Brain, Behavior, and Immunity 134 (2026) 106464

Contents lists available at ScienceDirect

Brain Behavior and Immunity

journal homepage: www.elsevier.com/locate/ybrbi

From non-specific biomarker to targeted action: transdiagnostic and sex-specific drivers of high-CRP status in severe mental illness across the FondaMental Advanced Centers of Expertise (FACE) cohorts

Tim Rietberg [a,b], Susana Barbosa [c], Pieter Meysman [d,e], Bruno Aouizerate [f,g], Frank Bellivier [f,h], Fabrice Berna [f,i], Delphine Capdevielle [f,j], Isabelle Chereau [f,k], Philippe Courtet [f,l], Clément Dondé [f,m,n,o,p], Caroline Dubertret [f,q,r], Bruno Etain [f,h], Wissam El Hage [f,s], Mona Gerentes [f,t], Jérôme Holtzmann [f,p], Eric Fakra [f,u,ab], Emmanuel Haffen [f,v], Noomane Bouaziz [f,w,x,ac], Antoine Lefrere [f,y], Pierre-Michel Llorca [f,k], Jasmina Mallet [f,z], David Misdrahi [f,aa,ab], Emilie Olié [f,l], Agnès Pelletier [f,t], Giulia Piva [f,ac], Baptiste Pignon [f,t,ad], Mircea Polosan [f,ae], Romain Rey [f,ab], Paul Roux [f,af,ag], Ludovic Samalin [f,k], Franck Schürhoff [f,t,ad], Raymund Schwan [f,ah], Florian Stephan [f,ai], David Szekely [f,aj], Mathieu Urbach [f,af,ag], Guillaume Vaiva [f,ak], Etienne Very [f,al,am], Antoine Yrondi [f,an,ao], Marion Leboyer [f,t,ad], Livia De Picker [a,b,*,1], Ophélia Godin [f,t,ad,1]

[a] Collaborative Antwerp Psychiatric Research Institute (CAPRI), University of Antwerp, Antwerp, Belgium
[b] Scientific Initiative of Neuropsychiatric and Psychopharmacological Studies (SINAPS), University Psychiatric Centre Campus Duffel, Duffel, Belgium
[c] Sorbonne Université, Inserm, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP), Epidémiologie Sociale, Santé Mentale et Addictions (ESSMA), 75012 Paris, France
[d] Antwerp Unit for Data Analysis and Computation in Immunology and Sequencing (AUDACIS), University of Antwerp, Antwerp, Belgium
[e] ADReM Data Lab, Department of Mathematics and Computer Science, University of Antwerp, Antwerp, Belgium
[f] Fondation FondaMental, Créteil, France
[g] Pôle de Psychiatrie Générale et Universitaire, Centre Hospitalier Charles Perrens, Laboratoire de Nutrition et Neurobiologie intégrée, NutriNeuro (UMR INRAe 1286), Université de Bordeaux, Bordeaux, France
[h] Université Paris Cité, INSERM, Optimisation Thérapeutique en Neuropharmacologie, OPTEN U1144, Département de Psychiatrie et de Médecine Addictologique, Hôpitaux Lariboisière-Fernand Widal, GHU APHP.Nord - Université Paris Cité, Paris F-75010, France
[i] Hôpitaux Universitaires de Strasbourg, Université de Strasbourg, INSERM U1329, Fédération de Médecine Translationnelle de Strasbourg, Strasbourg, France
[j] Universitary Department of Adult Psychiatrie, CHU Montpellier, IGF, Univ. Montpellier, CNRS, INSERM, Montpellier, France
[k] Department of Psychiatry, CHU Clermont-Ferrand, University of Clermont Auvergne, CNRS, Clermont Auvergne INP, Institut Pascal (UMR 6602), Clermont-Ferrand, France
[l] Department of Emergency Psychiatry and Acute Care, Lapeyronie Hospital, Centre Hospitalier Universitaire de Montpellier, Institute of Functional Genomics, University of Montpellier, National Centre for Scientific Research, French National Institute of Health and Medical Research, Montpellier, France
[m] Univ. Grenoble Alpes, F-38000 Grenoble, France
[n] Inserm U1216, F-38000 Grenoble, France
[o] Psychiatry Department, CH Alpes Isere, F-38000 Grenoble, France
[p] Psychiatry Department, CHU Grenoble F-38000 Grenoble, France
[q] AP-HP, Groupe Hospitalo-universitaire AP-HP Nord, DMU ESPRIT, Service de Psychiatrie et Addictologie, Hôpital Louis Mourier, Colombes, France
[r] Université de Paris Inserm UMR1266, Sorbonne Paris Cité, Faculté de Médecine, France
[s] Université de Tours, INSERM, Imaging Brain & Neuropsychiatry iBraiN U1253, 37032 Tours, France
[t] Assistance Publique Hôpitaux de Paris, Hôpitaux Universitaires Henri Mondor, Département Médico-Universitaire de Psychiatrie et d'Addictologie, France
[u] Department of Psychiatry, University Hospital of Saint-Etienne, Saint-Etienne, France
[v] Université Marie et Louis Pasteur, UMR INSERM LINC 1322, Besançon, Service de Psychiatrie de l'Adulte, CIC-1431 INSERM, CHU de Besançon, France
[w] Université Sorbonne Paris Nord 93017, Bobigny Cedex, France

Abbreviations: AUROC, area under the receiver operating characteristic curve; BD, bipolar disorder; BMI, body mass index; CRP, C-reactive protein; CTQ, Childhood Trauma Questionnaire; DBP, diastolic blood pressure; FACE, FondaMental Academic Centres of Expertise; HDL, high-density lipoprotein; CRP, high-sensitivity C-reactive protein; KNN, k-nearest neighbours; LDL, low-density lipoprotein; MDD, major depressive disorder; OR, odds ratio; PCA, principal component analysis; PLR, penalised logistic regression; PSQI, Pittsburgh Sleep Quality Index; RF, random forest; SBP, systolic blood pressure; SMI, severe mental illness; SZ, schizophrenia; UMAP, uniform manifold approximation and projection; VIP, variable inclusion probability.

* Corresponding author at: Collaborative Antwerp Psychiatric Research Institute (CAPRI), University of Antwerp, Antwerp, Belgium.
E-mail address: livia.depicker@uantwerpen.be (L.D. Picker).

https://doi.org/10.1016/j.bbi.2026.106464
Received 27 October 2025; Received in revised form 5 January 2026; Accepted 25 January 2026
Available online 30 January 2026

[x] Département de recherche clinique, Pole universitaire93G03, EPS Ville Evrard Neuilly sur Marne, 93330, France
[y] Pôle de Psychiatrie, Assistance Publique Hôpitaux de Marseille, Marseille, France
[z] Service de Psychiatrie, Centre Hospitalier Universitaire d'Orléans, EPSM du Loiret, B-CLINE, Laboratoire Interdisciplinaire pour l'Innovation et la Recherche en Santé d'Orléans (LI2RSO), Université d'Orléans, Orléans, France
[aa] University of Bordeaux, Aquitaine Institute for Cognitive and Integrative Neuroscience (CNRS UMR 5287-INCIA), Bordeaux, France
[ab] Department of Adult Psychiatry, Charles Perrens Hospital, Bordeaux, France
[ac] INSERM U1028, CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon, Université Claude Bernard Lyon 1, Equipe PSYR2, Centre Hospitalier Le Vinatier, Pole Est, 95 bd Pinel, BP 30039, 69678 Bron Cedex, France
[ad] Université Paris Est Créteil (UPEC), INSERM, Institut Henri Mondor de Recherche Biomédicale, Laboratoire Neuro-Psychiatrie Translationnelle, Créteil, France
[ae] Université Grenoble Alpes, Inserm, U1216, CHU Grenoble Alpes, Grenoble Institut Neurosciences, Grenoble, France
[af] Versailles Hospital, Department of Adult Psychiatry and Addictology, Centre Hospitalier de Versailles, 177 rue de Versailles, 78157 Le Chesnay, France
[ag] MOODS Team, INSERM UMR1018, CESP, Faculté de Médecine Paris-Saclay, Université de Versailles Saint-Quentin-En-Yvelines, Université Paris-Saclay, Villejuif, France
[ah] Université de Lorraine, Centre Psychothérapique de Nancy, Inserm U1254, Nancy, France
[ai] Service Hospitalo-Universitaire de Psychiatrie Générale et de Réhabilitation Psycho Sociale, 29G01 Et 29G02, Fondation FondaMental, EA 7479 CHU Brest, France
[aj] Service de Psychiatrie, Centre Hospitalier Princesse Grace, 9800, Monaco
[ak] Inserm-U1172–LilNCog–Lille Neuroscience & Cognition, Centre National de Ressources & Résilience pour les Psychotraumatismes (Cn2r Lille Paris), Université de Lille (CHU Lille), 59000 Lille, France
[al] Department of Psychiatry, Psychotherapies, Art-therapy, Toulouse University Hospital, France
[am] Equipe de Recherche sur les Rationalités Philosophiques et les Savoirs - EA3051, Université de Toulouse - Jean Jaurès, Toulouse, France
[an] Service de Psychiatrie et de Psychologie Médicale, Centre Expert Dépression Résistante FondaMental, CHU de Toulouse, Hôpital Purpan, ToNIC Toulouse, NeuroImaging Centre, Toulouse, France
[ao] Université de Toulouse, INSERM, UPS, Toulouse, France

ARTICLE INFO

Keywords:
C-reactive protein
Immunopsychiatry
Stratification
Machine learning
Data-driven modelling
Precision Psychiatry

ABSTRACT

Background and objectives: Low-grade systemic inflammation contributes to the pathophysiology of severe mental illness (SMI) in a substantial subset of patients, who often experience greater disease burden and poorer treatment response. Elevated C-reactive protein (CRP), defined as CRP $\geq$ 3 mg/L, has been proposed to identify this group, but its non-specificity limits the biomarker's ability to guide targeted intervention. We aimed to determine the most consistent drivers of high CRP across bipolar disorder (BD), schizophrenia (SZ) and major depressive disorder (MDD), and to translate these into clinically actionable intervention targets using robust data-driven methods.

