Immunometabolic Blood Biomarkers of Developmental Trajectories of Depressive Symptoms: Findings From the ALSPAC Birth Cohort

Running title: Immunometabolic markers of depression trajectories

Authors:

<table>
<thead>
<tr>
<th>Name</th>
<th>Degree</th>
<th>Email</th>
<th>ORCID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ruby S. M. Tsang</td>
<td>PhD</td>
<td><a href="mailto:ruby.tsang@bristol.ac.uk">ruby.tsang@bristol.ac.uk</a></td>
<td>0000-0002-2520-526X</td>
</tr>
<tr>
<td>Daniel Stow</td>
<td>PhD</td>
<td><a href="mailto:d.stow@qmul.ac.uk">d.stow@qmul.ac.uk</a></td>
<td>0000-0002-9534-4521</td>
</tr>
<tr>
<td>Alex S. F. Kwong</td>
<td>PhD</td>
<td><a href="mailto:akwong@ed.ac.uk">akwong@ed.ac.uk</a></td>
<td>0000-0003-1953-2771</td>
</tr>
<tr>
<td>Nicholas A. Donnelly</td>
<td>PhD, MRC Psych</td>
<td><a href="mailto:nick.donnelly@bristol.ac.uk">nick.donnelly@bristol.ac.uk</a></td>
<td>0000-0003-2234-8545</td>
</tr>
<tr>
<td>Holly Fraser</td>
<td>MSc</td>
<td><a href="mailto:holly.fraser@bristol.ac.uk">holly.fraser@bristol.ac.uk</a></td>
<td>0000-0001-9752-0042</td>
</tr>
<tr>
<td>Inês A. Barroso</td>
<td>PhD</td>
<td><a href="mailto:ines.barroso@exeter.ac.uk">ines.barroso@exeter.ac.uk</a></td>
<td>0000-0001-5800-4520</td>
</tr>
<tr>
<td>Peter A. Holmans</td>
<td>PhD</td>
<td><a href="mailto:holmanspa@cardiff.ac.uk">holmanspa@cardiff.ac.uk</a></td>
<td>0000-0003-0870-9412</td>
</tr>
<tr>
<td>Michael J. Owen</td>
<td>PhD, FRC Psych</td>
<td><a href="mailto:owenmj@cardiff.ac.uk">owenmj@cardiff.ac.uk</a></td>
<td>0000-0003-4798-0862</td>
</tr>
<tr>
<td>Megan L. Wood</td>
<td>PhD</td>
<td><a href="mailto:m.l.wood@leeds.ac.uk">m.l.wood@leeds.ac.uk</a></td>
<td>0000-0003-1882-2355</td>
</tr>
<tr>
<td>LINC Consortium*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marianne B. M. van den Bree</td>
<td>PhD</td>
<td><a href="mailto:vandenbreemb@cardiff.ac.uk">vandenbreemb@cardiff.ac.uk</a></td>
<td>0000-0002-4426-3254</td>
</tr>
<tr>
<td>Nicholas J. Timpson</td>
<td>PhD</td>
<td><a href="mailto:n.j.timpson@bristol.ac.uk">n.j.timpson@bristol.ac.uk</a></td>
<td>0000-0002-7141-9189</td>
</tr>
<tr>
<td>Golam M. Khandaker</td>
<td>PhD, FRC Psych</td>
<td><a href="mailto:golam.khandaker@bristol.ac.uk">golam.khandaker@bristol.ac.uk</a></td>
<td>0000-0002-4935-9220</td>
</tr>
</tbody>
</table>

Affiliations:

1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
2 Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
3 Centre for Academic Mental Health, Population Health Sciences, University of Bristol, Bristol, UK

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Wolfson Institute of Population Health, Queen Mary University of London, London, UK
Division of Psychiatry, University of Edinburgh, Edinburgh, UK
Avon and Wiltshire NHS Mental Health Partnership NHS Trust, Bristol, UK
Exeter Centre of Excellence for Diabetes Research, University of Exeter, UK
Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK
Neuroscience and Mental Health Innovation Institute, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK
School of Psychology, University of Leeds, Leeds, UK
NIHR Bristol Biomedical Research Centre, Bristol, UK

* A list of members and their affiliations appear at the end of the paper.

Corresponding author:
Dr Ruby Tsang
Population Health Sciences, Bristol Medical School
Oakfield House, Oakfield Grove
Bristol BS8 2BN
United Kingdom
ruby.tsang@bristol.ac.uk
Abstract

Studies of longitudinal trends of depressive symptoms in young people could provide insight into aetiologic mechanism, heterogeneity and origin of common cardiometabolic comorbidities for depression. Depression is associated with immunological and metabolic alterations, but immunometabolic characteristics of developmental trajectories of depressive symptoms remain unclear. Using depressive symptoms scores measured on 10 occasions between ages 10 and 25 years in the Avon Longitudinal Study of Parents and Children (n=7302), we identified four distinct trajectories: low-stable (70% of the sample), adolescent-limited (13%), adulthood-onset (10%) and adolescent-persistent (7%). We examined associations of these trajectories with: i) anthropometric, cardiometabolic and psychiatric phenotypes using multivariable regression (n=1709-3410); ii) 67 blood immunological proteins and 57 metabolomic features using empirical Bayes moderated linear models (n=2059 and n=2240 respectively); and iii) 28 blood cell counts and biochemical measures using multivariable regression (n=2256). Relative to the low-stable group, risk of depression and anxiety in adulthood was higher for all other groups, especially in the adolescent-persistent (OR_{depression}=22.80, 95% CI 15.25-34.37; OR_{GAD}=19.32, 95% CI 12.86-29.22) and adulthood-onset (OR_{depression}=7.68, 95% CI 5.31-11.17; OR_{GAD}=5.39, 95% CI 3.65-7.94) groups. The three depression-related trajectories vary in their immunometabolic profile, with evidence of little or no alterations in the adolescent-limited group. The adulthood-onset group shows widespread classical immunometabolic changes (e.g., increased immune cell counts and insulin resistance), while the adolescent-persistent group is characterised by higher BMI both in childhood and adulthood with few other immunometabolic changes. These findings point to distinct mechanisms and intervention opportunities for adverse cardiometabolic profile in different groups of young people with depression.
Introduction

