Full-length Article

Machine learning based prediction and the influence of complement - coagulation pathway proteins on clinical outcome: results from the NEURAPRO trial

Subash Raj Susai, David Mongan, Colm Healy, Mary Cannon, Gerard Cagney, Kieran Wynne, Jonah F. Byrne, Connie Markulev, Miriam R. Schäfer, Maximus Berger, Nilufar Mossaheb, Monika Schlögelhofer, Stefan Smesny, Ian B. Hickie, Gregor E. Berger, Eric Y.H. Chen, Lieuwe de Haan, Dorien H. Nieman, Merete Nordentoft, Anita Riecher-Rössler, Swapna Verma, Rebekah Street, Andrew Thompson, Alison Ruth Yung, Barnaby Nelson, Patrick D. McGorry, Melanie Föcking, G. Paul Amminger, David Cotter

PII: S0889-1591(22)00081-2
DOI: https://doi.org/10.1016/j.bbi.2022.03.013
Reference: YBRBI 4819

To appear in: *Brain, Behavior, and Immunity*

Received Date: 10 November 2021
Revised Date: 21 February 2022
Accepted Date: 21 March 2022

Please cite this article as: Raj Susai, S., Mongan, D., Healy, C., Cannon, M., Cagney, G., Wynne, K., Byrne, J.F., Markulev, C., Schäfer, M.R., Berger, M., Mossaheb, N., Schlögelhofer, M., Smesny, S., Hickie, I.B., Berger, G.E., Chen, E.Y.H., de Haan, L., Nieman, D.H., Nordentoft, M., Riecher-Rössler, A., Verma, S., Street, R., Thompson, A., Ruth Yung, A., Nelson, B., McGorry, P.D., Föcking, M., Paul Amminger, G., Cotter, D., Machine learning based prediction and the influence of complement - coagulation pathway proteins on clinical outcome: results from the NEURAPRO trial, *Brain, Behavior, and Immunity* (2022), doi: https://doi.org/10.1016/j.bbi.2022.03.013

Title: Machine learning based prediction and the influence of complement - coagulation pathway proteins on clinical outcome: results from the NEURAPRO trial

Authors: Subash Raj Susai*[a], David Mongan*[a], Colm Healy[a], Mary Cannon[a], Gerard Cagney[b], Kieran Wynne[b,r], Jonah F. Byrne[a], Connie Markulev[c,d], Miriam R. Schäfer[c,d], Maximus Berger[c,e], Nilufar Mossaheb[f], Monika Schlögelhofer[g], Stefan Smesny[h], Ian B. Hickie[i], Gregor E. Berger[j], Eric Y.H. Chen[k], Lieuwe de Haan[l], Dorien H. Nieman[l], Merete Nordentoft[m], Anita Riecher-Rössler[n], Swapna Verma[o], Rebekah Street[c,d], Andrew Thompson[c,d], Alison Ruth Yung [c,p,q], Barnaby Nelson[c,d] , Patrick D. McGorry [c,d], Melanie Föcking [a], G. Paul Amminger[# c,d] and David Cotter[# a].

[a] Department of Psychiatry, RCSI University of Medicine and Health Sciences, Dublin Ireland.

[b]School of Biomolecular and Biomedical Science, Conway Institute, University College Dublin, Dublin, Ireland

[c] Centre for Youth Mental Health, The University of Melbourne, Melbourne, Victoria, Australia.

[d] Orygen, 35 Poplar Rd, Parkville, 3052, Australia.

[e] Department of Psychiatry and Psychotherapy, University of Bern, Switzerland.

[f] Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria.

[g] BioPsyC-Biopsychosocial corporation – Non-Profit association for research funding, Vienna, Austria.

[h] Department of Psychiatry and Psychotherapy, Jena University Hospital, Jena, Germany.

[i] Brain and Mind Centre, The University of Sydney, Sydney, New South Wales, Australia.

[j] Child and Adolescent Psychiatric Service of the Canton of Zurich, Zürich, Switzerland.

[k] Department of Psychiatry, University of Hong Kong, Pok Fu Lam, Hong kong

[l] Department of Psychiatry, Academic Medical Center, Amsterdam, The Netherlands.

[m] Mental Health Center Copenhagen, Department of clinical Medicine, Copenhagen University Hospital, Denmark.

[n] Medical Faculty, University of Basel, Switzerland.

[o] Institute of Mental Health, Singapore, Singapore.

[p] Institute for Mental and Physical Health and Clinical Translation (IMPACT), Deakin University, Geelong, Australia

[q] School of Health Sciences, University of Manchester, UK

[r] Systems Biology Ireland, University College Dublin, Dublin, Ireland

* Joint first authors & # Joint senior authors.

Corresponding Authors:

1. Subash Raj Susai, Department of Psychiatry, RCSI University of Medicine and Health Sciences, Dublin, Ireland. Electronic address: subashrajsusai@rcsi.ie
2. David Cotter, Department of Psychiatry, RCSI University of Medicine and Health Sciences, Dublin, Ireland. Electronic address: drcotter@rcsi.ie

Word count:

- Abstract: 356
- Main article: 4228

Abstract:

Background: Functional outcomes are important measures in the overall clinical course of psychosis and individuals at clinical high-risk (CHR), however, prediction of functional outcome remains difficult based on clinical information alone. In the first part of this study, we evaluated whether a combination of biological and clinical variables could predict future functional outcome in CHR individuals. The complement and coagulation pathways have previously been identified as being of relevance to the pathophysiology of psychosis and have been found to contribute to the prediction of clinical outcome in CHR participants. Hence, in the second part we extended the analysis to evaluate specifically the relationship of complement and coagulation proteins with psychotic symptoms and functional outcome in CHR.

Materials and methods: We carried out plasma proteomics and measured plasma cytokine levels, and erythrocyte membrane fatty acid levels in a sub-sample (n=158) from the NEURAPRO clinical trial at baseline and 6 months follow up. Functional outcome was measured using Social and Occupational Functional assessment Score (SOFAS) scale. Firstly, we used support vector machine learning techniques to develop predictive models for functional outcome at 12 months. Secondly, we developed linear regression models to understand the association between 6-month follow-up levels of complement and coagulation proteins with 6-month follow-up measures of positive symptoms summary (PSS) scores and functional outcome.

Results and conclusion: A prediction model based on clinical and biological data including the plasma proteome, erythrocyte fatty acids and cytokines, poorly predicted functional outcome at 12 months follow-up in CHR participants. In linear regression models, four complement and coagulation proteins (coagulation protein X, Complement C1r subcomponent like protein, Complement C4A & Complement C5) indicated a significant association with functional outcome; and two proteins (coagulation factor IX and complement C5) positively associated with the PSS score. Our study does not provide support for the utility of cytokines, proteomic or fatty acid data for prediction of functional outcomes in individuals at high-risk for psychosis. However, the association of complement protein levels with clinical outcome suggests a role for the complement system and the activity of its related pathway in the functional impairment and positive symptom severity of CHR patients.

Key words: clinical high risk, functional outcome, prediction models, schizophrenia, psychosis.

1. Introduction:

Psychosis research is increasingly focusing on those in the clinical high risk (CHR) population who experience early signs of emerging psychosis (1). The CHR criteria comprise of the attenuated psychotic symptom (APS) criterion, the brief limited intermittent psychotic symptom (BLIPS) criterion, and the genetic risk and functional decline criterion (2). The functional impairment of CHR participants substantially impacts personal, familial and social well-being (3-5) and responds poorly to currently available treatments (6-8). The association of early functional deterioration with the development of psychotic symptoms indicates that functional measures could be used to improve early intervention strategies in psychosis (9-17).

Previous studies involving CHR participants have investigated baseline predictors of transition to psychosis and found that factors such as social dysfunction, neurocognitive measures, duration of untreated psychosis and severity of attenuated psychotic symptoms predict later transition to psychosis (18-37). The biological aspects of psychosis have been increasingly studied in relation to the clinical symptoms in the CHR state. Thus far, biological parameters such as neuroimaging data and electrophysiological indicators have provided some valuable prediction of functional outcome and transition to psychosis in CHR populations (26, 38-42). Studies of immune markers (43, 44) and membrane phospholipids (45-47) have also been undertaken in CHR participants. Although some alterations have been found to be associated with the development of psychosis (43, 44, 48-57), the clinical implication of these findings in terms of prediction of functional outcome in CHR individuals has not been specifically studied (52, 58).

Blood based biological marker studies have focused on predicting the development of psychosis in CHR participants (59, 60). Mongan et al. used mass spectrometry based proteomic data to predict clinical outcomes in a longitudinal CHR study (61). Combined clinical and proteomic data predicted the development of psychosis better than clinical data alone (61). In addition to the prediction of psychosis, the proteomic variables also predicted functional outcome with an AUC of 0.76 at two years follow-up in 133 CHR participants (61). In this prediction model the most abundant proteins that significantly predicted functional outcome were complement and coagulation proteins. (62). Similarly, other studies have found clinical and demographic features such as duration of treatment or untreated psychosis and poor cognition to be the predictors of later functional decline in the CHR (10, 13, 14, 17, 39, 63-68).

In the current study we attempted to investigate the combined predictive ability of blood based biological markers including inflammatory cytokines, erythrocyte membrane fatty acids and the plasma proteome on functional outcome. Using machine learning we sought to develop two prediction models, one using baseline clinical data alone and another using both baseline clinical and biological data. We developed these models in a subsample of the NEURAPRO clinical trial, which tested the potential preventive role of omega-3 fatty acids in CHR participants (69). Our team has previously reported dysregulation of complement and coagulation pathway proteins in relation to development of psychotic symptoms and functional decline in high-risk population (49, 61). These results supported the findings of Sekar et al., suggesting that the complement related activity might be involved in the development of clinical symptoms in the early stage of schizophrenia (70, 71). Hence, in the current study we extended our analysis to explore the individual relationship of complement associated proteins with positive symptoms and functional status. Based on our previous findings, we hypothesized that baseline biological data along with clinical parameters would predict functional improvement in CHR participants better than the clinical model alone. In addition, we also hypothesised that higher complement and coagulation proteins would associate with poor clinical outcomes.

2. Materials Methods

2.1. The NEURAPRO clinical trial

The NEURAPRO clinical trial was registered with the Australian New Zealand Clinical Trial Registry as ACTRN 12608000475347. The trial aimed to investigate the role of omega-3 fatty acids (FAs) on prevention of psychosis in CHR participants (72). The study was conducted between March 2010 and the end of September 2014, in accordance with the Declaration of Helsinki and consistent with the International Council for Harmonization of Good Clinical Practice with appropriate ethical approval obtained from each site before the trial commenced. Ethical approval for the plasma biomarker analysis presented in this study was obtained from the research ethics committee of the Royal College of Surgeons in Ireland [REC-No. 1699].

The inclusion criteria include participants aged between 13 and 40 years who fulfilled one of the criteria for at-risk state defined by the Comprehensive Assessment of At-Risk Metal State (CAARMS)(2). The exclusion criteria were: history of psychotic episodes of seven days or longer;

any current symptoms of organic brain disease or developmental disorder; abnormal coagulation profile; thyroid abnormalities; physical illness with psychotropic effect, if not stabilized; current treatment with any mood stabilizers or recreational use of ketamine; past neuroleptic exposure equivalent to a total lifetime haloperidol dose of >50 mg; a diagnosis of a serious developmental disorder; premorbid IQ less than 70; current acute suicidality /self-harm or aggression/dangerous behavior; pregnancy; or intake of more than 4 weeks of supplementation with omega-3 FAs (73).

2.2. Participants:

A total of 170 CHR participants who provided baseline and 12-month follow-up plasma samples and who had clinical outcome data available at 12 months were considered for this plasma biomarker analysis study.

2.3. Clinical measures:

Baseline psychopathological scores of CHR participants were measured using the Brief Psychiatric Rating scale (BPRS) (74), Scale for the Assessment of Negative Symptoms (SANS) (75), Youth Maniac Rating Scale (YMRS) (76), Montgomery-Åsberg Depression Rating Scale (MADRS) (77), Social and Occupational Functional assessment Score (SOFAS) (78), Global functioning Social (GF:S) (79) and Global functioning Role (GF:R) (80) were used for machine learning.