Methods: We pooled and harmonised data from three large French national SMI cohorts (n = 7149: 4797 bipolar disorder, 1958 schizophrenia and 394 resistant major depression) and classified participants by CRP $\geq$ 3 mg/L, as well as an alternative cut-off of 5 mg/L. We applied penalised logistic regression (PLR), random forests (RF) and unsupervised clustering, using 28 biopsychosocial variables to identify robust drivers of high-CRP status. We then grouped these into actionable targets and assessed relative dominance.

Results: In total, 30.16% of participants had CRP $\geq$ 3 mg/L. PLR identified female sex (OR [95% CI]: 1.60 [1.27, 1.93]), higher BMI (OR: 1.09 [1.07, 1.13]), current nicotine dependence (OR: 1.05 [1.02, 1.09]), lower HDL cholesterol (OR: 0.57 [0.44, 0.73]) and smoking (ex-smoker status OR: 0.84 [0.66, 0.98]) as consistent drivers. RF highlighted a similar set of key drivers, also including waist circumference, triglycerides and cardiovascular comorbidities. Clustering of the high-CRP group was almost entirely driven by smoking status and nicotine dependence. When grouped into actionable targets, the identified drivers accounted for 16% of variance in CRP status, with obesity emerging as most dominant contributor. This pattern was most pronounced in females; in males it was more diffuse, with a more prominent role for smoking.

Conclusions: We propose a decision tree framework where CRP can serve as a first-line screening marker for inflammation in SMI, with subsequent steps focusing on the main contributing factors to guide targeted interventions. Priority should be given to targeting obesity and metabolic dysregulation. Among females, hyperuricemia represents the next most appropriate target, whereas in males, smoking warrants greater attention. This stepwise approach provides a route from a non-specific biomarker to targeted treatment strategies and should be validated in prospective studies.

1. Introduction

Low-grade systemic inflammation has emerged as an important biological process implicated in the pathophysiology of all types of severe mental illness (SMI), including major depressive disorder (MDD) (Jha et al., 2025; Milaneschi et al., 2021; Miller and Raison, 2016), bipolar disorder (BD) (Dargél et al., 2015; Jones et al., 2021; Pereira et al., 2021), and schizophrenia (SZ) (Upthegrove and Khandaker, 2020; Müller, 2018). Across diagnostic groups, it has been associated with more severe symptom profiles, poorer treatment response, and increased physical comorbidity in a substantial subset of psychiatric patients (e.g. Goldsmith et al., 2016; Khandaker et al., 2017; Leboyer et al., 2021; Orsolini et al., 2023; Strawbridge et al., 2015). This subgroup has therefore been proposed as a target for adjunctive, anti-inflammatory interventions (Miller et al., 2025).

Among biomarkers that aid in the identification of this group, C-reactive protein (CRP) has received particular attention due to its reliability, availability, and reproducibility in clinical settings (Baysak et al. 2022; Jha et al. 2025; Penninx et al., 2025; Orsolini et al., 2023). However, CRP is a non-specific biomarker and elevations can reflect a wide range of underlying factors. Its clinical utility could therefore be greatly improved by moving beyond CRP alone to identify the pathways and drivers that contribute to its elevation.

Previous studies have often focused on hypothesis-driven individual drivers, such as obesity, smoking, or childhood adversity, or have restricted their analyses to single diagnostic groups (e.g. Berk et al., 2013; Danese et al., 2007; Hu et al., 2025; Jiang et al., 2025; Ninla-Aesong et al., 2023; Saitoh et al., 2025; Van den Noortgate et al., 2025), limiting the generalisability and clinical utility of their findings.

[1] Both authors contributed equally.

In contrast, a comprehensive, transdiagnostic and data-driven approach could more representatively capture the multifactorial nature of inflammation in SMI. Moreover, identifying clinically meaningful and actionable targets that most dominantly contribute to CRP elevations could facilitate the development of decision tools to guide intervention strategies.

In this study, we use a large, transdiagnostic psychiatric sample of French patients with SMI to identify the most consistent drivers of elevated CRP through a comprehensive, data-driven approach. These drivers are then grouped into clinically actionable targets to assess their relative dominance. We further explore how these patterns differ by sex and smoking status. Our aim is to advance stratified intervention strategies by mapping the heterogeneous pathways that lead to inflammation in psychiatric SMI populations.

2. Methods

2.1. Study design and participants

This cross-sectional study utilised data collected by the French networks of Expert centres, created and coordinated by the French non-profit foundation FondaMental. The FondaMental Academic Centres of Expertise (FACE) networks comprise FACE-BD (n = 4797 with bipolar disorder), FACE-SZ (n = 1958 with schizophrenia or schizoaffective disorder) and FACE-DR (n = 394 with resistant major depressive disorder), which collect standardised longitudinal data. Only data from each participant's initial evaluation visit were included.

Clinically stable outpatients aged ≥ 16 years with a DSM-IV-TR diagnosis of schizophrenia, schizoaffective disorder, BD, or MDD were included between 2009 and 2023. Stability was defined as no hospitalisation or major treatment modifications within the four weeks prior to assessment. Standardised clinical, neuropsychological, and biological assessments were conducted at inclusion by trained personnel. Detailed assessment procedures are described elsewhere (Leboyer et al., 2022; Schürhoff et al., 2019; Yrondi et al., 2017).

All participants provided a signed non-opposition form. The protocols were approved by the relevant ethics committee (CPP-Ile de France IX, CPP-Ile de France IX, January 18, 2010) and data protection authority (CNIL; DR-2012–157).

2.2. Drivers of high-CRP status

We selected variables related to demographics, lifestyle, psychiatric symptoms and history, medical comorbidities and biomarkers and retained only those which were available for all three diagnostic groups, excluding two variables with > 25% missing data. This resulted in 28 potential drivers that were investigated, detailed below and in Supplementary Table S1. Supplementary Fig. S9 provides an overview of missingness across these variables in the resulting dataset.

2.2.1. Demographics and lifestyle

Demographic and lifestyle variables included sex, age, birth season (winter, defined as month 1, 2, 11 or 12), waist circumference (cm), body mass index (BMI, kg/m^2), smoking status (current, former, or never), nicotine dependence assessed with the Fagerström Questionnaire (0–10; Heatherton et al., 1991), and lifetime cannabis use (yes/no).

2.2.2. Psychiatric symptoms and history

Psychiatric variables comprised lifetime (yes/no) and number of hospitalisations, lifetime (yes/no) and number of psychotic episodes, number of depressive episodes, presence of a current depressive episode (yes/no), lifetime (yes/no) and number of suicide attempts, history of childhood trauma assessed with the Childhood Trauma Questionnaire (CTQ, total score 25–125), and sleep disturbances assessed with the Pittsburgh Sleep Quality Index (PSQI, total score 0–21). Current psychotropic treatments (first- and second-generation antipsychotics, antidepressants, mood stabilisers, lithium and anticonvulsants) were also recorded.

2.2.3. Biomarkers

Biological markers included diastolic and systolic blood pressure (mmHg, measured lying down), total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, urate, and glucose (all in mmol/L). Extreme values for LDL cholesterol, HDL cholesterol, triglycerides, and glucose were winsorised at the 97.5th percentile to limit the influence of outliers.

2.2.4. Non-psychiatric comorbidities or risk factors

Non-psychiatric comorbidities were coded as present or absent. Autoimmune or inflammatory conditions included self-reported allergies, asthma, lupus, psoriasis, type 1 diabetes, and rheumatoid arthritis. Cardiovascular comorbidity included self-reported coronary heart disease, hypertension, myocardial infarction, type 2 diabetes, hypercholesterolemia, hypertriglyceridemia, and other cardiovascular diseases. In addition, cardiovascular conditions were also coded based on objective criteria: hypertension (systolic or diastolic blood pressure ≥ 140/90 mmHg), type 2 diabetes (glucose ≥ 7 mmol/L), hypertriglyceridemia (triglycerides ≥ 150 mg/dL), and hypercholesterolemia (LDL ≥ 160 mg/dL, or ≥ 130 mg/dL in the presence of risk factors such as age > 55 years, type 2 diabetes, hypertension, or HDL < 1 mmol/L), regardless of self-report.

2.3. Outcome: CRP status

We dichotomised the C-reactive protein (CRP) biomarker using two threshold values: CRP ≥ 3 mg/L and CRP ≥ 5 mg/L. CRP values were provided by the various hospitals within each FACE network, some being derived from standard and others from high-sensitivity assays. As we focused on binary classification, this distinction is unlikely to have meaningfully affected our analyses, although we acknowledge reduced measurement precision at the lower end of the distribution. Values reported as "< 5 mg/L" were excluded from the binary variable representing the 3 mg/L threshold but retained for the 5 mg/L threshold when applicable. Values reported as "< 3 mg/L" were changed to 2.99 to appropriately allocate these individuals to the CRP < 3 mg/L group, but were omitted when calculating the mean concentration of the group. Individuals with missing CRP values (n = 1706) were excluded from analyses.

The 3 mg/L cutoff was selected *a priori* based on consistency with prior clinical and research applications in immunopsychiatry, where it is frequently used to define low-grade systemic inflammation (Jha et al., 2025). The alternative threshold of 5 mg/L was included as exploratory secondary analysis to assess whether a more stringent definition of inflammation would yield differential effects on model performance or on the identification of clinically relevant variables. Additionally, this cut-off provided a larger sample size due to the inclusion of individuals with recorded CRP values of "< 5 mg/L".

2.4. Statistics

All analyses were performed in R (version 4.3.1; R Core Team, 2023) and RStudio (version 2023.6.9.421; Posit Team, 2023).

