Whether coagulation activity is increased independently or because of complement upregulation is unclear. However, the complement and coagulation systems share many intimate connections with much molecular cross-talk (218) such that it may be not be possible or indeed helpful to parse out their effects. Coagulation dysfunction has previously been hypothesised to have a role in schizophrenia pathology, as supported by evidence of low activity of key anticoagulant proteins including tissue plasminogen activator and protein S in patients with schizophrenia (219). Schizophrenia may be associated epidemiologically with enhanced thromboembolic risk, although this risk may be shared with other mental disorders (220). Furthermore, lifestyle and other environmental factors, such as side-effects of antipsychotic drugs, may confound this association (221).

The proteomic dysregulation preceding FEP in Chapter 3 is broadly in keeping with a hypercoagulant phenotype. Notably, the most predictive finding in relation to transition to FEP in CHR was decreased levels of A2M. A2M has a diverse range of functions, including inhibition of pro-inflammatory cytokines, but is also a key inhibitor of coagulation (164). Intriguingly, the effects of fibrin in the brain have been implicated in neurodegenerative diseases such as multiple disease (222). Deposition of blood-derived fibrin in the brain, facilitated by blood-brain barrier dysfunction, can cause localised neuroinflammation (222). However, the degree to which coagulation dysfunction is distinct from or related to complement over-activity, and whether hyper-coagulation is causal or simply a marker of concurrent complement dysfunction, requires further assessment.

7) Are the proteomic findings specific to psychosis?

Immune dysfunction has been evidenced in other psychiatric disorders including depressive disorder, bipolar affective disorder, anxiety disorders, substance use disorders and autism spectrum disorder (223-227). Complement C4 has been found to be increased in patients with depressive disorder relative to controls (228), and increased complement factor H has been associated with depression in older-age adults (229). In a preliminary study, patients with chronic bipolar disorder had reduced levels of serum complement biomarkers compared to controls (the authors hypothesised that this was possibly due to over-consumption) (230). Dysfunction of

the coagulation system has also been hypothesised to be involved in the pathophysiology of anxiety and depressive disorders as well as psychosis (231). One study found evidence for a prothrombotic phenotype in patients with depressive disorder which was modulated by antidepressant treatment (232). Thus, complement and coagulation abnormalities may not be entirely specific to psychotic disorders.

Future studies could assess the specificity of the predictive biomarkers in Chapter 3 in terms of prediction of other psychiatric disorders, since CHR individuals are at higher risk of a range of disorders beyond psychosis alone (40). In recognition of this, there have been calls for adaptation of the clinical high-risk concept towards a ó c l i n i c a l s t a g i n g ô a p p r o a c h t h a t e n c o m p a s s e s and outcomes (233). Others have proposed that psychiatric disorders may be at least partially explained using a unidimensional factor of general psychopathology known as the *p* factor (234). The validity and clinical utility of *p* is debated (235), but it may prove fruitful to assess the relationship between complement and coagulation dysfunction and a measure of *p* in large-scale longitudinal general population studies.

8) Is modulation of the complement and coagulation cascade a viable preventative or therapeutic target in psychosis?

Several drugs are already available to modulate complement system activity. For example, eculizimab is a monoclonal antibody against complement C5 (236) and the C1 esterase inhibitor conestat alfa is used for treatment of hereditary angioedema (237). Further drugs targeting different components of the complement system are in active development (238, 239). Challenges to the use of complement therapeutics for neuropsychiatric disorders include the potential side-effects of immunosuppression (240) and ensuring enough drug reaches the brain (241). Extensive preclinical studies are required to evaluate whether modulation of the complement system influences the course of psychotic disorders and, if it does, which point in the cascade represents the optimal target. With respect to coagulation, hypotheses of hypercoagulation in schizophrenia have been supported by case reports of psychotic remission in patients prescribed long-term anticoagulants for other reasons, but this data is preliminary (242). These agents may also be associated with important side-effects, particularly increased risk of

bleeding. Further studies are required to justify preclinical or clinical trials of antithrombotic agents for prevention or treatment of psychotic disorders.

9) Is the preventative effect of DHA on psychosis risk mediated through modulation of the complement and coagulation cascade?

Chapter 4 demonstrated evidence for a potential preventative effect of DHA in adolescence on risk of psychosis in early adulthood. Effects on synaptogenesis are hypothesised based on evidence of increased complement expression and synaptic elimination by microglia in n-3 deficient mice (187), but the precise mechanisms in humans are unconfirmed. Several further mechanistic pathways have been proposed, including reduction of oxidative stress, resolution of neuroinflammation and modulation of neuronal membrane functions (137). Omega-3 fatty acids also have effects on the coagulation system and lipid metabolism (243, 244) although the relevance of these in psychosis pathophysiology is unclear. Further preclinical studies (such as in mice models of psychosis), with careful consideration of the timing and duration of PUFA exposure and correlation with key markers of complement and coagulation function, may be informative in determining the mediating biological mechanisms by which DHA can influence risk of psychosis.

Such preclinical research may further justify trials of DHA in adolescence (whether through dietary changes or supplementation) with long-term follow-up to assess potential impacts in relation to psychosis risk. Challenges to this include the relative rarity of psychotic disorders in general population samples such that large sample sizes will be required. However, such trials may be valuable in informing public health guidelines and recommendations regarding adolescent n-3 intake as part of universal psychosis prevention strategies. Studies assessing longitudinal trajectories of DHA across childhood and adolescence in relation to psychosis outcomes may be helpful in delineating the optimal period or periods during which to intervene with DHA supplementation with the aim of preventing psychosis.

10) Are the predictive models valid in other cohorts?

A critical next step following the development of the predictive proteomic models in Chapter 3 is to determine whether the characteristic proteomic signature is replicated in other cohorts, and to test the external validity of the predictive models. This will require investigations using similar mass spectrometric methods in multiple CHR

cohorts. With further data, the predictive models may be refined and tested, initially retrospectively, but eventually prospectively to assess predictive accuracy where transition outcome is not yet known (for example integrated within a prospective clinical trial).

11) Does stratification based on predictive biomarkers enable improved outcomes in psychosis?

If validated retrospectively in further CHR cohorts, it would be valuable to test predictive models in a prospective context by, for example, integrating biomarker measurement at the beginning of a randomised controlled trial of interventions in CHR participants. Besides testing prospective predictive validity, this would enable investigation of whether CHR individuals receiving certain treatments (such as low-dose antipsychotic therapy, omega-3 fatty acid supplementation and/or psychosocial interventions) achieve positive outcomes in terms of prevention of transition and other adverse clinical outcomes. Investigations are also required to determine if proteomic biomarkers may aid prediction of treatment response in CHR and FEP (for example, whether there are theranostic biomarkers predictive of response to a particular antipsychotic therapy). Such investigations would helpfully inform precision medicine-based approaches that seek to tailor prognosis and management strategies to the individual patient.

12) How could biomarker-informed preventative approaches be implemented in clinical practice?

CHR clinics are not yet widely available across the globe. The predictive biomarkers in this thesis imply the utility of CHR clinics in terms of providing a risk enrichment strategy that, in combination with proteomic biomarkers, enables accurate prediction of FEP. However, mass spectrometry is not readily available in most standard clinical laboratories. Further testing and collaboration with industry is required to determine the optimal method of biomarker measurement (either using mass spectrometry or other techniques) that is tractable without severely compromising predictive accuracy.

There are also wider challenges to the implementation of novel clinical biomarkers in this setting. Would use of predictive biomarkers in this way be acceptable to patients,

their families, clinicians, service managers, industry partners, laboratory scientists, health economists and commissioners? Qualitative studies would be helpful to gather opinions across a range of possible stakeholders and to gauge engagement, suitability and sustainability. There may be ethical concerns requiring careful consideration, such as the scenario of informing someone already labelled as being at clinical risk of psychosis that they are at even greater risk (which may have insurance consequences) as well as the inevitable issues arising from false positive or false negative results (245). As with any service development initiative, an implementation framework would need to be carefully considered to determine the potential impacts and opportunities of a biomarker-informed preventative approach.

5.5 Methodological considerations

Methodological strengths and limitations of this thesis will now be considered.

