Title Page

Article Type: Original Investigation

Study Type: Cohort Study

Title

Blood Immuno-metabolic Biomarker Signatures of Depression and Affective Symptoms in Young Adults

Authors

Nicholas A Donnelly[1,2,3] MRCPsych, PhD (corresponding author, nick.donnelly@bristol.ac.uk)

Ruby SM Tsang[1,2], PhD

Éimear M Foley[1,2], MSc

Holly Fraser[1.2], PhD

Aimee L Hanson[1], PhD

Golam M Khandaker[1,2,3,4] FRCPsych, PhD

Affiliations

[1]MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol, UK

[2]Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

[3]Avon and Wiltshire Mental Health Partnership NHS Trust, Bristol, UK

[4] NIHR Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, UK

Word Count

Total word count: 3000 words

Key Points

Question: Depression is linked to immuno-metabolic dysfunction, but what is the precise nature of these associations at biomarker and symptom level, and can we predict depression using immuno-metabolic biomarkers?

Findings: Using 93 blood immuno-metabolic biomarkers and depression measures from up-to 4196 participants, we identified distinct clusters/groupings of immuno-metabolic biomarkers and depressive and anxiety symptoms which are differently associated with subsequent symptom persistence. These biomarkers predict specific symptom profiles better than others within people with depression.

Meaning: Heterogeneity in the associations of immuno-metabolic biomarkers with mood and anxiety symptoms is relevant for prognosis and could aid better stratification/prediction of depression.

Abstract

Importance

Depression is associated with alterations in blood immuno-metabolic biomarkers, but it remains unclear whether these alterations are limited to select measures or represent broader patterns and can predict depression diagnosis.

Objective

To examine immuno-metabolic biomarker changes in depression, pattern of effect at the symptom and symptom-dimension level, and prediction of depression diagnosis.

Design, Setting, and Participants

Case-control and cohort-wide analyses of ICD-10 depression, depressive and anxiety symptoms based on up to N=4161 participants (2363 female) aged 24 years from the Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort.

Exposures

Blood-based immunological and metabolic biomarkers (n=93) comprising inflammatory proteins, cell count, lipids, hormones, and metabolites.

Main Outcomes and Measures

ICD-10 diagnosis of depression, 11 individual depressive and anxiety symptoms, and four domain scores were used as outcomes. Confounders included sex assigned at birth, body mass index, smoking, and alcohol use.

Results

After adjusting for potential confounders and multiple testing, depression was associated with changes in concentrations of specific immuno-metabolic markers (IL-6, CDCP1, neutrophil count, and insulin), and greater number of extreme-valued inflammatory markers. We identified three distinct affective symptom-related biomarker clusters, including one comprising inflammatory cytokines, chemokines and cells which was positively associated with somatic and mood

symptoms, and one comprising liver-related biomarkers which was negatively associated with anxiety symptoms. Then using Partial Least Squares regression we identified two latent variables that capture the biomarker-symptom associations (Component 1: Somatic-Depressive-Inflammation and Component 2: Anxiety-Hepatic). Higher Component 1 score was associated with higher depressive symptom severity consistently over subsequent five years. Immuno-metabolic biomarkers performed poorly in predicting ICD-10 depression (0.569 Balanced Accuracy). However, within depression cases the addition of immuno-metabolic biomarkers improved the prediction of depression *with* high levels of mood (0.720 Balanced Accuracy) or anxiety symptoms (0.636 Balanced Accuracy).

Conclusion and Relevance

Depression is associated with disruption in immuno-metabolic homeostasis. Specific patterns of immuno-metabolic biomarkers are associated with differing subsets of affective symptoms, which are potentially relevant for poor depression prognosis. Immuno-metabolic biomarkers improve predictions of high levels of mood symptoms within people with depression, highlighting the symptom-level heterogeneity of depression and opportunities for immuno-metabolic biomarker-based subtyping, prediction, and targeted intervention.

Introduction

Evidence of immunological and metabolic alterations in depression is widely replicated, notably changes in concentrations of blood inflammatory (e.g., cytokines, acute phase proteins, immune cell count)[1–4] and metabolic biomarkers (e.g., lipids, metabolites[5,6]). Existing case-control studies have typically focused on selected common inflammatory proteins. However, studies involving a broader set of biomarkers are required to ascertain whether immuno-metabolic alterations in depression are restricted to a few biomarkers or reflect wider system-wide disruption in immuno-metabolic homeostasis. Group mean comparisons often overlook substantial individual variability in biomarkers, which can provide important additional information regarding the overall pattern of change[4]. Recent neuroimaging findings suggest that depression and other psychiatric disorders are characterised by increased total number of extreme deviations from normative models of grey matter volume[7], but it is unclear whether such extreme variability is also evident for immuno-metabolic biomarkers.

Depression is a syndrome comprising various mood and anxiety symptoms, and it is possible that immuno-metabolic processes have differing contributions across the spectrum of observed symptomatology[8]. Symptom-level investigations have hitherto involved few specific inflammatory markers, namely interleukin 6 (IL-6) and C-reactive protein (CRP), showing that these biomarkers are particularly associated with somatic symptoms of depression (e.g., fatigue, altered sleep[8,9]). While these candidate biomarker studies are informative, data driven approaches examining the heterogeneity of the effects of a wide variety of immuno-metabolic biomarkers across the affective spectrum (individual symptoms and symptom domains) is needed. This could identify distinct biomarker signatures and clustering of depressive symptoms along the immuno-metabolic axis, providing insight into potential pathophysiologic drivers and could inform biomarker-based stratification and prediction of depression.

We present an in-depth investigation of the relationship of blood immuno-metabolic biomarkers (93 inflammatory proteins, cell count, lipids, and hormones) and depression measures (ICD-10 diagnosis, 11 symptoms, 4 symptom domains) based on data from young adults aged 24 years from the Avon Longitudinal Study of Parents and Children[10], a UK birth cohort. In particular, we

aimed to: i) examine the pattern of immuno-metabolic biomarker changes associated with ICD-10 depression compared to controls, including mean concentrations and extreme values of biomarkers; ii) identify distinct biomarker clusters and latent components underlying the relationship between affective symptoms and immuno-metabolic variables; and iii) investigate whether machine learning models trained on immuno-metabolic variables can predict ICD-10 diagnoses of depression and specific symptom profiles within depression, and compare model performance with models trained on sociodemographic and clinical data.