2.4.1. Descriptive statistics

Categorical variables were summarised as percentages, and continuous variables as means with standard deviations. Using the demoGraphic R-package (Robinson, 2019), we compared sociodemographic, clinical, and biomarker characteristics between the two groups defined by the dichotomised CRP levels, using chi-square tests for categorical variables and Student's t-tests for continuous variables.

2.4.2. Modelling frameworks to identify key drivers

Penalised logistic regression (PLR) – To assess predictive performance and identify robust drivers, we first implemented penalised (LASSO) logistic regression using the *CARET* R-package (Kuhn, 2008). LASSO was chosen as it performs automatic variable selection by shrinking less informative coefficients to zero, which is particularly advantageous in settings with a large number of potentially correlated predictors, as in the present study. This approach balances model complexity and interpretability while mitigating overfitting. We applied a non-parametric bootstrapping approach, generating 500 resampled datasets. Resampling was employed to obtain robust estimates of model performance and variable stability. Within each bootstrap replicate, the data were randomly split into training and out-of-bootstrap (OOB) test sets. Missing values were imputed separately within each training and OOB set using k-nearest neighbours to avoid data leakage (*VIM* R-package; Kowarik and Templ, 2016) ($k = 5$). KNN imputation was selected due to its simplicity, flexibility, and ability to preserve complex multivariable relationships without imposing distributional assumptions, which is appropriate given the heterogeneity of the data. Binary missingness indicators were included in the model to account for potentially informative missing data. The LASSO penalty was tuned using 10-fold cross-validation repeated 10 times within each bootstrap training set, with log-loss used as the optimisation criterion.

Random Forest Classification (RF) – Secondly, we applied random forest classification using the *ranger* package in R (Wright and Ziegler, 2017). Random forests were chosen as a complementary modelling approach due to their ability to capture complex nonlinear interactions without prior specification, which may not be fully captured by the linear form of logistic regression. Multiple imputed datasets ($n = 10$) were generated using predictive mean matching via the *missRanger* package (Mayer, 2024) (100 trees per imputation). Multiple imputation was used to robustly account for uncertainty in missing data, while predictive mean matching preserves the distributional characteristics of the original data. Each imputed dataset was used to train a separate random forest model using the Extremely Randomised Trees (Extra-Trees) method with 500 trees, a commonly adopted choice that provides stable out-of-bag performance estimates and variable importance rankings while balancing computational efficiency.

Unsupervised Clustering – Third, we conducted clustering on a subset restricted to participants with elevated CRP. Unsupervised clustering was employed to explore potential subgroups with distinct clinical or biological profiles among individuals with elevated inflammation, without imposing a priori diagnostic categories. After imputation using *missForest* (Stekhoven, 2022), we scaled (z-scored) and dichotomised the continuous and categorical variables respectively. Imputation methods were chosen to match the analytic approach: for clustering, a single imputation via missForest preserves multivariate relationships in mixed-type data without introducing stochastic variation across imputations. Other analyses (PLR, RF) used imputation methods tailored to their respective modelling frameworks. Next, principal component analysis (PCA) was used to extract the top 10 components explaining the most variance, thereby reducing dimensionality while retaining most of the relevant information. Uniform Manifold Approximation and Projection (UMAP; *uwot* R-package; Melville, 2025) was then used for nonlinear dimensionality reduction, as it can capture complex data structure and preserve local and global relationships more effectively than linear methods alone. K-means clustering was applied to the UMAP output, with the optimal number of clusters selected via the elbow method. We merged cluster assignments back with the original dataset for interpretation.

2.4.3. Importance of individual drivers

We derived measures of importance from each modelling approach to identify drivers most strongly associated with high-CRP status. In PLR, variable importance was quantified using the Variable Inclusion Probability (VIP), defined as the proportion of bootstrap iterations in which a variable's coefficient was non-zero. Variables with VIP values of 85% or higher were considered to represent robust drivers. To provide an indication of statistical significance in the absence of conventional p-values, which are not readily applicable to regularised models, we computed bootstrap-based confidence intervals (CI) for the model coefficients. Drivers whose CI did not include zero were considered statistically significant.

In the RF models, importance was assessed through permutation-based importance scores, which measure the decrease in model performance when a given driver's values are randomly permuted. These scores were averaged across all imputations to produce stable estimates of each driver's relative contribution to prediction.

Finally, for the unsupervised clustering analysis, importance was inferred by examining the distribution of drivers across clusters to identify those features most strongly differentiating the identified subgroups.

2.4.4. Evaluation of overall predictive performance

To evaluate the predictive performance of the penalised logistic regression (PLR) and random forest (RF) models, we selected multiple complementary metrics, to ensure a comprehensive evaluation. For PLR, apparent performance was calculated on the bootstrap samples, with predictive performance assessed on the corresponding OOB test-sets, each of which was imputed independently of the training data to avoid information leakage. We assessed the Area Under the Receiver Operating Characteristic Curve (AUROC), the F1-score, and the Brier score. AUROC provides a measure of discriminative ability, indicating how well the model distinguishes between individuals with and without elevated CRP levels. Values closer to 1 indicate better performance. The F1-score, calculated both at a fixed probability threshold of 0.5 and at the ROC-optimal threshold, balances sensitivity and precision, which is particularly relevant in the presence of class imbalance. The Brier score quantifies the accuracy of predictions by measuring the mean squared difference between predicted probabilities and actual outcomes, with lower values indicating better calibration.

For the RF models, we estimated performance using the algorithm's internal OOB procedure, averaged across 10 multiply imputed datasets to provide robust AUROC and F1 estimates at the 0.5 threshold, as these metrics effectively summarise the models' discrimination and classification performance in both balanced and imbalanced outcome settings.

2.4.5. Sensitivity analyses

We performed several sensitivity analyses to assess the robustness of results:

- Diagnostic cohort-balanced analysis: We repeated PLR and RF on a subset of 900 individuals, including 300 patients each from MD, BD, and SZ cohorts.
- Medication adjustment: Additional models included binary indicators of psychotropic medication use (first– and second-generation antipsychotics, antidepressants, anxiolytics, lithium, mood stabilisers) to assess their influence on predictive models.

2.4.6. From drivers to targets

To enhance clinical interpretability, we translated consistently important drivers from our models into composite scores representing broader, clinically actionable targets. This step marks a shift from identifying statistical drivers to defining practical targets for stratified intervention; a necessary bridge between data-driven discovery and clinical decision-making. In line with how risk is typically conceptualised in practice, these composites grouped individual variables under broader domains: *Obesity* (BMI, waist circumference), *Dyslipidemia* (inverse HDL cholesterol, triglycerides), *Nicotine Exposure* (smoking status, Fagerström score), and *Blood Pressure* (systolic and diastolic measures). Composites were calculated as the mean of scaled component measures.

These domain-level targets, along with sex, were entered into a binomial logistic regression model to predict high-CRP status. Based on secondary analyses, the model was extended to include urate concentration and cardiovascular comorbidity. Overall model fit was evaluated using Nagelkerke's R2 to approximate explained variance. To assess the relative importance of each target, we conducted dominance analysis using the dominanceanalysis R-package (Bustos Navarrete and Coutinho Soares, 2024), which quantifies each target's contribution to the total model variance explained. This approach provides a simplified, clinically grounded representation of the most influential targetable domains, reducing complexity while aligning statistical findings with real-world intervention strategies.

2.4.7. Post-hoc exploratory split analyses

To explore differential predictive patterns across demographic subgroups, we repeated PLR, RF, and dominance analysis after stratifying by:

- Sex (male vs. female)
- Smoking status (smokers vs. non– & ex-smokers)

3. Results

3.1. Descriptive statistics

In total, 30.16% of participants had CRP $\geq$ 3 mg/L (n = 1434), with a mean concentration of 8.34 mg/L, compared to 1.04 mg/L for the CRP < 3 mg/L group (n = 3320). Diagnostic cohorts were equally represented across CRP groups; specifically, 29.81% of individuals with BD, 29.70% with SZ, and 35.71% MDD patients fell into the high-CRP category. Individuals in the high CRP group were significantly older, with more females and smokers. This group also exhibited increased BMI, waist circumference, blood pressure, prevalence of cardiovascular comorbidities and higher levels of total cholesterol, triglycerides, glucose and urate. In terms of psychiatric characteristics, the high CRP group showed poorer sleep quality, a higher incidence and frequency of past psychiatric hospitalisations and more individuals with a suicide attempt. Full descriptive statistics are presented in Table 1; results using the alternative CRP $\geq$ 5 mg/L threshold are provided in Table S2.

Overall, 25.31% of participants had CRP $\geq$ 5 mg/L. Group differences were largely consistent with those observed at the lower cut-off, with some additional distinctions: the high-CRP group scored higher on the Childhood Trauma Questionnaire (CTQ), reported a greater number of lifetime suicide attempts, included more individuals currently experiencing depressive symptoms, but fewer with a history of psychosis and lower rates of cannabis use.

3.2. Key drivers identified by predictive modelling

Table 2 presents variables that were robustly and significantly associated with CRP $\geq$ 3 mg/L in the penalised logistic regression. Factors associated with higher odds of elevated CRP included female sex (OR [95% CI]: 1.60 [1.27, 1.93], VIP: 0.998), higher BMI (OR: 1.09 [1.07, 1.13], VIP: 0.998), and current nicotine dependence (OR: 1.05 [1.02, 1.09], VIP: 0.996). Conversely, HDL cholesterol (OR: 0.56 [0.44, 0.73], VIP: 0.998), ex-smoker status (OR: 0.84 [0.66, 0.98], VIP: 0.86) and absence of suicide attempts (OR: 0.90 [0.77, 0.99], VIP: 0.92) were associated with lower odds of high CRP, though the latter was not retained in balanced sensitivity analyses (see Table S3 for full results and S4 for informative missingness indicators). These findings were supported by random forest permutation-based variable importance (Fig. 1), which highlighted a similar set of key drivers. Additional contributors identified in this model included waist circumference, triglycerides, cardiovascular comorbidities, and urate. However, urate was not consistently retained in balanced sensitivity analyses.

