

Article

Discrimination of Bipolar Disorder and Schizophrenia Patients Based on LC-HRMS Lipidomics

Milan R. Janković ¹, Nataša Avramović ^{2,*} , Zoran Miladinović ³ , Milka B. Jadranin ⁴ , Marija Takić ⁵ , Gordana Krstić ¹, Aleksandra Gavrilović ⁶, Čedo Miljević ⁷, Maja Pantović ⁸, Zorana Andrić ⁹ , Savvas Radević ⁹ , Danica Savić ⁴ , Stefan Lekić ⁴, Vele Tešević ¹  and Boris Mandić ^{1,*}

¹ University of Belgrade—Faculty of Chemistry, Studentski Trg 12–16, 11000 Belgrade, Serbia; mr.m.jankovic@gmail.com (M.R.J.); gkrstic@chem.bg.ac.rs (G.K.); vtsevic@chem.bg.ac.rs (V.T.)

² Institute of Medical Chemistry, University of Belgrade—Faculty of Medicine, Višegradska 26, 11000 Belgrade, Serbia

³ Institute of General and Physical Chemistry, Studentski Trg 12–16, 11158 Belgrade, Serbia; zmiladinovic@iofh.bg.ac.rs

⁴ Department of Chemistry, University of Belgrade—Institute of Chemistry, Technology and Metallurgy, Njegoševa 12, 11000 Belgrade, Serbia; milka.jadranin@ihtm.bg.ac.rs (M.B.J.); danica.savic@ihtm.bg.ac.rs (D.S.); stefan.lekic@ihtm.bg.ac.rs (S.L.)

⁵ Group for Nutrition and Metabolism, Center of Research Excellence for Nutrition and Metabolism, University of Belgrade—Institute for Medical Research, National Institute of Republic of Serbia, Tadeuša Košćuška 1, 11000 Belgrade, Serbia; marija.takic@imi.bg.ac.rs

⁶ Special Hospital for Psychiatric Diseases “Kovin”, Cara Lazara 253, 26220 Kovin, Serbia; gavrilovicaleksandra74@gmail.com

⁷ Institute of Mental Health, University of Belgrade—Faculty of Medicine, Milana Kašanina 3, 11000 Belgrade, Serbia; cedo.miljevic@imh.org.rs

⁸ Clinic of Psychiatry, University Clinical Center of Serbia, Pasterova 2, 11000 Belgrade, Serbia; pantovicm83@gmail.com

⁹ Blood Transfusion Institute of Serbia, Svetog Save 39, 11000 Belgrade, Serbia; z.andric@nbt.org.rs (Z.A.); s.radevic@nbt.org.rs (S.R.)

* Correspondence: natasa.avramovic@med.bg.ac.rs (N.A.); borism@chem.bg.ac.rs (B.M.)

Abstract

Background/Objectives: Schizophrenia (SCH) and bipolar disorder (BD) share overlapping symptoms and genetic factors, making differential diagnosis challenging and often leading to misdiagnosis. This study aimed to identify potential lipid biomarkers of serum capable of distinguishing BD from SCH. **Methods:** Lipid profiles of serum from 30 SCH and 31 BD patients were analyzed in triplicates using liquid chromatography–high-resolution mass spectrometry (LC-HRMS). Chemometric analysis was applied, including class and gender identifiers. Orthogonal partial least squares (OPLS) models with 1000 cross-validations were used to validate feature subsets. **Results:** The chemometric analysis included the most relevant metabolites in the comparison between all samples of SCH and BD patients, identifying five key biomarkers (LPC 16:0, SM 33:1, SM 32:1, compound C30H58O3, and PC 30:0) with VIP scores > 1 for distinguishing BD from SCH. Gender-specific models revealed five biomarkers in males (SM 32:1, SM 33:1, PC 32:1, PC 30:0, and FA 16:1) and two in females (LPC 16:0 and C30H58O3). These biomarkers primarily belonged to glycerophospholipids (GPs) and sphingophospholipids (SPs). **Conclusions:** Comparative lipid profiling between SCH and BD, including gender-specific subgroups, enabled identification of potential diagnosis-specific biomarkers. Elevated levels of GPs and SPs in SCH patients suggest lipid metabolism differences that may support improved diagnostic accuracy and personalized treatment strategies.

Keywords: lipidomics; biomarkers; mass spectrometry; schizophrenia; bipolar disorder; serum



Received: 12 November 2025

Revised: 31 December 2025

Accepted: 8 January 2026

Published: 12 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Schizophrenia (SCH) and bipolar disorder (BD) are complex psychiatric diseases characterized by profound cognitive, behavioral, and emotional impairments [1–3]. They affect approximately 1% of the world's population and their economic burden in the United States is estimated to exceed 300 billion dollars annually [4–7]. Based on the fact that the estimated burden of SCH in the United States doubled between 2013 and 2019 [8], effective strategies and treatment options are essential to improve the management of this disease. Schizophrenia and bipolar disorder arise from intricate interactions between genetic and environmental factors, and their underlying biological mechanisms are still not fully understood. The criteria for their psychiatric diagnoses are entirely based on observable symptoms, relying on the subjective assessments of psychiatrists following a structured clinical interview and patient or caregiver reports. This process is not only time-consuming but also limited by the clinician's judgment. The shared symptoms among SCH and BD (with psychotic symptoms), such as delusions and hallucinations, further complicate accurate diagnosis. In addition to gaps in understanding and diagnosing these disorders, there is an urgent need for personalized, targeted approaches to treatment, including individualized medical therapies and reliable methods for monitoring treatment effectiveness [9,10].