Methods

Datasets

Cohort Profile

We used data from the Avon Longitudinal Study of Parents and Children (ALSPAC), a UK general population-based longitudinal birth cohort[10–12]. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and Local Research Ethics Committees. All participants provided written informed consent. Biological samples were collected with informed consent and in accordance with the UK Human Tissue Act (2004).

Mental Health Data

Our main outcomes were ICD-10 depression diagnosis and specific symptoms or symptom domains derived from the Clinical Interview Schedule – Revised (CIS-R) at age 24[13]. Given previous research suggesting that depressive symptoms can be best described by a "psychological" and a "somatic" factor[14], we derived a domain score for psychological (0–4 scale) and somatic (0-9 scale) symptoms. For comparison to previous studies[8,15,16] we also created atypical (0–5 scale) and anxiety (0–7 scale) domain scores. Additionally, we used CIS-R items to derive specific symptom scores for: Depressed Mood, Depressive Cognitions, Fatigue, Sleep Problems, Impaired Concentration and Forgetfulness, Aches and Pains, Anxiety, Irritability, Panic, Phobias, and Worry.

As positive and negative controls, respectively, in relation to depression, we included sum scores for the Community Assessment of Psychic Experiences (CAPE[17]) and the Standardised Assessment of Personality – Abbreviated Scale (SAPAS[18]) questionnaires (**Supplementary Methods**). Because some negative symptoms of psychosis are phenomenologically like depressive symptoms, they could associate with immuno-metabolic biomarkers while personality difficulty should not.

For depression and anxiety symptoms at follow-up, we used the Short Mood and Feeling Questionnaire (SMFQ) total score, completed by ALSPAC participants at four timepoints between ages 25 and 30.

Blood Proteomic and Other Biomarker Analysis

These included immune proteins measured by the Olink Target 96 Inflammation Panel[19], blood cell counts, lipids, hormones and metabolites measured at age 24 (**Supplementary Table 1**). See **Supplementary Methods** for full details on sample processing. A total of 93 immuno-metabolic variables were included based on i) >50% of samples above the Olink Limit of Detection; ii) >90% complete data. For multiple imputation analysis we used Olink data from 3005 samples of the same variables from age 9 (**Supplementary Methods)**.

Statistical Analysis

Software

Statistical Analysis was carried out using R Statistical Software (version 4.3[20]).

Training/Testing Split

To preserve a held-out set of data for model evaluation, before analysis we split our dataset into training (80%) and testing sets (20%), stratified by sex, depression diagnosis and BMI quartile.

Univariate Statistical Models

We identified potential covariates using a directed acyclic graph (**Supplementary Methods**): sex assigned at birth (reference group, male), BMI at age 24, daily smoking, and alcohol use (measured via the AUDIT-C questionnaire). For each immuno-metabolic variable, we fit three statistical models: (1) unadjusted model: depression diagnosis and the immuno-metabolic variable, (2) adjusted model including sex, alcohol use and smoking; and (3) fully adjusted model additionally including BMI at age 24. Models were fit using the *lm* function, and average marginal effects (AME) calculated using the *marginaleffects* package. We corrected for multiple comparisons using the Benjamini-Hochberg method[21].

Symptom Models

We fit models to all pairs of depression measure (diagnosis, domain scores, and symptom scores) and immuno-metabolic measure, with the full set of covariates (above). Models were fit with a binomial distribution using *glm*. As a sensitivity analysis we examined the relationship between a subset of immuno-metabolic variables and antidepressant medication use at age 24.

Variable Clustering

We applied hierarchical clustering to the matrix of average marginal effect (AME) values for all pairs of symptoms and immuno-metabolic variables. The dendrogram was then cut into clusters, which were labelled after inspecting their constituent variables.

Extreme Value Analysis

We determined extreme-valued immuno-metabolic variables by taking non-depressed participants and calculating a normative model for each variable, using our full set of covariates. Then, for each participant (both non-depressed and depressed) a standardised residual was calculated using the model of non-depressed individuals. A residual was classified as “extreme” if the absolute value was greater than 2.6, based on a study applying a similar approach to brain imaging[7]. We then took the sum of extreme values across all immuno-metabolic variables.

Models predicting the count of extreme values from depression diagnosis or depression symptoms were fit with all covariates using a Poisson *glm*.

Partial Least Squares

We fit partial least squares models (PLS) to 11 depressive symptom scores and 38 hepatic and immune cluster biomarker variables using the mixOmics package *spls* function[22]. We selected 2 latent components for further analysis based on inspection of the plot of the explained variance for 1–5 potential components and used cross-validation to select the number of variables selected for each component. We extracted PLS component scores for all components from all individuals and regressed scores against covariates, childhood trauma scores[23] and total sum score on subsequent SMFQ measurements (**Supplementary Methods**).

Machine Learning Model Fitting

We fit elastic net regression models with the *glmnet* R package to predict ICD-10 depression diagnosis in all participants. In addition, we assessed the prediction of three symptom profiles within participants with ICD-10 depression. These were cases with high levels of: (1) *mood* (CIS-R Depressive Cognitions and Depressed Mood items score $\geq$ 6), (2) *anxiety* (CIS-R Anxiety and Worry items score $\geq$ 6), and (3) *somatic* (CIS-R Aches/Pains, Fatigue and Sleep items score $\geq$ 6)

symptoms. The cut-offs were chosen to ensure adequate symptom severity while maintaining statistical power based on the distribution of these symptom scores in the sample.

We used four sets of predictors: (1) *Immuno-metabolic*: immune and hepatic cluster variables, plus immune cluster extreme value counts and covariates; (2) *Sociodemographic*: all covariates, plus maternal occupational social class and history of adverse childhood events aged 0–17[24]; (3) *Mental Health History*: previous CIS-R depression diagnosis at age 18, SMFQ total score at ages 10 to 23, Strengths and Difficulties (SDQ) Total Difficulties Score ages 7 and 9, and Maternal Edinburgh Postnatal Depression Scale Score; (4) *Combined*: all variables included in previous models. Nested cross-validation was used to tune model hyperparameters (**Supplementary Methods**).