Table 1
Descriptive Statistics of the Low vs. High CRP groups.

Variable	CRP < 3 n = 3320	CRP ≥ 3 n = 1434	*p*	SMD
Diagnostic Group [n(%)]			0.093	0.068
BD	2232 (67.23)	948 (66.11)		
MD	198 (5.96)	110 (7.67)		
SZ	890 (26.81)	376 (26.22)		
Females [n(%)]	1682 (50.72)	795 (55.48)	0.003**	−0.095
Age [mean(SD)]	38.18 (13.30)	39.61 (13.17)	0.001**	0.108
Born in Winter [n(%)]	1049 (31.60)	458 (31.94)	0.839	−0.007
BMI [mean(SD)]	24.96 (4.57)	28.80 (6.29)	< 0.001***	0.698
Waist Circumference [mean(SD)]	90.45 (14.00)	99.93 (16.45)	< 0.001***	0.621
Diastolic Blood Pressure [mean(SD)]	73.65 (10.91)	76.37 (11.29)	< 0.001***	0.245
Systolic Blood Pressure [mean(SD)]	120.24 (14.90)	124.02 (15.99)	< 0.001***	0.245
Smoke Status [n(%)]			0.002**	0.116
Non-smokers	1355 (42.42)	531 (38.26)		
Ex-smokers	408 (12.77)	155 (11.17)		
Current smokers	1431 (44.80)	702 (50.58)		
Nicotine Dependence [mean(SD)]	1.95 (2.79)	2.52 (3.11)	< 0.001***	0.195
Cannabis use [n(%)]	672 (24.53)	285 (23.19)	0.377	−0.031
PSQI [mean(SD)]	7.30 (3.78)	7.74 (3.91)	0.001**	0.115
CTQ [mean(SD)]	42.44 (14.07)	43.21 (14.44)	0.102	0.054
Cardiovascular Comorbidity [n(%)]	1308 (39.47)	809 (56.57)	< 0.001***	0.347
Auto-immune or Inflammatory Comorbidity [n(%)]	850 (27.60)	381 (27.93)	0.827	0.007
Hospitalised during Lifetime [n(%)]	2166 (76.29)	1015 (81.14)	0.001**	0.118
No. of Hospitalisations [mean(SD)]	2.65 (2.82)	3.27 (3.27)	< 0.001***	0.201
Suicide Attempt Lifetime [n(%)]	1088 (34.89)	554 (40.65)	< 0.001***	0.119
Number of Attempts [mean(SD)]	2.59 (3.45)	2.52 (2.64)	0.642	−0.024
Psychotic Episode Lifetime [n(%)]	1334 (46.45)	575 (46.56)	0.973	0.002
No. of Psychotic Epi. [mean(SD)]	1.22 (1.93)	1.35 (2.27)	0.062	0.066
No. of Depressed Epi. [mean(SD)]	5.19 (5.02)	5.58 (5.31)	0.063	0.074
Current Depressed Episode [n(%)]	993 (31.26)	470 (33.76)	0.098	0.054
Total Cholesterol [mean (SD)]	5.00 (1.15)	5.25 (1.19)	< 0.001***	0.216
HDL Cholesterol [mean (SD)]	1.61 (4.86)	1.44 (4.59)	0.254	−0.036
LDL Cholesterol [mean (SD)]	3.11 (5.59)	3.23 (1.01)	0.226	0.032
Triglycerides [mean(SD)]	1.33 (1.14)	1.90 (6.16)	0.001**	0.128
Glucose [mean(SD)]	4.97 (1.75)	5.16 (2.46)	0.013*	0.087
Urate [mean(SD)]	291.07 (83.93)	308.63 (88.11)	< 0.001***	0.204
CRP [mean(SD)]	1.04 (0.82)	8.34 (12.38)	< 0.001***	1.003

Note. Total N = 4754. *P*-values calculated by $\chi 2$ and *t*-test for categorical and continuous data respectively.

Table 2
Results from Bootstrapped Penalised Logistic Regression for CRP ≥ 3.

Variable	Estimate	95%CI	OR	VIP
Sex [Female]	0.47	[0.24, 0.66]	1.60	0.998
BMI	0.09	[0.07, 0.12]	1.09	0.998
Waist Circumference	0.01	[0, 0.01]	1.01	0.93
Smoking Status [Ex]	−0.18	[-0.41, −0.02]	0.84	0.862
Nicotine Dependence	0.05	[0.02, 0.09]	1.05	0.996
PSQI Score	0.01	[0, 0.03]	1.01	0.888
Lifetime Suicide Attempt [N]	−0.11	[-0.26, −0.01]	0.90	0.924
Lifetime Suicide Attempt [Y]	0.00	[0, 0.12]	1.00	0.852
No. of Suicide Attempts	−0.02	[-0.06, 0]	0.98	0.858
Total cholesterol	0.06	[0, 0.14]	1.06	0.918
HDL Cholesterol	−0.56	[-0.81, −0.32]	0.57	0.998

Note. VIP (Variable Inclusion Probability) is the proportion of models in which the estimate was nonzero, only variables with VIP > 0.85 are shown. Variables in bold are considered statistically significant, i.e. confidence intervals do not include 0. See Table S3 for full results.

3.3. Predictive performance in primary analysis

Using all available variables, PLR and RF models showed comparable performance in classifying elevated CRP (≥3 mg/L). PLR yielded a mean AUROC of 0.70 [95% CI: 0.67–0.73] with an F1 score of 0.36 at the 0.5 threshold, improving to 0.54 using the optimal ROC-based threshold (Youden's J). The corresponding Brier score was 0.19. RF models achieved an AUROC of 0.70 and an F1 score of 0.30. Model performance remained consistent when diagnostic groups were balanced or when psychotropic medication use was included. This level of discrimination supports risk stratification and prioritisation of modifiable inflammatory drivers at the group level, rather than individual-level prediction.

3.4. Alternate CRP threshold (≥5 mg/L)

At the higher CRP threshold, model discrimination declined (PLR AUROC: 0.56 [0.51–0.61]; RF AUROC: 0.62). While F1 score was higher at the default 0.5 threshold (0.72), it dropped to 0.58 with optimal thresholding, suggesting decreased model calibration. RF performance was particularly affected (F1: 0.10). This threshold included a larger sample due to the inclusion of individuals with recorded CRP values of "<5 mg/L", which are less precise and may include borderline cases, potentially introducing misclassification. Although the high-CRP category remained similar in size, the expanded low-CRP group likely increased variability, reducing the signal-to-noise ratio and resulting in lower discriminative performance despite the larger overall sample size.

Nonetheless, PLR retained similar key drivers, and also included diastolic blood pressure (OR: 1.01, VIP: 1.00) and waist circumference (OR: 1.01, VIP: 1.00) (Table S5). RF similarly highlighted increased importance of blood pressure-related variables (Fig. S2).

3.5. Sex-stratified models

Given the strong effect of sex, we repeated analyses separately for males and females. Predictive performance was better in females (PLR AUROC: 0.73 [0.69–0.77], F1: 0.58) than in males (AUROC: 0.67 [0.63–0.71], F1: 0.50). RF performance showed similar trends (females AUROC: 0.74, F1: 0.44; males AUROC: 0.65, F1: 0.11).

In females, urate emerged as a strong contributor. In males, the pattern was more diffuse, with nicotine dependence showing a greater relative influence. Variable importance estimates for males were less stable, with wider confidence intervals and more variability in RF importance scores (Tables S6-7, Figs. S3-4).

3.6. Smoking-stratified models

Unsupervised clustering within the high CRP group revealed two distinct subgroups, primarily differentiated by smoking status and nicotine dependence (Table S8). Re-clustering non-smokers or Cluster 1 participants yielded no further substructure, suggesting smoking was the dominant driver.

This prompted stratified post-hoc analyses by smoking status. However, predictive models in smokers versus non– and ex-smokers did not yield substantial differences beyond the clustering results (Table S9-10, Figs. S5-6).

3.7. From drivers to targets

A summary of the most influential drivers across all analytic methods

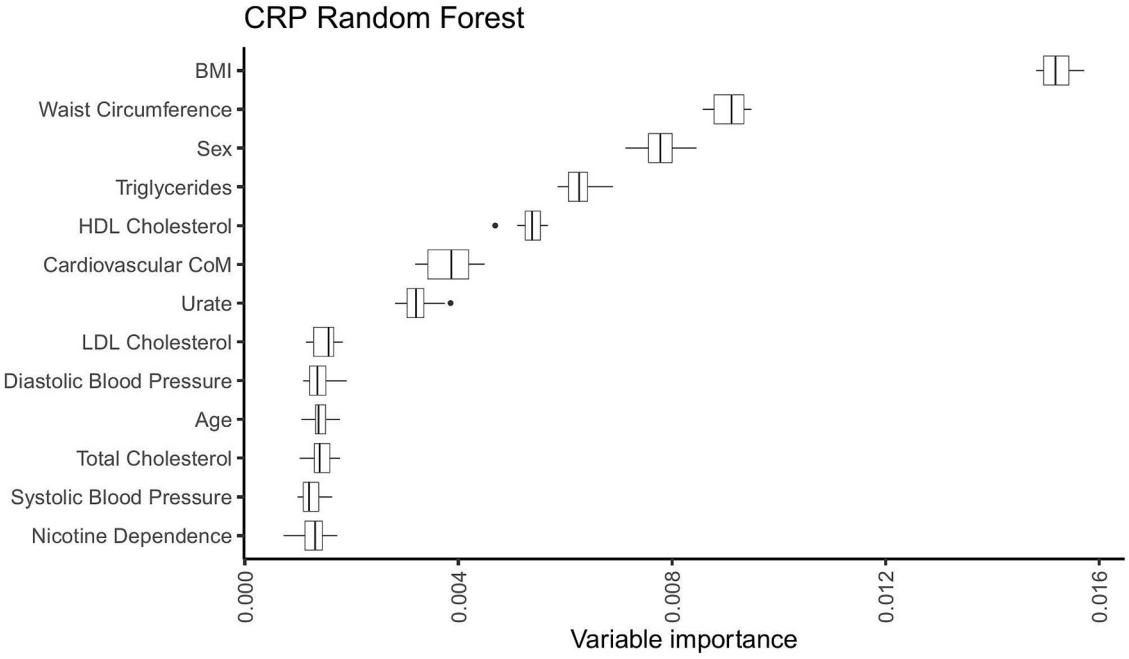

Fig. 1. Random Forest permutation-based relative importance of the first half of variables when predicting CRP ≥ 3. See Fig. S1 for plot including all variables.

is shown in Fig. 2. To enhance clinical utility, these were translated into composite variables representing actionable intervention targets. When entered into a simplified binomial model, these targets collectively explained 16% of the variance in CRP status (Nagelkerke's R2), with obesity emerging as the dominant contributor, accounting for over 57% of the explained variance (Fig. 3).