The first two decades of life represent a critical epoch for human neurodevelopment when most serious mental illnesses of adult life first emerge. Half of all lifetime cases of common mental disorders including depression and anxiety start by 14 years and 75% by 24 years. The first onset of clinically recognised depressive episodes typically occurs between the ages of 12 and 15 years and the increase in new onset of depression peaks between the ages of 15 and 18 years. Depressive symptoms in childhood and adolescence, including those below diagnostic thresholds, are associated with an elevated risk of depression and other psychiatric diagnoses subsequently in adulthood. These findings highlight the need for studying depressive symptoms during early life.

Characterisation of longitudinal profiles of depressive symptoms during development could help understand the pathogenesis and heterogeneity of later depression, as different individuals may arrive at the same destination via different routes. There is growing evidence to suggest characteristic depression trajectories in childhood and adolescence are differentially associated with risk factors and outcomes. Existing studies have reported associations of a ‘high’ or ‘increasing’ depression trajectory with female sex, lower socioeconomic status, stressful life events, conduct issues, substance use, and parental psychopathology. Trajectories with higher symptom burden have been associated with subsequent depression and other psychiatric diagnoses, lower educational attainment, income and poorer psychosocial adjustment. However, less is known about underlying biological correlates of depression trajectories, including blood-based biomarker signatures. A better understanding of the biological correlates may help uncover mechanistic insights and identify accessible predictive markers for depression.

Existing literature suggests that depression and specific symptoms or symptom dimensions of depression are associated with immunometabolic dysfunction, but there is limited work on immunometabolic correlates of depression trajectories. Depression is associated with immunometabolic alterations such as chronic low-grade inflammation, neuroendocrine dysregulations, as well as less favourable metabolic and lipid profiles. Overall effect sizes
for some of these associations are inconsistent, which could be partly due to clinical or phenotypic heterogeneity within cross-sectional studies.\textsuperscript{12} For instance, immunometabolic alterations appear to be more pronounced or common in individuals endorsing atypical energy-related symptoms of depression (e.g., hyperphagia, weight gain, hypersomnia, or leaden paralysis) as opposed to melancholic symptoms.\textsuperscript{12, 15} At the symptom level, inflammatory markers are particularly associated with somatic and neurovegetative symptoms of depression (e.g., fatigue, altered sleep and appetite) as opposed to psychological symptoms (e.g., hopelessness, excessive/inappropriate guilt).\textsuperscript{16, 17} Some of these findings are supported by Mendelian randomization analyses reporting a potentially causal link between inflammatory markers (e.g., C-reactive protein (CRP) or interleukin 6 (IL-6)) and fatigue, anhedonia, sleep problems, appetite and psychomotor changes.\textsuperscript{18, 19}

The accumulation of risk model for chronic diseases posits that cumulative exposures across the life course result in diverging health trajectories and widening health inequalities as people age.\textsuperscript{20} By characterising depression trajectories, developmental windows when trajectories begin to diverge can be identified, and we can then examine potential factors driving such divergence and biological dysregulations linked to subsequent disease risk. By studying the biomarker signatures of depression trajectories, we may also gain further insight into the origins of higher levels of cardiometabolic multimorbidity in individuals with depression.\textsuperscript{21, 22}

The aims of the current study were threefold: (i) to model depression trajectories from childhood to early adulthood to classify individuals into more homogeneous subgroups, (ii) to examine associations between these subgroups and risk of psychiatric and cardiometabolic outcomes in early adulthood, and (iii) to examine associations of these subgroups with clinical and blood immunometabolic markers including proteomic, metabolomic and biochemical measures in early adulthood. By examining the broader biomarker signature across different domains including the immune proteome, metabolome, and clinical biochemistry, we aim to provide more comprehensive insights into biological pathways and systems possibly involved in the development and persistence of depressive symptoms in young people.
Materials and Methods

Description of cohort

This study uses data from the Avon Longitudinal Study of Parents and Children (ALSPAC). Pregnant women resident in the former county of Avon, United Kingdom (UK) with expected dates of delivery between 1st April 1991 and 31st December 1992 were invited to take part in the ALSPAC study. The initial recruitment enrolled 14541 pregnancies, which resulted in 14062 live births and 13988 infants still alive at 12 months. Further recruitment of eligible participants took place when the oldest children were approximately seven years of age; the total sample size for analyses using any data collected after the age of seven is therefore 15447 pregnancies; of these, 14901 children were alive at 12 months of age.23, 24

The study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool: http://www.bristol.ac.uk/alspac/researchers/our-data/

Data

Sociodemographic and health variables

Sociodemographic characteristics used to characterise the identified depressive symptom trajectories include sex, ethnicity, maternal education, maternal occupational social class, socioeconomic deprivation, and family adversity during pregnancy. Health characteristics examined include smoking, at-risk drinking, carotid intima-media thickness, carotid-femoral pulse wave velocity, metabolic syndrome and its components, obesity, and psychiatric outcomes and medications. Detailed description of these variables as well as those included as covariables in the biomarker analyses are presented in Methods S1.

Depressive symptoms

Self-reported depressive symptoms were assessed using the 13-item Short Mood and Feelings Questionnaire (SMFQ).25 We used data collected on 10 occasions between the ages of 10 and 25 years (ages 10, 12, 13, 16, 17, 18, 21, 22, 23, 25), ending with the last questionnaire
administered in 2017-2018, prior to the start of the COVID-19 pandemic (see Table S1).