Strengths

The systematic review and meta-analysis in Chapter 2 adhered to recommended methodological guidelines to provide a thorough assessment of the existing published literature (246).

The prediction studies in Chapter 3 each used sophisticated discovery-based proteomic methods in an unbiased approach assessing the entire proteome. Several findings (including differential expression of A2M) were validated using alternative immunoassay-based methods. The longitudinal nature of both studies allows assessment of the predictive nature of proteomic biomarkers, rather than cross-sectional analyses which dominate the literature. The use of machine learning analyses facilitated development of algorithms for individualised prediction rather than solely analysis of group-level differences.

The study in Chapter 4 was based on a large and well-characterised sample from the general population, using biomarkers of PUFA status as opposed to self-reported dietary intake. Missing data was imputed using multiple imputation (rather than adopting complete-case analyses which may be prone to bias). The use of repeated measures facilitated longitudinal analyses with a focus on adolescence and early adulthood.

Limitations

The search strategy in Chapter 2 focused on studies specifically examining complement function *a priori*. This was so as to focus on hypothesis-driven investigations, but may have resulted in exclusion of studies adopting a more discovery-based approach. Data sources were not included in the search strategy. Some studies of complement function in schizophrenia may have been missed if not published in English-language peer-reviewed journals.

A key limitation of Chapter 3 is that the proteomic prediction models require external validation in other CHR cohorts. This is to be performed in a recently funded programme of work. Data on trauma and cognition were not available which meant that the relationships of these factors to proteomic biomarkers of transition status or their relative contribution to outcome prediction were unable to be examined. This may also be addressed in future studies. Furthermore, the lack of comparison to healthy control participants meant that it was not possible to evaluate how or if CHR individuals differ to the general population from a proteomic perspective, as has been evidenced in lipidomic analyses (247). More generally, the CHR paradigm has been criticised from the perspective that CHR subpopulations may not be representative of most individuals who present to mental health services with FEP (52, 53), although others have interpreted this as a basis to advocate for resources to increase accessibility and availability of CHR services (55).

In Chapter 4, sample attrition along a socio-economic gradient was noted which may limit generalisability. Unobserved factors (such as physical activity and other dietary content) may residually confound the observed associations. The number of psychosis cases in the longitudinal sample was relatively small. Plasma PUFA levels (rather than erythrocyte membrane composition) were used, which may be subject to higher metabolic turnover (248). Finally, plasma PUFA levels in adolescence may, at least in part, reflect PUFA exposure over a longer period. This will need to be examined in studies that assess longitudinal trajectories of PUFA exposure across early periods of development in relation to psychosis outcomes.

5.6 Conclusion

In conclusion, this thesis presents evidence concerning specific aspects of the molecular pathophysiology of psychosis. A systematic literature review and meta-analysis provided little consistent evidence of complement system dysfunction in chronic schizophrenia relative to controls. However, proteomic studies prior to development of FEP in a CHR sample and PEs in a general population sample suggested that dysregulation of the complement and coagulation systems occurs early in individuals vulnerable to psychosis and that, in combination with machine learning techniques, this may be used to enable more accurate prediction of psychosis risk. Finally, higher plasma levels of the omega-3 fatty acid DHA in late adolescence were found to be associated with reduced odds of psychosis in early adulthood, consistent with a potential role for DHA in relation to psychosis prevention. These findings have implications for the understanding of the pathophysiology of psychotic disorders and provide novel insights to inform prediction and prevention strategies. With further confirmation, validation and investigation, it is hoped that the findings of this and subsequent research will contribute to improving outcomes for individuals with psychosis.

References for Chapter 1 and Chapter 5

1. Velthorst E, Fett A-KJ, Reichenberg A, Perlman G, van Os J, Bromet EJ, et al. The 20-Year Longitudinal Trajectories of Social Functioning in Individuals With Psychotic Disorders. American Journal of Psychiatry. 2016;174(11):1075-85.
2. World Health Organisation. ICD-10 : international statistical classification of diseases and related health problems : tenth revision. 2nd ed ed. Geneva: World Health Organization; 2004.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders : DSM-5. American Psychiatric A, American Psychiatric Association DSMTF, editors. Arlington, VA: American Psychiatric Association; 2013.
4. Maj M, van Os J, De Hert M, Gaebel W, Galderisi S, Green MF, et al. The clinical characterization of the patient with primary psychosis aimed at personalization of management. World Psychiatry. 2021;20(1):4-33.
5. Moreno-Küstner B, Martín C, Pastor L. Prevalence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. PLoS One. 2018;13(4):e0195687-e.
6. Simeone JC, Ward AJ, Rotella P, Collins J, Windisch R. An evaluation of variation in published estimates of schizoph systematic literature review. BMC Psychiatry. 2015;15:193.
7. McGrath J, Saha S, Chant D, Welham J. Schizophrenia: a concise overview of incidence, prevalence, and mortality. Epidemiol Rev. 2008;30:67-76.
8. Charlson FJ, Ferrari AJ, Santomauro DF, Diminic S, Stockings E, Scott JG, et al. Global Epidemiology and Burden of Schizophrenia: Findings From the Global Burden of Disease Study 2016. Schizophr Bull. 2018;44(6):1195-203.
9. Laursen TM, Nordentoft M, Mortensen PB. Excess early mortality in schizophrenia. Annu Rev Clin Psychol. 2014;10:425-48.
10. World Health Organisation. Excess mortality in persons with severe mental disorders. Geneva, Switzerland; 2016.
11. Dickens AM, Sen P, Kempton MJ, Barrantes-Vidal N, Iyegbe C, Nordentoft M, et al. Dysregulated Lipid Metabolism Precedes Onset of Psychosis. Biol Psychiatry. 2020.
12. Perry BI, Stochl J, Upthegrove R, Zammit S, Wareham N, Langenberg C, et al. Longitudinal Trends in Childhood Insulin Levels and Body Mass Index and Associations With Risks of Psychosis and Depression in Young Adults. JAMA Psychiatry. 2021.
13. Kahn RS, Sommer IE, Murray RM, Meyer-Lindenberg A, Weinberger DR, Cannon TD, et al. Schizophrenia. Nat Rev Dis Primers. 2015;1:15067.
14. Lutgens D, Gariepy G, Malla A. Psychological and psychosocial interventions for negative symptoms in psychosis: Systematic review and meta-analysis. British Journal of Psychiatry. 2017;210(5):324-32.
15. Cocchi A, Cerati G, Lora A, Meneghelli A, Monzani E, Percudani M, et al. Patients With First-Episode Psychosis are Not a Homogeneous Population: Implications for Treatment. Clin Pract Epidemiol Ment Health. 2014;10:1-8.
16. Hodgekins J, Birchwood M, Christopher R, Marshall M, Coker S, Everard L, et al. Investigating trajectories of social recovery in individuals with first-episode psychosis: A latent class growth analysis. British Journal of Psychiatry. 2015;207(6):536-43.