Stratified analyses further underscored sex-specific patterns in these pathways. In females, the identified targets explained 22% of total variance, with obesity accounting for more than 55% of this explained variance. In males, the same targets explained substantially less variance overall (10%), indicating that CRP heterogeneity is less well captured despite a broadly similar contribution pattern across domains. Importantly, smoking accounted for almost 15% of explained variance in males but only ~ 0.5% in females, while urate explained ~ 9% in females and ~ 2% in males (Figs. S7, S8).

4. Discussion

This is the first study to apply a comprehensive, transdiagnostic, and data-driven approach to elucidate the key drivers underlying high-CRP status in patients with severe mental illness (SMI). A major strength of this study lies in its use of robust statistical methods and large transdiagnostic cohort, which is particularly relevant given the increased recognition of shared biological mechanisms across psychiatric disorders. By integrating penalised regression, random forests, and unsupervised clustering, we sought to move beyond isolated risk factors or single diagnostic categories, and instead delineate the complex pathways driving systemic inflammation. These drivers were then aggregated into clinically actionable targets, reflecting how risk is conceptualised and managed in clinical practice. Collectively, these findings lay the foundation for a framework that could support stratified intervention through the development of structured, evidence-based decision tools.

4.1. Drivers of CRP elevation

Consistent with previous reports, approximately 30% of patients in our sample exhibited CRP levels exceeding 3 mg/L. Across analytic methods, a relatively stable set of underlying drivers emerged. Sex was among the strongest and most consistent, with females significantly more likely to exhibit high CRP status. Body mass index (BMI) and waist circumference were similarly dominant contributors, reaffirming the central role of adiposity in driving systemic inflammation. Low HDL cholesterol and elevated triglycerides also featured prominently, alongside markers of nicotine exposure, including smoking status and Fagerström scores. These findings converge with prior work linking metabolic dysregulation and smoking to inflammatory burden, but extend this literature by demonstrating their stability as predictors within a large, heterogeneous, transdiagnostic psychiatric cohort.

Notably, some expected drivers were less clearly implicated. For instance, inflammatory or autoimmune comorbidities, as assessed by self-report of conditions including allergies, asthma, lupus, psoriasis, type 1 diabetes, or rheumatoid arthritis, did not predict high CRP themselves. Instead, the presence of recorded data regarding these conditions was associated with high CRP, suggesting that the act of documentation may function as a proxy for underlying risk. This phenomenon of "informative missingness" reflects real-world clinical practice, where comorbidities are more likely to be recorded if present, and highlights both the utility and limitations of secondary clinical datasets.

4.2. From drivers to targets

To enhance clinical utility, the consistent individual drivers identified across models were aggregated into broader, actionable targets. BMI and waist circumference were combined into an obesity composite, while low HDL cholesterol and high triglycerides formed a dyslipidemia composite. Smoking status and Fagerström scores were summarised under nicotine exposure, and systolic and diastolic measures were combined as blood pressure. This process mirrors how risk is typically conceptualised in clinical practice, bridging the gap between data-driven discovery and practical implementation.

When entered into a simplified model, these composites explained 16% of the variance in CRP status across the full transdiagnostic sample. Obesity emerged as the dominant contributor, accounting for the majority of explained variance and underscoring its role as a central intervention target, while dyslipidemia and cardiovascular comorbidities represent additional, clinically meaningful avenues for risk reduction.

Sex also contributed strongly, but is not an actionable target itself.

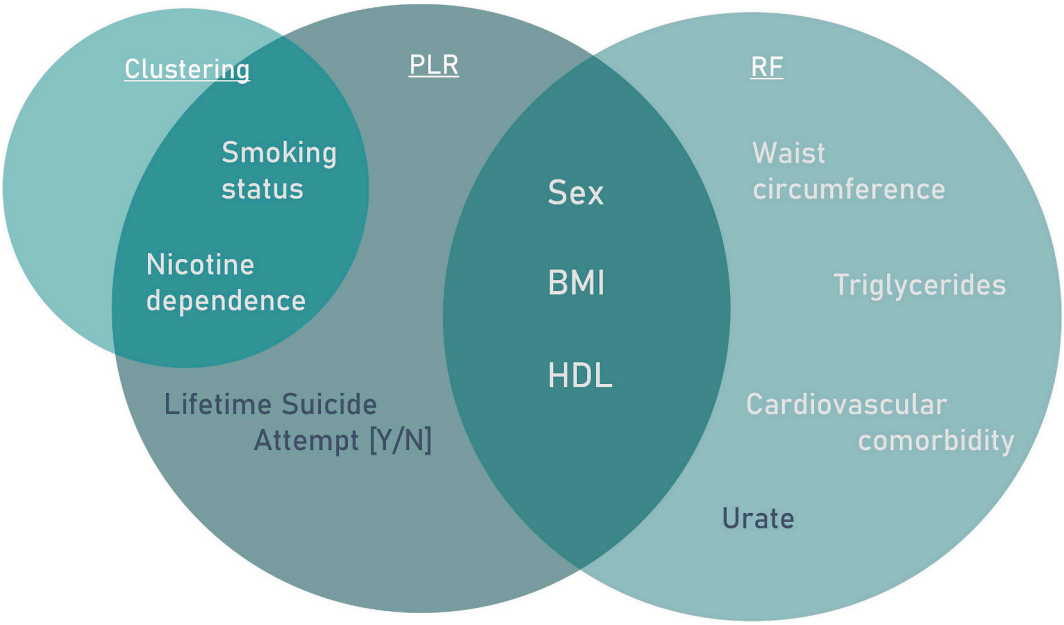

Fig. 2. Venn diagram of most important variables across methods. Variables in grey were not robust to sensitivity analyses.

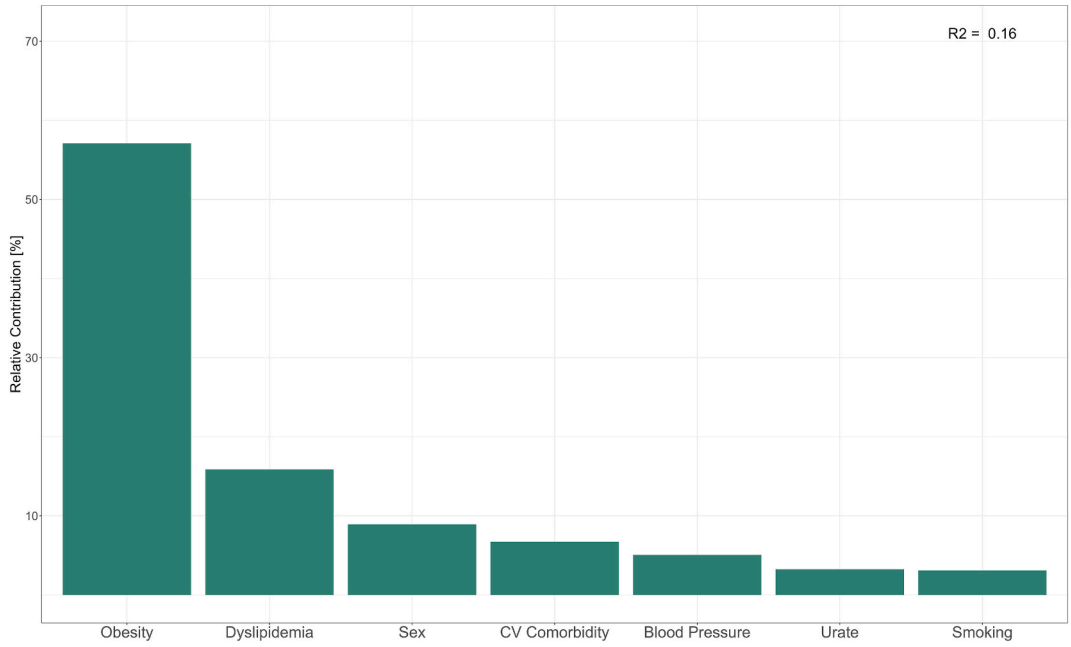

Fig. 3. Variance explained by the various clinical domains (Nagelkerke's R^2). CV = cardiovascular, Obesity = BMI & waist circumference, Dyslipidemia = inverse HDL cholesterol & triglycerides, Smoking = smoking status & Fagerström score, and Blood Pressure = systolic & diastolic measures. See Figs. S7&S8 for sex-stratified results.

Rather, this provided a strong rationale for conducting sex-stratified analyses to examine how the relative importance of other targets differs between males and females. In doing so, these analyses offered more nuanced insights into the findings from the overall model, most notably, the surprisingly minor contribution of smoking.