Questions were answered based on the two-weeks prior to completing the questionnaire. Each SMFQ item is scored as 0 = “not true”, 1 = “sometimes true” and 2 = “always true”, resulting in a total SMFQ sum score 0-26 (higher score reflects more symptoms). For individuals who had missing data on fewer than three questions, score was imputed to the median value for missing items. For each time-point, those with missing data on more than three questions had their total score recoded as missing.

Circulating blood biomarkers

For this analysis, blood biomarkers were assayed in blood samples collected at the face-to-face research clinic undertaken at 24 years. A total of 92 circulating inflammatory proteins were measured using the Olink Target 96 Inflammation panel (Olink Analysis Service, Uppsala, Sweden); proteins with ≥50% values below the limit of detection (LOD) were excluded leaving 67 proteins to be included (Table S2). Over 220 metabolomic features (148 metabolites and 77 ratios) were quantified using a high-throughput ¹H-NMR spectroscopy-based platform (Nightingale Health, Helsinki, Finland) using a standardised protocol and parameters described elsewhere.²⁶-²⁸ Lipoprotein subclasses were excluded from the analysis to minimise redundancy of information, leaving a subset of 57 metabolomic features (9 cholesterol measures, 12 apolipoproteins and lipids measures, 3 lipoprotein particle sizes, 16 fatty acids and saturation measures, 3 glycolysis-related metabolites, 8 amino acids, 3 ketone bodies, 2 fluid balance-related measures and 1 inflammation-related measure) to be included in the analysis (Table S3).

All 26 blood count and chemistry measures collected at the same clinic interaction were included (Table S4). Additionally, we computed the aspartate aminotransferase/alanine aminotransferase (AST/ALT) ratio and the Homeostatic Model Assessment for Insulin Resistance (HOMA-IR).

Further information on data collection and processing of all blood biomarkers are presented in Methods S1.

Covariables
Prior to statistical analysis, we plotted a directed acyclic graph (DAG) showing theoretical relationships between depression trajectories (independent variable), immunometabolic markers (dependent variables) and important covariables based on the literature (Figure S1). The minimum adjustment set of confounders included in models was sex at birth, maternal education, maternal occupational social class, and body mass index (BMI) at age 10.

**Statistical analysis**

Characterisation of depression trajectories

Latent class trajectory modelling was performed using the \textit{lcm}m R package\textsuperscript{29} to identify subgroups with distinct SMFQ trajectories. This type of modelling seeks to identify homogenous groups of individuals with similar trajectories within a heterogeneous population by combining a latent class model and a mixed model. Models are estimated within the maximum likelihood framework.\textsuperscript{29} The \textit{lcm}m package distinguishes time of measurement and occasion, so individuals with missing data can still be included; we included those with at least three measurements for better modelling of non-linear trajectories. A multi-step approach adapted from the model selection framework suggested by Lennon et al.\textsuperscript{30} and van der Nest et al.\textsuperscript{31} was used, with the order of steps changed to address potential overextraction of latent classes from model under-specification as reported in the simulation literature\textsuperscript{32, 33}. The steps followed:

1. Scope literature and inspect plots to inform polynomial order and potential number of classes. We modelled smfq \textasciitilde age + age\textsuperscript{2} and estimated models up to six latent classes.

2. Estimate growth mixture models (GMM) with random intercepts and class-specific proportional random-effect variance-covariance matrix with increasing number of classes. Select the most appropriate number of classes \textit{k} based on model convergence, model fit (Bayesian information criterion (BIC), Integrated Completed Likelihood (ICL), and relative entropy), smallest class size \geq 5\% and visual inspection of the trajectories.

3. Test alternative model structures with \textit{k} classes – GMM with random intercepts and common random-effect variance-covariance matrix, and group-based trajectory models (GBTM); compare model fit indices, smallest class size and visually inspect trajectories as above, and
assess model adequacy (average posterior probabilities $\geq 0.7$ and odds of correct
classification $\geq 5$ for all classes).

4. Refine trajectory shape by testing up to second-degree fractional polynomials including $(-2, -1, -0.5, 0, 0.5, 1, 2, 3)$ where 0 refers to log X and repeated polynomials refer to $(X^i + X^{i+1} \cdot \log X)$. Select final model based on model convergence and model fit.

Age (in years) was used as the time variable in all models. No covariables were included in these
latent class mixed models as the aim is to describe the trajectories; covariables were accounted
for in the next step when testing for associations with phenotypes of interest and biomarkers. For
each model, an automatic grid search with 50 sets of random initial values and up to 10 iterations
was run to reduce the odds of the model converging towards a local maximum and then up to
500 iterations were allowed for the final estimation. Using the selected model, posterior
probabilities for class membership were then estimated and individuals were assigned to the
class of highest posterior probability in the entire sample using the predictClass function. We first
performed the latent class mixed modelling on the subsample with three or more data points and
then predicted class membership in the entire sample to reduce uncertainty in the modelling
stage and to maximise sample sizes in the subsequent analyses. Additional information on the
modelling is presented in Supplementary Methods S1. R scripts are provided in
Supplementary Methods S2. The reporting of this study adheres to the Guidelines for Reporting
on Latent Trajectory Studies (GRoLTS)$^{34}$ (Table S5).

Associations with clinical and sociodemographic variables
Sociodemographic characteristics are stratified by trajectory class and summarised using mean
(SD), median [interquartile range] or count (%) as appropriate, with differences between
trajectories tested with chi-square or Kruskal-Wallis tests. Associations of trajectory membership
with psychiatric or cardiometabolic outcomes of interest at age 24 or 28 years were tested using
multivariable linear or logistic regressions, using the largest trajectory class as the reference
group and adjusting for sex, maternal education, occupational social class and BMI at age 10.
These variables are described in detailed in Supplementary Methods S1.
Associations with immunometabolic biomarkers

For both proteomic and metabolomic data, associations between depression trajectories and markers were evaluated using multiple linear models fitted in the *limma* R package.\(^{35}\) *Limma* uses an empirical Bayes method to moderate the standard errors of the estimated log-fold changes by borrowing strength from linear models of the other analytes and allowing for different variability between analytes and between samples. Planned contrasts of each of the intermediary trajectories against the trajectory with the most individuals were conducted. With the blood count and clinical chemistry data, linear regressions were fitted with blood markers as dependent variables and SMFQ trajectory class as the independent variable.