'Omic' research, which analyzes a wide range of biomolecules within biological samples, has significant potential for uncovering new diagnostic markers that could pave the way for better identification and treatment of SCH and BD. These biomarker sets could be developed using advanced techniques such as nuclear magnetic resonance (NMR), liquid chromatography-mass spectrometry (LC-MS) and gas chromatography-mass spectrometry (GC-MS) [11–17]. Metabolomics and lipidomics analyses, combined with chemometric assessments of serum and urine samples from both patients and healthy individuals, could provide valuable insights [12–20]. In recent years, chemometric methods have been increasingly developed to help extract useful information from lipidomic data, playing a significant role in the identification and validation of potential biomarkers and contributing to improved disease diagnosis and monitoring [18–26].

Lipids comprise a heterogeneous group of molecular subclasses that exert essential regulatory functions in neuronal development and activity, modulate membrane-associated neuronal plasticity [27–29], and influence brain energy metabolism [30–32].

Previous LC-MS studies on schizophrenia [18,19,21–23] and bipolar disorder [20,24–26] indicate that potential lipid biomarkers may belong to various serum and plasma lipid categories, including glycerophospholipids (GP), sphingolipids (SP), glycerolipids (GL), sterol lipids (ST), and fatty acyls (FA). Several recent lipidomics studies suggest that antipsychotic treatments in SCH patients could lead to the alteration of serum lipid profiles, including the downregulation of shorter-chain triacylglycerols [21,33,34]. Diverse lipidomics studies have highlighted the strong influence of lipid metabolism in SCH and BD [18–26]. For instance, abnormalities in phospholipid metabolism have been strongly linked to SCH [18,19,21–23], while lower levels of omega-3 polyunsaturated fatty acids have been observed in individuals with both SCH and BD, suggesting possible therapeutic applications [35,36]. However, no biomarkers have yet been approved for diagnosing psychiatric disorders.

Our research group recently studied lipid profiles of SCH (30) and BD (31) patients compared to their controls (SCH-C 31 and BD-C 31), applying liquid chromatography coupled with high-resolution mass spectrometry (LC-HRMS) [19,20], and established the alteration of the majority of lipid classes: GP, SP, GL, fatty acids (FA), and cholesterol esters (CE), indicating a key role of lipid pathways in the pathogenesis of SCH and BD. These findings about the alterations of lipid profiles in SCH and BD could enable the discovery of

potential biomarkers, enhance SCH and BD diagnosis, and support the development of more effective treatments [18–26].

The aim of this work is to compare the lipid profiles of sera from SCH and BD patients and explore the potential for their differentiation by combining obtained LC-HRMS-based lipid data with chemometric analysis.

2. Materials and Methods

2.1. Sampling

This study followed ethical guidelines from Kovin Psychiatric Hospital and the University of Belgrade Faculty of Chemistry. Blood samples were collected from 30 SCH patients and 31 BD patients at Kovin Hospital. All participants or their guardians gave written consent. No statistically significant differences were found between the two groups of patients in terms of age, gender, or BMI. SCH patients (100%) were treated with anxiolytics, while 6.7% used first-generation antipsychotics, 50% second-generation, and 43.3% both. BD patients were treated with first-generation antipsychotics (6.45%), second-generation antipsychotics (77.42%), and anxiolytics (16.13%). Blood serum samples were collected after at least 8 h of fasting and each sample was prepared in triplicate for analysis.

2.2. Sample Preparation

Blood samples were kept on ice for one hour, then centrifuged. Serum was collected and stored at $-80\text{ }^{\circ}\text{C}$ for up to two weeks. Lipids were extracted in triplicate following O'Brien et al. (2019) [37]. LC-HRMS analysis and data processing of lipid extracts of bipolar disorder (BD) and schizophrenia (SCH) patients, together with extraction blanks and pool samples, were performed with slight modifications as described in Jadranin et al. (2023) [20].

2.3. Software

Data processing and chemometric analysis were performed using MATLAB 9.7 (MathWorks, Natick, MA, USA) [38] with suitable toolboxes and in-house routines. The subroutine for the Borda count method was adapted from Python's *mlpy* module version 3.0 [39], while LC-HRMS chromatograms and spectra were processed using the R package XCMS 3.16.1 [40] and then imported into MATLAB workspace. The Robust PCA (ROBPCA) method [41,42] from the LIBRA package [43] was used alongside classical PCA for efficient outlier detection. Preprocessing and chemometric analysis were conducted with PLS Toolbox 8.9.1 [44].

2.4. Reading in Data

After the cleaning of background noise, and isotopic and unrelated ions, the initial dataset included 183 chromatograms: 93 triplicates from 31 individuals with BD and 90 triplicates from 30 individuals with SCH, and 183 variables (featuring or one or more ions per metabolite with characteristic m/z and retention time values) belonging to the 129 metabolites for combined negative–positive ions. Additional categorical variables, such as class, gender, triplicate identifiers, and grouping, were included in final dataset structure [19,45,46]. Replicated samples were used to identify different acquired LC-MS chromatograms originating from the same individuals, which also facilitated feature validation.