2.4. Gas-chromatography based Erythrocyte membrane fatty acid measures:

Fasting plasma samples were collected at baseline and 6 months follow-up. The erythrocytes were separated from the plasma using an automated high-throughput method described in (81). The molecular percentage of Total omega-3 FAs, total omega-6 FAs, docosapentaenoic acid (DPA), docosahexaenoic acid (DHA), linoleic acid (LA), arachidonic acid (ARA), omega-3 index (EPA+DHA), omega-6:omega-3 ratio, Alpha-linolenic acid (ALA), eicosapentaenoic acid (EPA) were measured using gas chromatography (73, 82). The Phosphatidyl-ethanolamine (PE) fraction was used to determine the omega-3 FA content, because of their high abundancy in the lipid raft (56, 83).

2.5. Mass spectrometry based proteomic measures:

Plasma samples of baseline and follow-up time points were processed according to the manufacturer's instructions (PreOmics iST kit, no.iST 96x). Briefly, 4 µl of individual samples were solubilized in 50 µL of "Lyse" buffer (containing Tris-HCl, sodium deoxycholate (SDC), 0.1% sodium dodecyl sulfate (SDS), tris (2-carboxyethyl) phosphine (TCEP), and 2-chloroacetamide and heated to 95 °C for 10 min. 50 µL of the resulting denatured, reduced, and alkylated solution was transferred to the reaction tube. Enzyme (LysC and trypsin) was added, and samples were hydrolysed at 37°C for 1.5 hours. The resulting peptide mixture was washed and eluted as per the manufacturer's instructions. The eluted peptides were vacuum-dried and dissolved in 100 µl of LC Load buffer. The reconstituted digested peptide mixture [200 ng/ µl] was then eluted using Evotips and injected using Evosep One (Evosep, Odense, Denmark (84). The digested samples were run on a Bruker timeTof Pro mass spectrometer connected to a Evosep One liquid chromatography system. The mass spectrometry was operated in positive ion mode with a capillary voltage of 1500 V, dry gas flow of 3 l/min and a dry temperature of 180^{o}C. Trapped ions were selected for ms/ms using parallel accumulation serial fragmentation (PASEF). A scan range of (100-1700 m/z) was performed at a rate of 10 PASEF MS/MS frames to 1 MS scan with a cycle time of 1.15s (85, 86). The MS raw files were then processed with MaxQuant (87) version 1.6.17.0 as described in (86) and the peptide data were further annotated and interpreted using the Perseus platform (V 1.6.7, www.maxquant.net/perseus/) (88). FDR was set at 0.01 to global protein identification level. Proteins that were identified in less than 70% of the total samples were not taken forward for analysis. Log_2 transformed values of LFQ intensities were used for statistical analysis. Missing values of mass spectrometry based proteomic data (corresponding to values below the level of detection) were imputed with minimum values.

2.6. Multi-plex assay-based estimation of plasma immune markers:

Plasma levels of Granulocyte-macrophage colony-stimulating factor (GM-CSF), Interleukin (IL) -1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p40, IL-12p70, IL-15, Tumor necrosis factor-α (TNF-α), Interferon gamma (INF-γ), Inter-cellular Adhesion Molecule (ICAM)-1 and Vascular cell adhesion molecule (VCAM)-1 were measured at baseline using the Pro-inflammatory Panel 1, Cytokine Panel 1 and Vascular Injury Panel 2 v-PLEX® multiplex immunoassay kits (Mesoscale Discovery Systems) according to the manufacturer's instructions. A Sector Imager 2400 plate reader was used to quantify concentrations of each marker (Meso Scale Diagnostics). The

concentrations of plasma markers that were expressed in at least 70% of all samples with a coefficient of variation of a maximum of 20% were taken forward. These included: IL12p40, IL15, IFN-gamma, IL6, IL8, IL10, TNF-alpha, CRP, sVCAM-1 and sICAM-1. Log2-transformed concentrations for these cytokines were used for the statistical analysis.

2.7. Clinical outcome measure:

Functional outcome was assessed using the SOFAS scale at baseline, 6-month and 12-month follow-up. For the machine learning model, we investigated prediction of SOFAS score as a continuous variable with a minimum possible score of 0 points and a maximum possible score of 100 points. For the linear regression model a positive symptom summary (PSS) score was used along with the SOFAS score. The PSS score was derived from CAARMS symptom severity score by summing up the product of global rating scale score (0-6) and frequency (0-6) of the four subscales (89).

2.8. Statistical analysis:

2.8.1. Machine learning models

Samples with clinical and biological data available at baseline and SOFAS outcome data available at 12-month follow-up were used for the machine learning analysis. Missing values of mass spectrometry based proteomic data were imputed with minimum values. Missing values of the remaining data were imputed using k nearest neighbours' imputation (k=7). All continuous measures were standardised to z scores and winsorised within +/- 4z.

Models were developed using a support vector machine (SVM) approach based on the LIBSVM algorithm. SVM is a computationally efficient form of supervised machine learning that has been used previously in multiple different contexts within psychiatry (90-92). SVM methods can integrate hyperparameter optimization to reduce propensity for over-fitting. Neurominer version 1.0 (https://github.com/neurominer-git) for MatLab 2018a (Math Works Inc) was used to develop SVM models with nested cross-validation. For a detailed description of nested cross-validation, see (93). The data were first divided into 5 random folds in the 'outer loop'. For each cycle of cross-validation, data from each fold were held out and the rest of the data moved into the 'inner loop'. Within the inner loop, we used 5 non-overlapping folds with iterative training-test cycles. Models were trained and tested with a range of values for the regularisation

parameter and the best-performing models tested against the held-out data in the outer loop to derive the optimal model.

Model 1: clinical predictors

Firstly, we generated a model for prediction of 12-month SOFAS score based on data for the 11 clinical predictors. We used the LIBSVM algorithm with a linear kernel where mean squared error was used as the performance criterion. The regularisation parameter was applied across a range of 7 values (0.015625, 0.03125, 0.0625, 0.125, 0.25, 0.5,1 1) and the epsilon parameter across a range of 6 values (0.05, 0.075, 0.1, 0.125, 0.15, 0.2).

Model 2: biomarker predictors

Secondly, we generated a model for prediction of 12-month SOFAS score based on data for the 177 potential biomarkers. This model was also developed using the LIBSVM algorithm with linear kernel, mean squared error as the performance criterion, and regularisation parameter optimisation as for Model 1.

Model 3: clinical and biomarker predictors

Thirdly, we generated a model for prediction of 12-month SOFAS score based on data for the 11 clinical predictors plus data for the 177 potential biomarkers. This model was also developed using the LIBSVM algorithm with linear kernel, mean squared error as the performance criterion, and regularisation parameter optimisation as for Model 1.

For all three models, random-label permutation analysis with 1000 permutations was used to derive *p*-values for model significance and mean feature weights. Presented performance metrics include the mean squared error, Pearson's r, coefficient of determination, mean absolute error and normalised root mean square deviation. We also present classification-based performance metrics (such as sensitivity and specificity) for each model based on a SOFAS threshold of 70 points (70 points and below reflects some, moderate or major functional impairment, whereas 71 points and above reflects no more than slight functional impairment).

2.8.2. Linear regression models:

The participants who had proteomic data available at baseline and 6 month follow-up along with SOFAS and CAARMS score were considered for the secondary analysis. Linear regression models

were developed to assess the relationship between protein levels at 6-month follow up with clinical scores (positive symptom summery score and functional score) at 6-month follow-up and the model was adjusted for age, sex, BMI along with corresponding baseline complement protein levels and baseline clinical scores. The level of significance was set to 0.05.

3. Results:

3.1. Sample characteristics

Out of 170 participants, 158 participants have baseline clinical and biological measures and functional outcome at 12 months follow-up. The mean age of the study sample participants was 18 years (SD 4) with an average BMI of 24 kg/m^2 (SD 6). 58% of the study participants were females. The baseline demographic, clinical and biological characteristics of the study participants are given in Table 1.

3.2. Predictive models

The clinical predictor pool comprised 11 features in total (4 demographic variables including sex, age, smoking status, BMI; and 7 symptom scale scores). The biomarker predictor pool comprised 177 features in total (10 cytokines; 157 proteomic markers; and 10 fatty acid markers). The full list of features is provided in Supplementary Table 1.

Model 1: Clinical predictors

Model 1 demonstrated poor predictive performance with mean squared error of 239.00 (p<0.001 on permutation analysis). Pearson's r was 0. 30 (95% confidence interval 0.13 – 0.45), coefficient of determination 8.9%, mean absolute error 13.0 and normalised root mean square deviation 22.4. Further performance metrics, including classification performance based on a threshold of 70 points, are presented in Table 2. Observed versus predicted SOFAS values are plotted in Figure 1. Features are ranked by mean feature weight in Table 3.

Model 2: Biomarker predictors

Model 2 demonstrated poor predictive performance with mean squared error of 256.2 (p<0.001 on permutation analysis). Pearson's r was 0.25 (95% confidence interval 0.08 – 0.40), coefficient of determination 6.2%, mean absolute error 13.4 and normalised root mean square deviation 23.2. Further performance metrics, including classification performance based on a threshold of 70

points, are presented in Table 2. Observed versus predicted SOFAS values are plotted in Figure 2. The highest-weighted 10% of predictors ranked by mean feature weight are provided in Table 3.

Model 3: Clinical and biomarker predictors

Model 3 demonstrated poor predictive performance with mean squared error of 250.0 (p=0.023 on permutation analysis). Pearson's r was 0.22 (95% confidence interval 0.05 – 0.38), coefficient of determination 5.0%, mean absolute error 13.4 and normalised root mean square deviation 22.9. Further performance metrics, including classification performance based on a threshold of 70 points, are presented in Table 2. Observed versus predicted SOFAS values are plotted in Figure 3. The highest-weighted 10% of predictors ranked by mean feature weight are provided in Table 3.

3.3. Associations between complement proteins and functional outcome:

A total of 114 participants had proteomic and functional data at baseline and 6 months follow-up. In a linear regression analysis using 6-month SOFAS score, Coagulation protein Factor X at 6-month follow-up showed a positive association with functional outcome [β coef (95% CI)= 2.6(0.1to5.2), p value= 0.04], whereas complement proteins Complement C1r subcomponent like protein, C4A and C5 expressed an inverse association with functional outcome [β coef (95% CI)= -2.7 (-5.3to-0.2), -3.1(-5.8to-0.5) & -2.9(-5.6to-0.3), p value= 0.04, 0.02 & 0.03, respectively] (Table 4). The complement C5 and coagulation factor IX associated positively with the positive symptom score after adjusting for age, sex, BMI and baseline clinical score [β coef (95% CI) =2.6(0.2to5.0) and 2.6(0.1to5.0); p value= 0.034 and 0.043, respectively] (Figure 4).

4. Discussion:

In this study, we attempted to develop a machine learning model to predict 12-month functional outcome in a CHR population using baseline clinical data (symptom and sociodemographic measures) and baseline levels of plasma biomarkers (fatty acids, immune markers and proteomic measures). We hypothesized that baseline biological and clinical measures would collectively show better prediction of functional outcome than clinical measures alone. The clinical model (Model 1) had poor predictive performance in relation to functional outcome at 12 months follow-up with mean squared error of 239.0 (Area Under the receiver-operating characteristic Curve [AUC] 0.63). A model based on biomarker data from several modalities (Model 2) showed poor predictive performance with mean squared error of 256.2 (AUC 0.62). A model based on

combined clinical and biomarker data (Model 3) also showed poor predictive performance (mean squared error 250.0, AUC 0.58). Hence our results did not support the hypothesis that biomarkers would improve prediction of functional outcome at 12-month follow-up. However, in regression analysis, several complement and coagulation proteins at 6-month follow-up associated with psychotic symptoms and functional outcome at follow-up. In particular, an increased level of complement C5 and coagulation protein factor IX at 6-month follow-up associated with high positive symptoms at 6-month follow-up after adjusting for their corresponding baseline clinical and proteomic measures. Similarly, an increase in complement proteins C1r subcomponent like protein, C4A and C5 associated with decrease in functional outcome, while coagulation protein Factor X associated inversely with functional outcome.

Previous studies of CHR individuals have investigated the role of plasma-based biological markers in the prediction of transition to psychosis in the CHR population. In the North American Prodrome Longitudinal Study, 15 selected plasma analytes not only distinguished the CHR participants from healthy controls, but also successfully differentiated CHR participants who developed psychosis from those who did not (59). Similarly, a combination of 26 plasma biomarkers which were found to be differentially expressed in schizophrenia patients compared to controls, predicted the development of psychosis within two years follow-up. In this model, addition of clinical parameters increased the performance of this prediction model (60). However, the ability of these models to predict functional outcome in CHR was not evaluated.