Model performance was evaluated in previously held-out test data. Confidence intervals were calculated for final model performance by drawing bootstrap samples (n=1000) from the training dataset and calculating classification performance in each sample, then calculating the 95% percentile interval of these. Variable importance was calculated from final model coefficients.

Results

The training dataset contained 2357 individuals (Male N=862, Female N=1,495), of whom 245 (10%) had ICD-10 depression at age 24 (**Table 1** and **Supplementary Figure 1B-G**).

Peripheral Blood Immuno-metabolic Biomarker Changes in Depression

We compared the mean concentration of immuno-metabolic variables between depression cases and controls. In unadjusted analysis, after correction for multiple comparisons (n=93 immuno-metabolic variables), 9 immuno-metabolic variables were altered in depression (**Table 2**): alanine aminotransferase (ALT) (Average Standardised Mean Difference [AME]=-0.198, 95% Confidence Interval [-0.332, -0.064], P_{FDR}=0.048); CUB domain-containing protein 1 (0.272 [0.138, 0.405], P_{FDR}=0.002); Haemoglobin (-0.205 [-0.339, -0.07], P_{FDR}=0.045); Haematocrit (-0.256 [-0.391, -0.122], P_{FDR} = 0.004); Interleukin-6 (0.275 [0.141, 0.409], P_{FDR}=0.002); Insulin (0.239 [0.107, 0.371], P_{FDR}=0.007); Neutrophil Count (0.282 [0.147, 0.418], P_{FDR}=0.002), Red Blood Cell Count (-0.196 [-0.331, -0.061]), P_{FDR} = 0.046 and White Blood Cell Count (0.25 [0.115, 0.386], P_{FDR}=0.007).

Evidence for association remained for IL-6, neutrophil count, CDCP1, and insulin after adjustment for sex at birth, smoking, alcohol use, and BMI, though the respective P-values did not survive correction for multiple testing. See **Supplementary Table 2** for results for all biomarkers.

As a sensitivity analysis we used multiple imputation for missing data and fit the same models to the imputed data. The results showed similar patterns of association (**Supplementary Figure 2 and Supplementary Table 3)**.

Symptom Level Heterogeneity in Immuno-metabolic Biomarker Associations

Next we investigated where our immuno-metabolic variables were correlated with 11 individual depression and anxiety symptoms, four symptom domains scores, psychosis negative symptom score and personality difficulty score (**Supplementary Methods and Supplementary Figure 3)**.

We found largely positive associations of immuno-metabolic biomarkers and core depressive symptoms (depressed mood, depressive cognitions) and somatic symptoms of depression (e.g.,

fatigue, altered sleep), and negative associations with anxiety symptoms (**Figure 1A**). As predicted, immuno-metabolic biomarkers were largely positively associated with psychosis negative symptoms but not with personality difficulty.

Affective Symptom-Associated Immuno-metabolic Biomarker Clusters

Hierarchical clustering applied to all pairwise associations of symptoms and immuno-metabolic variables identified three clusters, which were distinct with regards to their relationship to affective symptoms (**Figure 1B-D** and **Supplementary Table 1**). Cluster 3 (*immune*) comprised 33 biomarkers which were mainly inflammatory cytokines (e.g., IL-6, IL-17C), chemokines (e.g., CCL4, CCL11, CCL19, CCL20, CCL25) and immune cell counts (e.g., neutrophil, monocytes and white blood cells). This cluster showed positive associations with mood and somatic symptoms (sleep problems, depressive mood, depressive cognitions, and fatigue) (**Figure 1D**). Cluster 1 (*hepatic*) comprised 5 biomarkers, which were largely liver-related including CRP, ALT, aspartate aminotransferase (AST), and gamma-glutamyl transpeptide (γGT), and showed predominantly negative associations with most symptoms, especially anxiety and worry (**Figure 1C**). Cluster 2 (*low-association*) comprised 55 biomarkers and showed weak associations with all symptoms.

Extreme Values in Immuno-metabolic Biomarkers in Depression

We calculated the total number of extreme-valued immuno-metabolic variables for each participant. All biomarkers had at least one participant with an extreme value (**Supplementary Figure 4A**). Participants with depression had higher total extreme-valued biomarkers (adjusted AME=0.371 [0.190, 0.552], P<0.001, **Supplementary Figure 4B-C**).

Next, we fit a model including immuno-metabolic cluster identity. We found an interaction between depression and cluster on the number of extreme values ($\chi^2_{df=2}$=7.711, P=0.021). The number of extreme-valued biomarkers was higher in people with depression for biomarkers belonging to the immune cluster (AME=0.242 [0.085, 0.399], P=0.003), but not the other two clusters (AME for the low-association cluster=0.141 [-0.060, 0.341], P=0.169, and AME for hepatic cluster= -0.004 [-0.042, 0.034], P=0.854) (**Supplementary Figure 4D**).

Latent Components Underlying Affective Symptoms and Immuno-metabolic Biomarker Associations

We used Partial Least Squares (PLS) regression to identify latent variables that capture associations between immuno-metabolic biomarkers and affective symptoms. We fit a PLS model to 11 symptoms and the 38 affective symptom-associated immune and hepatic cluster biomarkers in our training dataset, which identified 2 latent components explaining 21.4% of variation in immuno-metabolic variables and 68.9% of variance in the depression variables. Component 1 ("*Somatic-depressive-inflammation"*) loaded on somatic and depressive symptoms (Fatigue, Sleep, Depressed Mood, Aches/Pain, and Depressive Cognitions) and inflammatory markers including neutrophil count, WBC count and IL-6; Component 2 ("*Anxiety-hepatic"*) loaded on Worry and Anxiety symptoms and ADA and ALT (**Figure 2A-C**). Component scores correlated differently with BMI (**Figure 2F**). Component 1 score was consistently associated with higher depressive symptoms severity at follow-ups between ages 25 and 30 (**Figure 2G**).

We validated our results by extracting the same PLS components from the held-out test dataset not used for PLS model development (N=833). Test scores had similar relationships to covariates (although the magnitude of effects was reduced), and Component 1 scores continued to be associated with future SMFQ score (**Supplementary Figure 5**).