2.5. Data Pretreatment (Preprocessing)

Before further processing, data pretreatment steps were necessary, including handling missing data through removal or imputation, data transformation, normalization, and the final centering and scaling of all variables. Since centering and scaling depend on sample

size, these steps were applied before model assembly and cross-validation. Excluding observations with missing values results in further dataset reduction; hence, replacement or imputation represent a more reliable approach in situations where a relatively low number of observations are present in the dataset [19]. Linear interpolation of neighboring non-missing values was most suitable for this dataset [47], though a moving median (over a 3-point window) produced similar outcomes. As result, all identified missing values were handled using linear interpolation.

For LC-MS metabolomics data, performing logarithmic transformation (e.g., replacing each value x with $\log_{10}(x)$ or $\log_2(x)$) [48,49] reduced extreme values, producing homoscedastic and near-normal residuals. Figure S1a (Supplementary Materials) shows boxplots of $\log_{10}$ -transformed data grouped by classes ("SCH" for schizophrenia and "BD" for bipolar disorder), while Figure S1b (Supplementary Materials) presents the same data, mean-centered after logarithmic transformation.

Scaling is often paired with centering, with "autoscaling" being the most common method. Autoscaling involves mean centering each metabolite's data (subtracting the sample mean) and dividing by the standard deviation [48], ensuring a unit standard deviation for unbiased comparisons. The $\log_{10}$ -transformed results in Figure S1a (Supplementary Materials) show significantly reduced skewness compared to untransformed data. Additionally, Figure S1b (Supplementary Materials) highlights the benefits of mean centering and scaling to the standard deviation for future modeling.

Hence, this study applies consistent $\log_{10}$ transformation followed by autoscaling (mean centering and scaling to standard deviation) to all variables in the dataset.

2.6. Data Normalization

Principal Component Analysis (PCA) of the samples processed across four consecutive batches (Figure S2, Supplementary Materials) revealed no significant drift. Additionally, the comparison of intra-day and inter-day coefficients of variation (CV) for filtered m/z (rt) values in quality control (QC) samples showed CVs below 30%. Therefore, fluctuations in instrument sensitivity during measurements were deemed negligible, and no further normalization of the samples was performed.

2.7. PCA and Outlier Detection/Identification

Classical PCA was applied to the initial datasets (overall and gender-specific subsets), but it revealed no significant variation between groups or classes. However, classical PCA, combined with robust PCA (ROBPCA), was used for outlier identification. ROBPCA outlier maps, based on score and orthogonal distances, identified "bad leverage points" as relevant outliers. Screeplots (LIBRA package) determined the number of PCA components for each dataset for the whole dataset, seven samples (33, 48, 110, 111, 130, 175, and 176) were flagged as outliers (Figure S3a, Supplementary Materials). Outlier maps for gender-specific subsets (Supplementary Materials Figure S3b,c) revealed outliers: three samples (33, 65, and 66) in the male subset, and three samples (3, 91, and 92) in the female subset. All identified outliers based on this method (robust PCA) for gender-specific subsets match with corresponding outliers identified for the complete dataset. Further analysis with classical PCA identified one additional outlier in each gender-specific subset. This way, a total of seven distinctive outliers were identified in the main dataset, with four outliers in each of the gender-specific subsets (male and female participants). Figure S4a,b show the skewness and kurtosis results before and after removing the seven outliers from the main dataset.

The results in Figure S4 (Supplementary Materials) show a significant reduction in skewness and kurtosis after removing the identified outliers, particularly in the m/z range

of 675–850. Consequently, all identified outliers were excluded from the corresponding datasets and subsets prior to further LC-MS data analysis.

2.8. Cross-Validation (CV) and Feature Validation Dataset Partition

In this work, a 5-fold contiguous block CV method was used, keeping triplicates from the same patients in the same block. CV blocks were formed by splitting the dataset and randomly shuffling samples within each block while maintaining triplicate structures [45]. For thorough feature validation, an additional observation subsampling approach was applied. For each CV fold group, one triplicate per sample was randomly selected. The subsampling process involved three steps: (1) partitioning the dataset into folds, grouping samples with their triplicates (repeated *iter1* times); (2) for each fold group, selected samples are gathered samples from all other remaining folding groups to create *n* groups of observations for further analysis; and (3) randomly sub-selecting one triplicate per sample in each fold (repeated *iter2* times). The final matrix of logical indices of selected observation had dimensions (*iter1* × *iter2* × *fold*, *N*), where *N* is the total number of observations of each subset, and *fold* represents the number of used fold groups. For the current investigation following parameter settings were consistently employed: *iter1* = 100, *iter2* = 500, and *fold* was set to 5 in all cases. Full details are available in [19] and the Supplementary Materials of the current work.

2.9. Classification Performance Metrics for Model Evaluation

To compare classification performance across models, various performance metrics can be utilized. Selecting and interpreting the right metrics is crucial for accurate model evaluation, as relying on a single metric might not suffice. Instead, combining metrics provides better insights for model fine-tuning during development.

In this study, the OPLS-DA model was evaluated using a confusion matrix derived from cross-validated predicted class labels. Metrics such as accuracy, sensitivity, and specificity were calculated from the matrix:

Accuracy: The ratio of correctly classified samples to total samples, especially effective for near equally balanced datasets.

Sensitivity: Depends on correctly identifying the SCH group as positive samples.

Specificity: Relies on accurately classifying the BD group as negative samples.