A recent study from our team investigated the predictive ability of plasma proteome on the development of psychotic disorder among the CHR (61). The proteomic data successfully predicted the development of psychosis with and without clinical parameters. Moreover the baseline proteomic data along with clincial variables also predicted functional outcome at 2 years follow-up in CHR participants, albeit more weakly than models predicting transition outcome (61). In contrast, the current study did not predict the functional outcome at short term follow-up (12 months) using biological and clinical markers together. The current study investigated a wider array of biological predictors including plasma inflammatory markers measured using multiplex assays and erythrocyte membrane fatty acid assessed by gas chromatography levels but found no evidence of significant predictive performance. This finding could be due to the presence of

masking effects of multiple biological variables such as plasma proteins that are not directly related to the functional outcome.

The membrane phospholipid hypothesis has specified the potential involvement of fatty acid imbalance in the development of psychosis (43, 44, 50, 51, 53, 94-97). However, very few clinical studies have investigated the biological relationship of omega-3 FAs with functional outcomes such as social, role functioning and occupational functioning in CHR participants (58, 98-101). These studies suggest that there is a weak cross-sectional association between omega-3 FAs and functional outcome, and longitudinal analyses in the same samples have not shown evidence for strong relationships (58). Considering the limited knowledge of omega-3 FAs and plasma immune markers with functional outcome, the negative results of our study may suggest that more investigations are required to understand the therapeutic and prognostic ability of these fatty acid biomarkers in terms of functional status to consider them in the prediction models. For instance, a recent study has indicated that plasma levels of docosahexaenoic acid were associated cross-sectionally and longitudinally with psychotic disorder in early adulthood in the general population (54) but whether this extends to general functioning in the wider population is not yet known.

Apart from biological markers, previous studies have identified demographic, clinical and neuro-anatomical markers as reliable predictors of functional outcome in CHR psychosis. Another combined machine learning approach in CHR participants by Koutsouleris et al., revealed that social functioning impairment can be predicted using both clinical and neuro-anatomical measures (38). In this latter study, the authors also showed that the combination of neuroimaging models with clinical prediction models increased the performance by 1.9 fold compared to models based on the clinical measures alone (38). Moreover, among the clinical measures, neurocognition and functioning at baseline provided a strong link with functional outcome and provided a basis for domain-wise prediction in functional outcome (10, 13, 14, 17, 63-68). For instance, baseline processing speed and social functioning predicted social functioning at follow-up whereas baseline verbal memory and role functioning predicted role functioning at follow-up (102). In contrast, in our current study, blood-based biomarkers were not able to match the predictive ability of neuroanatomical parameters and domain specific cognitive measures (38).

The results of linear regression analyses revealed an important relationship between complement and coagulation proteins and functional outcome. Increased complement proteins at follow-up

were associated cross-sectionally with increased positive symptoms and decreased functional symptoms at follow-up in CHR participants. Several complement pathway proteins are involved in synaptic pruning at early developmental stages and previous genetic and preclinical investigations have revealed the importance of complement related activity in relation to schizophrenia (51, 52, 71, 103, 104). In a population-based study, Föcking et al, (2021) reported an increase in complement and related proteins significantly associated with future development of psychotic symptoms. In this study the authors reported an upregulation of six proteins including C1r subcomponent like protein, and C5 associated with future psychotic symptoms (50). Furthermore, Mongan et al. also observed that the complement proteins were among the top weighted predictive features of functional outcome and transition to psychosis in machine learning models (105). In line with our previous findings in the current study we evaluated the association of individual complement proteins with functional status and found that higher C1r subcomponent like protein, C4A and C5 complement proteins at follow-up were cross sectionally associated with poor clinical outcomes such as high positive symptoms and poor functional outcome. This supports an inflammatory association with clinical outcome among participants with early psychosis and adds further support to the importance of complement associated changes in the pathophysiology of schizophrenia. Our proteomic findings are in line with genetic results of GWAS study which found that an association of increased risk of schizophrenia with an increased expression C4A gene (70, 71). These findings together suggest that an increase in baseline complement pathway protein levels predispose towards pathological changes on functioning at follow-up during the early stage of psychosis. The results also indicate a group level association of a few complement proteins with functional status even though individualized prediction was not achieved in machine learning based approach. The findings open up new avenues for understanding the molecular mechanisms through which complement and coagulation proteins might influence functional outcome among subjects in the CHR. Furthermore, it is possible that a subgroup of individuals vulnerable to psychosis with dysregulated complement activity may benefit from modulation of the complement pathway (for example through pharmacological interventions targeting complement activity) but this hypothesis would require extensive preclinical testing before human trials.

Strengths and limitations

Our study has several strengths. First our study utilised unique and in-depth biological data which included proteomic, inflammatory cytokine, membrane FA measures and in-depth clinical measures from a valuable CHR population. Secondly, the analysis focused on functional outcome among the CHR. This is an area that has been under investigated in the past. Thirdly our study is unique in quantifying erythrocyte omega-3 markers and plasma complement protein levels at both baseline and follow-up time points in CHR trajectory. Furthermore, we were able to adjust the linear models for potential confounders. This statistical approach allowed us to take inter-individual heterogeneity into account. The limitations of our study include: i) use of relative quantification methods such as discovery proteomics and semi-quantitative biological assays such as multi-plex ELISA assays; ii) absence of some potentially relevant measures such as neuro-imaging data that has successfully predicted the functional outcome in CHR participants in the past; iii) a relatively small number of samples (n=158) compared to those who contributed to the NEURAPRO clinical trial as a whole (n=304), due to non-participation in some aspects of the study; iv) the large number of predictors relative to the sample size may give rise to concern regarding overfitting and v) in linear regression analyses, considering the use of discovery based proteomic data and clinically homogenous samples, we did not adjust the results for multiple correction.

Our study suggests that in CHR participants, addition of baseline plasma biomarker data involving proteomic markers, erythrocyte membrane FA levels and plasma cytokine levels did not improve prediction of 12-month functional outcome beyond baseline clinical data alone. However statistical analysis found an association between increased complement pathway proteins and worsening of clinical outcome such as increased positive symptoms and poor functional outcome in CHR participants. These findings point to a need of further studies exploring and validating the association of complement and related pathway activity with clinical outcome in psychosis. Furthermore, the machine learning models point to a need for a deeper understanding contribution of other types of biological and clinical markers to improve prognostication in CHR individuals.

5. Acknowledgement

We would like to thank the participants and their families. We express our thanks to Mr. John Butler from Mesoscale diagnostics for his guidance and support to run the multiplex assays. We

thank Eugène Dillon from University College Dublin for helping us in obtaining proteomic data. We express our sincere thanks to Hok Pan Yuen from ORYGEN for his contribution in clinical data acquisition for the NEURAPRO study.

6. Funding

This work was supported by an Irish Health Research Board research grant to DC (HRB ILP POR 2019-005). This publication has emanated from research supported by Health Research Board (HRB) [to 411 DC, MF] under grant number HRB/HRA/PHR/2015-1293. The research was funded in 412 part by a research grant from Science Foundation Ireland (SFI) under Grant Number 413 16/RC/3948 and co-funded under the European Regional Development Fund and by 414 FutureNeuro industry partners. The immune marker study was supported by Orygen, Australia. 415 DM is a Fellow on the Irish Clinical Academic Training (ICAT) Programme which is supported 416 by the Wellcome Trust and the Health Research Board (Grant Number 203930/B/16/Z), the 417 Health Service Executive National Doctors Training and Planning and the Health and Social 418 Care, Research and Development Division, Northern Ireland. JFB was supported by a Wellcome Flagship Innovations Award (220438Z/20/Z). BN was supported by an NHMRC Senior Research Fellowship (1137687) and a University of Melbourne Dame Kate Campbell Fellowship. MB was supported by a MRFF RART Fellowship. Research reported in this publication was supported by The Comprehensive Molecular Analytical Platform (CMAP) under The SFI Research Infrastructure Programme, reference 18/RI/5702.

8. Bibliography

1. McGorry PD. "A stitch in time" ... the scope for preventive strategies in early psychosis. Eur Arch Psychiatry Clin Neurosci. 1998;248(1):22-31.
2. 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. The Australian and New Zealand journal of psychiatry. 2005;39(11-12):964-71.
3. Wu EQ, Birnbaum HG, Shi L, Ball DE, Kessler RC, Moulis M, et al. The economic burden of schizophrenia in the United States in 2002. J Clin Psychiatry. 2005;66(9):1122-9.
4. Hutchinson G, Bhugra D, Mallett R, Burnett R, Corridan B, Leff J. Fertility and marital rates in first-onset schizophrenia. Soc Psychiatry Psychiatr Epidemiol. 1999;34(12):617-21.
5. Horan WP, Subotnik KL, Snyder KS, Nuechterlein KH. Do recent-onset schizophrenia patients experience a "social network crisis"? Psychiatry. 2006;69(2):115-29.

6. Ventura J, Subotnik KL, Guzik LH, Hellemann GS, Gitlin MJ, Wood RC, et al. Remission and recovery during the first outpatient year of the early course of schizophrenia. Schizophr Res. 2011;132(1):18-23.
7. Robinson DG, Woerner MG, McMeniman M, Mendelowitz A, Bilder RM. Symptomatic and functional recovery from a first episode of schizophrenia or schizoaffective disorder. Am J Psychiatry. 2004;161(3):473-9.
8. Tohen M, Hennen J, Zarate CM, Jr., Baldessarini RJ, Strakowski SM, Stoll AL, et al. Two-year syndromal and functional recovery in 219 cases of first-episode major affective disorder with psychotic features. Am J Psychiatry. 2000;157(2):220-8.
9. Bora E, Pantelis C. Theory of mind impairments in first-episode psychosis, individuals at ultra-high risk for psychosis and in first-degree relatives of schizophrenia: systematic review and meta-analysis. Schizophr Res. 2013;144(1-3):31-6.
10. Carrión RE, McLaughlin D, Goldberg TE, Auther AM, Olsen RH, Olvet DM, et al. Prediction of functional outcome in individuals at clinical high risk for psychosis. JAMA Psychiatry. 2013;70(11):1133-42.
11. Fusar-Poli P, Deste G, Smieskova R, Barlati S, Yung AR, Howes O, et al. Cognitive functioning in prodromal psychosis: a meta-analysis. Arch Gen Psychiatry. 2012;69(6):562-71.
12. Giuliano AJ, Li H, Mesholam-Gately RI, Sorenson SM, Woodberry KA, Seidman LJ. Neurocognition in the psychosis risk syndrome: a quantitative and qualitative review. Curr Pharm Des. 2012;18(4):399-415.
13. Lin A, Wood SJ, Nelson B, Brewer WJ, Spiliotacopoulos D, Bruxner A, et al. Neurocognitive predictors of functional outcome two to 13 years after identification as ultra-high risk for psychosis. Schizophr Res. 2011;132(1):1-7.
14. Meyer EC, Carrión RE, Cornblatt BA, Addington J, Cadenhead KS, Cannon TD, et al. The relationship of neurocognition and negative symptoms to social and role functioning over time in individuals at clinical high risk in the first phase of the North American Prodrome Longitudinal Study. Schizophr Bull. 2014;40(6):1452-61.
15. Salokangas RK, Nieman DH, Heinimaa M, Svirskis T, Luutonen S, From T, et al. Psychosocial outcome in patients at clinical high risk of psychosis: a prospective follow-up. Soc Psychiatry Psychiatr Epidemiol. 2013;48(2):303-11.
16. Salokangas RK, Heinimaa M, From T, Löyttyniemi E, Ilonen T, Luutonen S, et al. Short-term functional outcome and premorbid adjustment in clinical high-risk patients. Results of the EPOS project. Eur Psychiatry. 2014;29(6):371-80.
17. Burton CZ, Tso IF, Carrión RE, Niendam T, Adelsheim S, Auther AM, et al. Baseline psychopathology and relationship to longitudinal functional outcome in attenuated and early first episode psychosis. Schizophr Res. 2019;212:157-62.
18. Cannon TD, Cadenhead K, Cornblatt B, Woods SW, Addington J, Walker E, et al. Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. Arch Gen Psychiatry. 2008;65(1):28-37.
19. Buchy L, Perkins D, Woods SW, Liu L, Addington J. Impact of substance use on conversion to psychosis in youth at clinical high risk of psychosis. Schizophrenia Research. 2014;156(2-3):277-80.
20. Fusar-Poli P, Byrne M, Valmaggia L, Day F, Tabraham P, Johns L, et al. Social dysfunction predicts two years clinical outcome in people at ultra high risk for psychosis. Journal of psychiatric research. 2010;44(5):294-301.
21. Dragt S, Nieman DH, Veltman D, Becker HE, van de Fliert R, de Haan L, et al. Environmental factors and social adjustment as predictors of a first psychosis in subjects at ultra high risk. Schizophr Res. 2011;125(1):69-76.