17. Lally J, Ajnakina O, Stubbs B, Cullinane M, Murphy KC, Gaughran F, et al. Remission and recovery from first-episode psychosis in adults: systematic review and meta-analysis of long-term outcome studies. Br J Psychiatry. 2017;211(6):350-8.
18. Fusar-Poli P, Cappucciati M, Rutigliano G, Heslin M, Stahl D, Brittenden Z, et al. Diagnostic Stability of ICD/DSM First Episode Psychosis Diagnoses: Meta-analysis. Schizophr Bull. 2016;42(6):1395-406.
19. Golay P, Alameda L, Mebdouhi N, Baumann P, Ferrari C, Solida A, et al. Age at the time of onset of psychosis: A marker of specific needs rather than a determinant of outcome? European Psychiatry. 2017;45:20-6.
20. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S-J, Dick B, Ezeh AC, et al. Adolescence: a foundation for future health. The Lancet. 2012;379(9826):1630-40.
21. Jääskeläinen E, Juola P, Hirvonen N, McGrath JJ, Saha S, Isohanni M, et al. A systematic review and meta-analysis of recovery in schizophrenia. Schizophr Bull. 2013;39(6):1296-306.
22. Harrison G, Hopper K, Craig T, Laska E, Siegel C, Wanderling J, et al. Recovery from psychotic illness: a 15- and 25-year international follow-up study. Br J Psychiatry. 2001;178:506-17.
23. Suvisaari J, Mantere O, Keinänen J, Mäntylä T, Rikandi E, Lindgren M, et al. Is It Possible to Predict the Future in First-Episode Psychosis? Frontiers in Psychiatry. 2018;9(580).
24. Marshall M, Lewis S, Lockwood A, Drake R, Jones P, Croudace T. Association Between Duration of Untreated Psychosis and Outcome in Cohorts of First-Episode Patients: A Systematic Review. Archives of General Psychiatry. 2005;62(9):975-83.
25. Howes OD, Whitehurst T, Shatalina E, Townsend L, Onwordi EC, Mak TLA, et al. The clinical significance of duration of untreated psychosis: an umbrella review and random-effects meta-analysis. World Psychiatry. 2021;20(1):75-95.
26. Birchwood M, Todd P, Jackson C. Early intervention in psychosis: The critical period hypothesis. British Journal of Psychiatry. 1998;172(S33):53-9.
27. Kuipers E, Yesufu-Udechuku A, Taylor C, Kendall T. Management of psychosis and schizophrenia in adults: summary of updated NICE guidance. British Medical Journal. 2014;348:g1173.
28. Klosterkötter J, Schultze-Lutter F, Ruhrmann S. Kraepelin and psychotic prodromal conditions. European Archives of Psychiatry and Clinical Neuroscience. 2008;258(2):74.
29. McGorry PD, Edwards J, Mihalopoulos C, Harrigan SM, Jackson HJ. EPPIC: an evolving system of early detection and optimal management. Schizophr Bull. 1996;22(2):305-26.
30. Yung AR, Yuen HP, McGorry PD, Phillips LJ, Kelly D, Dell'Olio M, et al. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. Aust N Z J Psychiatry. 2005;39(11-12):964-71.
31. Miller TJ, McGlashan TH, Rosen JL, Cadenhead K, Cannon T, Ventura J, et al. Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. Schizophr Bull. 2003;29(4):703-15.
32. Riecher-Rössler A, Aston J, Ventura J, Merlo M, Borgwardt S, Gschwandtner U, et al. [The Basel Screening Instrument for Psychosis (BSIP): development, structure, reliability and validity]. Fortschr Neurol Psychiatr. 2008;76(4):207-16.

33. Fusar-Poli P, Borgwardt S, Bechdolf A, Addington J, Riecher-Rössler A, Schultze-Lutter F, et al. The psychosis high-risk state: a comprehensive state-of-the-art review. JAMA Psychiatry. 2013;70(1):107-20.
34. Fusar-Poli P, Cappucciati M, Borgwardt S, Woods SW, Addington J, Nelson B, et al. Heterogeneity of Psychosis Risk Within Individuals at Clinical High Risk: A Meta-analytical Stratification. JAMA Psychiatry. 2016;73(2):113-20.
35. Fusar-Poli P, De Micheli A, Signorini L, Baldwin H, de Pablo GS, McGuire P. Real-world long-term outcomes in individuals at clinical risk for psychosis: The case for extending duration of care. EClinicalMedicine. 2020;28.
36. Fusar-Poli P, Salazar de Pablo G, Correll CU, Meyer-Lindenberg A, Millan MJ, Borgwardt S, et al. Prevention of Psychosis: Advances in Detection, Prognosis, and Intervention. JAMA Psychiatry. 2020;77(7):755-65.
37. 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. Archives of General Psychiatry. 2012;69(3):220-9.
38. Yung AR, Phillips LJ, Yuen HP, Francey SM, McFarlane CA, Hallgren M, et al. Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. Schizophr Res. 2003;60(1):21-32.
39. Cotter J, Drake RJ, Bucci S, Firth J, Edge D, Yung AR. What drives poor functioning in the at-risk mental state? A systematic review. Schizophr Res. 2014;159(2-3):267-77.
40. Lin A, Wood SJ, Nelson B, Beavan A, McGorry P, Yung AR. Outcomes of nontransitioned cases in a sample at ultra-high risk for psychosis. Am J Psychiatry. 2015;172(3):249-58.
41. Fusar-Poli P, Cappucciati M, Rutigliano G, Schultze-Lutter F, Bonoldi I, Borgwardt S, et al. At risk or not at risk? A meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction. World Psychiatry. 2015;14(3):322-32.
42. Addington J, Farris M, Devoe D, Metzak P. Progression from being at-risk to psychosis: next steps. npj Schizophrenia. 2020;6(1):27.
43. Riecher-Rössler A, Studerus E. Prediction of conversion to psychosis in individuals with an at-risk mental state: a brief update on recent developments. Curr Opin Psychiatry. 2017;30(3):209-19.
44. Malda A, Boonstra N, Barf H, de Jong S, Aleman A, Addington J, et al. Individualized Prediction of Transition to Psychosis in 1,676 Individuals at Clinical High Risk: Development and Validation of a Multivariable Prediction Model Based on Individual Patient Data Meta-Analysis. Frontiers in Psychiatry. 2019;10(345).
45. Montemagni C, Bellino S, Bracale N, Bozzatello P, Rocca P. Models Predicting Psychosis in Patients With High Clinical Risk: A Systematic Review. Frontiers in Psychiatry. 2020;11:223-.
46. Studerus E, Ramyead A, Riecher-Rössler A. Prediction of transition to psychosis in patients with a clinical high risk for psychosis: a systematic review of methodology and reporting. Psychol Med. 2017;47(7):1163-78.
47. Rosen M, Betz LT, Schultze-Lutter F, Chisholm K, Haidl TK, Kambeitz-Ilankovic L, et al. Towards clinical application of prediction models for transition to psychosis: A systematic review and external validation study in the PRONIA sample. Neurosci Biobehav Rev. 2021;125:478-92.
48. Mei C, van der Gaag M, Nelson B, Smit F, Yuen HP, Berger M, et al. Preventive interventions for individuals at ultra high risk for psychosis: An updated and extended meta-analysis. Clinical Psychology Review. 2021:102005.

49. Fusar-Poli P, Davies C, Solmi M, Brondino N, De Micheli A, Kotlicka-Antczak M, et al. Preventive Treatments for Psychosis: Umbrella Review (Just the Evidence). Frontiers in Psychiatry. 2019;10(764).
50. Davies C, Cipriani A, Ioannidis JPA, Radua J, Stahl D, Provenzani U, et al. Lack of evidence to favor specific preventive interventions in psychosis: a network meta-analysis. World Psychiatry. 2018;17(2):196-209.
51. Bosnjak Kuharic D, Kekin I, Hew J, Rojnic Kuzman M, Puljak L. Interventions for prodromal stage of psychosis. Cochrane Database of Systematic Reviews. 2019(11).
52. van Os J, Guloksuz S. A critique of the "ultra-high risk" and "transition" paradigm. World Psychiatry. 2017;16(2):200-6.
53. Ajnakina O, David AS, Murray RM. ïsïAt risk an idea whose time has come ï and gone! Psychological Medicine. 2019;49(4):529-34.
54. Ajnakina O, Morgan C, Gayer-Anderson C, Oduola S, Bourque F, Bramley S, et al. Only a small proportion of patients with first episode psychosis come via prodromal services: A retrospective survey of a large UK mental health programme. BMC Psychiatry. 2017;17.
55. Yung AR, Wood SJ, Malla A, Nelson B, McGorry P, Shah J. The reality of at risk mental state services: a response to recent criticisms. Psychological Medicine. 2019:1-7.
56. Shah JL, Crawford A, Mustafa SS, Iyer SN, Joober R, Malla AK. Is the Clinical High-Risk State a Valid Concept? Retrospective Examination in a First-Episode Psychosis Sample. Psychiatr Serv. 2017;68(10):1046-52.
57. van Os J, Reininghaus U. Psychosis as a transdiagnostic and extended phenotype in the general population. World psychiatry. 2016;15(2):118-24.
58. van Os J, Linscott RJ, Myin-Germeys I, Delespaul P, Krabbendam L. A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. Psychol Med. 2009;39(2):179-95.
59. Healy C, Brannigan R, Dooley N, Coughlan H, Clarke M, Kelleher I, et al. Childhood and adolescent psychotic experiences and risk of mental disorder: a systematic review and meta-analysis. Psychol Med. 2019;49(10):1589-99.
60. Dominguez MDG, Wichers M, Lieb R, Wittchen H-U, van Os J. Evidence That Onset of Clinical Psychosis Is an Outcome of Progressively More Persistent Subclinical Psychotic Experiences: An 8-Year Cohort Study. Schizophr Bull. 2009;37(1):84-93.
61. Yates K, Lång U, Cederlöf M, Boland F, Taylor P, Cannon M, et al. Association of Psychotic Experiences With Subsequent Risk of Suicidal Ideation, Suicide Attempts, and Suicide Deaths: A Systematic Review and Meta-analysis of Longitudinal Population Studies. JAMA Psychiatry. 2019;76(2):180-9.
62. Scott A, Khan KM, Cook JL, Duronio V. Wha to move beyond Celsus? British Journal of Sports Medicine. 2004;38(3):248.
63. Khandaker GM, Cousins L, Deakin J, Lennox BR, Yolken R, Jones PB. Inflammation and immunity in schizophrenia: implications for pathophysiology and treatment. Lancet Psychiatry. 2015;2(3):258-70.
64. Goldsmith DR, Rapaport MH, Miller BJ. A meta-analysis of blood cytokine network alterations in psychiatric patients: comparisons between schizophrenia, bipolar disorder and depression. Molecular Psychiatry. 2016;21(12):1696-709.

65. Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-Analysis of Cytokine Alterations in Schizophrenia: Clinical Status and Antipsychotic Effects. Biological Psychiatry. 2011;70(7):663-71.
66. Upthegrove R, Manzanares-Teson N, Barnes NM. Cytokine function in medication-naive first episode psychosis: A systematic review and meta-analysis. Schizophr Res. 2014;155(1-3):101-8.
67. Khoury R, Nasrallah HA. Inflammatory biomarkers in individuals at clinical high risk for psychosis (CHR-P): State or trait? Schizophr Res. 2018;199:31-8.
68. Park S, Miller BJ. Meta-analysis of cytokine and C-reactive protein levels in high-risk psychosis. Schizophr Res. 2020;226:5-12.
69. Metcalf S, Jones P, Nordstrom T, Timonen M, Maki P, Miettunen J, et al. Serum C-reactive protein in adolescence and risk of Schizophrenia in adulthood: A prospective birth cohort study from Finland. Brain, Behavior, and Immunity. 2017;66:e2-e3.
70. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB. Association of Serum Interleukin 6 and C-Reactive Protein in Childhood With Depression and Psychosis in Young Adult Life. JAMA Psychiatry. 2014;71(10):1121.
71. 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. 2014;41(2):419-28.
72. Föcking M, Dicker P, Lopez LM, Cannon M, Schafer MR, McGorry PD, et al. Differential expression of the inflammation marker IL12p40 in the at-risk mental state for psychosis: a predictor of transition to psychotic disorder? BMC Psychiatry. 2016;16(1) (no pagination).
73. Dunkelberger JR, Song W-C. Complement and its role in innate and adaptive immune responses. Cell Research. 2010;20(1):34-50.
74. Sarma JV, Ward PA. The complement system. Cell and Tissue Research. 2010;343(1):227-35.
75. Foley JH. Examining coagulation-complement crosstalk: complement activation and thrombosis. Thrombosis Research. 2016;141:S50-S4.
76. Elvington M, Liszewski MK, Atkinson JP. Evolution of the complement system: from defense of the single cell to guardian of the intravascular space. Immunol Rev. 2016;274(1):9-15.
77. Popescu A, H. Kao A. Neuropsychiatric Systemic Lupus Erythematosus. Current Neuropharmacology. 2011;9(3):449-57.
78. Ingram G, Hakobyan S, Robertson NP, Morgan BP. Complement in multiple sclerosis: its role in disease and potential as a biomarker. Clinical & Experimental Immunology. 2009;155(2):128-39.
79. Michailidou I, Naessens DMP, Hametner S, Guldenaar W, Kooi E-J, Geurts JJG, et al. Complement C3 on microglial clusters in multiple sclerosis occur in chronic but not acute disease: Implication for disease pathogenesis. Glia. 2016;65(2):264-77.
80. Kopczynska M, Zelek WM, Vespa S, Touchard S, Wardle M, Loveless S, et al. Complement system biomarkers in epilepsy. Seizure. 2018;60:1-7.
81. VanItallie TB. Alzheimer's disease: Innate immunity gone awry? Metabolism. 2017;69:S41-S9.
82. Hong S, Beja-Glasser VF, Nfonoyim BM, Frouin A, Li S, Ramakrishnan S, et al. Complement and microglia mediate early synapse loss in Alzheimer mouse models. Science. 2016;352(6286):712-6.

83. Nimgaonkar VL, Prasad KM, Chowdari KV, Severance EG, Yolken RH. The complement system: a gateway to gene-environment interactions in schizophrenia pathogenesis. Molecular Psychiatry. 2017;22(11):1554-61.
84. Woo JJ, Pouget JG, Zai CC, Kennedy JL. The complement system in schizophrenia: where are we now and what's next? Molecular Psychiatry. 2020;25(1):114-30.
85. Irish Schizophrenia Genomics Consortium and the Wellcome Trust Case Control Consortium 2. Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. Biological Psychiatry. 2012;72(8):620-8.
86. 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. 2009;460(7256):748-52.
87. Corvin A, Morris DW. Genome-wide association studies: findings at the major histocompatibility complex locus in psychosis. Biological Psychiatry. 2014;75(4):276-83.
88. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. Nature. 2016;530(7589):177-83.
89. 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.
90. Stephan AH, Barres BA, Stevens B. The Complement System: An Unexpected Role in Synaptic Pruning During Development and Disease. Annual Review of Neuroscience. 2012;35(1):369-89.
91. Osimo EF, Beck K, Reis Marques T, Howes OD. Synaptic loss in schizophrenia: a meta-analysis and systematic review of synaptic protein and mRNA measures. Molecular Psychiatry. 2018.
92. Shcherbakova I, Neshkova E, Dotsenko V, Platonova T, Shcherbakova E, Yarovaya G. The possible role of plasma kallikrein-kinin system and leukocyte elastase in pathogenesis of schizophrenia. Immunopharmacology. 1999;43(2-3):273-9.
93. Hakobyan S, Boyajyan A, Sim RB. Classical pathway complement activity in schizophrenia. Neuroscience Letters. 2005;374(1):35-7.
94. 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 Research. 1997;66(1):1-11.
95. Santos Soria L, Moura Gubert C, Cereser KM, Gama CS, Kapczinski F. Increased serum levels of C3 and C4 in patients with schizophrenia compared to eutymic patients with bipolar disorder and healthy. Revista Brasileira de Psiquiatria. 2012;34(1):119-20.
96. Ali FT, Abd El-Azeem EM, Hamed MA, Ali MAM, Abd Al-Kader NM, Hassan EA. Redox dysregulation, immuno-inflammatory alterations and genetic variants of BDNF and MMP-9 in schizophrenia: Pathophysiological and phenotypic implications. Schizophr Res. 2017;188:98-109.
97. Hong W, Zhao M, Li H, Peng F, Wang F, Li N, et al. Higher Plasma S100B Concentrations in Schizophrenia Patients, and Dependently Associated with Inflammatory Markers. Scientific Reports. 2016;6:27584-.