Prediction of Depression and Specific Symptom Profiles using Immuno-metabolic Biomarkers

We built machine learning models (elastic net regression) to predict ICD-10 depression. Informed by our symptom-based analyses, we also generated three additional outcomes to extend our investigation to specific symptom types within ICD-10 depression: depression *with* high levels of mood (n=151/264), anxiety (n=128/264), or somatic (sleep, fatigue, and aches/pains) symptoms (n=178/264, **Supplementary Figure 6**). We evaluated classification performance in the held-out test data (n=651 for ICD-10 depression diagnosis, n=73 for symptoms within individuals with ICD-10 depression).

Models based on immuno-metabolic variables performed better than chance in predicting ICD-10 depression diagnosis (Balanced Accuracy=0.569; 95%CI = [0.508, 0.634]); **Figure 3, Supplementary Table 3.** Performance was similar for models trained using socio-demographic factors (0.604 [0.543, 0.663]), while a model trained with past mental health history (0.663 [0.605, 0.724]) and a model containing all variables (0.688 [0.631, 0.744]) performed better.

Within ICD-10 depression, combining immuno-metabolic biomarkers with past mental health history improved the predictive performance for depression *with* high levels of mood symptoms (0.72 [0.592, 0.831]) or anxiety symptoms (0.636 [0.518, 0.739]), whereas other combinations of predictors and symptoms did not perform above chance.

Variable importance varied widely between models. For the model predicting highest levels of mood symptoms, smoking, WBC, monocyte count, ADA, LDL, HGF, CX3CL1, neutrophil count, platelets and cholesterol were the 10 most important features (**Supplementary Figure 7**).

Discussion

We show that young adults with depression have immunological and metabolic alterations reflected by changes in mean levels and number of extreme valued biomarkers. The relationship of biomarkers with individual affective symptoms varies, with inflammatory and hepatic biomarkers showing distinct, opposing associations with depressive/somatic and anxiety symptoms respectively. We identified two latent components explaining the relationship between immuno-metabolic biomarkers and affective symptoms, one correlated with mood, somatic symptoms and inflammation-, metabolism-related biomarkers, while the other related to anxiety, worry and hepatic biomarkers. The somatic-depressive-inflammatory component was consistently associated with higher depressive symptoms in subsequent follow-ups. Finally, we show that machine learning models trained on immuno-metabolic biomarkers predict ICD-10 depression diagnosis better than chance, and similar to models trained on sociodemographic data, but at an absolute level their performance would be considered poor[26]. However, the addition of immuno-metabolic biomarkers to clinical history improved the prediction of depression with greater mood or anxiety symptoms.

Unlike most previous studies[4,6,8,9], we examined a large number and variety of immuno-metabolic biomarkers (93 in total) including inflammatory proteins, cell count, lipids, and metabolites, which helped identify key depression-related biomarkers (e.g., IL-6, neutrophil count, and insulin). In addition, we show that depression is associated with greater variability in immune biomarker levels (more extreme valued biomarkers), consistent with a system-wide disruption in immuno-metabolic homeostasis in young adults at age 24, well before the typical age of onset of cardiometabolic diseases.

Our symptom-level analyses provide novel insight into heterogenous relationships between immuno-metabolic biomarkers and individual features of depression. We identified a positive relationship of inflammatory proteins with somatic and core mood-related symptoms of depression and a negative relationship of hepatic biomarkers with anxiety symptoms. While this study is not focused on the issue of causality, the potential aetiological relevance of these divergent effects requires further investigation.

We identified two clusters of inflammation and metabolism related biomarkers, and similar latent components, which related differently to depressive/somatic and anxiety symptoms. The convergence of specific depressive symptoms and immuno-metabolic biomarkers is potentially relevant for prognosis. Individuals with higher somatic/depressive symptoms and altered inflammatory/metabolic biomarkers had higher severity of depressive symptoms subsequently at follow-up.

As well as corroborating previously reported biomarkers (e.g., IL-6, neutrophil count, insulin), we identified several novel biomarkers which are part of the depressive-immuno-metabolic axis (e.g., CCL25, FGF19, FGF21). Expressed in the gut, CCL25 has been reported to be associated with bipolar disorder[27,28], post-traumatic stress disorder[29], and with bowel inflammation[30] and impaired intestinal barrier function[31] in schizophrenia. FGF19 and 21 are linked to glucose and glycogen metabolism[32,33]; FGF21 is associated with the risk of type-two diabetes[34]. Taken together, our findings support the idea of immuno-metabolic subtype(s) of depression[35].

We found that immuno-metabolic biomarkers perform poorly in predicting ICD-10 depression diagnosis, similar to the predictive performance of neuroimaging biomarkers[25]. However, we also show that the addition of immuno-metabolic biomarkers markedly improves the prediction of depression *with* high levels of mood or anxiety symptoms over and above past mental health history (the strongest predictor of depression diagnosis). Taken together, our findings support the idea that immuno-metabolic dysfunction is associated with specific subtype(s) of depression[35] and suggest that combining immuno-metabolic biomarkers with symptoms and clinical history could be a useful strategy for depression stratification/prediction.

Regarding limitations, the ALSPAC study is not diverse in terms of ethnicity. Individuals who have continued to participate into early adulthood tend to be women, come from more educated and affluent backgrounds[15], and have fewer depressive symptoms[38]. Therefore, generalisability of findings to other populations could be an issue. Immuno-metabolic biomarkers were assayed at age 24. It is possible that longitudinal trends in biomarker levels derived from repeat measures could yield more accurate prediction, but such datasets remain scarce.

In conclusion, depression is associated with disruption in immuno-metabolic homeostasis. Specific patterns of immuno-metabolic biomarkers are associated with differing subsets of affective symptoms, which is potentially relevant for prognosis. While immuno-metabolic biomarkers perform poorly at predicting depression diagnosis, prediction varies by symptom profile, performing noticeably better for cases with high levels of mood symptoms, highlighting the heterogeneity of depression and opportunities for disease subtyping, prediction and targeted intervention.