4.3. Sex-specific patterns

Sex-stratified analyses revealed distinct profiles of contributors to CRP elevation. Among females, obesity dominated the explained variance, with dyslipidemia and urate levels as secondary but clinically relevant contributors. Urate may promote systemic inflammation by acting as a danger-associated molecular pattern (DAMP), activating the NLRP3 inflammasome and stimulating the production of pro-inflammatory cytokines such as IL-1β (Braga et al., 2017). In males, a broadly similar pattern emerged, yet the same targets accounted for substantially less variance overall, consistent with greater heterogeneity and lower predictive performance. Here, urate was negligible, whereas smoking emerged as a substantially stronger driver.

This shift provides crucial context for the seemingly minor role of smoking in the combined model. Rather than indicating an weak association overall, aggregation masked a sex-specific signal: nicotine use meaningfully contributed to inflammation in males but not in females. Although higher smoking prevalence among men with SMI likely contributes to this effect, broader metabolic and hormonal differences may also help explain these divergent patterns, ranging from sex-specific fat distribution, differences in urate and lipid metabolism to the influence of oestrogen, testosterone and other hormones on inflammatory pathways. Taken together, these findings reinforce the importance of sex as primary stratification step in any future decision framework, shaping which targets should be prioritised for risk reduction in males versus females.

4.4. Toward a stratified decision tool

The identification of consistent, clinically interpretable targets lays the groundwork for a structured framework to guide intervention strategies in SMI. Rather than treating CRP as a standalone biomarker, our findings suggest that elevations can be meaningfully parsed into underlying targets.

We propose a decision tree framework whereby elevated CRP serves as an initial screening step, followed by sex-specific and domain-specific stratification (Fig. 4). For example, in females with elevated CRP, metabolic factors such as obesity, dyslipidemia, and hyperuricemia levels represent primary targets for intervention. In males, while metabolic factors remain relevant, smoking may warrant greater emphasis. Such stratified decision pathways could ultimately enable more precise, individualised treatment strategies, tailoring interventions to the dominant contributors of inflammation in each patient.

While CRP is not yet routinely applied as a clinical decision tool in psychiatry, it is increasingly recognised as a stratification marker to identify patients for immunotherapy trials (e.g. Wessa et al., 2024; Wessa et al., 2025 (in press); Treadway et al., 2024) and clinically in specialised immunopsychiatry clinics. Our findings suggest that, while awaiting the outcomes of such trials, clinicians could already leverage CRP to identify individuals who might benefit from targeted interventions, particularly weight management (e.g. exercise, GLP–1 receptor agonists), lipid optimisation, and smoking cessation (e.g. bupropion) where relevant.

For research purposes, the framework we propose also allows for iterative refinement. As additional biomarkers or pathways are identified, they can be layered onto the tool to improve specificity and identify subgroups with different inflammatory pathways. At the same time, it is important to recognise the trade-offs of this approach. Multi-omics integration is often assumed to solve the problem of heterogeneity, but in practice, it can introduce additional noise and obscure clinically useful signals.

Our model demonstrates that even a focused set of routinely available clinical measures can meaningfully explain certain cases of increased inflammation, supporting a pragmatic, stepwise approach: starting with clear, actionable targets and adding specialist assessments (e.g. lumbar puncture, auto-antibodies) only when it demonstrably enhances predictive or clinical value. Ultimately, the goal is to translate these insights into a practical, evidence-based decision tool that clinicians can realistically implement in routine care, forming the basis for future refinement and validation.

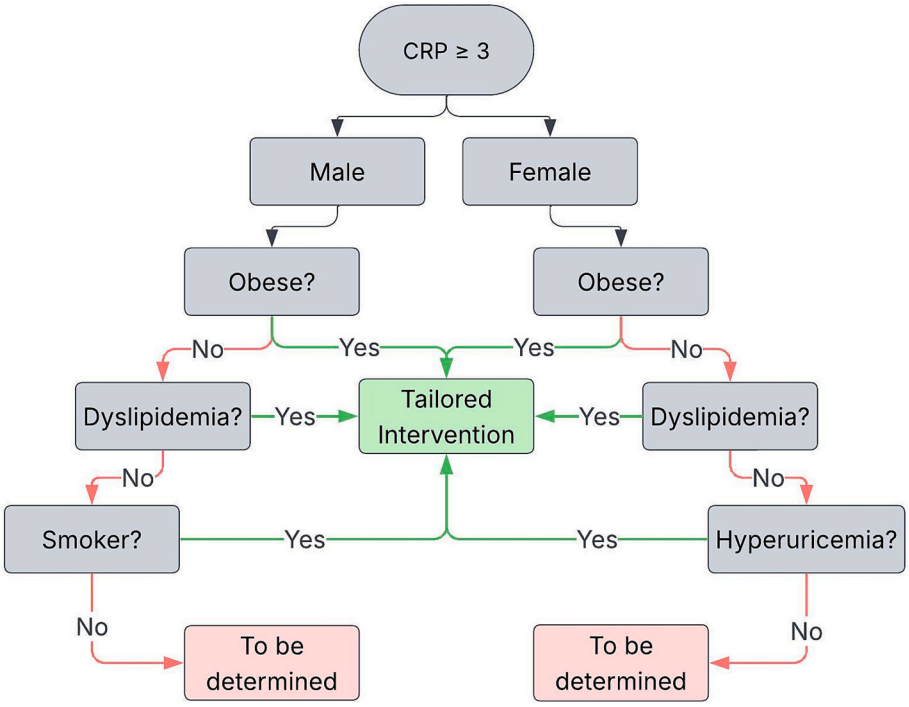

Fig. 4. Proposed decision tree framework. Following CRP screening at the ≥ 3 mg/L threshold, patients are stratified into clinically actionable pathways. The framework should be refined and extended by identifying the dominant drivers of elevated CRP within the residual group remaining at each step.

4.5. Limitations and future directions

Several methodological considerations should be acknowledged. The current set of predictors was limited to variables available across all diagnostic cohorts, which precluded the inclusion of potentially relevant factors such as alcohol consumption, age of disease onset, physical activity, dietary patterns and socio-economic status. This limitation reflects the inherent challenges of repurposing even standardised prospective clinical datasets for transdiagnostic research and highlights the need for harmonised data collection protocols. Failure to include relevant variables may have restricted the total explanatory power of our models. While the simplified models explained 16% of CRP variance overall and 22% in females, which is substantial given the extreme heterogeneity of the cohort, discriminatory ability still ranged from poor to moderate. This partly reflects the inherent challenge of identifying meaningful subgroups using a non-specific biomarker like CRP. Numerous biological, behavioural, and environmental factors contribute to elevated inflammation, and many of these remain unmeasured here. Some of these factors may act upstream of the included predictors, influencing CRP indirectly. Consequently, while our models identify key correlates of high-CRP status, they capture only part of the broader causal network, and effective targeting of these proximate drivers may require consideration of the more distal, unmeasured contributors as well.

It should further be noted that permutation-based variable importance can be sensitive to correlated predictors, such as BMI, waist circumference, and lipid measures, which may lead to over- or underestimation of their relative contributions. Nonetheless, the consistency of key drivers across complementary methods provides convergent support for the identified targets.

In addition, although models were extensively internally validated, we did not perform external validation in an independent cohort. Consequently, generalisability beyond the studied population remains to be established.

Looking forward, these observations highlight several priorities for future research. First, improving data collection frameworks with harmonised measures across diagnostic groups and a broader range of clinically and behaviourally relevant variables will be essential for deepening the explanatory model.

Second, while additional data modalities (e.g. multi-omics profiles) could help to more comprehensively elucidate the multifactorial pathways underlying inflammation in psychiatric populations, these should be integrated with caution. More biomarkers can easily increase noise and reduce clinical utility unless carefully structured around interpretable, actionable targets.

Third, longitudinal study designs will be essential to capture dynamic and temporal relationships, including repeated CRP measurements, which would allow assessment of the stability of inflammation over time and whether interventions on obesity, dyslipidemia, or other pathways meaningfully reduce inflammation, an aspect that could not be addressed in our cross-sectional design. Importantly, while our analyses identify factors that are predictive of high-CRP status and suggest clinically actionable targets, these findings are inherently associative rather than causal. Intervening on these predictors may not reduce inflammation unless other unmeasured causal pathways are also addressed. Future studies could formalise these relationships using a directed acyclic graph (DAG) or similar causal framework.

Finally, further research is warranted to validate and extend the identified pathways and to test whether differential inflammatory biomarkers, psychiatric symptomatology or treatment response profiles emerge at each step of a decision tree framework. A key next step will be to move toward an iterative, stratified approach that mirrors clinical decision-making. While the present study establishes a foundation for such stratification, future studies should refine this model by focusing on under-characterised subgroups, such as individuals with elevated CRP in the absence of obesity or smoking. Identifying the dominant contributors within such residual groups may reveal novel targets for intervention or highlight the need for additional biomarkers. Large-scale, transdiagnostic datasets such as the one used here are well suited for this kind of exploratory analysis and can serve as a valuable basis for generating hypotheses. However, prospective studies specifically designed to stratify participants from the outset, with harmonised

clinical assessments and more granular biological sampling, will be essential to validate these pathways and determine their relevance for clinical decision-making.

4.6. Conclusion

Our findings provide initial evidence that clinically meaningful, sex-specific patterns underlie high-CRP status in patients with SMI, with obesity, metabolic dysregulation, and smoking emerging as dominant contributors. Despite relying on routine variables, these actionable clinical targets explained a substantial proportion of CRP variance, especially in females, highlighting their potential utility for stratified risk assessment. While we did not assess psychiatric outcomes directly, elevated CRP is a well-established marker of cardiovascular risk. Identifying inflammation-associated subgroups in SMI may therefore offer a critical opportunity to intervene early and improve long-term health.