98. Mayilyan KR, Dodds AW, Boyajyan AS, Soghoyan AF, Sim RB. Complement C4B protein in schizophrenia. World Journal of Biological Psychiatry. 2008;9(3):225-30.
99. Schwarz E, Guest PC, Rahmoune H, Harris LW, Wang L, Leweke FM, et al. Identification of a biological signature for schizophrenia in serum. Molecular Psychiatry. 2012;17(5):494-502.
100. Kopczynska M, Zelek W, Touchard S, Gaughran F, Di Forti M, Mondelli V, et al. Complement system biomarkers in first episode psychosis. Schizophr Res. 2019;204:16-22.
101. English JA, Lopez LM, O'Gorman A, Focking M, Hryniewiecka M, Scaife C, et al. Blood-Based Protein Changes in Childhood Are Associated With Increased Risk for Later Psychotic Disorder: Evidence From a Nested Case-Control Study of the ALSPAC Longitudinal Birth Cohort. Schizophr Bull. 2018;44(2):297-306.
102. Severance EG, Gressitt KL, Buka SL, Cannon TD, Yolken RH. Maternal complement C1q and increased odds for psychosis in adult offspring. Schizophr Res. 2014;159(1):14-9.
103. Anderson KK, Norman R, MacDougall A, Edwards J, Palaniyappan L, Lau C, et al. Effectiveness of Early Psychosis Intervention: Comparison of Service Users and Nonusers in Population-Based Health Administrative Data. American Journal of Psychiatry. 2018;175(5):443-52.
104. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89-95.
105. Selleck MJ, Senthil M, Wall NR. Making Meaningful Clinical Use of Biomarkers. Biomarker Insights. 2017;12:1177271917715236-.
106. de Bock R, Mackintosh AJ, Maier F, Borgwardt S, Riecher-Rössler A, Andreou C. EEG microstates as biomarker for psychosis in ultra-high-risk patients. Translational Psychiatry. 2020;10(1):300.
107. van Tricht MJ, Ruhrmann S, Arns M, Müller R, Bodatsch M, Velthorst E, et al. Can quantitative EEG measures predict clinical outcome in subjects at Clinical High Risk for psychosis? A prospective multicenter study. Schizophrenia Research. 2014;153(1):42-7.
108. Ranlund S, Nottage J, Shaikh M, Dutt A, Constante M, Walshe M, et al. Resting EEG in psychosis and at-risk populations--a possible endophenotype? Schizophr Res. 2014;153(1-3):96-102.
109. Hamilton HK, Boos AK, Mathalon DH. Electroencephalography and Event-Related Potential Biomarkers in Individuals at Clinical High Risk for Psychosis. Biological Psychiatry. 2020;88(4):294-303.
110. Seidman LJ, Giuliano AJ, Meyer EC, Addington J, Cadenhead KS, Cannon TD, et al. Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. Arch Gen Psychiatry. 2010;67(6):578-88.
111. Keefe RS, Perkins DO, Gu H, Zipursky RB, Christensen BK, Lieberman JA. A longitudinal study of neurocognitive function in individuals at-risk for psychosis. Schizophr Res. 2006;88(1-3):26-35.
112. Brewer WJ, Francey SM, Wood SJ, Jackson HJ, Pantelis C, Phillips LJ, et al. Memory impairments identified in people at ultra-high risk for psychosis who later develop first-episode psychosis. Am J Psychiatry. 2005;162(1):71-8.
113. Jahshan C, Heaton RK, Golshan S, Cadenhead KS. Course of neurocognitive deficits in the prodrome and first episode of schizophrenia. Neuropsychology. 2010;24(1):109-20.

114. Bang M, Kim KR, Song YY, Baek S, Lee E, An SK. Neurocognitive impairments in individuals at ultra-high risk for psychosis: Who will really convert? Australian & New Zealand Journal of Psychiatry. 2014;49(5):462-70.
115. De Herdt A, Wampers M, Vancampfort D, De Hert M, Vanhees L, Demunter H, et al. Neurocognition in clinical high risk young adults who did or did not convert to a first schizophrenic psychosis: A meta-analysis. Schizophr Res. 2013;149(1):48-55.
116. Andreou C, Borgwardt S. Structural and functional imaging markers for susceptibility to psychosis. Molecular Psychiatry. 2020;25(11):2773-85.
117. Carletti F, Woolley JB, Bhattacharyya S, Perez-Iglesias R, Fusar Poli P, Valmaggia L, et al. Alterations in white matter evident before the onset of psychosis. Schizophr Bull. 2012;38(6):1170-9.
118. Bloemen OJ, de Koning MB, Schmitz N, Nieman DH, Becker HE, de Haan L, et al. White-matter markers for psychosis in a prospective ultra-high-risk cohort. Psychol Med. 2010;40(8):1297-304.
119. Garner B, Pariante CM, Wood SJ, Velakoulis D, Phillips L, Soulsby B, et al. Pituitary volume predicts future transition to psychosis in individuals at ultra-high risk of developing psychosis. Biological Psychiatry. 2005;58(5):417-23.
120. Du Y, Fryer SL, Lin D, Sui J, Yu Q, Chen J, et al. Identifying functional network changing patterns in individuals at clinical high-risk for psychosis and patients with early illness schizophrenia: A group ICA study. NeuroImage: Clinical. 2018;17:335-46.
121. Koutsouleris N, Meisenzahl EM, Davatzikos C, Bottlender R, Frodl T, Scheuerecker J, et al. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. Arch Gen Psychiatry. 2009;66(7):700-12.
122. Kambeitz-Ilankovic L, Meisenzahl EM, Cabral C, von Saldern S, Kambeitz J, Falkai P, et al. Prediction of outcome in the psychosis prodrome using neuroanatomical pattern classification. Schizophr Res. 2016;173(3):159-65.
123. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, Rosen M, Ruef A, Dwyer DB, et al. Prediction Models of Functional Outcomes for Individuals in the Clinical High-Risk State for Psychosis or With Recent-Onset Depression: A Multimodal, Multisite Machine Learning Analysis. JAMA Psychiatry. 2018;75(11):1156-72.
124. Walker E, Mittal V, Tessner K. Stress and the hypothalamic pituitary adrenal axis in the developmental course of schizophrenia. Annu Rev Clin Psychol. 2008;4:189-216.
125. Day FL, Valmaggia LR, Mondelli V, Papadopoulos A, Papadopoulos I, Pariante CM, et al. Blunted Cortisol Awakening Response in People at Ultra High Risk of Developing Psychosis. Schizophr Res. 2014;158(1):25-31.
126. Walker EF, Trotman HD, Pearce BD, Addington J, Cadenhead KS, Cornblatt BA, et al. Cortisol levels and risk for psychosis: initial findings from the North American prodrome longitudinal study. Biological Psychiatry. 2013;74(6):410-7.
127. Karanikas E, Garyfallos G. Role of cortisol in patients at risk for psychosis mental state and psychopathological correlates: A systematic review. Psychiatry Clin Neurosci. 2015;69(5):268-82.
128. Chaumette B, Kebir O, Mam-Lam-Fook C, Morvan Y, Bourgin J, Godsil BP, et al. Salivary cortisol in early psychosis: New findings and meta-analysis. Psychoneuroendocrinology. 2016;63:262-70.
129. Mongan D, Ramesar M, Föcking M, Cannon M, Cotter D. Role of inflammation in the pathogenesis of schizophrenia: A review of the evidence,