For each of these markers, the basic model included sex, maternal education and maternal occupational social class as covariables and the adjusted model further included BMI at age 10 as a covariable. Correction for multiple testing was performed for each set of models using the Benjamini-Hochberg procedure, using a false discovery rate (FDR) q-value threshold of <0.1. This threshold was chosen due to the large number of biomarkers tested, a relatively small sample size and the exploratory nature of this work. R scripts for these analyses are provided in *Supplementary Methods S2.*

Sensitivity analyses

To address potential error carried over from the probabilistic latent class assignment into the association analyses, we performed two sets of sensitivity analyses, the first set by restricting the sample to individuals who had a modal posterior probability \(\geq 0.7\), and the second set by using the individuals’ posterior probabilities for each latent class as separate terms in the models.

Data extraction and initial data cleaning was performed in StataMP version 17.\(^{36}\) Further data preparation and statistical analyses were conducted in R versions 4.1.1 and 4.2.1,\(^ {37}\) using packages *tidyverse* (v2.0.0), *lcmm* (v2.0.2), *LCTMtools* (v0.1.3), *tableone* (version 0.13.2), *knitr*
Latent class trajectory modelling was performed on data from 7302 participants who had SMFQ scores available from at least three time-points between ages 10 and 25 years. Once the best-fitting model was identified, posterior probabilities and class membership were estimated in the entire sample, and 9595 individuals were assigned class membership. Of these 9595 individuals, 2256 had sufficient biomarker and complete covariable data to be included in the biomarker analyses (Figure S2).

Depressive symptom trajectories from childhood to early adulthood

Following comparison of model fit and adequacy statistics and visual inspection of trajectory plots (Table S6 and Figures S2-S3), a four-class group-based trajectory model was identified as best describing the data. As shown in Figure 1, the four identified depressive symptom trajectories from childhood to early adulthood can be described as follows: low-stable – those who consistently had no or low levels of depressive symptoms (69.6%, n=6680), adolescent-limited – those who had elevated depressive symptoms in childhood/adolescence that decreased over time (13.3%, n=1280), adolescent-persistent – those who had elevated depressive symptoms in childhood/adolescence that remained high into adulthood (7.0%, n=672) and adulthood-onset –
those who started with low levels of depressive symptoms that increased in late adolescence/early adulthood (10.0%, n=973).

Characteristics of depression trajectories

Descriptive statistics for characteristics of these individuals, stratified by trajectory, are presented in Table 1 below. There were more women in all three depression-related trajectories: adolescent-limited (66.3%), adolescent-persistent (76.5%), and adulthood-onset trajectories (64.6%). Additionally, the adolescent-persistent trajectory was associated with lower maternal education and greater family adversity during pregnancy. Descriptive statistics for the same characteristics of the subset of individuals who were included in the biomarker analyses are presented in Table S7.
Table 1. Characteristics of depressive symptom trajectories in the ALSPAC birth cohort.

<table>
<thead>
<tr>
<th></th>
<th>Low-stable (n=6680)</th>
<th>Adolescent-limited (n=1280)</th>
<th>Adolescent-persistent (n=672)</th>
<th>Adulthood-onset (n=963)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex: Female</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>3064 (45.9%)</td>
<td>849 (66.3%)</td>
<td>514 (76.5%)</td>
<td>622 (64.6%)</td>
<td></td>
</tr>
<tr>
<td><strong>Ethnicity: Non-white</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>240 (4.1%)</td>
<td>44 (3.9%)</td>
<td>36 (6.2%)</td>
<td>40 (4.8%)</td>
<td></td>
</tr>
<tr>
<td><strong>Maternal education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.026</td>
</tr>
<tr>
<td>CSE or none</td>
<td>894 (15.1%)</td>
<td>159 (13.9%)</td>
<td>107 (17.9%)</td>
<td>120 (14.1%)</td>
<td></td>
</tr>
<tr>
<td>Vocational</td>
<td>560 (9.5%)</td>
<td>91 (8.0%)</td>
<td>53 (8.9%)</td>
<td>64 (7.5%)</td>
<td></td>
</tr>
<tr>
<td>O-level</td>
<td>2064 (34.9%)</td>
<td>406 (35.6%)</td>
<td>231 (38.6%)</td>
<td>293 (34.5%)</td>
<td></td>
</tr>
<tr>
<td>A-level</td>
<td>1483 (25.1%)</td>
<td>308 (27.0%)</td>
<td>141 (23.6%)</td>
<td>230 (27.1%)</td>
<td></td>
</tr>
<tr>
<td>Degree</td>
<td>911 (15.4%)</td>
<td>178 (15.6%)</td>
<td>66 (11.0%)</td>
<td>142 (16.7%)</td>
<td></td>
</tr>
<tr>
<td><strong>Maternal occupational social class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.556</td>
</tr>
<tr>
<td>I – highest</td>
<td>356 (7.0%)</td>
<td>63 (6.6%)</td>
<td>28 (5.5%)</td>
<td>45 (6.3%)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>1714 (33.8%)</td>
<td>335 (35.0%)</td>
<td>161 (31.8%)</td>
<td>250 (34.9%)</td>
<td></td>
</tr>
<tr>
<td>III (non-manual)</td>
<td>2137 (42.2%)</td>
<td>392 (41.0%)</td>
<td>216 (42.7%)</td>
<td>302 (42.1%)</td>
<td></td>
</tr>
<tr>
<td>III (manual)</td>
<td>350 (6.9%)</td>
<td>66 (6.9%)</td>
<td>34 (6.7%)</td>
<td>57 (7.9%)</td>
<td></td>
</tr>
<tr>
<td>IV or V¹ – lowest</td>
<td>509 (10.0%)</td>
<td>100 (10.5%)</td>
<td>67 (13.2%)</td>
<td>63 (8.8%)</td>
<td></td>
</tr>
<tr>
<td><strong>English IMD 2000 quintile</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.723</td>
</tr>
<tr>
<td>1 – least deprived</td>
<td>1665 (32.1%)</td>
<td>355 (35.3%)</td>
<td>146 (30.7%)</td>
<td>242 (33.3%)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1044 (20.1%)</td>
<td>188 (18.7%)</td>
<td>100 (21.0%)</td>
<td>143 (19.7%)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>956 (18.4%)</td>
<td>174 (17.3%)</td>
<td>84 (17.6%)</td>
<td>140 (19.3%)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>790 (15.2%)</td>
<td>153 (15.2%)</td>
<td>77 (16.2%)</td>
<td>115 (15.8%)</td>
<td></td>
</tr>
<tr>
<td>5 – most deprived</td>
<td>739 (14.2%)</td>
<td>137 (13.6%)</td>
<td>69 (14.5%)</td>
<td>86 (11.8%)</td>
<td></td>
</tr>
<tr>
<td><strong>Family Adversity Index</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>1 [0, 2]</td>
<td>1 [0, 2]</td>
<td>1 [0, 3]</td>
<td>1 [0, 2]</td>
<td></td>
</tr>
<tr>
<td>BMI at age 10</td>
<td>17.43 [15.97, 19.73]</td>
<td>17.63 [16.07, 19.96]</td>
<td>18.25 [16.24, 21.00]</td>
<td>17.43 [15.92, 19.60]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Number of SMFQ measurements</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