22. Tarbox SI, Addington J, Cadenhead KS, Cannon TD, Cornblatt BA, Perkins DO, et al. Premorbid functional development and conversion to psychosis in clinical high-risk youths. Dev Psychopathol. 2013;25(4 Pt 1):1171-86.
23. Ziermans T, Wit Sd, Schothorst P, Sprong M, Engeland Hv, Kahn R, et al. Neurocognitive and clinical predictors of long-term outcome in adolescents at ultra-high risk for psychosis: a 6-year follow-up. PloS one. 2014;9(4):e93994.
24. Riecher-Rössler A, Pflueger MO, Aston J, Borgwardt SJ, Brewer WJ, Gschwandtner U, et al. Efficacy of using cognitive status in predicting psychosis: a 7-year follow-up. Biological psychiatry. 2009;66(11):1023-30.
25. Pukrop R, Ruhrmann S, Schultze-Lutter F, Bechdolf A, Brockhaus-Dumke A, Klosterkötter J. Neurocognitive indicators for a conversion to psychosis: comparison of patients in a potentially initial prodromal state who did or did not convert to a psychosis. Schizophr Res. 2007;92(1-3):116-25.
26. Koutsouleris N, Davatzikos C, Bottlender R, Patschurek-Kliche K, Scheuerecker J, Decker P, et al. Early recognition and disease prediction in the at-risk mental states for psychosis using neurocognitive pattern classification. Schizophrenia bulletin. 2012;38(6):1200-15.
27. Hoffman RE, Woods SW, Hawkins KA, Pittman B, Tohen M, Preda A, et al. Extracting spurious messages from noise and risk of schizophrenia-spectrum disorders in a prodromal population. The British Journal of Psychiatry. 2007;191(4):355-6.
28. Nieman DH, Ruhrmann S, Dragt S, Soen F, van Tricht MJ, Koelman JH, et al. Psychosis prediction: stratification of risk estimation with information-processing and premorbid functioning variables. Schizophr Bull. 2014;40(6):1482-90.
29. Lencz T, Smith CW, McLaughlin D, Auther A, Nakayama E, Hovey L, et al. Generalized and specific neurocognitive deficits in prodromal schizophrenia. Biol Psychiatry. 2006;59(9):863-71.
30. Cornblatt BA, Carrion RE, Auther A, McLaughlin D, Olsen RH, John M, et al. Psychosis Prevention: A Modified Clinical High Risk Perspective From the Recognition and Prevention (RAP) Program. Am J Psychiatry. 2015;172(10):986-94.
31. Corcoran CM, Keilp JG, Kayser J, Klim C, Butler PD, Bruder GE, et al. Emotion recognition deficits as predictors of transition in individuals at clinical high risk for schizophrenia: a neurodevelopmental perspective. Psychological medicine. 2015;45(14):2959-73.
32. Thompson A, Nelson B, Yung A. Predictive validity of clinical variables in the “at risk” for psychosis population: international comparison with results from the North American Prodrome Longitudinal Study. Schizophrenia research. 2011;126(1-3):51-7.
33. Zimmermann R, Gschwandtner U, Wilhelm FH, Pflueger MO, Riecher-Rössler A, Fuhr P. EEG spectral power and negative symptoms in at-risk individuals predict transition to psychosis. Schizophrenia research. 2010;123(2-3):208-16.
34. Albert N, Melau M, Jensen H, Hastrup LH, Hjorthøj C, Nordentoft M. The effect of duration of untreated psychosis and treatment delay on the outcomes of prolonged early intervention in psychotic disorders. NPJ schizophrenia. 2017;3(1):34.
35. Miglietta E, Lasalvia A, Bonetto C, Comacchio C, Cristofalo D, Tosato S, et al. Pathways to care, DUP, and types of interventions over 5 years following psychosis onset: findings from a naturalistic study conducted in routine generalist mental health services. Soc Psychiatry Psychiatr Epidemiol. 2020;55(2):175-86.
36. Tsuang MT, Van Os J, Tandon R, Barch DM, Bustillo J, Gaebel W, et al. Attenuated psychosis syndrome in DSM-5. Schizophrenia research. 2013;150(1):31-5.
37. 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.
38. 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.
39. Koutsouleris N, Riecher-Rössler A, Meisenzahl EM, Smieskova R, Studerus E, Kambeitz-Ilankovic L, et al. Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. Schizophr Bull. 2015;41(2):471-82.
40. 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(1-3):42-7.
41. Ramyead A, Studerus E, Kometer M, Uttinger M, Gschwandtner U, Fuhr P, et al. Prediction of psychosis using neural oscillations and machine learning in neuroleptic-naive at-risk patients. World J Biol Psychiatry. 2016;17(4):285-95.
42. Van Tricht MJ, Nieman DH, Koelman JHTM, Van Der Meer JN, Bour LJ, De Haan L, et al. Reduced parietal P300 amplitude is associated with an increased risk for a first psychotic episode. Biological Psychiatry. 2010;68(7):642-8.
43. Muller N, Weidinger E, Leitner B, Schwarz MJ. The role of inflammation in schizophrenia. Frontiers in neuroscience. 2015;9:372.
44. Müller N. Inflammation in Schizophrenia: Pathogenetic Aspects and Therapeutic Considerations. Schizophr Bull. 2018;44(5):973-82.
45. Hsu MC, Huang YS, Ouyang WC. Beneficial effects of omega-3 fatty acid supplementation in schizophrenia: possible mechanisms. Lipids Health Dis. 2020;19(1):159.
46. Marrone MC, Coccurello R. Dietary Fatty Acids and Microbiota-Brain Communication in Neuropsychiatric Diseases. Biomolecules. 2019;10(1).
47. 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.
48. 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. Schizophrenia bulletin. 2018;44(2):297-306.
49. Focking 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):326.
50. Focking 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. 2019.
51. 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. 2021:1-14.
52. 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;222:58-72.
53. Cannon M, Clarke MC, Cotter DR. Priming the brain for psychosis: maternal inflammation during fetal development and the risk of later psychiatric disorder. Am J Psychiatry. 171. United States2014. p. 901-5.
54. 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. Transl Psychiatry. 2021;11(1):321.

55. 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. Transl Psychiatry. 2017;7(8):e1220.
56. Smesny S, Milleit B, Hipler UC, Milleit C, Schäfer MR, Klier CM, et al. Omega-3 fatty acid supplementation changes intracellular phospholipase A2 activity and membrane fatty acid profiles in individuals at ultra-high risk for psychosis. Molecular Psychiatry. 2014;19(6):317-24.
57. Amminger GP, Schafer MR, Papageorgiou K, Klier CM, Cotton SM, Harrigan SM, et al. Long-chain omega-3 fatty acids for indicated prevention of psychotic disorders: a randomized, placebo-controlled trial. Arch Gen Psychiatry. 2010;67(2):146-54.
58. 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.
59. 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.
60. 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. Transl Psychiatry. 2015;5:e601.
61. Mongan D. Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence. JAMA Psychiatry (in press)2020.
62. Mongan D, Föcking M, Healy C, Susai SR, Heurich M, Wynne K, et al. Development of Proteomic Prediction Models for Transition to Psychotic Disorder in the Clinical High-Risk State and Psychotic Experiences in Adolescence. JAMA Psychiatry. 2020.
63. Niendam TA, Bearden CE, Zinberg J, Johnson JK, O'Brien M, Cannon TD. The course of neurocognition and social functioning in individuals at ultra high risk for psychosis. Schizophr Bull. 2007;33(3):772-81.
64. Bryson G, Bell MD. Initial and final work performance in schizophrenia: cognitive and symptom predictors. J Nerv Ment Dis. 2003;191(2):87-92.
65. Bowie CR, Leung WW, Reichenberg A, McClure MM, Patterson TL, Heaton RK, et al. Predicting schizophrenia patients' real-world behavior with specific neuropsychological and functional capacity measures. Biol Psychiatry. 2008;63(5):505-11.
66. Re: McFarlane, W. R., et al: clinical and functional outcomes after 2 years in the early detection and intervention for the prevention of psychosis multisite effectiveness trial. Schizophr Bull. 2015 Jan;41(1):30–43. Schizophr Bull. 412015. p. 532.
67. McFarlane WR, Levin B, Travis L, Lucas FL, Lynch S, Verdi M, et al. Clinical and functional outcomes after 2 years in the early detection and intervention for the prevention of psychosis multisite effectiveness trial. Schizophr Bull. 2015;41(1):30-43.
68. Allen P, Chaddock CA, Egerton A, Howes OD, Barker G, Bonoldi I, et al. Functional outcome in people at high risk for psychosis predicted by thalamic glutamate levels and prefronto-striatal activation. Schizophr Bull. 2015;41(2):429-39.
69. Markulev C, McGorry PD, Nelson B, Yuen HP, Schaefer M, Yung AR, et al. NEURAPRO-E study protocol: a multicentre randomized controlled trial of omega-3 fatty acids and cognitive-behavioural case management for patients at ultra high risk of schizophrenia and other psychotic disorders. Early intervention in psychiatry. 2017;11(5):418-28.
70. Sekar A, Bialas A, De Rivera H, Hammond T, Davis A, Presumey J, et al. Complex structural variation in the MHC locus influences schizophrenia risk by shaping expression of complement component 4. European Neuropsychopharmacology. 2017;27:S142.
71. 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.

72. McGorry P, Markulev C, Nelson B, Yuen HP, Schaefer M, Yung AR, et al. The neurapro-E study: a multicenter rct of omega-3 fatty acids and cognitive-behavioural case management for patients at ultra high risk of schizophrenia and other psychotic disorders. Schizophrenia bulletin. 2015;41:S322-S3.
73. Markulev C, McGorry PD, Nelson B, Yuen HP, Schaefer M, Yung AR, et al. NEURAPRO-E study protocol: a multicentre randomized controlled trial of omega-3 fatty acids and cognitive-behavioural case management for patients at ultra high risk of schizophrenia and other psychotic disorders. Early intervention in psychiatry. 2017;11(5):418-28.
74. Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. Psychological Reports. 1962;10(3):799-812.
75. Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. The British journal of psychiatry Supplement. 1989(7):49-58.
76. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. The British journal of psychiatry : the journal of mental science. 1978;133:429-35.
77. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. The British journal of psychiatry : the journal of mental science. 1979;134:382-9.
78. Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. The American journal of psychiatry. 1992;149(9):1148-56.
79. Auther AM SC, Cornblatt BA. Global Functioning: Social Scale (GF: Social). Glen Oaks, NY: Zucker Hillside Hospital. 2006.
80. Niendam TA, Bearden CE, Jognson JK, Cannon TD. Global Functioning: Role Scale (GF: Role). Los Angeles, CA: University of California, Los Angeles. 2006.
81. Alqarni A, McIntyre KJ, Brown SHJ, Meyer BJ, Mitchell TW. A High-Throughput Method for the Analysis of Erythrocyte Fatty Acids and the Omega-3 Index. Lipids. 2018;53(10):1005-15.
82. McLaverty A, Allott KA, Berger M, Hester R, McGorry PD, Nelson B, et al. Omega-3 fatty acids and neurocognitive ability in young people at ultra-high risk for psychosis. Early Interv Psychiatry. 2020.
83. Kim SW, Jhon M, Kim JM, Smesny S, Rice S, Berk M, et al. Relationship between Erythrocyte Fatty Acid Composition and Psychopathology in the Vienna Omega-3 Study. PLoS One. 2016;11(3):e0151417.
84. Bache N, Geyer PE, Bekker-Jensen DB, Hoerning O, Falkenby L, Treit PV, et al. A Novel LC System Embeds Analytes in Pre-formed Gradients for Rapid, Ultra-robust Proteomics. Mol Cell Proteomics. 2018;17(11):2284-96.
85. Beck S, Michalski A, Raether O, Lubeck M, Kaspar S, Goedecke N, et al. The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics. Mol Cell Proteomics. 2015;14(7):2014-29.
86. Meier F, Brunner AD, Koch S, Koch H, Lubeck M, Krause M, et al. Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. Mol Cell Proteomics. 2018;17(12):2534-45.
87. 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):1367-72.
88. Malipatil N, Fachim HA, Siddals K, Geary B, Wark G, Porter N, et al. Data Independent Acquisition Mass Spectrometry Can Identify Circulating Proteins That Predict Future Weight Loss with a Diet and Exercise Programme. Journal of clinical medicine. 2019;8(2):141.
89. Morrison AP, French P, Stewart SL, Birchwood M, Fowler D, Gumley AI, et al. Early detection and intervention evaluation for people at risk of psychosis: multisite randomised controlled trial. BMJ (Clinical research ed). 2012;344:e2233.
90. Dwyer D, Koutsouleris N. Annual Research Review: Translational machine learning for child and adolescent psychiatry. Journal of Child Psychology and Psychiatry. 2022;n/a(n/a).

91. Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annual review of clinical psychology. 2018;14:91-118.
92. Shatte ABR, Hutchinson DM, Teague SJ. Machine learning in mental health: a scoping review of methods and applications. Psychological medicine. 2019;49(9):1426-48.
93. Koutsouleris N, Kahn RS, Chekroud AM, Leucht S, Falkai P, Wobrock T, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. The lancet Psychiatry. 2016;3(10):935-46.
94. Ohara K. The n-3 polyunsaturated fatty acid/dopamine hypothesis of schizophrenia. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2007;31(2):469-74.
95. Das UN. Biological significance of essential fatty acids. J Assoc Physicians India. 2006;54:309-19.
96. Kiecolt-Glaser JK, Belury MA, Andridge R, Malarkey WB, Glaser R. Omega-3 supplementation lowers inflammation and anxiety in medical students: a randomized controlled trial. Brain Behav Immun. 2011;25(8):1725-34.
97. Kiecolt-Glaser JK, Belury MA, Andridge R, Malarkey WB, Hwang BS, Glaser R. Omega-3 supplementation lowers inflammation in healthy middle-aged and older adults: a randomized controlled trial. Brain Behav Immun. 2012;26(6):988-95.
98. Pawelczyk T, Grancow M, Kotlicka-Antczak M, Trafalska E, Gebski P, Szemraj J, et al. Omega-3 fatty acids in first-episode schizophrenia - a randomized controlled study of efficacy and relapse prevention (OFFER): rationale, design, and methods. BMC psychiatry. 2015;15:97.
99. Cadenhead KS, Minichino A, Kelsven S, Addington J, Bearden C, Cannon TD, et al. Metabolic abnormalities and low dietary Omega 3 are associated with symptom severity and worse functioning prior to the onset of psychosis: findings from the North American Prodrome Longitudinal Studies Consortium. Schizophrenia research. 2018.
100. Amminger GP, Chanen AM, Ohmann S, Klier CM, Mossaheb N, Bechdolf A, et al. Omega-3 fatty acid supplementation in adolescents with borderline personality disorder and ultra-high risk criteria for psychosis: a post hoc subgroup analysis of a double-blind, randomized controlled trial. Can J Psychiatry. 2013;58(7):402-8.
101. Berger M, McGorry PD, Nelson B, Markulev C, Yuen HP, Schaefer M, et al. Cross-sectional association of membrane fatty acid composition and psychopathology in the NEURAPRO-E study. Schizophrenia Bulletin. 2018;44:S116.
102. Cornblatt BA, Carrión RE, Addington J, Seidman L, Walker EF, Cannon TD, et al. Risk factors for psychosis: impaired social and role functioning. Schizophr Bull. 2012;38(6):1247-57.
103. Pouget J, Gonc¸alves V, Misener V, Kennedy J, Knight J. Hypothesis-driven genome-wide association study (GWAS) analysis highlights the role of immune genes in the extended major histocompatibility complex in schizophrenia. Schizophrenia research. 2014;153:S239.
104. Pouget JG. The Emerging Immunogenetic Architecture of Schizophrenia. Schizophrenia bulletin. 2018;44(5):993-1004.
105. Mongan D, Föcking M, Healy C, Susai SR, Heurich M, Wynne K, et al. 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.

TABLES

Table 1: Baseline Characteristics of the participants

	Individuals with baseline and follow-up data		Individuals with only baseline data		Total	
N	158		146		304	
	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Age in years	16.32	(±3.12)	19.67	(±4.5)	18.98	(±4.49)
BMI in Kg/m^2	24.33	(±4.77)	23.92	(±4.03)	23.97	(±5.46)
Sex-Females (%)	80 (50.6 %)		76 (50 %)		156 (46.8%)	
Clinical Characteristics						
BPRS Total	41.95	(±10.32)	37.11	(±9.64)	41.07	(±9.72)
SANS Total	19.53	(±15.33)	13.11	(±11.1)	17.98	(±12.8)
YMRS Total	3.32	(±2.38)	1.89	(±2.47)	3.25	(±3.02)
MADRS Total	19.63	(±10.42)	13.67	(±11.7)	19.27	(±8.97)
SOFAS	53.53	(±10.72)	65.00	(±15.8)	53.51	(±11.9)
Global Functioning - Social	6.11	(±1.24)	7.33	(±1.22)	6.51	(±1.21)
Global Functioning - Role	5.58	(±1.83)	6.89	(±1.36)	5.96	(±1.54)
Omega-3 Fatty acid levels						
EPA in %	0.98	(±0.31)	0.97	(±0.35)	0.98	(±0.33)
DHA in %	6.23	(±1.25)	6.64	(±1.78)	6.44	(±1.61)
Omega-3 INDEX in %	7.22	(±1.38)	7.62	(±1.97)	7.41	(±1.77)
Total Omega-3 in %	12.01	(±1.66)	12.10	(±2.11)	12.03	(±2.01)

Table 2. Performance metrics for Model 1 (clinical predictors), Model 2 (biomarker predictors) and Model 3 (clinical and biomarker predictors)

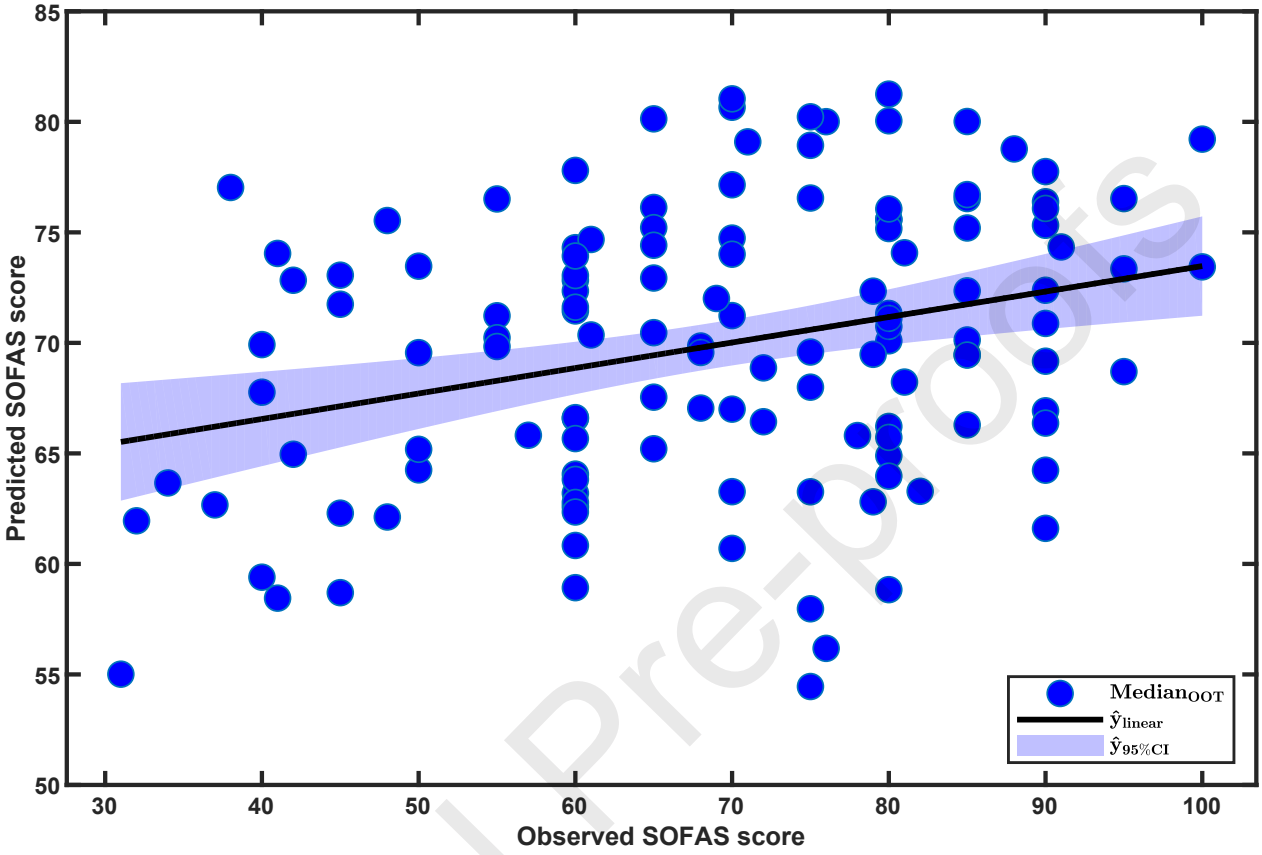
	Model 1	**Model 2**	**Model 3**
Regression performance metrics			
Coefficient of determination (R^2), %	8.9	6.2	5.0
Pearson's r	0.30	0.25	0.22
Mean absolute error	13.0	13.4	13.4
Mean squared error	239.0	256.2	250.0
Normalised root mean square deviation	22.4	23.2	22.9
Classification performance metrics (≤70 vs > 70 points)			
True positives, n	41	34	45
True negatives, n	33	42	31
False positives, n	27	18	29
False negatives, n	30	37	26
Sensitivity, %	57.7	47.9	63.4
Specificity, %	55.0	70.0	51.7
Balanced accuracy, %	56.4	58.9	57.5
Area under the curve	0.63	0.62	0.58
Positive predictive value	60.3	65.4	60.8
Negative predictive value	52.4	53.2	54.4
Positive likelihood ratio	1.3	1.6	1.3
Negative likelihood ratio	0.8	0.7	0.7

Table 3. Mean feature weighting in each model (top 10% of features shown for Models 2 and 3)

Model 1		Model 2 (top 10% features)		Model 3 (top 10% features)	
Feature	**Mean weight**	**Feature**	**Mean weight**	**Feature**	**Mean weight**
Log BMI	-0.53	Immunoglobulin heavy constant delta	0.20	GFS score	-0.24
Log Age	0.40	Beta-Ala-His dipeptidase	-0.17	SOFAS score	-0.17
MADRS score	-0.38	Biotinidase	-0.15	Immunoglobulin heavy variable 2-26	-0.17
GFS score	0.36	Actin, cytoplasmic 1	0.15	Monocyte differentiation antigen CD14	0.16
SANS score	0.21	Platelet factor 4 variant	-0.14	Biotinidase	0.15
YMRS score	-0.18	Monocyte differentiation antigen CD14	-0.14	MADRS score	0.14
Male sex	-0.16	Complement C4-A	0.14	Linoelic acid	-0.13
SOFAS score	0.16	Immunoglobulin heavy variable 3-15	-0.14	Immunoglobulin lambda variable 3-21	0.13
BPRS score	-0.07	Total Omega-6	0.13	Serum paraoxonase/arylesterase 1	-0.12
GFR score	0.07	TNF-alpha	-0.13	Complement C1r subcomponent-like protein	-0.12
Smoking	-0.07	Secretoglobin family 3A member 1	0.13	Immunoglobulin heavy variable 3-9	0.12
		Prothrombin	-0.13	Immunoglobulin lambda variable 1-36	0.12
		Actin, alpha skeletal muscle	0.12	Apolipoprotein E	0.12
		Filamin A-interacting protein 1-like	0.12	Immunoglobulin heavy constant mu	-0.12
		Immunoglobulin heavy constant alpha 1	-0.12	Immunoglobulin heavy variable 3-15	0.12
		Alpha-2-macroglobulin	0.11	Log Age	-0.11
		Complement C4-B	0.11	Transthyretin	0.11
				sICAM1	0.10