proposed mechanisms and implications for treatment. Early Intervention in Psychiatry. 2020;14(4):385-97.
130. 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. 2015;41(2):419-28.
131. Chan MK, Krebs MO, 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. Translational Psychiatry. 2015;5(7):e601-e.
132. Mendelson T, Eaton WW. Recent advances in the prevention of mental disorders. Social Psychiatry and Psychiatric Epidemiology. 2018;53(4):325-39.
133. Godos J, Currenti W, Angelino D, Mena P, Castellano S, Caraci F, et al. Diet and Mental Health: Review of the Recent Updates on Molecular Mechanisms. Antioxidants. 2020;9(4).
134. McGrath J, Brown A, St Clair D. Prevention and schizophrenia--the role of dietary factors. Schizophr Bull. 2011;37(2):272-83.
135. Jacka FN, Ystrom E, Brantsaeter AL, Karevold E, Roth C, Haugen M, et al. Maternal and early postnatal nutrition and mental health of offspring by age 5 years: a prospective cohort study. J Am Acad Child Adolesc Psychiatry. 2013;52(10):1038-47.
136. Schmitz G, Ecker J. The opposing effects Progress in Lipid Research. 2008;47(2):147-55.
137. Hsu M-C, Huang Y-S, Ouyang W-C. Beneficial effects of omega-3 fatty acid supplementation in schizophrenia: possible mechanisms. Lipids in Health and Disease. 2020;19(1):159.
138. Weiser MJ, Butt CM, Mohajeri MH. Docosahexaenoic Acid and Cognition throughout the Lifespan. Nutrients. 2016;8(2):99-.
139. Hoen WP, Lijmer JG, Duran M, Wanders RJ, van Beveren NJ, de Haan L. Red blood cell polyunsaturated fatty acids measured in red blood cells and schizophrenia: a meta-analysis. Psychiatry Res. 2013;207(1-2):1-12.
140. Lin PY, Huang SY, Su KP. A meta-analytic review of polyunsaturated fatty acid compositions in patients with depression. Biological Psychiatry. 2010;68(2):140-7.
141. Ross BM. Omega-3 polyunsaturated fatty acids and anxiety disorders. Prostaglandins Leukot Essent Fatty Acids. 2009;81(5-6):309-12.
142. Jacka FN, Pasco JA, Williams LJ, Meyer BJ, Digger R, Berk M. Dietary intake of fish and PUFA, and clinical depressive and anxiety disorders in women. Br J Nutr. 2013;109(11):2059-66.
143. Li F, Liu X, Zhang D. Fish consumption and risk of depression: a meta-analysis. Journal of Epidemiology and Community Health. 2016;70(3):299.
144. Smith KJ, Sanderson K, McNaughton SA, Gall SL, Dwyer T, Venn AJ. Longitudinal associations between fish consumption and depression in young adults. Am J Epidemiol. 2014;179(10):1228-35.
145. Astorg P, Bertrais S, Alessandri JM, Guesnet P, Kesse-Guyot E, Linard A, et al. Long-chain n-3 fatty acid levels in baseline serum phospholipids do not predict later occurrence of depressive episodes: a nested case-control study within a cohort of middle-aged French men and women. Prostaglandins Leukot Essent Fatty Acids. 2009;81(4):265-71.

146. Persons JE, Robinson JG, Ammann EM, Coryell WH, Espeland MA, Harris WS, et al. Omega-3 fatty acid biomarkers and subsequent depressive symptoms. Int J Geriatr Psychiatry. 2014;29(7):747-57.
147. Firth J, Teasdale SB, Allott K, Siskind D, Marx W, Cotter J, et al. The efficacy and safety of nutrient supplements in the treatment of mental disorders: a meta-review of meta-analyses of randomized controlled trials. World Psychiatry. 2019;18(3):308-24.
148. Amminger GP, Schäfer MR, Schlögelhofer M, Klier CM, McGorry PD. Longer-term outcome in the prevention of psychotic disorders by the Vienna omega-3 study. Nature Communications. 2015;6(1):7934.
149. McGorry PD, Nelson B, Markulev C, Yuen HP, Schäfer MR, Mossaheb N, et al. Effect of ω-3 Polyunsaturated Fatty Acids in Young People at Ultrahigh Risk for Psychotic Disorders: The NEURAPRO Randomized Clinical Trial. JAMA Psychiatry. 2017;74(1):19-27.
150. Amminger GP, Nelson B, Markulev C, Yuen HP, Schäfer MR, Berger M, et al. The NEURAPRO Biomarker Analysis: Long-Chain Omega-3 Fatty Acids Improve 6-Month and 12-Month Outcomes in Youths at Ultra-High Risk for Psychosis. Biological Psychiatry. 2020;87(3):243-52.
151. McGrath JJ, Saha S, Al-Hamzawi AO, Alonso J, Andrade L, Borges G, et al. Age of Onset and Lifetime Projected Risk of Psychotic Experiences: Cross-National Data From the World Mental Health Survey. Schizophr Bull. 2016;42(4):933-41.
152. Ramsay H, Barnett JH, Murray GK, Miettunen J, Mäki P, Järvelin M-R, et al. Cognition, psychosis risk and metabolic measures in two adolescent birth cohorts. Psychological Medicine. 2018;48(15):2609-23.
153. Thompson AD, Jones HJ, Heron J, Hibbeln J, Sullivan S, Zammit S. Omega-3 and Omega-6 fatty acids and risk of psychotic outcomes in the ALSPAC birth cohort. Schizophr Res. 2020;224:108-15.
154. Berger ME, Smesny S, Kim SW, Davey CG, Rice S, Sarnyai Z, et al. Omega-6 to omega-3 polyunsaturated fatty acid ratio and subsequent mood disorders in young people with at-risk mental states: a 7-year longitudinal study. Translational Psychiatry. 2017;7(8):e1220.
155. Berger M, Nelson B, Markulev C, Yuen HP, Schäfer MR, Mossaheb N, et al. Relationship Between Polyunsaturated Fatty Acids and Psychopathology in the NEURAPRO Clinical Trial. Frontiers in Psychiatry. 2019;10:393-.
156. European Network of National Networks studying Gene-Environment Interactions in Schizophrenia, van Os J, Rutten BP, Myin-Germeys I, Delespaul P, Viechtbauer W, et al. Identifying gene-environment interactions in schizophrenia: contemporary challenges for integrated, large-scale investigations. Schizophr Bull. 2014;40(4):729-36.
157. Kraan TC, Velthorst E, Themmen M, Valmaggia L, Kempton MJ, McGuire P, et al. Child Maltreatment and Clinical Outcome in Individuals at Ultra-High Risk for Psychosis in the EU-GEI High Risk Study. Schizophr Bull. 2018;44(3):584-92.
158. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol. 2013;42(1):111-27.
159. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. Int J Epidemiol. 2013;42(1):97-110.

160. Gallego JA, Blanco EA, Morell C, Lencz T, Malhotra AK. Complement component C4 levels in the cerebrospinal fluid and plasma of patients with schizophrenia. Neuropsychopharmacology. 2020.
161. Dickinson D, Pratt DN, Giangrande EJ, Grunnagle M, Orel J, Weinberger DR, et al. Attacking Heterogeneity in Schizophrenia by Deriving Clinical Subgroups From Widely Available Symptom Data. Schizophr Bull. 2018;44(1):101-13.
162. Schwarz E, van Beveren NJM, Ramsey J, Leweke FM, Rothermundt M, Bogerts B, et al. Identification of subgroups of schizophrenia patients with changes in either immune or growth factor and hormonal pathways. Schizophr Bull. 2014;40(4):787-95.
163. Kroken RA, Sommer IE, Steen VM, Dieset I, Johnsen E. Constructing the Immune Signature of Schizophrenia for Clinical Use and Research; An Integrative Review Translating Descriptives Into Diagnostics. Frontiers in Psychiatry. 2019;9(753).
164. Borth W. Alpha 2-macroglobulin, a multifunctional binding protein with targeting characteristics. FASEB Journal. 1992;6(15):3345-53.
165. Noris M, Remuzzi G. Overview of complement activation and regulation. Semin Nephrol. 2013;33(6):479-92.
166. Kopczynska M, Zelek W, Touchard S, Gaughran F, Di Forti M, Mondelli V, et al. Complement system biomarkers in first episode psychosis. Schizophr Res. 2019;204:16-22.
167. Laskaris L, Zalesky A, Weickert CS, Di Biase MA, Chana G, Baune BT, et al. Investigation of peripheral complement factors across stages of psychosis. Schizophr Res. 2019;204:30-7.
168. English JA, Lopez LM, OôGorman A, Fºcking al. Blood-Based Protein Changes in Childhood Are Associated With Increased Risk for Later Psychotic Disorder: Evidence From a Nested Caseï Control Study of the ALSPAC Longitudinal Birth Cohort. Schizophr Bull. 2018;44(2):297-306.
169. Gogtay N, Vyas NS, Testa R, Wood SJ, Pantelis C. Age of onset of schizophrenia: perspectives from structural neuroimaging studies. Schizophr Bull. 2011;37(3):504-13.
170. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. Nature. 2016;530(7589):177-83.
171. 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 [11C]PBR28 PET Brain Imaging Study. Am J Psych. 2015;173(1):44-52.
172. Hafizi S, Da Silva T, Gerritsen C, Kiang M, Bagby RM, Prce I, et al. Imaging Microglial Activation in Individuals at Clinical High Risk for Psychosis: an In Vivo PET Study with [(18)F]FEPPA. Neuropsychopharmacology. 2017;42(13):2474-81.
173. Osimo EF, Beck K, Reis Marques T, Howes OD. Synaptic loss in schizophrenia: a meta-analysis and systematic review of synaptic protein and mRNA measures. Molecular Psychiatry. 2019;24(4):549-61.
174. Schmidt A, Cappucciati M, Radua J, Rutigliano G, Rocchetti M, Dell'Osso L, et al. Improving Prognostic Accuracy in Subjects at Clinical High Risk for Psychosis: Systematic Review of Predictive Models and Meta-analytical Sequential Testing Simulation. Schizophr Bull. 2017;43(2):375-88.
175. Bourque J, Spechler PA, Potvin S, Whelan R, Banaschewski T, Bokde ALW, et al. Functional Neuroimaging Predictors of Self-Reported Psychotic Symptoms in Adolescents. Am J Psych. 2017;174(6):566-75.