In all cases, SCH was treated as the positive class and BD as the negative class. These metrics were specifically expressed for this study based on this classification framework:

$$Accuracy = \frac{T_{SCH} + T_{BD}}{T_{SCH} + T_{BD} + F_{SCH} + F_{BD}}$$

$$Sensitivity = \frac{T_{SCH}}{T_{SCH} + F_{BD}}, \text{ Specificity} = \frac{T_{BD}}{F_{SCH} + T_{BD}}$$

Correctly classified SCH samples are marked as T_{SCH} , while those misclassified as BD are F_{BD} . Similarly, BD samples are counted as T_{BD} if correctly classified, and F_{SCH} if misclassified.

Accuracy, though widely used, is sensitive to imbalanced datasets where one class outnumbers the other [50]. In this study, the SCH to BD ratio ranges from 1 (entire dataset) to 1.22 (male) and 0.86 (female), ensuring the least-populated class represents no less than 45% of the data. However, using metrics insensitive to class imbalance is a safer choice when multiple methods are evaluated together.

Apart from sensitivity and specificity, which are robust to imbalance, another complementary metric is class error (or Balance Error Rate/Half Total Error Rate) [50,51]. This uses $1 - \text{sensitivity}$ and $1 - \text{specificity}$ for evaluation.

$$\text{classerror} = 1 - \frac{(\text{sensitivity} + \text{specificity})}{2} = \frac{1}{2} \left(\frac{F_{SCH}}{F_{SCH} + T_{BD}} + \frac{F_{BD}}{T_{SCH} + F_{BD}} \right)$$

The AUROC (Area Under the ROC Curve) is a metric used to numerically compare classifiers by reducing the ROC curve to a single scalar value that represents classification performance [52]. In this study, AUROC was calculated using cross-validated predictions and reflects the probability that the OPLS-DA classifier ranks a randomly chosen SCH instance higher than a BD instance [48]. A perfect classifier yields an AUROC of 1.0, while an AUROC of 0.5 indicates random classification. No classifier realistically falls below 0.5 [50].

While PLS-DA is a classification model, it is fundamentally based on the PLS regression algorithm [53–55]. Additionally, RMSECV (Root Mean Square Error by Cross-Validation) was introduced for continuity with prior publications [19,45] and to compare its diagnostic utility alongside classification metrics in this study.

3. Results

3.1. OPLS-DA Models

In assembled PCA models, no distinctive patterns were observed among scores for the class of interest, consistent with their unsupervised nature and limited utility for biomarker identification [48] (results for gender-specific datasets are depicted in Figure 1). PCA performs best when among-group variability exceeds within-group variation. In contrast, supervised methods like PLS-DA and OPLS-DA better highlight class distinctions by linking predictors to outcomes [53,56,57], emphasizing within-class similarities, and at the same time, between-class differences. To avoid overfitting, models were validated via cross-validation and external testing [56,58]. OPLS-DA parameter optimization often involves mutually combining different CV methods and model performance metrics. To convert the obtained output of OPLS-DA self-predicted continuous values of responses into reliable binary labels, thresholds for each class were determined using a Bayesian method [51,59,60].

In addition to the developed cross-validation (CV) method (detailed explanation in the Experimental section) for datasets containing triplicates, venetian blinds and random subset-selected CV methods [61] were also applied. Determining the optimal model complexity, such as the number of latent variable (LV) components, is a critical step for PLS-DA modeling. For this study, a random sampling and further sub-selection method was used to identify the most reliable number of LV components for model assessment. The procedure can be summarized as follows: (1) Randomly partition the main dataset into 5-fold groups of observations, maintaining triplicate structures. Sub-partitioning by male and female groups was also performed during optimization. (2) Independently sub-select one triplicate per sample within each fold, repeating this 1000 times.

For each subset, OPLS-DA models with up to 10 LV components were generated. The optimal number of LV components was determined based on the minimum RMSECV (Root Mean Square of Error by Cross-Validation), minimum class error, and maximum AUROC values. Figure S5 in Supplementary Material presents histograms for each fold, showing sub-selected samples with one triplicate and “the venetian blinds” selection method applied during CV. As shown in Supplementary Materials Figure S5, three LV components were most frequently selected, providing optimal model performance across the selected metrics. Additionally, permutation tests confirmed this result, indicating that models with more

than three LV components are prone to overfitting (see Supplementary Materials Figure S6 for the OPLS-DA model in Figure 2). Consequently, all OPLS-DA models in this study were constructed using three LV components.