¹ Categories have been collapsed due to small cell counts in subsequent analyses.

Notes: Numbers presented as mean (SD), median [IQR], or n (%). Percentages are column percentages and computed based on the number of individuals with available data on each variable. Group comparisons were conducted using Chi-square tests for categorical variables and Kruskal-Wallis tests for non-normal continuous variables.
Abbreviations: BMI – body mass index; CSE – Certificate of Secondary Education; O-level – Ordinary level; A-level – Advanced level; IMD – Index of Multiple Deprivation; SMFQ – Short Mood and Feelings Questionnaire
Table 2. Associations of depressive symptom trajectories with anthropometric and cardiometabolic outcomes at 24 years and psychiatric outcomes at 24 and 28 years

<table>
<thead>
<tr>
<th>Outcomes assessed at age 24</th>
<th>Adolescent-limited</th>
<th>Adolescent-persistent</th>
<th>Adulthood-onset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Adjusted unstandardised regression coefficient (SE)</td>
<td>Adjusted unstandardised regression coefficient (SE)</td>
</tr>
<tr>
<td>cIMT (continuous)</td>
<td>1565</td>
<td>-0.0002 (0.0033)</td>
<td>0.0099 (0.0050)</td>
</tr>
<tr>
<td>cfPWV (continuous)</td>
<td>1708</td>
<td>-0.0578 (0.0708)</td>
<td>0.1429 (0.1109)</td>
</tr>
<tr>
<td>Smoking</td>
<td>2731</td>
<td>1.66 (1.32-2.09)</td>
<td>2.66 (1.93-3.67)</td>
</tr>
<tr>
<td>AUDIT-C score ≥5</td>
<td>2704</td>
<td>0.97 (0.79-1.21)</td>
<td>0.65 (0.48-0.89)</td>
</tr>
<tr>
<td>BMI ≥30kg/m²</td>
<td>2731</td>
<td>1.41 (0.99-2.01)</td>
<td>1.93 (1.19-3.06)</td>
</tr>
<tr>
<td>Elevated waist circumference</td>
<td>2707</td>
<td>1.21 (0.93-1.56)</td>
<td>1.24 (0.84-1.80)</td>
</tr>
<tr>
<td>Triglycerides ≥1.7mmol/L</td>
<td>2241</td>
<td>1.37 (0.86-2.11)</td>
<td>1.14 (0.53-2.23)</td>
</tr>
<tr>
<td>HDL &lt;1.0mmol/L</td>
<td>2241</td>
<td>0.86 (0.61-1.20)</td>
<td>1.13 (0.71-1.77)</td>
</tr>
<tr>
<td>SBP ≥130mmHg</td>
<td>2751</td>
<td>0.85 (0.58-1.22)</td>
<td>0.60 (0.29-1.12)</td>
</tr>
<tr>
<td>DBP ≥85mmHg</td>
<td>2751</td>
<td>1.32 (0.69-2.69)</td>
<td>0.73 (0.17-2.21)</td>
</tr>
<tr>
<td>Fasting glucose ≥5.6mmol/L</td>
<td>2241</td>
<td>0.86 (0.65-1.14)</td>
<td>0.92 (0.58-1.42)</td>
</tr>
<tr>
<td>Metabolic syndrome</td>
<td>2759</td>
<td>0.97 (0.59-1.55)</td>
<td>0.78 (0.35-1.54)</td>
</tr>
<tr>
<td>ICD-10 depressive episode</td>
<td>2729</td>
<td>4.03 (2.73-5.95)</td>
<td>22.80 (15.25-34.37)</td>
</tr>
<tr>
<td>ICD-10 GAD</td>
<td>2719</td>
<td>3.59 (2.41-5.33)</td>
<td>19.32 (12.86-29.22)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes assessed at age 28</th>
<th>Prescribed antidepressants in past 5y</th>
<th>Prescribed anxiolytics in past 5y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2792</td>
<td>2828</td>
</tr>
<tr>
<td></td>
<td>3.02 (2.28-3.99)</td>
<td>2.31 (1.25-4.16)</td>
</tr>
<tr>
<td></td>
<td>7.29 (5.22-10.17)</td>
<td>6.76 (3.74-12.05)</td>
</tr>
<tr>
<td></td>
<td>5.12 (3.88-6.75)</td>
<td>4.16 (2.39-7.18)</td>
</tr>
</tbody>
</table>

Notes: Regression models were adjusted for sex, maternal education, maternal occupational social class, and BMI at age 10. Effect estimates presented are unstandardised regression coefficients for continuous outcomes and odds ratios (95% confidence intervals) for binary outcomes. The reference group for all analyses is the low-stable trajectory. Text in bold indicates evidence for the association after FDR correction of p-values.