These findings lay the groundwork for developing decision tools that go beyond CRP as a standalone marker, ultimately supporting more targeted, preventive care in this high-risk population. Future research should validate and extend these findings in clinical settings and, ultimately, determine whether tailoring interventions based on such decision tools can improve psychiatric and physical health outcomes in this vulnerable subgroup.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used the ChatGPT AI language model from OpenAI in order to refine language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

CRediT authorship contribution statement

Tim Rietberg: Writing – review & editing, Writing – original draft, Methodology, Formal analysis. **Susana Barbosa:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Pieter Meysman:** Formal analysis. **Bruno Aouizerate:** Investigation. **Frank Bellivier:** Investigation. **Fabrice Berna:** Investigation. **Delphine Capdevielle:** Investigation. **Isabelle Chereau:** Investigation. **Philippe Courtet:** Investigation. **Clément Dondé:** Investigation. **Caroline Dubertret:** Investigation. **Bruno Etain:** Investigation. **Wissam El Hage:** Investigation. **Mona Gerentes:** Investigation. **Jérôme Holtzmann:** Investigation. **Eric Fakra:** Investigation. **Emmanuel Haffen:** Investigation. **Noomane Bouaziz:** Investigation. **Antoine Lefrere:** Investigation. **Pierre-Michel Llorca:** Investigation. **Jasmina Mallet:** Investigation. **David Misdrahi:** Investigation. **Emilie Olié:** Investigation. **Agnès Pelletier:** Investigation. **Giulia Piva:** Investigation. **Baptiste Pignon:** Investigation. **Mircea Polosan:** Investigation. **Romain Rey:** Investigation. **Paul Roux:** Investigation. **Ludovic Samalin:** Investigation. **Franck Schürhoff:** Investigation. **Raymund Schwan:** Investigation. **Florian Stephan:** Investigation. **David Szekely:** Investigation. **Mathieu Urbach:** Investigation. **Guillaume Vaiva:** Investigation. **Etienne Very:** Investigation. **Antoine Yrondi:** Investigation. **Marion Leboyer:** Writing – review & editing, Resources, Conceptualization. **Livia De Picker:** Writing – review & editing, Supervision, Resources, Conceptualization. **Ophélia Godin:** Writing – review & editing, Supervision, Methodology.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Unrelated to the submitted work, LDP reports honoraria for consultancy and presentations from Boehringer-Ingelheim and Janssen R&D. LDP and ML are members of the ECNP Immuno-NeuroPsychiatry Network.

Acknowledgements

This study was completed with support of the Tournesol mobility project by the Flemish fund for scientific research (FWO), the ECNP Immuno-NeuroPsychiatry Network, AP-HP, Université Paris-Est-Créteil (UPEC), Fondation Fondamental, Institut National de la Santé et de la Recherche Médicale (INSERM), UPCD, the University of Antwerp, and by the Investissements d'Avenir program managed by the Agence National de la Recherche (ANR) under reference ANR-11-IDEX-0004-02 and ANR-10-COHO-10-01. We thank all patients and hospital staff who have participated in this study, and express our gratitude to the clinical teams of AP-HP and UPCD, as well as all FACE collaborators:

FACE-BD collaborators

Llorca PM, Bensalem S, Kulakovskyi A, Dansou Y, Altunlu, E, Hotier S, Pelletier A, Macone A, Dominique C, Pin C , C. Lam, J. Lantier, F. Harter, V. Nicolaidis, Hennion V, Marlinge E, Meheust J, Richard A, Carminati M, Scemama D, V. Feuga, S. Frileux, D. Niculcea, P. Roux, L. Brehon, L. Durand, E. Vaux, C. Borssard, C. Fourmont, A. Lefevre, Mazer Nicolas, Caroline Dubertret, Pauline Laurent, Frédérique Dousset, S. Gard, B. Aouizerate, C. Elkaël, K. M'Bailara, N. Da Ros, I. Minois, J. Sportich, F. Hoorelbeke, P.Courtet, B.Deffinis, D.Ducasse, M.Gachet, E. Olié, L.Boukhobza,E.Palmeri, G.Tarquini, E.Michel, M. Benramdane, A. Lefrere, E. Moreau, F.Groppi, H.Polomeni, J.Baurberg, L.Lescalier, I. Muraccioli, A.Suray, Anamaria Bogdan, Paula Martinez, Isabelle Medecin, Marie Agnes Dupont, M Polosan, A Pouchon, A Bertrand, J Pastol, E Winiarski, B Fredembach, P Reboud, M. Veloso Da Fonseca, L Samalin, M. Gambier, T. Bonnet, PM Llorca, C. Garcin, C. Beal, O Blanc, D Januel, Sarak Ateb, I. Belkhatir, N. Bouaziz, S. Joanny, G. Gracieux, Antoine Yrondi, Racan Abidine, Christine Sarramon, Christophe Arbus, Sandrine Ivon.

FACE-SZ collaborators

Andre M, Andrieu-Haller C, Aouizerate B, Berna F, Blanc O, Capdevielle D, Cléry M, Chereau-Boudet I, Clauss-Kobayashi, J,Dassing R, Dorey JM, Dubertret C, D'Amato T, Esselin A, Fakra E, Fond G, Gabayet F, Goze T, Jarroir M, Leboyer M, Leignier S, Llorca PM, Mallet J, Marongiu Z, Metairie E, Michel T, Misdrahi D, Oriol N, Passerieux C, Pignon B, Peri P, Portalier C, Rey R, Roman C, Roux P, Schorr B, Schürhoff F, Szöke A, Tessier A, Urbach M, Very E, Zinetti-Bertschy A.

FACE-DR collaborators

A. Yrondi, R. Abidine, L Schmitt, C. Arbus, E. Haffen, D. Bennabi, I. Ryff, E. Beuchet, G. Tio, C. Cappe and E. Clerc; L. Samalin, PM Llorca; B. Etain, V. Hennion, F. Bellivier, M. Carminati; M. Polosan, T. Bougerol, J. Holtzmann, B. Fredembach, AS Nguon, C. Berthaud and P. Reboud; W. El Hage , L.Warck and V. Camus; T. D'Amato, R. Rey, JF. Costemale-Lacoste, T. Bastelica, C. Dubien and J. Brunelin; B. Aouizerate, D. Etchecopar-Etchart, S. Boukoulo, F. Hoorelbeke; C. Lancon, T. Korchia, G Fond, C Andrieu, N Bochu, P. Peri, E Metairie, J Bourgon and M. Faugere; P. Courtet, E Olie, R Strumila, A Lengevnyte and N. Benachir; F. Stephan, M. Walter and R. Loaec; G. Vaiva, A. Amad and M Horn; D. Januel, H. Benrejeb

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbi.2026.106464.

Data availability

Data will be made available on request.

References

Baysak, E., Guden, D.S., Aricioglu, F., Halaris, A., 2022. C-reactive protein as a potential biomarker in psychiatric practice: are we there yet? World J. Biol. Psychiatry 23 (4), 243–256. https://doi.org/10.1080/15622975.2021.1961502.

Berk, M., Williams, L.J., Jacka, F.N., O'Neil, A., Pasco, J.A., Moylan, S., Allen, N.B., Stuart, A.L., Hayley, A.C., Byrne, M.L., Maes, M., 2013. So depression is an