FIGURES

Figure 1. Observed vs. predicted SOFAS score for Model 1



SOFAS: Social and Occuptional Functioning Scale

Figure 2. Observed vs. predicted SOFAS score for Model 2 (biomarker predictors)

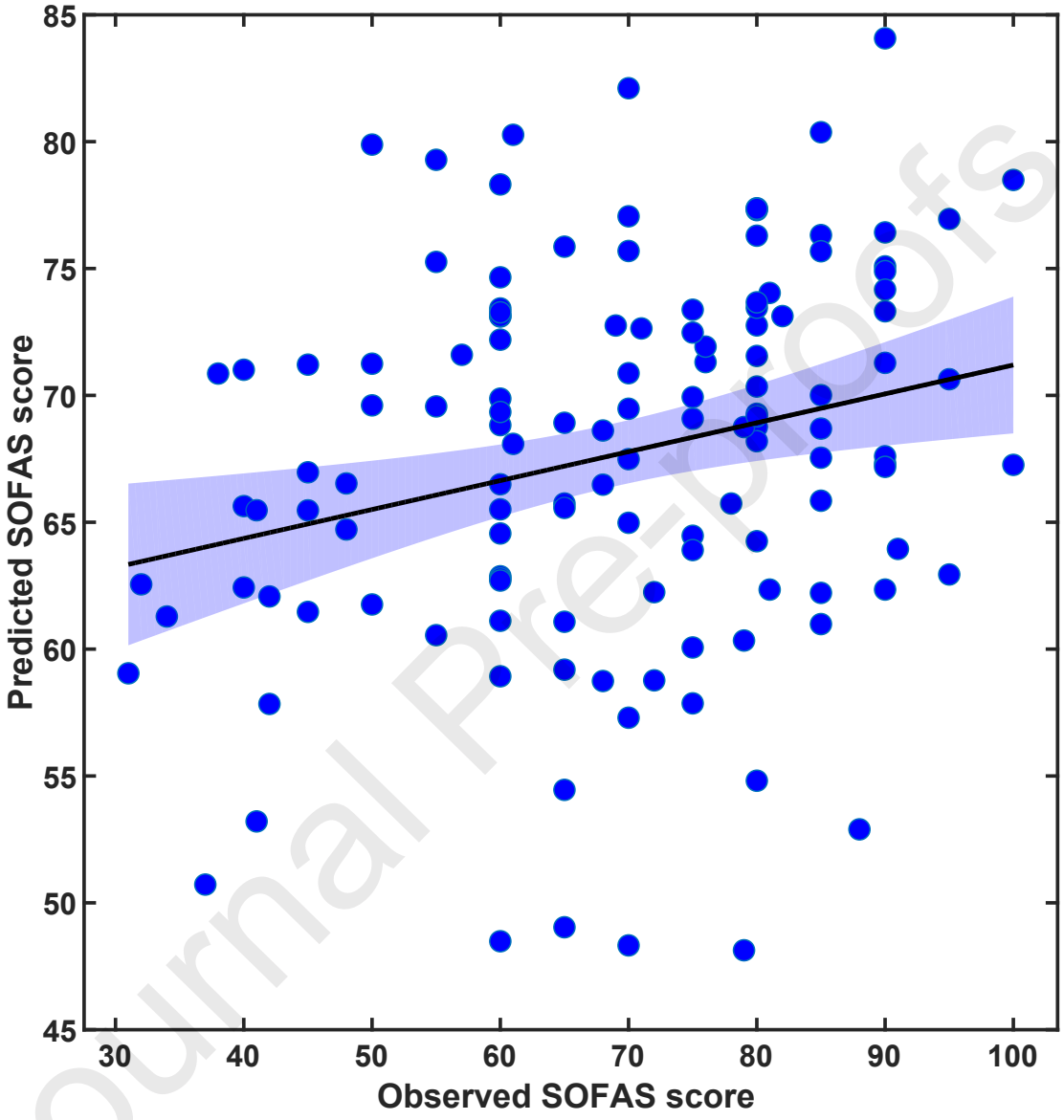
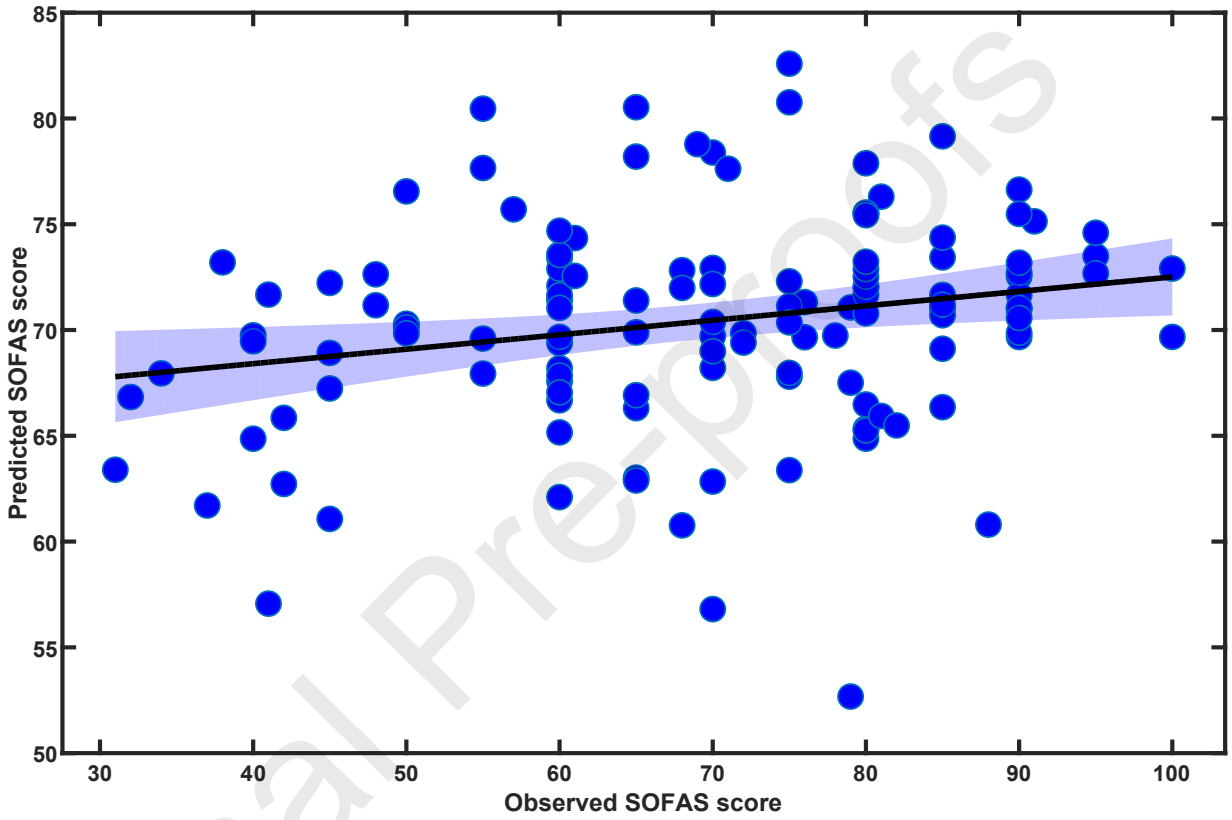


Figure 3. Class predictions based on mean algorithm score for Model 3 (clinical and biomarker predictors)



SOFAS: Social and Occupational Functioning Scale

Figure 4: Scatter plot depicting the relationship of complement and coagulation proteins with clinical outcomes at follow-up (SOFAS and PSS)

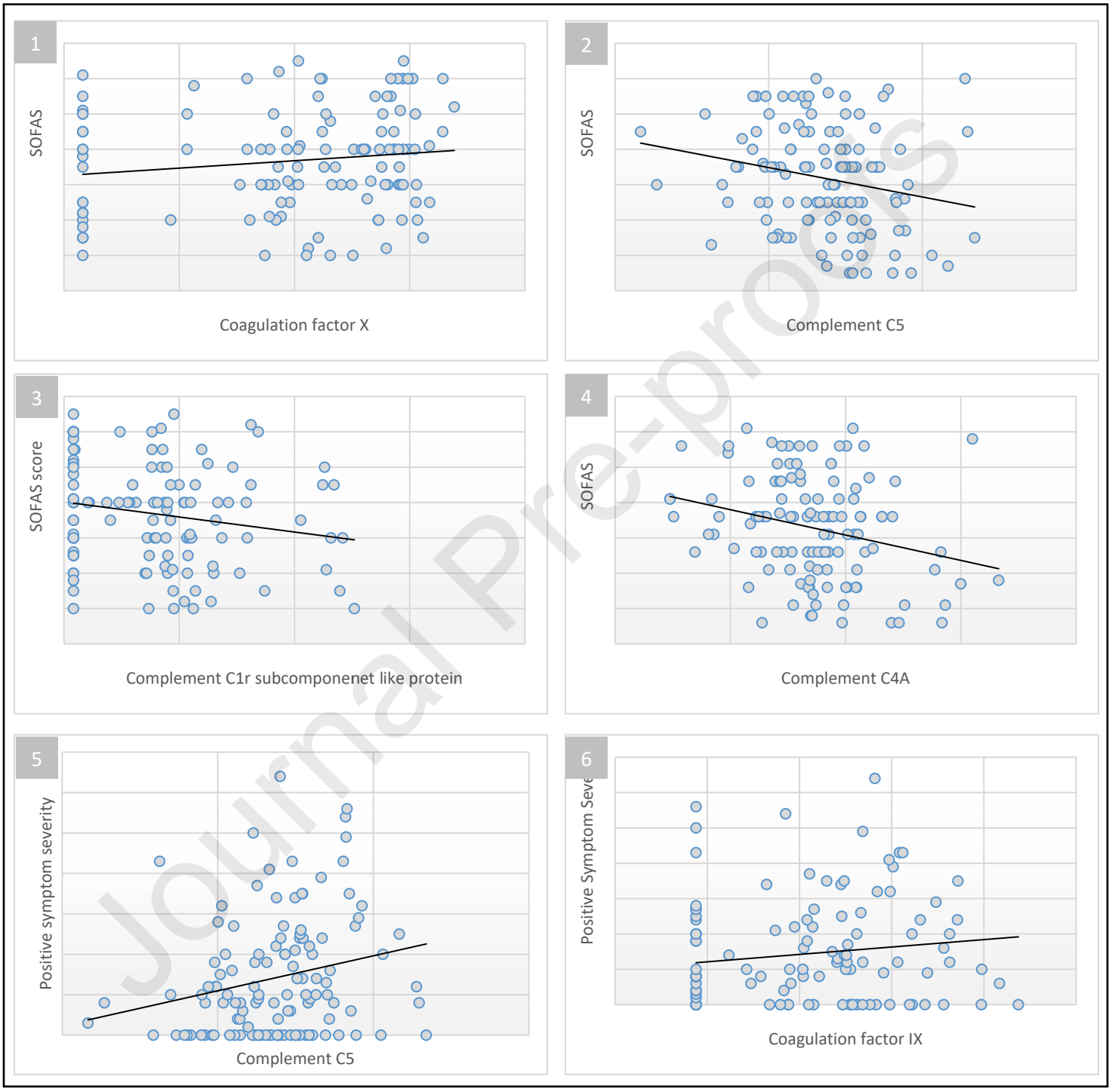


Table 4: The results of linear regression analysis between follow-up complement and coagulation proteins and follow-up SOFAS score adjusting for age, sex, baseline protein levels and baseline SOFAS score