176. Linscott RJ, van Os J. An updated and conservative systematic review and meta-analysis of epidemiological evidence on psychotic experiences in children and adults: on the pathway from proneness to persistence to dimensional expression across mental disorders. Psychol Med. 2013;43(6):1133-49.
177. Deane KHO, Jimoh OF, Biswas P, O'Brien A, Hanson S, Abdelhamid AS, et al. Omega-3 and polyunsaturated fat for prevention of depression and anxiety symptoms: systematic review and meta-analysis of randomised trials. Br J Psychiatry. 2021;218(3):135-42.
178. van Reekum R, Streiner DL, Conn DK. Applying Bradford Hill's Criteria for Causation to Neuropsychiatry. The Journal of Neuropsychiatry and Clinical Neurosciences. 2001;13(3):318-25.
179. Calder PC. Omega-3 fatty acids and inflammatory processes. Nutrients. 2010;2(3):355-74.
180. McNamara RK, Vannest JJ, Valentine CJ. Role of perinatal long-chain omega-3 fatty acids in cortical circuit maturation: Mechanisms and implications for psychopathology. World J Psychiatry. 2015;5(1):15-34.
181. Cutuli D. Functional and Structural Benefits Induced by Omega-3 Polyunsaturated Fatty Acids During Aging. Curr Neuropharmacol. 2017;15(4):534-42.
182. Tanaka K, Farooqui AA, Siddiqi NJ, Alhomida AS, Ong W-Y. Effects of docosahexaenoic Acid on neurotransmission. Biomol Ther (Seoul). 2012;20(2):152-7.
183. Che H, Fu X, Zhang L, Gao X, Wen M, Du L, et al. Neuroprotective Effects of n-3 Polyunsaturated Fatty Acid-Enriched Phosphatidylserine Against Oxidative Damage in PC12 Cells. Cell Mol Neurobiol. 2018;38(3):657-68.
184. Heras-Sandoval D, Pedraza-Chaverri J, Pérez-Rojas JM. Role of docosahexaenoic acid in the modulation of glial cells in Alzheimer's disease. J Neuroinflammation. 2016;13(1):61.
185. Dyall SC. Long-chain omega-3 fatty acids and the brain: a review of the independent and shared effects of EPA, DPA and DHA. Front Aging Neurosci. 2015;7:52.
186. Feinberg I. Schizophrenia: caused by a fault in programmed synaptic elimination during adolescence? J Psychiatr Res. 1982;17(4):319-34.
187. Madore C, Leyrolle Q, Morel L, Rossitto M, Greenhalgh AD, Delpech JC, et al. Essential omega-3 fatty acids tune microglial phagocytosis of synaptic elements in the mouse developing brain. Nature Communications. 2020;11(1):6133.
188. Cao D, Kevala K, Kim J, Moon H-S, Jun SB, Lovinger D, et al. Docosahexaenoic acid promotes hippocampal neuronal development and synaptic function. J Neurochem. 2009;111(2):510-21.
189. Deckelbaum RJ, Torrejon C. The Omega-3 Fatty Acid Nutritional Landscape: Health Benefits and Sources. The Journal of Nutrition. 2012;142(3):587S-91S.
190. Allswede DM, Yolken RH, Buka SL, Cannon TD. Cytokine concentrations throughout pregnancy and risk for psychosis in adult offspring: a longitudinal case-control study. Lancet Psychiatry. 2020;7(3):254-61.
191. Allswede DM, Cannon TD. Prenatal inflammation and risk for schizophrenia: A role for immune proteins in neurodevelopment. Development and Psychopathology. 2018;30(3):1157-78.
192. Radua J, Ramella-Cravaro V, Ioannidis JPA, Reichenberg A, Phiphopthatsanee N, Amir T, et al. What causes psychosis? An umbrella review of risk and protective factors. World Psychiatry. 2018;17(1):49-66.

193. Davies C, Segre G, Estradé A, Radua J, De Micheli A, Provenzani U, et al. Prenatal and perinatal risk and protective factors for psychosis: a systematic review and meta-analysis. Lancet Psychiatry. 2020;7(5):399-410.
194. Zwicker A, Denovan-Wright EM, Uher R. Gene-environment interplay in the etiology of psychosis. Psychol Med. 2018;48(12):1925-36.
195. Regal JF, Gilbert JS, Burwick RM. The complement system and adverse pregnancy outcomes. Mol Immunol. 2015;67(1):56-70.
196. Girardi G, Fraser J, Lennen R, Vontell R, Jansen M, Hutchison G. Imaging of activated complement using ultrasmall superparamagnetic iron oxide particles (USPIO)-conjugated vectors: an in vivo in utero non-invasive method to predict placental insufficiency and abnormal fetal brain development. Molecular Psychiatry. 2015;20(8):1017-26.
197. Semple DM, McIntosh AM, Lawrie SM. Cannabis as a risk factor for psychosis: systematic review. Journal of Psychopharmacology. 2005;19(2):187-94.
198. Gurillo P, Jauhar S, Murray RM, MacCabe JH. Does tobacco use cause psychosis? Systematic review and meta-analysis. Lancet Psychiatry. 2015;2(8):718-25.
199. Fett A-KJ, Lemmers-Jansen ILJ, Krabbendam L. Psychosis and urbanicity: a review of the recent literature from epidemiology to neurourbanism. Current opinion in psychiatry. 2019;32(3):232-41.
200. Fusar-Poli P, Tantardini M, De Simone S, Ramella-Cravaro V, Oliver D, Kingdon J, et al. Deconstructing Vulnerability for Psychosis: Meta-Analysis of Environmental Risk Factors for Psychosis in Subjects at Ultra High-Risk. European Psychiatry. 2017;40:65-75.
201. Rodrigues M, Wiener JC, Stranges S, Ryan BL, Anderson KK. The risk of physical multimorbidity in people with psychotic disorders: A systematic review and meta-analysis. Journal of Psychosomatic Research. 2021;140:110315.
202. Molloy C, Conroy RM, Cotter DR, Cannon M. Is Traumatic Brain Injury A Risk Factor for Schizophrenia? A Meta-Analysis of Case-Controlled Population-Based Studies. Schizophr Bull. 2011;37(6):1104-10.
203. Morgan C, Gayer-Anderson C. Childhood adversities and psychosis: evidence, challenges, implications. World Psychiatry. 2016;15(2):93-102.
204. Sideli L, Murray RM, Schimmenti A, Corso M, La Barbera D, Trotta A, et al. Childhood adversity and psychosis: a systematic review of bio-psycho-social mediators and moderators. Psychol Med. 2020;50(11):1761-82.
205. Aas M, Dazzan P, Mondelli V, Melle I, Murray R, Pariante C. A Systematic Review of Cognitive Function in First-Episode Psychosis, Including a Discussion on Childhood Trauma, Stress, and Inflammation. Frontiers in Psychiatry. 2014;4(182).
206. Rasmussen LJH, Moffitt TE, Arseneault L, Danese A, Eugen-Olsen J, Fisher HL, et al. Association of Adverse Experiences and Exposure to Violence in Childhood and Adolescence With Inflammatory Burden in Young People. JAMA Pediatrics. 2020;174(1):38-47.
207. Hovhannisyan LP, Mkrtchyan GM, Sukiasian SH, Boyajyan AS. Alterations in the complement cascade in post-traumatic stress disorder. Allergy, Asthma & Clinical Immunology. 2010;6(1):3.
208. Föcking M, Sabherwal S, Cates HM, Scaife C, Dicker P, Hryniewiecka M, et al. Complement pathway changes at age 12 are associated with psychotic experiences at age 18 in a longitudinal population-based study: evidence for a role of stress. Molecular Psychiatry. 2021;26(2):524-33.

209. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB. Association of Serum Interleukin 6 and C-Reactive Protein in Childhood With Depression and Psychosis in Young Adult Life: A Population-Based Longitudinal Study. JAMA Psychiatry. 2014;71(10):1121-8.
210. Miller BJ, Goldsmith DR. Inflammatory biomarkers in schizophrenia: Implications for heterogeneity and neurobiology. Biomarkers in Neuropsychiatry. 2019;1:100006.
211. Upthegrove R, Manzanares-Teson N, Barnes NM. Cytokine function in medication-naive first episode psychosis: a systematic review and meta-analysis. Schizophr Res. 2014;155(1-3):101-8.
212. Veerhuis R, Nielsen HM, Tenner AJ. Complement in the brain. Mol Immunol. 2011;48(14):1592-603.
213. Alexander JJ. Blood-brain barrier (BBB) and the complement landscape. Mol Immunol. 2018;102:26-31.
214. Pollak TA, Drndarski S, Stone JM, David AS, McGuire P, Abbott NJ. The blood-brain barrier in psychosis. Lancet Psychiatry. 2018;5(1):79-92.
215. Magdalon J, Mansur F, Teles e Silva AL, de Goes VA, Reiner O, Sertié AL. Complement System in Brain Architecture and Neurodevelopmental Disorders. Frontiers in Neuroscience. 2020;14(23).
216. Jacob A, Alexander JJ. Complement and blood-brain barrier integrity. Mol Immunol. 2014;61(2):149-52.
217. Griffa A, Baumann PS, Klauser P, Mullier E, Cleusix M, Jenni R, et al. Brain connectivity alterations in early psychosis: from clinical to neuroimaging staging. Translational Psychiatry. 2019;9(1):62.
218. Conway EM. Complement-coagulation connections. Blood Coagul Fibrinolysis. 2018;29(3):243-51.
219. Hoirisch-Clapauch S, Amaral OB, Mezzasalma MAU, Panizzutti R, Nardi AE. Dysfunction in the coagulation system and schizophrenia. Translational Psychiatry. 2016;6(1):e704-e.
220. Lin C-E, Chung C-H, Chen L-F, Chien W-C. Increased risk for venous thromboembolism among patients with concurrent depressive, bipolar, and schizophrenic disorders. General Hospital Psychiatry. 2019;61:34-40.
221. Arasteh O, Nomani H, Baharara H, Sadjadi SA, Mohammadpour AH, Ghavami V, et al. Antipsychotic Drugs and Risk of Developing Venous Thromboembolism and Pulmonary Embolism: A Systematic Review and Meta-Analysis. Curr Vasc Pharmacol. 2020;18(6):632-43.
222. Ryu JK, Rafalski VA, Meyer-Franke A, Adams RA, Poda SB, Rios Coronado PE, et al. Fibrin-targeting immunotherapy protects against neuroinflammation and neurodegeneration. Nat Immunol. 2018;19(11):1212-23.
223. Hodes GE, Kana V, Menard C, Merad M, Russo SJ. Neuroimmune mechanisms of depression. Nat Neurosci. 2015;18(10):1386-93.
224. Estes ML, McAllister AK. Immune mediators in the brain and peripheral tissues in autism spectrum disorder. Nat Rev Neurosci. 2015;16(8):469-86.
225. Hou R, Baldwin DS. A neuroimmunological perspective on anxiety disorders. Hum Psychopharmacol. 2012;27(1):6-14.
226. Loftis JM, Huckans M. Substance use disorders: psychoneuroimmunological mechanisms and new targets for therapy. Pharmacol Ther. 2013;139(2):289-300.
227. Fries GR, Walss-Bass C, Bauer ME, Teixeira AL. Revisiting inflammation in bipolar disorder. Pharmacology Biochemistry and Behavior. 2019;177:12-9.

228. Wei J, Liu Y, Zhao L, Yang X, Ni P, Wang Y, et al. Plasma complement component 4 increases in patients with major depressive disorder. Neuropsychiatr Dis Treat. 2017;14:37-41.
229. Shin C, Ham BJ, Ko YH, Pae CU, Park MH, Steffens DC, et al. Increased plasma complement factor H is associated with geriatric depression. Int Psychogeriatr. 2019;31(1):101-8.
230. Akcan U, Karabulut S, Ķsmail K¿-¿kali C, patients display reduced serum complement levels and elevated peripheral blood complement expression levels. Acta Neuropsychiatrica. 2018;30(2):70-8.
231. Hoirisch-Clapauch S, Nardi AE, Gris J-C, Brenner B. Coagulation and mental disorders. Rambam Maimonides Med J. 2014;5(4):e0036-e.
232. Lopez-Vilchez I, Serra-Millas M, Navarro V, Rosa Hernandez M, Villalta J, Diaz-Ricart M, et al. Prothrombotic platelet phenotype in major depression: downregulation by antidepressant treatment. J Affect Disord. 2014;159:39-45.
233. McGorry PD, Hartmann JA, Spooner R, Nelson B. Beyond the "at risk mental state" concept: transitioning to transdiagnostic psychiatry. World Psychiatry. 2018;17(2):133-42.
234. Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, et al. The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? Clin Psychol Sci. 2014;2(2):119-37.
235. Fried EI, Greene AL, Eaton NR. The p factor is the sum of its parts, for now. World Psychiatry. 2021;20(1):69-70.
236. Druart M, Le Magueresse C. Emerging Roles of Complement in Psychiatric Disorders. Frontiers in psychiatry. 2019;10:573.
237. Cruz MP. Conestat alfa (ruconest): first recombinant c1 esterase inhibitor for the treatment of acute attacks in patients with hereditary angioedema. Pharmacy and Therapeutics. 2015;40(2):109-14.
238. Ricklin D, Mastellos DC, Reis ES, Lambris JD. The renaissance of complement therapeutics. Nat Rev Nephrol. 2018;14(1):26-47.
239. Dalakas MC, Alexopoulos H, Spaeth PJ. Complement in neurological disorders and emerging complement-targeted therapeutics. Nature Reviews Neurology. 2020;16(11):601-17.
240. Harris CL. Expanding horizons in complement drug discovery: challenges and emerging strategies. Semin Immunopathol. 2018;40(1):125-40.
241. Morgan BP. Complement in the pathogenesis of Alzheimer's disease. Semin Immunopathol. 2018;40(1):113-24.
242. Hoirisch-Clapauch S, Nardi AE. Psychiatric remission with warfarin: Should psychosis be addressed as plasminogen activator imbalance? Medical Hypotheses. 2013;80(2):137-41.
243. Franceschini G, Calabresi L, Maderna P, Galli C, Gianfranceschi G, Sirtori CR. Omega-3 fatty acids selectively raise high-density lipoprotein 2 levels in healthy volunteers. Metabolism. 1991;40(12):1283-6.
244. Burillo E, Mateo-Gallego R, Cenarro A, Fiddyment S, Bea AM, Jorge I, et al. Beneficial effects of omega-3 fatty acids in the proteome of high-density lipoprotein proteome. Lipids in Health and Disease. 2012;11(1):116.
245. Lawrie SM, Fletcher-Watson S, Whalley HC, McIntosh AM. Predicting major mental illness: ethical and practical considerations. BJPsych Open. 2019;5(2):e30-e.
246. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of

studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.
247. Dickens AM, Sen P, Kempton MJ, Barrantes-Vidal N, Iyegbe C, Nordentoft M, et al. Dysregulated Lipid Metabolism Precedes Onset of Psychosis. Biological Psychiatry. 2021;89(3):288-97.
248. Sun Q, Ma J, Campos H, Hankinson SE, Hu FB. Comparison between plasma and erythrocyte fatty acid content as biomarkers of fatty acid intake in US women. Am J Clin Nutr. 2007;86(1):74-81.