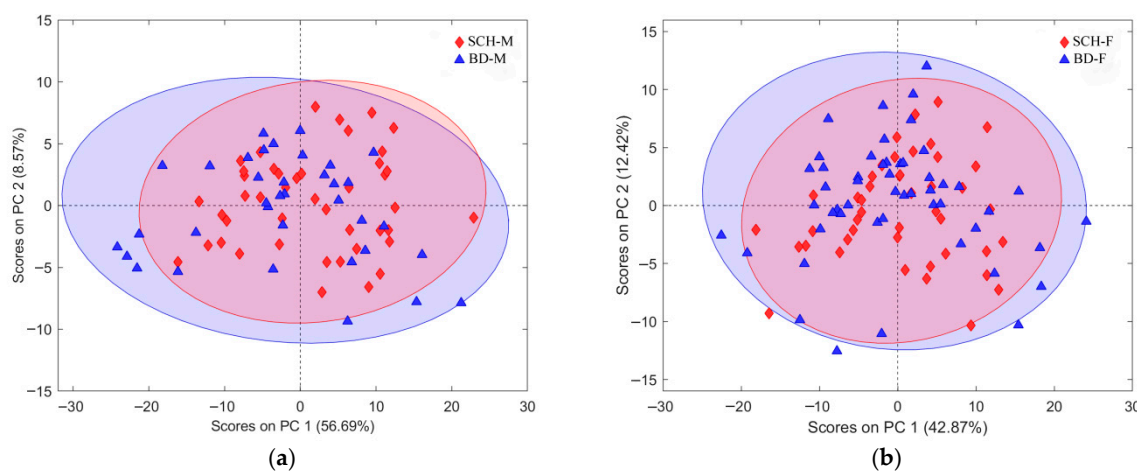


Figure 1. PCA model for dataset partitions with: (a) male subjects and all belonging features, using four PCA components; (b) female subjects and all belonging features, using four PCA components. Scores plots present first two PC components with assigned explained variability for each of them. Each group of patients was assigned a different color, BD as blue full triangles and SCH as red full diamonds.

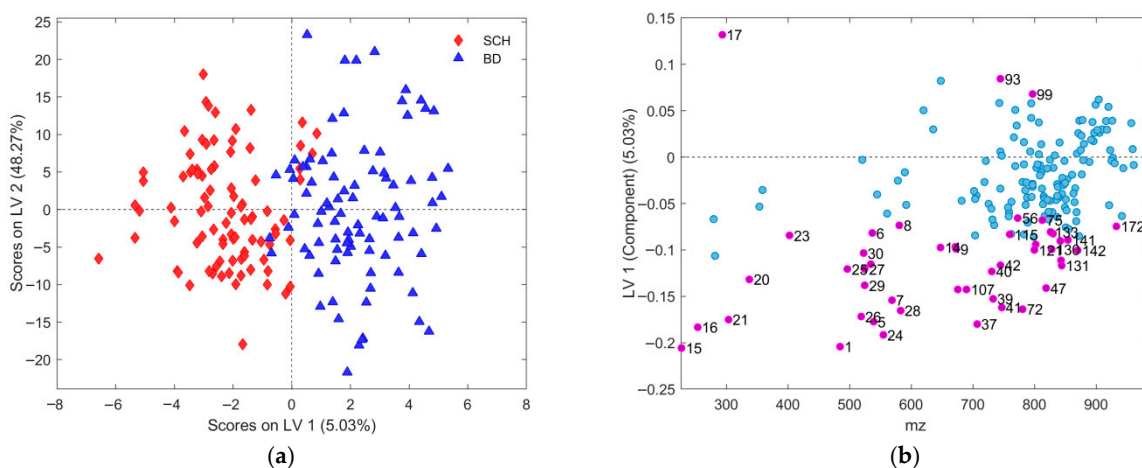


Figure 2. (a) Score plot of the OPLS-DA model composed of the dataset from all observations with excluded outliers, where discrimination classes belong to the BD (blue full triangles) and SCH (red full diamonds) groups of samples. (b) Loading plot of the predictive component LV1, where VIP scores > 1 determined from the model are presented in pink with assigned feature position numbers from the corresponding dataset positions. The accuracy was 0.8523, as well as the sensitivity and specificity, too, for this model. AUROC was determined to 0.9401.

Figure 2 presents the scores and loading plots of the predictive LV1 component for the best OPLS-DA model identified among 1000 models generated during CV, all using three latent variables (LV) and a dataset with all observation with outliers previously removed. The results for accuracy, sensitivity, and specificity for this model are 0.8523, whereas AUROC was 0.9401 at the same time. Variance captured along the predictive LV component was 5.03%, which, together with different performance metric distributions (details presented in Section 3.4), suggests the relatively poor performance of these models. Additionally, CV prediction results for both patient classes are presented in Figure 3.

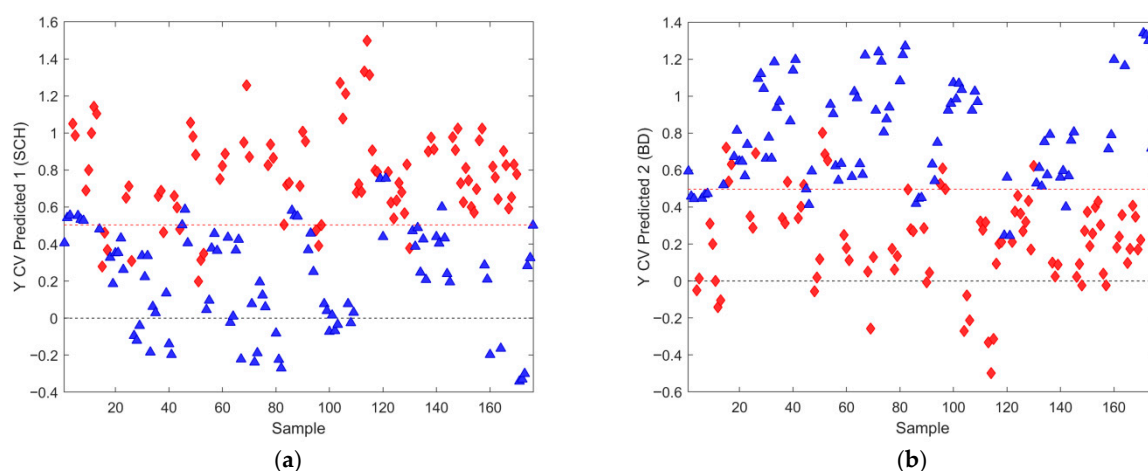


Figure 3. For the best selected OPLS-DA model presented in Figure 2, we present results for: (a) Cross-validated prediction results for class SCH (red filled diamonds), where the discrimination threshold was determined at 0.5036 (red dashed line) (b) Cross-validated prediction results for class BD (blue filled triangles), where the discrimination threshold was determined at 0.4964 (red dashed line).