Figures

Figure 1

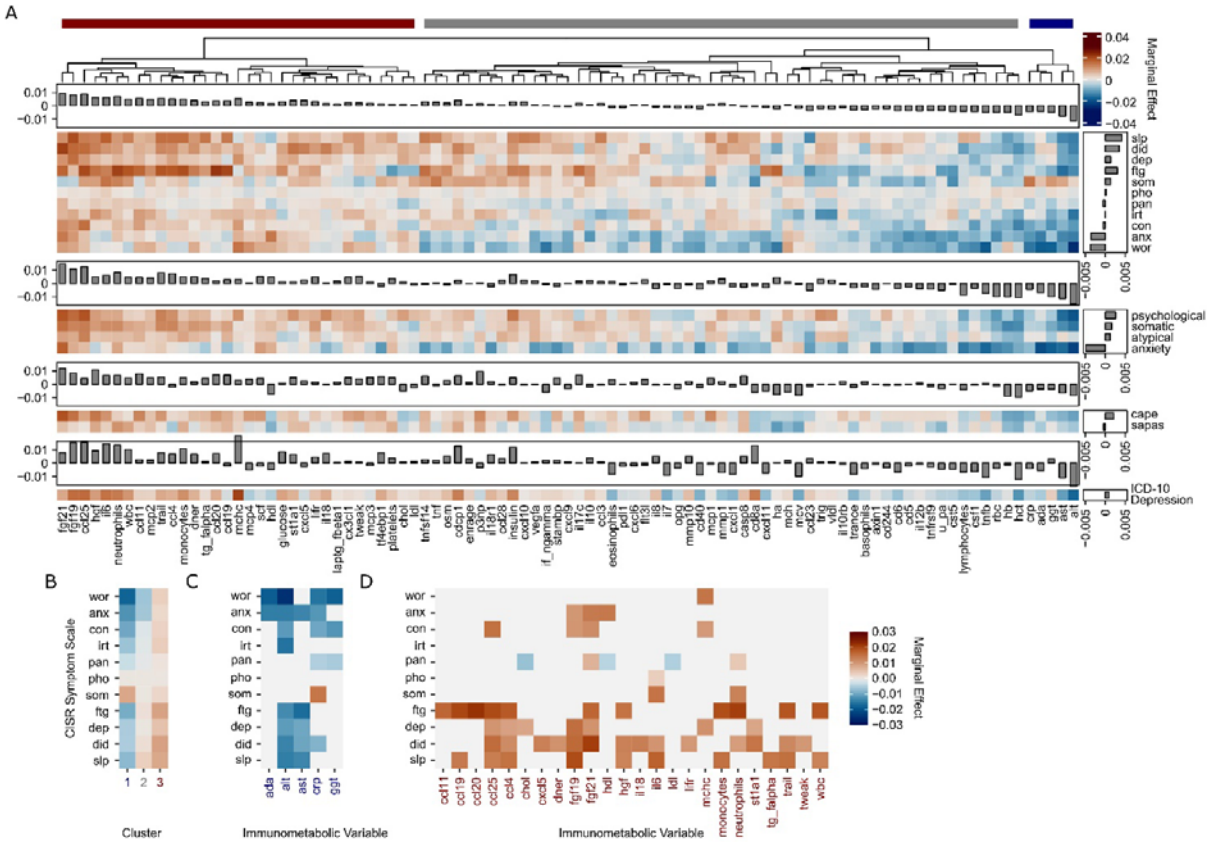

Figure 1: Associations of immuno-metabolic variables with depression and anxiety-related measures in the ALSPAC cohort

A: Matrix plots showing the average marginal effect of a 1 SD-unit change in an immuno-metabolic variable on the probability of endorsing a symptom. From top to bottom, plots are of 11 individual depressive/anxiety symptom scores (sleep problems [slp], depressive cognitions [did], depressed mood [dep], fatigue [ftg], aches and pains [som], phobias [pho], panic [pan], irritability [irt], impaired concentration [con], anxiety [anx] and worry [wor]); four depressive/anxiety symptom domain scores (psychological, somatic, atypical and anxiety); psychosis negative symptoms score (CAPE), personality difficulty score (SAPAS), and ICD-10 depression diagnosis. Within each plot, marginal histograms represent the averaged standardised effect over all rows for each column in the plot (i.e., average over all outcomes to give an overall effect for each immuno-metabolic variable), marginal histograms to the right represent the averaged AME over all columns for each row in the

plot (i.e. average over all immuno-metabolic variables for each outcome). At the top, the dendrogram from hierarchical clustering of the CISR symptom score matrix; immuno-metabolic variables are grouped based clustering of all mental health measures. Symptoms are also sorted by a similar clustering process. Variable names and descriptions are found in **Supplementary Table 1.** Note that all matrix plots are on the same colour scale.

B: Based on clustering of the immuno-metabolic variables, this plot shows the average of the average marginal effect of all immuno-metabolic-symptom associations within a cluster.

C: Depressive/anxiety symptom associations for each biomarker from Cluster 1 (hepatic cluster). Only associations that remained significant after FDR correction are included.

D: Results for biomarkers in Cluster 3 (immune cluster), presented as in Panel C.