	Coef.	p value	[95% Conf.Interval]	
Clusterin	2.087	0.112	-0.493	4.668
Coagulation factor IX	-2.264	0.089	-4.879	0.351
Coagulation factor V	1.624	0.230	-1.040	4.287
Coagulation factor X	**2.634**	**0.042**	**0.101**	**5.168**
Coagulation factor XII	0.598	0.682	-2.284	3.480
Coagulation factor XIII A chain	-1.348	0.311	-3.972	1.276
Coagulation factor XIII B chain	1.873	0.169	-0.809	4.556
Complement C1q subcomponent subunit B	0.997	0.468	-1.714	3.709
Complement C1q subcomponent subunit C	1.326	0.321	-1.309	3.962
Complement C1r subcomponent	-1.283	0.334	-3.903	1.338
Complement C1r subcomponent like protein	**-2.720**	**0.036**	**-5.263**	**-0.177**
Complement C1s subcomponent	-0.299	0.824	-2.949	2.352
Complement C2	0.038	0.977	-2.574	2.649
Complement C3	-1.637	0.227	-4.307	1.034
Complement C4A	**-3.132**	**0.020**	**-5.753**	**-0.511**
Complement C4B	-1.251	0.404	-4.211	1.709
Complement C5	**-2.936**	**0.030**	**-5.580**	**-0.291**
Complement component C6	-0.522	0.693	-3.131	2.088
Complement component C7	-0.344	0.806	-3.110	2.421
Complement component C8 alpha chain	0.158	0.907	-2.521	2.837
Complement component C8 beta chain	-1.756	0.197	-4.437	0.924
Complement component C8 gamma chain	0.521	0.694	-2.100	3.142
Complement component C9	-0.715	0.600	-3.408	1.977
Complement factor B	-1.115	0.428	-3.893	1.663
Complement factor H	-1.086	0.429	-3.797	1.624
Complement factor I	-1.254	0.361	-3.965	1.456
Fibrinogen alpha chain	-1.496	0.251	-4.065	1.073
Fibrinogen beta chain B	-1.942	0.138	-4.516	0.631
Fibrinogen beta chain C	-0.836	0.525	-3.437	1.764
Fibrinogen gamma chain	-2.030	0.119	-4.593	0.533
Ficolin3	2.723	0.051	-0.014	5.460
Heparin co factor 2	-1.597	0.232	-4.233	1.039
Protein Z dependent protease inhibitor	0.051	0.970	-2.608	2.710
Prothrombin	1.131	0.429	-1.695	3.957
Vitronectin	1.304	0.379	-1.622	4.230

Table 5: The results of linear regression analysis between follow-up complement and coagulation proteins and follow-up Positive symptom summery score adjusting for age, sex, baseline protein levels and baseline Positive symptom summary score

	Coef.	P value	[95% Conf.Interval]	
Clusterin	-0.022	0.987	-2.646	2.603
Coagulation factor IX	**2.559**	**0.043**	**0.086**	**5.031**
Coagulation factor V	-0.881	0.504	-3.489	1.727
Coagulation factor X	-0.426	0.741	-2.985	2.132
Coagulation factor XII	0.424	0.761	-2.332	3.179
Coagulation factor XIII A chain	0.632	0.608	-1.803	3.068
Coagulation factor XIII B chain	-0.609	0.642	-3.203	1.984
Complement C1q subcomponent subunit B	1.011	0.421	-1.475	3.497
Complement C1q subcomponent subunit C	0.067	0.955	-2.300	2.434
Complement C1r subcomponent	0.165	0.894	-2.281	2.611
Complement C1r subcomponent like protein	-0.993	0.429	-3.476	1.490
Complement C1s subcomponent	1.479	0.248	-1.047	4.004
Complement C2	-0.576	0.623	-2.899	1.746
Complement C3	0.354	0.782	-2.180	2.889
Complement C4A	-0.092	0.942	-2.615	2.431
Complement C4B	-0.654	0.649	-3.495	2.186
Complement C5	**2.610**	**0.034**	**0.199**	**5.020**
Complement component C6	0.406	0.736	-1.982	2.790
Complement component C7	-0.319	0.806	-2.890	2.252
Complement component C8 alpha chain	0.825	0.511	-1.655	3.305
Complement component C8 beta chain	0.860	0.492	-1.614	3.333
Complement component C8 gamma chain	0.593	0.635	-1.877	3.063
Complement component C9	0.128	0.921	-2.428	2.684
Complement factor B	-1.182	0.427	-4.124	1.760
Complement facto H	-0.727	0.570	-3.255	1.801
Complement factor I	2.262	0.075	-0.231	4.755
Fibrinogen alpha chain A	-0.323	0.802	-2.873	2.227
Fibrinogen beta chain B	-0.479	0.706	-2.991	2.032
Fibrinogen beta chain C	-0.504	0.679	-2.915	1.907
Fibrinogen gamma chain	-0.088	0.946	-2.666	2.490
Ficolin 3	-0.330	0.806	-2.995	2.335
Heparin co factor 2	0.269	0.837	-2.310	2.847
Protein Z dependent protease inhibitor	2.139	0.092	-0.352	4.630
Prothrombin	1.399	0.302	-1.278	4.075
Vitronectin	-0.147	0.914	-2.845	2.550

Supplementary Table: list of 187 predictors

Sex
Smoking_Status
log_age
log_BMI
B_bprst
B_sanst
B_ymrst
B_madrst
B_sofas
B_gf_s
B_gf_r
logIL12p40_bl_nooutliersCV20
logIL15_bl_nooutliersCV20
logIFNy_bl_nooutliersCV20
logIL6_bl_nooutliersCV20
logIL8_bl_nooutliersCV20
logIL10_bl_nooutliersCV20
logTNFa_bl_nooutliersCV20
logCRP_bl_nooutliersCV20
logsVCAM1_bl_nooutliersCV20
logsICAM1_bl_nooutliersCV20
log2_B_P04217
log2_B_P01023
log2_B_P43652
log2_B_P01019
log2_B_P02765
log2_B_P02768
log2_B_P02760
log2_B_P02743
log2_B_P02647
log2_B_P02652
log2_B_P06727
log2_B_P04114
log2_B_P02655
log2_B_P02656
log2_B_P02649
log2_B_O147912
log2_B_O95445
log2_B_P25311

log2_B_P00736
log2_B_P09871
log2_B_P06681
log2_B_P01024
log2_B_P0C0L5
log2_B_P01031
log2_B_P13671
log2_B_P07357
log2_B_P07358
log2_B_P00751
log2_B_P08603
log2_B_P05156
log2_B_P10909
log2_B_P00450
log2_B_P00748
log2_B_P00488
log2_B_P00734
log2_B_P02671
log2_B_P02675
log2_B_P02679
log2_B_Q4L180
log2_B_P0275114
log2_B_P02774
log2_B_P06396
log2_B_P69905
log2_B_P68871
log2_B_P00738
log2_B_P00739
log2_B_P02790
log2_B_P04196
log2_B_P35858
log2_B_P01876
log2_B_P01857
log2_B_P01859
log2_B_P01860
log2_B_P01861
log2_B_P01871
log2_B_P01834
log2_B_P01615
log2_B_P0DOY3
log2_B_B9A064
log2_B_P19827
log2_B_P19823

log2_B_Q14624
log2_B_P03952
log2_B_P01042
log2_B_P02750
log2_B_P02763
log2_B_P19652
log2_B_P00747
log2_B_P27169
log2_B_P01009
log2_B_P01011
log2_B_P29622
log2_B_P08185
log2_B_P05543
log2_B_P01008
log2_B_P05546
log2_B_P36955
log2_B_P08697
log2_B_P05155
log2_B_P02787
log2_B_P02766
log2_B_P04004
log2_B_P60709
log2_B_P02654
log2_B_P05090
log2_B_P10643
log2_B_P15169
log2_B_P51884
log2_B_P02753
log2_B_O75460
log2_B_Q08380
log2_B_P07225
log2_B_P35542
log2_B_P51451
log2_B_P07360
log2_B_Q14520
log2_B_A0A0B4J1V2
log2_B_Q06033
log2_B_P0C0L4
log2_B_P02746
log2_B_O75636
log2_B_P02748
log2_B_Q9Y490
log2_B_P06702

log2_B_P01764
log2_B_P23142
log2_B_P06312
log2_B_Q96PD5
log2_B_Q92820
log2_B_P02749
log2_B_P01619
log2_B_P01782
log2_B_Q96QR1
log2_B_Q15582
log2_B_P08571
log2_B_P18428
log2_B_Q04756
log2_B_Q15848
log2_B_P01591
log2_B_P31944
log2_B_P01880
log2_B_P01817
log2_B_Q16610
log2_B_P00742
log2_B_P80748
log2_B_P01614
log2_B_Q6EMK4
log2_B_P02775
log2_B_A0A0B4J1X5
log2_B_P20742
log2_B_A0A0B4J1U3
log2_B_P02747
log2_B_A0A0B4J1V0
log2_B_P04003
log2_B_P10720
log2_B_P22792
log2_B_P43251
log2_B_P26927
log2_B_Q96KN2
log2_B_P12259
log2_B_P231424
log2_B_Q9UK55
log2_B_A0A0C4DH38
log2_B_P01780
log2_B_A0A0C4DH33
log2_B_Q9NZP8
log2_B_P04406

log2_B_P68133
log2_B_A0A0B4J1Y9
log2_B_P21333
log2_B_P05160
log2_B_A0A0C4DH31
log2_B_P00740
log2_B_Q9H4B7
log2_B_P16070
log2_B_P35579
log2_B_P0DJI8
Total_Omega_3_t1
DPA_t1
DHA_t1
Total_Omega_6_t1
LA_t1
ARA_t1
Omega_3_index_t1
Omega_6_3_ratio_t1
logALA_t1
logEPA_t1

Supplementary table: 1

Biological Measures	
C Reactive Protein in pg/ml (Mean±SD)	9223.4 ± 10146861.6
Iterferon γ in pg/ml (Mean±SD)	6.0 ± 18.3
Inter leukin-10 in pg/ml (Mean±SD)	0.4 ± 0.2
Inter leukin-12p40 in pg/ml (Mean±SD)	166.2 ± 75.3
Inter leukin-6 in pg/ml (Mean±SD)	1.0 ± 2.8
Inter leukin-8 in pg/ml (Mean±SD)	4.5 ± 4.3
Intercellular adhasion molecule-1 in pg/ml (Mean±SD)	9223.4 ± 9223.4
Vascular cell adhesion molecue-1 in pg/ml (Mean±SD)	9223.4 ± 9223.4
Tumor necrosis factor-α in pg/ml (Mean±SD)	2.3 ± 0.7
Eicosapentaenoic acid (20:5) in %	1.0 ± 0.3
Docosahexaenoic acid (22:5) in %	6.3 ± 1.3
Total Omega-3 fatty acids in %	11.9 ± 1.7
Palmitic acid (16:0) in %	32.6 ± 2.4
Margeric acid (17:0) in %	1.3 ± 0.2
Stearic acid (18:0) in %	12.9 ± 1.5
Oleic acid (18:1) in %	16.6 ± 2.1
Linoleic acid (18:2) in %	12.0 ± 2.1
Dihomo-γ-linolenic acid (20:3) in %	0.8 ± 0.3
Arachidonic acid (20:4) in %	8.4 ± 2.6
Eicosapentaenoic acid (20:5) in %	1.0 ± 0.3
Docosatetraenoic acid (22:4) in %	0.4 ± 0.3
Docosahexaenoic acid (22:5) in %	6.3 ± 1.3
Cervonic acid (22:6) in %	2.3 ± 1.0
Lignoceric acid (24:0) in %	4.5 ± 2.2
Nervonic acid (24:1) in %	6.4 ± 3.3
P04217, LFQ (Mean ± SD)	926379.2 ± 211751.5
P01023, LFQ (Mean ± SD)	5180773.4 ± 1556636.7
P43652, LFQ (Mean ± SD)	69980.1 ± 27880.8
P01019, LFQ (Mean ± SD)	422425.4 ± 210805.5
P02765, LFQ (Mean ± SD)	3290935.6 ± 1367534.3
P02768, LFQ (Mean ± SD)	50363651.9 ± 9957066.2
P02760, LFQ (Mean ± SD)	484019.4 ± 300487.7
P02743, LFQ (Mean ± SD)	240745.6 ± 105269.4
P02647, LFQ (Mean ± SD)	5971527.8 ± 1420031.2
P02652, LFQ (Mean ± SD)	501833.8 ± 694358.5
P06727, LFQ (Mean ± SD)	874553.5 ± 307956.5
P04114, LFQ (Mean ± SD)	1200241.3 ± 333413.4
P02655, LFQ (Mean ± SD)	505803.3 ± 237060.2
P02656, LFQ (Mean ± SD)	967881.2 ± 550408.9