3.2. Feature Selection and Validation

Selection of features (two-dimensional (m/z and retention time) LC/MS signals or ions) and their validation are essential for building efficient predictive models by identifying relevant variables and reducing dimensionality [62]. Variable Importance in Projection (VIP) is a commonly used method, with the “greater than one” rule ($VIP > 1$) serving as a typical selection threshold [63]. However, values between 0.83 and 1.21 may be more suitable in specific cases [64–66]. Stability can be improved through the bootstrap with replacement resampling technique, which produces multiple VIP estimates per variable [65]. Larger values for resampling enhance the handling of unsmooth and nonlinear data, but at the cost of increased computational time.

In our previous study [19], a similar approach with a slightly modified sub-sampling method based on k-fold partition of the dataset was introduced. For the present study, the procedure for feature selection and subsequent validation was slightly adopted and upgraded, detailed in the Supplementary Materials (*Features selection methods and algorithm*) [19,37,67,68].

Feature validation was extended by applying feature selection before each validation step. Features are ranked by VIP scores, the least informative ones are removed, and the remaining set is evaluated using an OPLS-DA model. This cycle repeats until optimal performance is reached [66]. The approach resembles stepwise backward elimination [66,69], with modifications that leverage the ranked VIP score list.

To evaluate OPLS-DA model performance in relation to the number of included features, feature validation plots were constructed using classification metrics derived from cross-validation (CV) confusion matrices and RMSECV. Accuracy and AUROC were used to estimate classification performance [48,50], while RMSECV and class error were plotted as loss curves. These curves provided insight into how model performance evolved by measuring classification error over cumulative feature subsets from the ranked list.

Validation curves were generated across cumulative subsets of ranked features. These typically exhibited a sharp initial performance increase [70], followed by an extreme point, maximum for accuracy and AUROC, minimum for RMSECV and class error, and a subsequent decline, often attributed to the inclusion of noise variables. Irregular patterns in the early curve regions, such as sporadic local minima or maxima, suggested the presence of less informative features and were considered during further subsetting refinement.

Model performance was evaluated across the full dataset, including both male and female subjects, using complete sets of features ranked by VIP scores (see *Features selection and validation* in Supplementary Materials). As shown in Figure 4, performance metrics exhibited a distinct peak around a subset of ~130 features, while a plateau was observed between ranks 17 and 43. Variables within this plateau were repositioned to the start of the sequence and re-evaluated, revealing their limited contribution to all model performance metrics. These features were subsequently excluded. The revised validation curves, shown in Figure 5, demonstrate a smoother and more consistent increase in performance across all metrics, with reduced sporadic disruption compared to the original ranked list. This suggests that careful feature selection and exclusion of low-impact variables can significantly enhance model robustness, on which validation curve results depend.