Figure 2

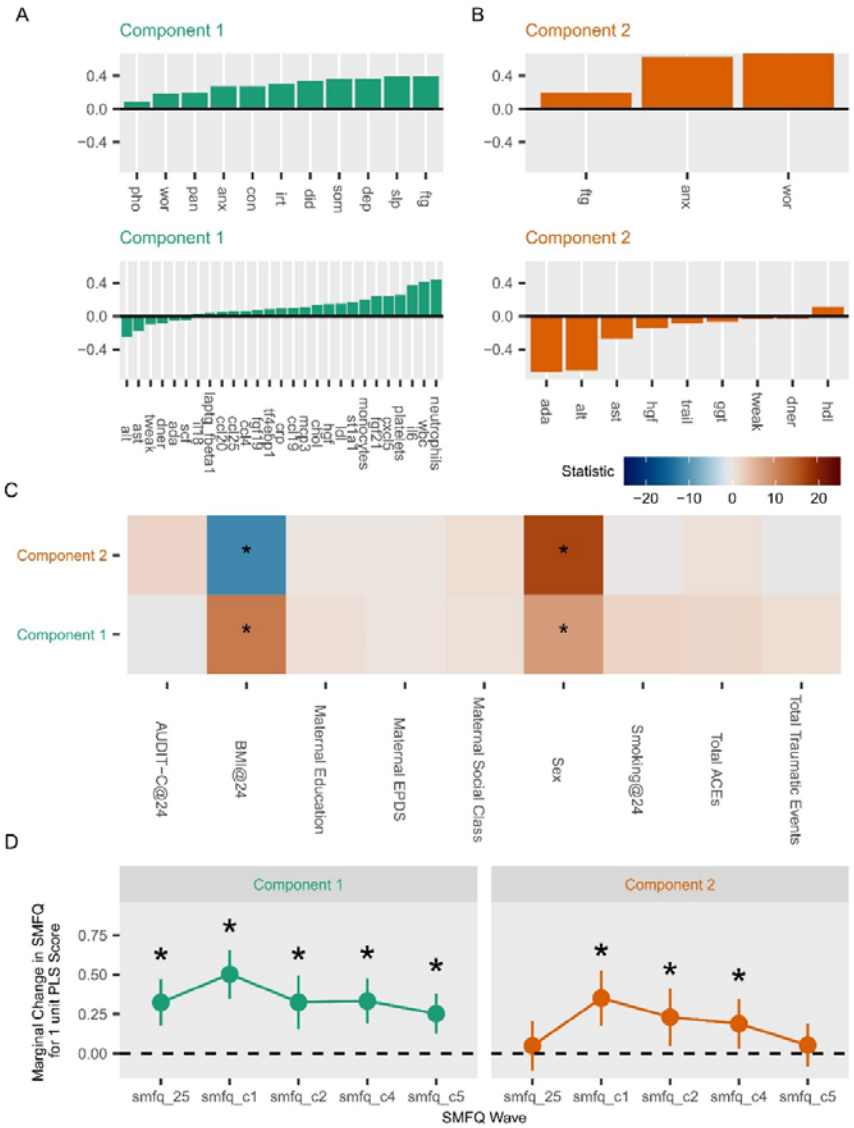

Figure 2: Latent components derived from Partial Least Squares analysis of immuno-metabolic variables and depressive symptoms

A: PLS Component 1 Loadings on symptoms (top) and immuno-metabolic variables (bottom). Component 1 had highest loadings from somatic symptoms, including fatigue (ftg), sleep problems (slp), aches/pains (som), as well as depressed mood (dep) and depressive cognitions (did). The highest loadings from immuno-metabolic variables were neutrophils and white blood cell counts and IL6.

B: PLS Component 2 Loadings on symptoms (top) and immuno-metabolic variables (bottom). Component 2 had highest loadings from anxiety and worry symptoms and the highest absolute value immuno-metabolic variables loading onto Component 2 were ALT, ADA, and AST.

C: Heatmap of pairwise test statistics of models correlating covariates and immuno-metabolic component scores. Both components correlated with sex and BMI.

D: Plot of Marginal change in SMFQ total score associated with a 1 unit increase in PLS Component score, with 95% confidence intervals. Asterix indicate $P_{FDR}<0.05$

Figure 3

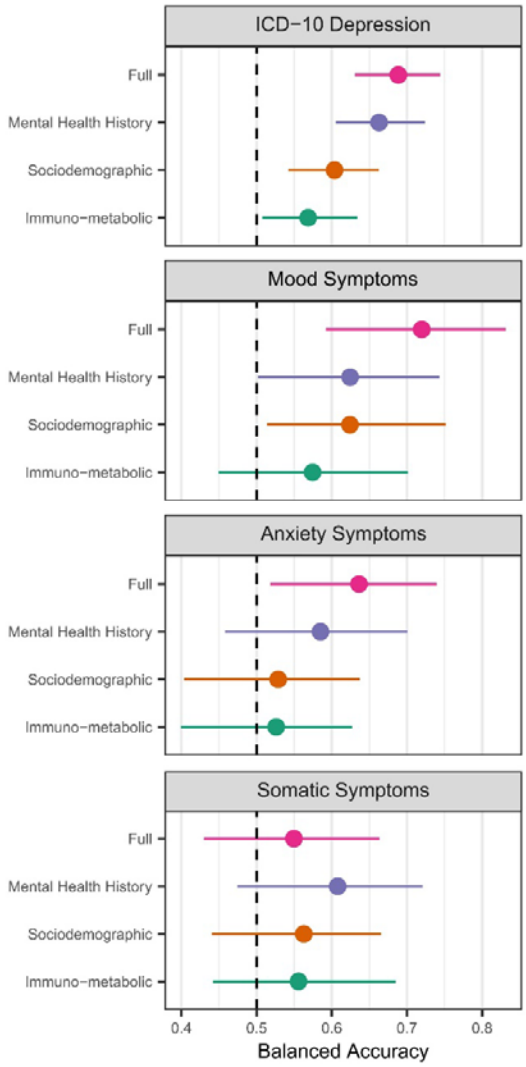

Figure 3: Performance of Machine Learning models predicting ICD-10 depression diagnosis and specific symptom profiles within individuals with ICD-10 depression

Classification performance (balanced accuracy) of models trained with selected immuno-metabolic variables predicting the presence of ICD-10 depression or high mood symptoms within ICD-10 depression, high anxiety symptoms within ICD-10 depression or high somatic symptoms within ICD-10 depression, with bootstrapped 95% confidence intervals

Tables

Table 1

Table 1: Characteristics of ALSPAC study sample used for model training

Variable	Overall, N = 2,357[1]	Sex at birth	
		Male, N = 862[1]	**Female**, N = 1,495[1]
ICD-10 Depression			
No Depression	2,112 (90%)	806 (94%)	1,306 (87%)
Depression	245 (10%)	56 (6.5%)	189 (13%)
BMI @ 24	23.7 (21.5, 26.8)	24.2 (21.8, 26.8)	23.4 (21.3, 26.8)
Missing	23	<5*	19
Daily Smoking @ 24	277 (12%)	98 (11%)	179 (12%)
Missing	13	5	8
AUDIT Alcohol Questionnaire Score @ 24	5.00 (4.00, 7.00)	6.00 (4.00, 7.00)	5.00 (3.00, 7.00)
Missing	27	11	16

[1] n (%); Median (IQR)

***Cell counts <5 are reported as "<5" in accordance with ALSPAC guidance on small cell counts**

Table 2

Table 2: Top hits for immuno-metabolic biomarker associations for ICD-10 depression at age 24 years in the ALSPAC cohort