Metabolic syndrome was defined based on the 2009 consensus definition from the International Diabetes Federation and the American Heart Association/National Heart, Lung, and Blood Institute, i.e. the presence of any three of the following five risk factors: elevated waist circumference (≥94cm for white men, ≥90cm for non-white men, ≥80cm for women); elevated triglycerides (≥1.7mmol/L), reduced high-density lipoprotein-cholesterol.
(<1.0mmol/L), elevated blood pressure (systolic ≥130mmHg or diastolic ≥85mmHg), and elevated fasting glucose (≥5.6mmol/L).

Abbreviations: AUDIT-C – Alcohol Use Disorders Identification Test for Consumption; BMI – body mass index; cfPWV – carotid-femoral pulse wave velocity; cIMT – carotid intima-media thickness; DBP – diastolic blood pressure; GAD – generalised anxiety disorder; HDL – high-density lipoprotein; ICD-10 – International Classification of Diseases 10th Revision; SBP – systolic blood pressure.
Adulthood cardiometabolic and psychiatric outcomes associated with depression trajectories

Compared to the low-stable trajectory, after adjusting for sex, maternal education, maternal occupational class and BMI at age 10, all three depression-related trajectories were associated with ICD-10 diagnosis of depression, generalised anxiety disorder, being prescribed an antidepressant or anxiolytic at ages 24 and 28 years. However, the magnitude of association varied between the trajectories, with the risk for these outcomes being the highest for the adolescent-persistent trajectory (approximately 20-fold risk), followed by the adulthood-onset and adolescent-limited trajectories. The adolescent-persistent trajectory was additionally associated with obesity (Table 2). Unadjusted model results are presented in Table S8.

Differentially abundant proteins associated with depression trajectories

Relative to the low-stable trajectory, after adjusting for sex, maternal education, maternal occupational class and BMI at age 10, one protein (C-C motif chemokine 25 [CCL25]) was upregulated in the adolescent-limited trajectory; four proteins (fibroblast growth factor 21 [FGF-21], hepatocyte growth factor [HGF], eukaryotic translation initiation factor [4E-BP1], and eotaxin-1 [CCL11]) were upregulated in the adolescent-persistent trajectory; and five proteins (FGF-21, fibroblast growth factor 19 [FGF-19], CUB domain-containing protein 1 [CDCP1], HGF, CCL11) were upregulated in the adulthood-onset trajectory (Figure 2). Full model results are presented in Tables S9-10.

Differentially abundant metabolites associated with depressive symptom trajectories

Relative to the low-stable trajectory, after adjusting for sex, maternal education and occupational class, and BMI at age 10, creatinine was decreased in the adolescent-persistent trajectory. Three metabolite ratios (omega-3 to total fatty acids [omega-3/FA], docosahexaenoic acid to total fatty acids [DHA/FA], polyunsaturated fatty acids to total fatty acids [PUFA/FA] ratios) were decreased whereas the monosaturated to total fatty acids (MUFA/FA) and apolipoprotein B to apolipoprotein A1 (ApoB/ApoA1) ratios were increased in the adult-onset trajectory (Figure 3). Full model results are presented in Tables S11-12.
Blood count and clinical chemistry markers

Relative to the low-stable trajectory, after adjusting for sex, maternal education and occupational class, and BMI at age 10, there was evidence for decreased AST levels in the adolescent-limited trajectory, and increased HOMA-IR, insulin, neutrophil and white blood cell (WBC) counts in the adulthood-onset trajectory (Figure 5). Full model results are presented in Tables S13-14.

Sensitivity analyses

The sensitivity analyses showed patterns of associations that are largely similar to those observed in the primary analyses, with consistent associations of depression-related trajectories with anthropometric, cardiometabolic and psychiatric outcomes, and top-ranking immunometabolic biomarkers with similar effect sizes in the same directions (Supplementary Methods S3).

Discussion

Depression is a complex heterogeneous disorder, which poses a challenge for discovering biomarkers associated with disease onset and/or progression. We have taken a longitudinal approach to identifying blood-based biomarkers for depression by examining longitudinal patterns of depressive symptoms in the population during the critical developmental epoch of childhood, adolescence and early adulthood. Using data from a prospective birth cohort, we have identified four longitudinal population subgroups based on repeated measures of depressive symptoms over a 15-year period from ages 10 to 25 years. We show that majority of participants (approximately 70%) have little or no depressive symptoms (low-stable group). We identified three depression-related groups which comprise a group with higher symptom levels during childhood and adolescence which later decrease (adolescent limited group, 13%), a group with symptoms emerging during puberty which persist throughout adolescence through to adulthood (adolescent persistent group, 7%), and a group with symptoms emerging during late adolescence/early adulthood and increasing thereafter (adulthood onset, 10%).
We have examined health phenotypes and blood biomarkers associated with these subgroups for greater insight into the developmental course of depression and associated biomarkers. Our analyses show that compared to the low-stable group, risk of depression and anxiety in adulthood is higher for all three depression-related groups. However, such risk is particularly elevated risk for the adolescent-persistent group (19 to 23-fold risk) followed by the adulthood-onset group (five to eight-fold risk). Interestingly, the group where higher levels of symptoms are mostly limited to adolescence, they still have a three-fold increased risk of depression in adulthood.