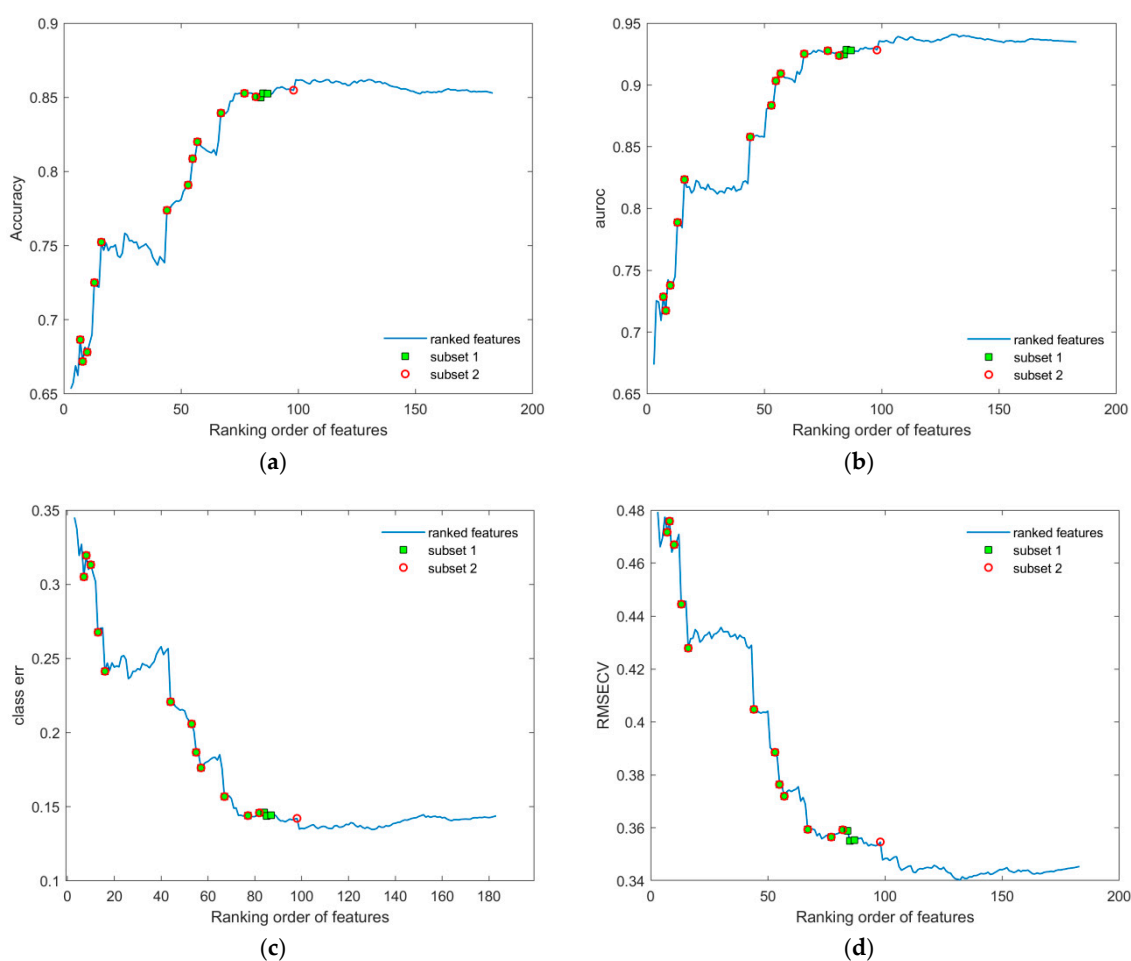


Figure 4. Several performance metrics are displayed in relation to the increasing sequence of sets of ranked features for the initial dataset (the number of features in the subsets that were created by adding features cumulatively, one by one, from the ranking list): (a) accuracy; (b) AUROC; (c) class error; and (d) RMSECV. Additionally, markers colored as green squares (subset 1) and red circles (subset 2) represent the smallest sets of features identified through the subsequent selection procedure (shown in Figure 6 and Table 1), and are included for comparison with the original feature list. The first two features were omitted, since the number of LV components was always 3 for each assembled OPLS-DA model during validation.

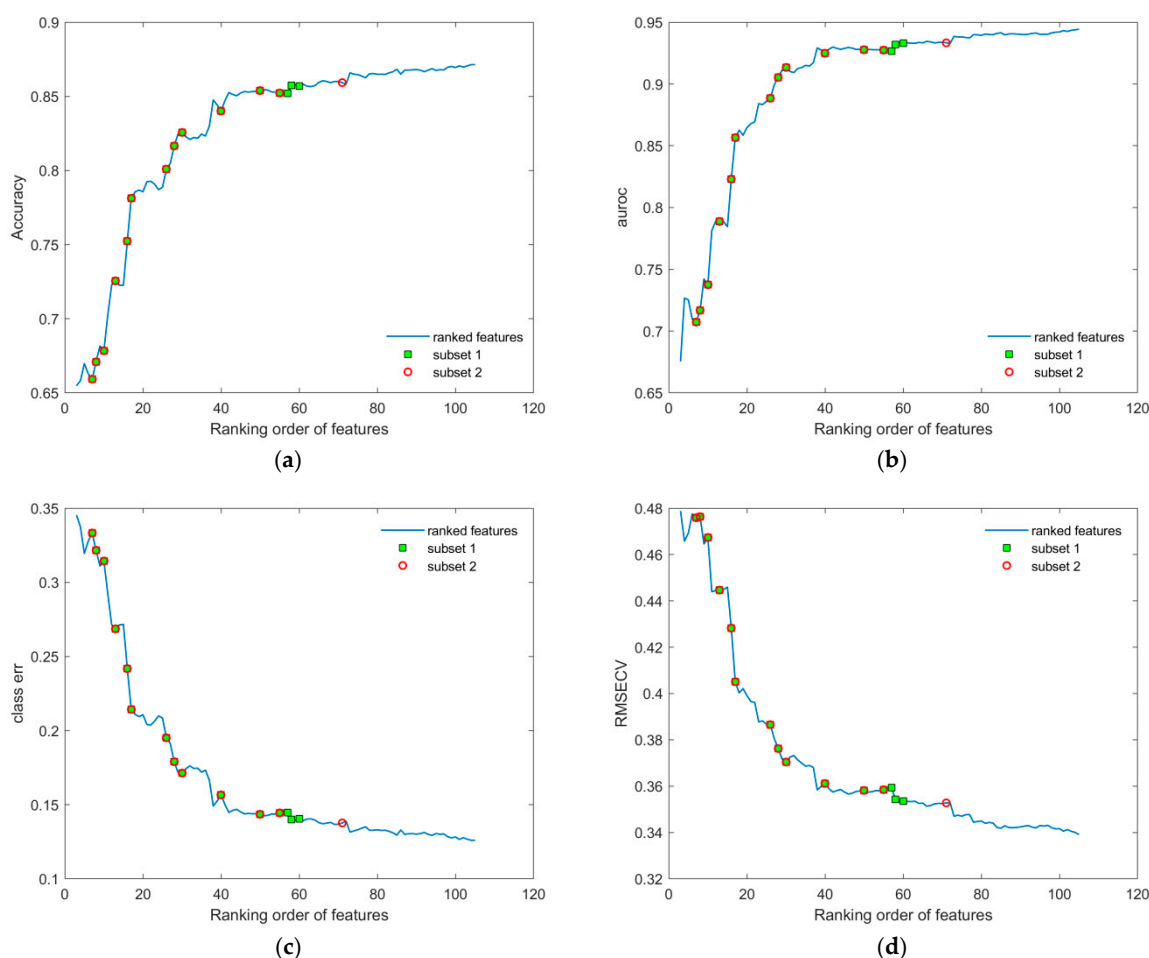


Figure 5. Performance metrics displayed in relation to the increasing sequence of sets of ranked features of the initial dataset: (a) accuracy; (b) AUROC; (c) class error; and (d) RMSECV. Additionally, markers colored as green squares (subset 1) and red circles (subset 2) represent the smallest sets of features identified through the subsequent selection procedure (shown in Figure 6 and Table 1), and are included for comparison with the original feature list. The first two features were omitted, since the number of LV components was always 3 for each assembled OPLS-DA model during validation.