Immunometabolic Variable	Average Marginal Effect (95% Confidence Interval) for Depression		
	Unadjusted	**Adjusted for sex at birth, smoking, and alcohol use**	**Additional adjustment for BMI**
Alanine aminotransferase	-0.198 [-0.332, -0.064], P_{FDR} = 0.043	-0.088 [-0.22, 0.044], P_{FDR} = 0.528	-0.158 [-0.285, -0.031], P_{FDR} = 0.259
CUB domain-containing protein 1	0.272 [0.138, 0.405], P_{FDR} = 0.002	0.224 [0.084, 0.365], P_{FDR} = 0.055	0.135 [0.002, 0.267], P_{FDR} = 0.437
Haemoglobin	-0.205 [-0.339, -0.07], P_{FDR} = 0.039	0.007 [-0.096, 0.11], P_{FDR} = 0.958	0.002 [-0.102, 0.105], P_{FDR} = 0.982
Haematocrit	-0.256 [-0.391, -0.122], P_{FDR} = 0.004	-0.053 [-0.162, 0.057], P_{FDR} = 0.72	-0.061 [-0.171, 0.049], P_{FDR} = 0.769
Interleukin-6	0.275 [0.141, 0.409], P_{FDR} = 0.002	0.242 [0.1, 0.384], P_{FDR} = 0.04	0.158 [0.027, 0.288], P_{FDR} = 0.259
Insulin	0.239 [0.107, 0.371], P_{FDR} = 0.006	0.246 [0.107, 0.386], P_{FDR} = 0.04	0.145 [0.026, 0.265], P_{FDR} = 0.259
Neutrophil Count	0.282 [0.147, 0.418], P_{FDR} = 0.002	0.209 [0.069, 0.349], P_{FDR} = 0.08	0.176 [0.038, 0.315], P_{FDR} = 0.259
Red blood Cell Count	-0.196 [-0.331, -0.061], P_{FDR} = 0.046	0.009 [-0.101, 0.118], P_{FDR} = 0.958	-0.017 [-0.125, 0.091], P_{FDR} = 0.913
White Blood Cell Count	0.25 [0.115, 0.386], P_{FDR}	0.17 [0.03, 0.31], P_{FDR} = 0.182	0.127 [-0.01, 0.264], P_{FDR} =

	= 0.006		0.536

Footnote: Biomarkers were selected based on FDR-corrected P Value (P_{FDR}) of <0.05 in the unadjusted model. Values displayed are Average Marginal Effect (95% Confidence Interval) and FDR-corrected P Value.

Acknowledgements

We are extremely grateful to all the families who took part in the ALSPAC cohort study, 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. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Funding

GMK acknowledges funding support from the UK Medical Research Council (MRC), grant no: MC_UU_00032/06, which forms part of the Integrative Epidemiology Unit at the University of Bristol. This grant also supports HF and ÉMF. GMK acknowledges additional funding from the Wellcome Trust (201486/Z/16/Z and 201486/B/16/Z), the MRC (MR/W014416/1; MR/S037675/1; and MR/Z50354X/1), and the UK National Institute of Health and Care Research (NIHR) Bristol Biomedical Research Centre (NIHR 203315). NAD was supported by an NIHR Clinical Lectureship in General Adult Psychiatry. The views expressed are those of the authors and not necessarily those of the UK NIHR or the Department of Health and Social Care. RSMT is supported by the Tackling Multimorbidity at Scale Strategic Priorities Fund programme (MR/W014416/1). ALM is supported by the Medical Research Council via the Integrative Epidemiology Unit (MC_UU_00032/01).

The UK Medical Research Council and Wellcome Trust (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding awarded to the ALSPAC study is available from ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). This publication is the work of the authors and NAD and GMK will serve as guarantors for the contents of this paper.

Preprint

An early version of this manuscript was deposited as a preprint to medRxiv

Data Availability

Data from the ALSPAC study is available via a managed open access system: in line with ALSPAC data access policy, all raw data used for this work can be accessed by other investigators by making a request to the ALSPAC study executive (see the study website: http://www.bristol.ac.uk/alspac/researchers/access/). Available variables can be browsed via the ALSPAC data dictionary (http://www.bristol.ac.uk/alspac/researchers/our-data/) and variable search (https://variables.alspac.bris.ac.uk/).

Analysis scripts are available online: https://github.com/NADonnelly/dep_infl

Competing Interests

The authors have no competing interests to declare.

CRediT authorship statement

Donnelly N: Conceptualisation; Study design; Software; Data Curation; Formal Analysis; Visualization; Writing – Original Draft; Writing – Reviewing & Editing

Tsang RSM: Writing – Reviewing & Editing; Resources

Foley E: Writing – Reviewing & Editing; Resources

Fraser H: Writing – Reviewing & Editing; Resources

Hanson A: Writing – Reviewing & Editing; Resources

Khandaker GM: Conceptualisation; Study design; Writing – Reviewing & Editing; Funding acquisition; Supervision

References

1. Foley, É. M., Parkinson, J. T., Mitchell, R. E., Turner, L. & Khandaker, G. M. Peripheral blood cellular immunophenotype in depression: a systematic review and meta-analysis. *Mol. Psychiatry* **28**, 1004–1019 (2023).
2. Lynall, M.-E. *et al.* Peripheral Blood Cell-Stratified Subgroups of Inflamed Depression. *Biol. Psychiatry* **88**, 185–196 (2020).
3. Goldsmith, D. R., Rapaport, M. H. & Miller, B. J. A meta-analysis of blood cytokine network alterations in psychiatric patients: comparisons between schizophrenia, bipolar disorder and depression. *Mol. Psychiatry* **21**, 1696–1709 (2016).
4. Osimo, E. F. *et al.* Inflammatory markers in depression: A meta-analysis of mean differences and variability in 5,166 patients and 5,083 controls. *Brain. Behav. Immun.* **87**, 901–909 (2020).
5. Bot, M. *et al.* Metabolomics Profile in Depression: A Pooled Analysis of 230 Metabolic Markers in 5283 Cases With Depression and 10,145 Controls. *Biol. Psychiatry* **87**, 409–418 (2020).
6. Davyson, E. *et al.* Metabolomic Investigation of Major Depressive Disorder Identifies a Potentially Causal Association With Polyunsaturated Fatty Acids. *Biol. Psychiatry* **94**, 630–639 (2023).
7. Segal, A. *et al.* Regional, circuit and network heterogeneity of brain abnormalities in psychiatric disorders. *Nat. Neurosci.* 1–17 (2023) doi:10.1038/s41593-023-01404-6.
8. Milaneschi, Y. *et al.* Association of inflammation with depression and anxiety: evidence for symptom-specificity and potential causality from UK Biobank and NESDA cohorts. *Mol. Psychiatry* (2021) doi:10.1038/s41380-021-01188-w.
9. Chu, A. L. *et al.* Longitudinal association between inflammatory markers and specific symptoms of depression in a prospective birth cohort. *Brain. Behav. Immun.* **76**, 74–81 (2019).
10. Northstone, K. *et al.* The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. *Wellcome Open Res.* **4**, 51 (2019).
11. 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.* **42**, 111–127 (2013).