P02649, LFQ (Mean ± SD)	111951.5 ± 44456.8
O14791, LFQ (Mean ± SD)	246066.4 ± 108181.4
O95445, LFQ (Mean ± SD)	272500.1 ± 126474.8
P25311, LFQ (Mean ± SD)	286765.8 ± 129124.0
P00736, LFQ (Mean ± SD)	138054.1 ± 48333.9
P09871, LFQ (Mean ± SD)	133487.7 ± 43612.9
P06681, LFQ (Mean ± SD)	86331.4 ± 20472.8
P01024, LFQ (Mean ± SD)	3126393.0 ± 736850.6
P0C0L5, LFQ (Mean ± SD)	1674898.2 ± 473435.8
P01031, LFQ (Mean ± SD)	611807.3 ± 159410.2
P13671, LFQ (Mean ± SD)	229613.6 ± 70526.4
P07357, LFQ (Mean ± SD)	312692.5 ± 92151.8
P07358, LFQ (Mean ± SD)	92111.0 ± 30765.1
P00751, LFQ (Mean ± SD)	698053.2 ± 203160.8
P08603, LFQ (Mean ± SD)	828811.1 ± 179926.5
P05156, LFQ (Mean ± SD)	156576.4 ± 53584.8
P10909, LFQ (Mean ± SD)	790484.4 ± 215506.2
P00450, LFQ (Mean ± SD)	2260714.4 ± 593770.3
P00748, LFQ (Mean ± SD)	255915.9 ± 104593.8
P00488, LFQ (Mean ± SD)	99887.5 ± 127954.9
P00734, LFQ (Mean ± SD)	218670.1 ± 106638.9
P02671, LFQ (Mean ± SD)	7868638.0 ± 4554649.4
P02675, LFQ (Mean ± SD)	6377647.5 ± 3641791.0
P02679, LFQ (Mean ± SD)	5895613.3 ± 3165533.7
Q4L180, LFQ (Mean ± SD)	107156.0 ± 52011.2
P027511, LFQ (Mean ± SD)	757063.8 ± 1125560.6
P02774, LFQ (Mean ± SD)	759986.5 ± 346360.4
P06396, LFQ (Mean ± SD)	612246.4 ± 194166.0
P69905, LFQ (Mean ± SD)	2996032.7 ± 1534224.3
P68871, LFQ (Mean ± SD)	2346903.7 ± 1182810.7
P00738, LFQ (Mean ± SD)	4505029.4 ± 1791915.5
P00739, LFQ (Mean ± SD)	364916.9 ± 235317.1
P02790, LFQ (Mean ± SD)	2140823.4 ± 553552.2
P04196, LFQ (Mean ± SD)	351694.1 ± 137271.4
P35858, LFQ (Mean ± SD)	207829.0 ± 56511.6
P01876, LFQ (Mean ± SD)	609347.8 ± 437131.3
P01857, LFQ (Mean ± SD)	25562246.8 ± 9716938.8
P01859, LFQ (Mean ± SD)	13778848.1 ± 7978493.8
P01860, LFQ (Mean ± SD)	8118640.5 ± 3032653.0
P01861, LFQ (Mean ± SD)	452901.9 ± 397949.0
P01871, LFQ (Mean ± SD)	2588777.5 ± 1387710.2

P01834, LFQ (Mean ± SD)	12143191.1 ± 3168508.4
P01615, LFQ (Mean ± SD)	440919.1 ± 260914.2
P0DOY3, LFQ (Mean ± SD)	1111800.3 ± 1744134.9
B9A064, LFQ (Mean ± SD)	613847.3 ± 290571.4
P19827, LFQ (Mean ± SD)	646335.3 ± 258491.7
P19823, LFQ (Mean ± SD)	1278065.5 ± 284617.1
Q14624, LFQ (Mean ± SD)	827661.2 ± 170778.7
P03952, LFQ (Mean ± SD)	112914.3 ± 34709.4
P01042, LFQ (Mean ± SD)	1365858.1 ± 272388.5
P02750, LFQ (Mean ± SD)	502177.8 ± 211514.2
P02763, LFQ (Mean ± SD)	7349527.2 ± 2174441.9
P19652, LFQ (Mean ± SD)	887286.0 ± 258657.5
P00747, LFQ (Mean ± SD)	774582.7 ± 169004.5
P27169, LFQ (Mean ± SD)	627053.2 ± 210113.2
P01009, LFQ (Mean ± SD)	9887256.3 ± 1951893.3
P01011, LFQ (Mean ± SD)	1019439.2 ± 259491.4
P29622, LFQ (Mean ± SD)	82565.4 ± 32204.3
P08185, LFQ (Mean ± SD)	369387.5 ± 125087.6
P05543, LFQ (Mean ± SD)	58999.0 ± 47301.5
P01008, LFQ (Mean ± SD)	1076794.7 ± 191550.8
P05546, LFQ (Mean ± SD)	228000.2 ± 82784.6
P36955, LFQ (Mean ± SD)	292096.0 ± 91292.7
P08697, LFQ (Mean ± SD)	1205121.2 ± 256794.4
P05155, LFQ (Mean ± SD)	1859243.5 ± 834119.1
P02787, LFQ (Mean ± SD)	4636338.6 ± 869870.0
P02766, LFQ (Mean ± SD)	4429880.1 ± 1105753.7
P04004, LFQ (Mean ± SD)	552299.9 ± 116630.4
P15636, LFQ (Mean ± SD)	2803067.6 ± 2575238.4
P60709, LFQ (Mean ± SD)	355878.7 ± 442117.2
P02654, LFQ (Mean ± SD)	95730.0 ± 91909.7
P05090, LFQ (Mean ± SD)	146637.1 ± 79555.4
P10643, LFQ (Mean ± SD)	226166.3 ± 128551.1
P15169, LFQ (Mean ± SD)	64560.3 ± 31863.4
P51884, LFQ (Mean ± SD)	86791.9 ± 39340.8
P02753, LFQ (Mean ± SD)	121092.5 ± 65430.3
O75460, LFQ (Mean ± SD)	487720.3 ± 241985.1
Q08380, LFQ (Mean ± SD)	153232.9 ± 69192.8
P07225, LFQ (Mean ± SD)	133708.0 ± 66554.4
P35542, LFQ (Mean ± SD)	152378.2 ± 96834.2
P51451, LFQ (Mean ± SD)	1246495.6 ± 719864.7
P07360, LFQ (Mean ± SD)	201121.3 ± 108212.7

Q14520, LFQ (Mean ± SD)	49573.6 ± 15528.6
A0A0B4J1V2, LFQ (Mean ± SD)	92695.8 ± 53465.4
Q06033, LFQ (Mean ± SD)	69707.0 ± 36693.6
P0C0L4, LFQ (Mean ± SD)	259380.0 ± 213745.6
P02746, LFQ (Mean ± SD)	753308.6 ± 1347829.9
O75636, LFQ (Mean ± SD)	46782.9 ± 22892.4
P02748, LFQ (Mean ± SD)	112312.2 ± 142523.7
Q9Y490, LFQ (Mean ± SD)	21745.6 ± 11162.7
P06702, LFQ (Mean ± SD)	88699.3 ± 644187.6
P01764, LFQ (Mean ± SD)	120193.9 ± 118078.1
P23142, LFQ (Mean ± SD)	28387.2 ± 20581.2
P06312, LFQ (Mean ± SD)	112357.3 ± 165125.5
Q96PD5, LFQ (Mean ± SD)	44333.3 ± 27338.2
Q92820, LFQ (Mean ± SD)	105851.0 ± 76675.7
P02749, LFQ (Mean ± SD)	97265.9 ± 91479.2
P01619, LFQ (Mean ± SD)	264531.3 ± 304933.2
P01782, LFQ (Mean ± SD)	68173.7 ± 74239.0
Q96QR1, LFQ (Mean ± SD)	41220.2 ± 32644.5
Q15582, LFQ (Mean ± SD)	20837.2 ± 14092.2
P08571, LFQ (Mean ± SD)	17145.2 ± 11300.2
P18428, LFQ (Mean ± SD)	17218.5 ± 16056.4
Q04756, LFQ (Mean ± SD)	47221.5 ± 35456.9
Q15848, LFQ (Mean ± SD)	33871.3 ± 23503.3
P01591, LFQ (Mean ± SD)	106243.8 ± 133385.4
P31944, LFQ (Mean ± SD)	50632.3 ± 71177.8
P01880, LFQ (Mean ± SD)	93563.3 ± 135165.9
P02676, LFQ (Mean ± SD)	77939.0 ± 120292.0
P01817, LFQ (Mean ± SD)	83955.1 ± 71659.8
Q16610, LFQ (Mean ± SD)	41353.9 ± 84882.6
P00742, LFQ (Mean ± SD)	28964.0 ± 24339.7
P80748, LFQ (Mean ± SD)	153427.4 ± 156940.0
P01614, LFQ (Mean ± SD)	53768.5 ± 50380.3
Q6EMK4, LFQ (Mean ± SD)	16083.3 ± 11734.5
P02775, LFQ (Mean ± SD)	66113.3 ± 63716.3
A0A0B4J1X5, LFQ (Mean ± SD)	19783.8 ± 24511.9
P20742, LFQ (Mean ± SD)	193488.8 ± 507297.4
A0A0B4J1U3, LFQ (Mean ± SD)	48484.2 ± 50882.7
P02747, LFQ (Mean ± SD)	36434.6 ± 30286.5
A0A0B4J1V0, LFQ (Mean ± SD)	40782.7 ± 34775.1
P04003, LFQ (Mean ± SD)	42664.8 ± 41170.1
P10720, LFQ (Mean ± SD)	32531.7 ± 34941.1

P22792, LFQ (Mean ± SD)	42792.0 ± 60629.0
P43251, LFQ (Mean ± SD)	14252.0 ± 14373.0
P26927, LFQ (Mean ± SD)	9320.5 ± 7766.1
Q96KN2, LFQ (Mean ± SD)	16516.7 ± 14285.7
P12259, LFQ (Mean ± SD)	50112.7 ± 42420.0
P231424, LFQ (Mean ± SD)	23006.2 ± 18314.4
Q9UK55, LFQ (Mean ± SD)	13918.7 ± 16722.8
A0A0C4DH38, LFQ (Mean ± SD)	34472.4 ± 39744.1
P01780, LFQ (Mean ± SD)	151323.9 ± 134699.9
A0A0C4DH33, LFQ (Mean ± SD)	14261.1 ± 27913.7
Q9NZP8, LFQ (Mean ± SD)	35350.7 ± 43068.8
P04406, LFQ (Mean ± SD)	14004.1 ± 24652.0
P68133, LFQ (Mean ± SD)	70577.0 ± 98559.6
A0A0B4J1Y9, LFQ (Mean ± SD)	13187.2 ± 15818.0
P21333, LFQ (Mean ± SD)	27941.8 ± 65925.6
P05160, LFQ (Mean ± SD)	27430.3 ± 58784.5
A0A0C4DH31, LFQ (Mean ± SD)	34929.5 ± 36954.7
P00740, LFQ (Mean ± SD)	10425.1 ± 16104.1
Q9H4B7, LFQ (Mean ± SD)	2758543.1 ± 2776277.6
P16070, LFQ (Mean ± SD)	5348.0 ± 4828.8
P35579, LFQ (Mean ± SD)	25449.8 ± 50149.2
P0DJI8, LFQ (Mean ± SD)	32355.5 ± 86921.2

Supplementary Table 2: Details of percentage of missing values

Variable	n missing	% missing
M12_sofas	0	0.0
male	0	0.0
Smoking_Status	0	0.0
log_age	0	0.0
log_BMI	9	6.9
B_bprst	0	0.0
B_sanst	0	0.0
B_ymrst	0	0.0
B_madrst	0	0.0
B_sofas	0	0.0
B_gf_s	0	0.0
B_gf_r	0	0.0
logIL12p40_bl_nooutliersCV20	3	2.3
logIL15_bl_nooutliersCV20	3	2.3
logIFNy_bl_nooutliersCV20	3	2.3
logIL6_bl_nooutliersCV20	2	1.5
logIL8_bl_nooutliersCV20	2	1.5
logIL10_bl_nooutliersCV20	0	0.0
logTNFa_bl_nooutliersCV20	3	2.3
logCRP_bl_nooutliersCV20	0	0.0
logsVCAM1_bl_nooutliersCV20	1	0.8
logsICAM1_bl_nooutliersCV20	2	1.5
Total_Omega_3_t1	2	1.5
DPA_t1	2	1.5
DHA_t1	2	1.5
Total_Omega_6_t1	2	1.5
LA_t1	2	1.5
ARA_t1	2	1.5
Omega_3_index_t1	2	1.5
Omega_6_3_ratio_t1	2	1.5
logALA_t1	5	3.8
logEPA_t1	2	1.5

Highlights:

- Biological markers did not improve machine learning prediction of clinical outcome in CHR
- Complement proteins (Factor X, C1r subcomponent, C4A & C5) associate inversely with functional outcome
- C 5 associate positively with positive symptoms severity