Having examined health phenotypic and blood proteomic, immunological and metabolic biomarker associations for these three groups, we show that the adolescent-limited group is distinct from the other two depression-related groups as it showed little immunometabolic alterations. In contrast, both the adolescent-persistent and adulthood-onset groups are associated with immunometabolic changes, but the exact pattern of associations varies between the two groups. The adolescent-persistent group was associated with higher BMI in childhood and adulthood, whereas the adulthood-onset group did not show this, but rather more widespread alterations in blood-based metabolic parameters including insulin resistance, insulin levels and changes in fatty acid ratios. Blood proteomic changes were largely similar between the two groups and involved proteins that mainly act as growth factors, cytokines and chemokines. While immunometabolic associations persisted given adjustment for childhood BMI in the adolescent-persistent group, the presence or absence of this BMI adjustment had an impact on some association estimates in the adolescent-persistent and adulthood-onset groups. While this may relate to the apparent differences in BMI by trajectories, it is difficult to distinguish this as an artefact of adjustment or a true impact of BMI.

Epidemiological studies consistently report a bidirectional relationship (both cross-sectionally and longitudinally) between obesity and depression, whereas Mendelian randomization studies support a causal role of BMI on major depressive disorder and depressive symptoms but not vice
versa. The comorbidity between obesity and depression is generally associated with poorer prognosis, with studies reporting associations with a more chronic course of depression in adulthood as well as poorer treatment response. Our findings add to this evidence by showing an association between higher childhood BMI and persistent depressive symptoms between adolescence and early adulthood.

Many of the alterations observed in the adulthood-onset trajectory are already well studied markers of cardiometabolic disease risk. The ApoB/ApoA1 ratio is associated with cardiovascular diseases and metabolic syndrome, and can be used to predict longer-term cardiovascular risk when measured in early life. Higher values of HOMA-IR are associated with an increased risk of developing type 2 diabetes mellitus (T2DM), systemic arterial hypertension and non-fatal major adverse cardiovascular events. Lower blood omega-3 fatty acid concentrations are associated with poorer cardiovascular outcomes and may also contribute to chronic systemic inflammation, whereas changes in MUFA, PUFA and DHA concentrations in early adulthood were associated with incident obesity, insulin resistance and elevated blood pressure 10 years later. Since dietary intake and supplementation are the main predictors of blood levels of omega-3 fatty acids, with lifestyle factors such as BMI, smoking and alcohol consumption also playing a role, this suggests that the adulthood-onset group may benefit from lifestyle intervention to prevent future cardiometabolic disease.

The overlapping proteomic signals observed between the adolescent-persistent and adulthood-onset trajectories potentially suggest shared underlying mechanisms (genetic or environmental) or shared biological responses to depression, which warrant further study for their roles in the pathophysiology of depression and cardiometabolic disease. FGF-21 is a novel regulator of glucose and lipid metabolism that mainly acts through an FGF receptor 1 (FGFR1)/β-klotho receptor complex and the Ras/Raf MAPK signalling pathway, which have been implicated in the pathophysiology of depression and therapeutic effects of antidepressants. Elevated circulating FGF-21 concentrations have also been shown to be associated with a range of cardiometabolic markers and diseases. HGF mediates inflammatory responses to tissue
injury and regulates cell growth and morphogenesis through the activation of the HGF/mesenchymal-epithelial transition factor (c-Met) signalling pathway, which has downstream effects on the Raf/MAPK and PI3K/Akt pathways. Altered HGF/c-Met signalling has been suggested to play a role in the pathogenesis of depression in adolescents through disrupting interneuron development. CCL11 is a chemokine involved in the selective recruitment of eosinophils into sites of inflammation and has been implicated in various allergic and inflammatory conditions. It can be transported across the blood-brain barrier and is an age-related systemic factor associated with reduced synaptic plasticity and impaired hippocampal-dependent learning and memory in mice. In humans, CCL11 levels increase with age, and there is emerging evidence to suggest that CCL11 levels are associated with psychiatric disorders. In summary, the overlapping proteomic signals between the adolescent-persistent and adulthood-onset trajectories highlight potential roles of physiological stress from lifestyle or environmental factors, disruptions in neurodevelopment, and neurogenesis and cellular senescence in the underlying vulnerability or biological response to depression, and may be key biomarkers relevant to illness pathogenesis.

An advantage of our work is that by examining depressive symptoms longitudinally using a latent class trajectory approach, we can account for population heterogeneity and obtain better characterisation of subgroups and their changes over time. Existing literature shows that depression is associated with alterations in various immunometabolic biomarkers, including increased inflammatory cytokines, WBC, neutrophils, T-lymphocytes and other immune cell counts, HOMA-IR, insulin, lipids and fatty acids. Using longitudinal data from young people, we add to this evidence base by showing that classical immunometabolic changes are particularly associated with an adulthood-onset trajectory, rather than other developmental subgroups of depressive symptoms, including one with persistent symptoms since adolescence.

This study has several limitations. Firstly, the approach of treating assigned class membership as discrete in assessing relationships with other variables has been shown to underestimate the strength of the relationships. However, as we are mainly interested in subpopulations with
different depressive symptom trajectories, this approach allows for easier interpretation and translation. Secondly, depression is episodic in nature and the use of polynomials cannot fully capture the dynamics of depressive symptom severity over time; however, this approach was chosen over other methods (e.g., splines) for model parsimony.

Furthermore, the sample size in this study is relatively small and may be underpowered to detect differences after stratifying individuals into four separate trajectories. For this reason, we did not further stratify our analyses by sex or potentially relevant variables (e.g., BMI). While we have adjusted our biomarker analyses for several potential confounders, residual confounding could still be an issue. For instance, we have not accounted for medication use or chronic disease etc., but these are likely to be uncommon in young people. As the biomarkers were measured at age 24, which is after the onset of depressive symptoms in many individuals, further research is required to assess the direction and causality of associations we have identified.