12. Fraser, A. *et al.* Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
13. Lewis, G., Pelosi, A. J., Araya, R. & Dunn, G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol. Med.* **22**, 465–486 (1992).
14. Thorp, J. G. *et al.* Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9. *Psychol. Med.* **50**, 2385–2396 (2020).
15. Donnelly, N. A., Perry, B. I., Jones, H. J. & Khandaker, G. M. Childhood immuno-metabolic markers and risk of depression and psychosis in adulthood: A prospective birth cohort study. *Psychoneuroendocrinology* **139**, 105707 (2022).
16. Lamers, F. *et al.* Depression profilers and immuno-metabolic dysregulation: Longitudinal results from the NESDA study. *Brain. Behav. Immun.* **88**, 174–183 (2020).
17. Mossaheb, N. *et al.* The Community Assessment of Psychic Experience (CAPE) questionnaire as a screening-instrument in the detection of individuals at ultra-high risk for psychosis. *Schizophr. Res.* **141**, 210–214 (2012).
18. Moran, P. *et al.* Standardised Assessment of Personality - Abbreviated Scale (SAPAS): preliminary validation of a brief screen for personality disorder. *Br. J. Psychiatry J. Ment. Sci.* **183**, 228–232 (2003).
19. Goulding, N. *et al.* Inflammation proteomics datasets in the ALSPAC cohort [version 1; peer review: 1 approved, 1 approved with reservations]. *Wellcome Open Res.* (2022) doi:https://doi.org/10.12688/wellcomeopenres.18482.
20. R Development Core Team. R: A Language and environment for statistical computing. *Vienna Austria R Found. Stat. Comput.* (2017).
21. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).
22. Rohart, F., Gautier, B., Singh, A. & Lê Cao, K.-A. mixOmics: An R package for ’omics feature selection and multiple data integration. *PLoS Comput. Biol.* **13**, e1005752 (2017).
23. Croft, J. *et al.* Association of Trauma Type, Age of Exposure, and Frequency in Childhood and Adolescence With Psychotic Experiences in Early Adulthood. *JAMA Psychiatry* **76**, 79–80 (2019).

24. Houtepen, L. C., Heron, J., Suderman, M. J., Tilling, K. & Howe, L. D. Adverse childhood experiences in the children of the Avon Longitudinal Study of Parents and Children (ALSPAC). *Wellcome Open Res.* **3**, 106 (2018).
25. Winter, N. R. *et al.* A Systematic Evaluation of Machine Learning-Based Biomarkers for Major Depressive Disorder. *JAMA Psychiatry* e235083 (2024) doi:10.1001/jamapsychiatry.2023.5083.
26. Alba, A. C. *et al.* Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. *JAMA* **318**, 1377–1384 (2017).
27. Göteson, A. *et al.* A serum proteomic study of two case-control cohorts identifies novel biomarkers for bipolar disorder. *Transl. Psychiatry* **12**, 55 (2022).
28. Poletti, S. *et al.* A peripheral inflammatory signature discriminates bipolar from unipolar depression: A machine learning approach. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **105**, 110136 (2021).
29. Zhang, L. *et al.* Potential chemokine biomarkers associated with PTSD onset, risk and resilience as well as stress responses in US military service members. *Transl. Psychiatry* **10**, 31 (2020).
30. Wurbel, M.-A., McIntire, M. G., Dwyer, P. & Fiebiger, E. CCL25/CCR9 interactions regulate large intestinal inflammation in a murine model of acute colitis. *PloS One* **6**, e16442 (2011).
31. Jensen, S. B. *et al.* Elevated Systemic Levels of Markers Reflecting Intestinal Barrier Dysfunction and Inflammasome Activation Are Correlated in Severe Mental Illness. *Schizophr. Bull.* **49**, 635–645 (2023).
32. Kir, S. *et al.* FGF19 as a Postprandial, Insulin-Independent Activator of Hepatic Protein and Glycogen Synthesis. *Science* **331**, 1621–1624 (2011).
33. BonDurant, L. D. & Potthoff, M. J. Fibroblast Growth Factor 21: A Versatile Regulator of Metabolic Homeostasis. *Annu. Rev. Nutr.* **38**, 173–196 (2018).
34. Post, A. *et al.* Circulating FGF21 Concentration, Fasting Plasma Glucose, and the Risk of Type 2 Diabetes: Results From the PREVEND Study. *J. Clin. Endocrinol. Metab.* **108**, 1387–1393 (2023).

35. Milaneschi, Y., Lamers, F., Berk, M. & Penninx, B. W. J. H. Depression Heterogeneity and Its Biological Underpinnings: Toward Immunometabolic Depression. *Biol. Psychiatry* **88**, 369–380 (2020).
36. Koutsouleris, N. *et al.* Toward Generalizable and Transdiagnostic Tools for Psychosis Prediction: An Independent Validation and Improvement of the NAPLS-2 Risk Calculator in the Multisite PRONIA Cohort. *Biol. Psychiatry* **90**, 632–642 (2021).
37. Koutsouleris, N. *et al.* Multimodal Machine Learning Workflows for Prediction of Psychosis in Patients With Clinical High-Risk Syndromes and Recent-Onset Depression. *JAMA Psychiatry* **78**, 195–209 (2021).
38. Depression or anxiety in adults, Great Britain: 22 September to 3 October 2021 - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/mentalhealth/adhocs/13844depressionoranxietyinadultsgreatbritain22septemberto3october2021.