inflammatory disease, but where does the inflammation come from? BMC Med. 11, 200. https://doi.org/10.1186/1741-7015-11-200.
Braga, T.T., Forni, M.F., Correa-Costa, M., Ramos, R.N., Barbuto, J.A., Branco, P., Castoldi, A., Hiyane, M.I., Davanso, M.R., Latz, E., Franklin, B.S., Kowaltowski, A.J., Camara, N.O.S., 2017. Soluble uric acid activates the NLRP3 inflammasome. Sci. Rep. 7, 39884. https://doi.org/10.1038/srep39884.
Bustos Navarrete, C., Coutinho Soares, F., 2024. dominanceanalysis: Dominance Analysis (Version 2.1.0) [R package]. CRAN. https://CRAN.R-project.org/package=dominanceanalysis.
Danese, A., Pariante, C.M., Caspi, A., Taylor, A., Poulton, R., 2007. Childhood maltreatment predicts adult inflammation in a life-course study. PNAS 104 (4), 1319–1324. https://doi.org/10.1073/pnas.0610362104.
Dargél, A.A., Godin, O., Kapczinski, F., Kupfer, D.J., Leboyer, M., 2015. C-reactive protein alterations in bipolar disorder: a meta-analysis. J. Clin. Psychiatry 76 (2), 142–150. https://doi.org/10.4088/JCP.14r09007.
Goldsmith, D.R., Rapaport, M.H., Miller, B.J., 2016. A meta-analysis of blood cytokine network alterations in psychiatric patients: Comparisons between schizophrenia, bipolar disorder and depression. Mol. Psychiatry 21 (12), 1696–1709. https://doi.org/10.1038/mp.2016.3.
Hu, Z., Cui, E., Chen, B., Zhang, M., 2025. Association between cigarette smoking and the risk of major psychiatric disorders: a systematic review and meta-analysis in depression, schizophrenia, and bipolar disorder. Front. Med. 12, 1529191. https://doi.org/10.3389/fmed.2025.1529191.
Jha, M.K., Leboyer, M., Pariante, C.M., Miller, A.H., 2025. Should inflammation be a specifier for major depression in the DSM-6? JAMA Psychiat. 82 (6), 549–550. https://doi.org/10.1001/jamapsychiatry.2025.0206.
Jiang, W., He, Y., Liu, Q., Peng, S., Ni, Y., Zhong, X., Guo, L., 2025. Associations between childhood maltreatment, peripheral immune biomarkers, and psychiatric symptoms in adults: a cohort study of over 138,000 participants. Brain Behav. Immun. 123, 840–850. https://doi.org/10.1016/j.bbi.2024.10.034.
Jones, G.H., Vecera, C.M., Pinjari, O.F., Machado-Vieira, R., 2021. Inflammatory signaling mechanisms in bipolar disorder. J. Biomed. Sci. 28 (1), 45. https://doi.org/10.1186/s12929-021-00742-6.
Khandaker, G.M., Dantzer, R., Jones, P.B., 2017. Immunopsychiatry: Important facts. Psychol. Med. 47 (13), 2229–2237. https://doi.org/10.1017/S0033291717000745.
Kowarik, A., Templ, M., 2016. Imputation with the R package VIM. J. Stat. Softw. 74 (7), 1–16. https://doi.org/10.18637/jss.v074.i07.
Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26. https://doi.org/10.18637/jss.v028.i05.
Leboyer, M., Godin, O., Llorca, P.M., Aubin, V., Bellivier, F., Belzeaux, R., Courtet, P., Costagliola, D., Dubertret, C., M'Bailara, K., Haffen, E., Henry, C., Laouamri, H., Passerieux, C., Pelletier, A., Polosan, M., Roux, P., Schwan, R., Samalin, L., Etain, B., 2022. Key findings on bipolar disorders from the longitudinal FondaMental Advanced Center of Expertise-Bipolar Disorder (FACE-BD) cohort. J. Affect. Disord. 307, 149–156. https://doi.org/10.1016/j.jad.2022.03.053.
Leboyer, M., Godin, O., Terro, E., Boukouaci, W., Lu, C.-L., Andre, M., Aouizerate, B., Berna, F., Barau, C., Capdevielle, D., Clauss-Kobayashi, J., Chereau, I., D Amato, T., Dubertret, C., Dubreucq, J., Fond, G., Laouamri, H., Leignier, S., Lancon, C., FACE-SZ (FondaMental Academic Centers of Expertise for Schizophrenia) Groups, 2021. Immune signatures of treatment-resistant schizophrenia: a FondaMental Academic Centers of Expertise for Schizophrenia (FACE-SZ) study. Schizophrenia Bulletin Open, 2(1), sgab012. Doi: 10.1093/schizbullopen/sgab012.
Mayer, M., 2024. missRanger: Fast imputation of missing values (Version 2.6.1) [R package]. CRAN. https://CRAN.R-project.org/package=missRanger.
Melville, J., 2025. uwot: The uniform manifold approximation and projection (UMAP) method for dimensionality reduction (Version 0.2.3) [R package]. CRAN. https://CRAN.R-project.org/package=uwot.
Milaneschi, Y., Kappelmann, N., Ye, Z., Lamers, F., Moser, S., Jones, P.B., Burgess, S., Penninx, B.W.J.H., Khandaker, G.M., 2021. Association of inflammation with depression and anxiety: evidence for symptom-specificity and potential causality from UK Biobank and NESDA cohorts. Mol. Psychiatry 26 (12), 7393–7402. https://doi.org/10.1038/s41380-021-01188-w.
Miller, A.H., Raison, C.L., 2016. The role of inflammation in depression: from evolutionary imperative to modern treatment target. Nat. Rev. Immunol. 16 (1), 22–34. https://doi.org/10.1038/nri.2015.5.
Miller, A.H., Berk, M., Bloch, G., Briquet-Laugier, V., Brouillon, C., Cuthbert, B.N., Dantzer, R., Davis, M.C., De Picker, L.J., Drevets, W.C., Eyre, H.A., Hack, L.M., Harrison, N.A., Krystal, A.D., Lombardo, G., Mondelli, V., Pariante, C.M., Pulvirenti, L., Salvadore, G., Leboyer, M., 2025. Advancing precision psychiatry and targeted treatments: Insights from immunopsychiatry. Brain Behav. Immun. 125, 319–329. https://doi.org/10.1016/j.bbi.2025.01.002.
Müller, N., 2018. Inflammation in schizophrenia: pathogenetic aspects and therapeutic considerations. Schizophr. Bull. 44 (5), 973–982. https://doi.org/10.1093/schbul/sby024.
Ninla-Aesong, P., Puangsri, P., Kietdumrongwong, P., Jongkrijak, H., Noipha, K., 2023. Being overweight and obese increases suicide risk, the severity of depression, and the inflammatory response in adolescents with major depressive disorders. Front. Immunol. 14, 1197775. https://doi.org/10.3389/fimmu.2023.1197775.
Orsolini, L., Pompili, S., Volpe, U., 2023. C-Reactive Protein (CRP): a potent inflammation biomarker in psychiatric disorders. Adv. Exp. Med. Biol. 1411, 135–160. https://doi.org/10.1007/978-981-19-7376-5_7.
Penninx, B.W.J.H., Lamers, F., Jansen, R., Berk, M., Khandaker, G.M., De Picker, L., Milaneschi, Y., 2025. Immuno-metabolic depression: from concept to implementation. The Lancet Regional Health. Europe 48, 101166. https://doi.org/10.1016/j.lanepe.2024.101166.
Pereira, A.C., Oliveira, J., Silva, S., Madeira, N., Pereira, C.M.F., Cruz, M.T., 2021. Inflammation in bipolar disorder (BD): Identification of new therapeutic targets. Pharmacol. Res. 163, 105325. https://doi.org/10.1016/j.phrs.2020.105325.
Posit Team, 2023. RStudio: Integrated development environment for R. Posit Software, PBC. http://www.posit.co/.
R Core Team, 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.
Robinson, L., 2019. demoGraphic: Providing demographic table with the p-value, standardized mean difference value (Version 0.1.0) [R package]. CRAN. https://CRAN.R-project.org/package=demoGraphic.
Saitoh, K., Lajnef, M., Boukouaci, W., Andre, M., Andrieu, C., Aouizerate, B., Barau, C., Bouassida, J., Capdevielle, D., Chereau, I., Clauss-Kobayashi, J., Coulon, N., Dorey, J.-M., Dubertret, C., Fakra, E., Fond, G., Goze, T., Le Corvoisier, P., Leignier, S., Tamouza, R., 2025. Pro-inflammatory processes and metabolic syndrome: combined risk of resistance to treatment in patients with schizophrenia from the FACE-SZ cohort. Psychiatry Res. 350, 116557. https://doi.org/10.1016/j.psychres.2025.116557.
Schürhoff, F., Fond, G., Berna, F., Bulzacka, E., Godin, O., Boyer, L., Misdrahi, D., Andrianarisoa, M., Brunel, L., Coulon, N., Aouizerate, B., Capdevielle, D., Chereau, I., D'Amato, T., Dubertret, C., Dubreucq, J., Faget, C., Gabayet, F., Mallet, J., FACE-SZ (FondaMental Academic Centers of Expertise for Schizophrenia) group, 2019. The 10-year findings from the FondaMental Academic Center of Expertise for Schizophrenia (FACE-SZ): review and recommendations for clinical practice. L'encephale 45 (1), 9–14. https://doi.org/10.1016/j.encep.2018.07.007.
Stekhoven, D.J., 2022. missForest: nonparametric missing value imputation using random forest. (Version 1.5) [R package]. CRAN. https://cran.r-project.org/web/packages/missForest/index.html.
Strawbridge, R., Arnone, D., Danese, A., Papadopoulos, A., Herane Vives, A., Cleare, A.J., 2015. Inflammation and clinical response to treatment in depression: a meta-analysis. Eur. Neuropsychopharmacol. 25 (10), 1532–1543. https://doi.org/10.1016/j.euroneuro.2015.06.007.
Treadway, M.T., Etuk, S.M., Cooper, J.A., Hossein, S., Hahn, E., Betters, S.A., Liu, S., Arulpragasam, A.R., DeVries, B.A.M., Irfan, N., Nuutinen, M.R., Wommack, E.C., Woolwine, B.J., Bekhbat, M., Kragel, P.A., Felger, J.C., Haroon, E., Miller, A.H., 2024. A randomized proof-of-mechanism trial of TNF antagonism for motivational deficits and related corticostriatal circuitry in depressed patients with high inflammation. Mol. Psychiatry. https://doi.org/10.1038/s41380-024-02751-x.
Upthegrove, R., Khandaker, G.M., 2020. Cytokines, oxidative stress and cellular markers of inflammation in schizophrenia. Curr. Top. Behav. Neurosci. 44, 49–66. https://doi.org/10.1007/7854_2018_88.
Van den Noortgate, M., Morrens, M., Foiselle, M., De Picker, L., 2025. Immune dysregulation in psychiatric disorders with and without exposure to childhood maltreatment: a transdiagnostic stratified meta-analysis. Brain Behav. Immun. 127, 193–204. https://doi.org/10.1016/j.bbi.2025.03.002.
Wessa, C., Janssens, J., Coppens, V., El Abdellati, K., Vergaelen, E., van den Ameele, S., Baeken, C., Zeeuws, D., Milaneschi, Y., Lamers, F., Penninx, B., Claes, S., Morrens, M., De Picker, L., 2024. Efficacy of inflammation-based stratification for add-on celecoxib or minocycline in major depressive disorder: Protocol of the INSTA-MD double-blind placebo-controlled randomised clinical trial. Brain, Behavior, & Immunity Health 41, 100871. https://doi.org/10.1016/j.bbih.2024.100871.
Wright, M.N., Ziegler, A., 2017. ranger: a fast implementation of random forests for high dimensional data in C++ and R. J. Stat. Softw. 77 (1), 1–17. https://doi.org/10.18637/jss.v077.i01.
Yrondi, A., Bennabi, D., Haffen, E., Garnier, M., Bellivier, F., Bourgerol, T., Camus, V., D'Amato, T., Doumy, O., Haesebaert, F., Holtzmann, J., Lançon, C., Vignaud, P., Moliere, F., Nieto, I., Richieri, R.M., Domenech, P., Rabu, C., Mallet, L., Aouizerate, B., 2017. Significant need for a French network of expert centers enabling a better characterization and management of treatment-resistant depression (Fondation FondaMental). Front. Psych. 8, 244. https://doi.org/10.3389/fpsyt.2017.00244.