In conclusion, we identified distinct developmental trajectories of depression from childhood to early adulthood, which show differential associations with cardiometabolic and psychiatric outcomes, and are characterised by distinct immunometabolic profiles. In particular, individuals with persistent depressive symptoms from childhood through to early adulthood were more likely to have higher BMI both in childhood and in adulthood and few other immunometabolic changes, whereas individuals who develop depressive symptoms towards early adulthood show classical immunometabolic alterations in immune cell counts, insulin resistance and fatty acid profiles. These findings point to distinct mechanisms and intervention opportunities for different groups of young people with depressive symptoms.
Acknowledgements

We are extremely grateful to all the families who took part in this 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. We would also like to thank members of the LINC study public advisory group for their contribution.

This work was funded by the Tackling Multimorbidity at Scale Strategic Priorities Fund programme (MR/W014416/1) delivered by the UK Medical Research Council (MRC) and the UK National Institute for Health Research (NIHR) in partnership with the UK Economic and Social Research Council and in collaboration with the UK Engineering and Physical Sciences Research Council. RSMT, DS and MLW are supported by this grant. ASFK is supported by a Wellcome Early Career Award (227063/Z/23/Z). NAD was supported by an NIHR Clinical Lectureship in General Adult Psychiatry. MBMvdB acknowledges additional funding support from the MRC (MR/W028395/1, MR/W020297/1, MR/T033045/1 and MR/S037667/1), NIMH (U01MH119758), the Wellcome Trust (226709/Z/22/Z) and Welsh Government (HS 22 04). NJT is a director of the MRC/ESRC/UKRI supported Population Research UK Coordination Hub (ES/Y008340/1), supported by a Wellcome Trust Investigator award (202802/Z/16/Z), is the PI of the Avon Longitudinal Study of Parents and Children (MRC & WT 217065/Z/19/Z), is supported by the University of Bristol NIHR Biomedical Research Centre (BRC-1215-2001), the MRC Integrative Epidemiology Unit (MC_UU_00011/1), CRUK and works within the CRUK Integrative Cancer Epidemiology Programme (C18281/A29019) and with support from CRUK (PRCPJT-May22\100028). GMK acknowledges funding support from the MRC (MC_UU_00032/06), which forms part of the MRC Integrative Epidemiology Unit at the University of Bristol. This grant also supports HF. 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 NIHR Bristol Biomedical Research Centre (NIHR 203315). The views expressed are those of the authors and not necessarily those of the UK NIHR or the Department of Health and Social Care.
The UK MRC and Wellcome (217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and RSMT, NJT and GMK will serve as guarantors for the contents of this paper. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf).

LINC Consortium members:

Marianne B. M. van den Bree, George Kirov, Michael J. Owen, James T. R. Walters, Peter A. Holmans, Jane Lynch, Ioanna K. Katzourou, Nabila Ali, Lowri O’Donovan (Cardiff University, UK)
David A. van Heel, Sarah Finer, Daniel Stow (Queen Mary University of London, UK)
Golam M. Khandaker, Nicholas J. Timpson, John A. A. Macleod, Julie P. Clayton, Ruby S. M. Tsang, Jane Sprackman, Shahid Khan (University of Bristol, UK)
Inês A. Barroso, Rupert A. Payne (University of Exeter, UK)
Mark Mon-Williams, Megan L. Wood (University of Leeds, UK)
Hilary C. Martin (Wellcome Sanger Institute, UK)
Thomas Werge, Andrés Ingason, Morteza Vaez, Lam O. Huang (Institute of Biological Psychiatry, Denmark)

Conflict of Interest

None
References


Figure 1. Predicted Marginal Mean Depressive Symptom Trajectories from Childhood to Early Adulthood in the ALSPAC Cohort

Lines showing predicted marginal mean depressive symptom trajectories with shaded areas representing 95% confidence intervals.

Figure 2. Volcano Plots Showing Differential Immune Protein Abundance Levels in Depressive Symptom Trajectories

Panels a and b show results from basic and adjusted models respectively. The reference group for all analyses is the low-stable trajectory.

Abbreviations: 4E-BP1 – eukaryotic translation initiation factor; CCL11 – eotaxin-1; CCL25 – C-C motif chemokine 25; CDCP1 – CUB domain-containing protein 1; FGF-19 – fibroblast growth factor 19; FGF-21 – fibroblast growth factor 21; HGF – hepatic growth factor

Figure 3. Volcano Plots Showing Differential Metabolite Abundance Levels in Depressive Symptom Trajectories

Panels a and b show results from basic and adjusted models respectively. The reference group for all analyses is the low-stable trajectory. Orange points indicate upregulation and blue points indicate downregulation.

Abbreviations: ApoB/ApoA1 – apolipoprotein B to apolipoprotein A1 ratio, DHA/FA – docosahexaenoic acid to total fatty acids ratio, MUFA/FA – monounsaturated fatty acids to total fatty acids ratio, Omega-3/FA – omega-3 fatty acids to total fatty acids ratio, PUFA/FA – polyunsaturated fatty acids to total fatty acids ratio

Figure 4. Differential Levels of Full Blood Count and Clinical Biochemistry Biomarkers in Different Depressive Symptom Trajectories

Dot-and-whisker plots showing effect estimates and 95% confidence intervals for each biomarker. The effect estimates and 95% confidence intervals from the models have been back-transformed.
into their original scale for ease of interpretation; effect estimates represent the percentage
difference in mean levels of each biomarker for respective trajectory, in relation to the low-stable
trajectory (reference group). Panels a and b show results from basic and adjusted models
respectively. Asterisks indicate evidence for the association after FDR correction of p-values.
Abbreviations: WBC – white blood count, RBC – red blood count, Hb – haemoglobin, Hct –
haematocrit, MCV – mean cell volume, MCH – mean cell haemoglobin, MCHC – mean
corpuscular haemoglobin concentration, HOMA-IR – Homeostatic Model Assessment for Insulin
Resistance, HDL – high-density lipoprotein, LDL – low-density lipoprotein, VLDL – very low-
density lipoprotein, CRP – C-reactive protein, GGT – gamma-glutamyl transpeptidase, ALT –
alanine aminotransferase, AST – aspartate aminotransferase, P3NP – procollagen-3 N-terminal
peptide