To improve the resulting performance, we applied additional subsetting to the ranked feature list by removing noisy or redundant features. This approach, illustrated in Figure 6, significantly reduced computational load, particularly given the high cost of validation (e.g., 1000 cross-validation iterations per point). The final point in the sequence with the corresponding performance metric within each reduced subset was used as a benchmark for evaluating overall performance for the current list of ranked features.

Iterative pruning revealed the features that consistently degraded performance. Their removal led to measurable improvements in all performance metrics and stability. However, in the cases where performance declined after the removal of a particular feature, reverting to the previous optimal configuration step was necessary. Furthermore, reordering features around critical thresholds exposed additional low-impact features, enabling further refinement of the feature set.

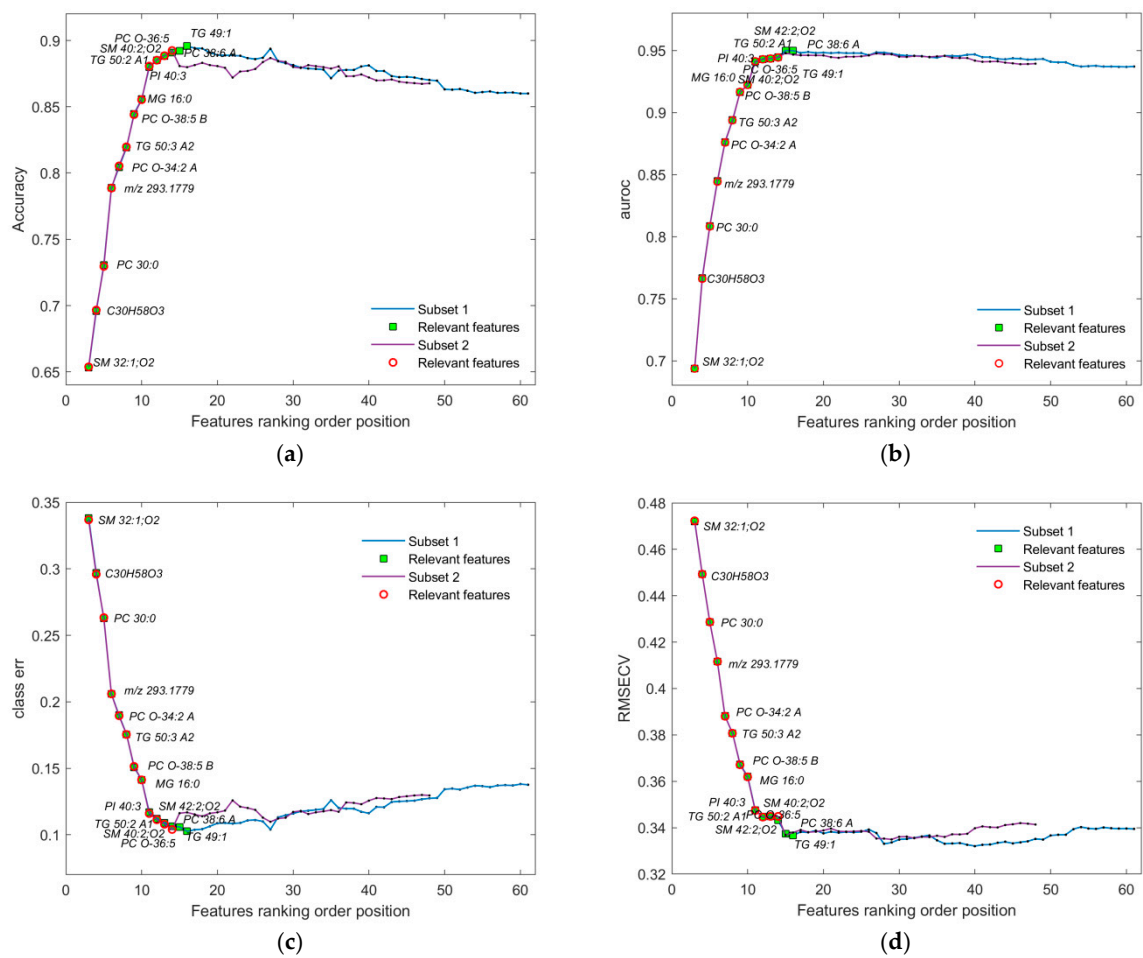


Figure 6. Performance metrics displaying the smallest subset of the identified subset of features for the initial dataset: (a) accuracy; (b) AUROC; (c) class error; and (d) RMSECV. Markers colored as green squares represent subset 1, consisting of the first 16 most relevant features out of the 61 total presented features on the plot. Markers colored as red circles represent subset 2, consisting of the first 14 most relevant features from the 48 total presented features. The first two features were omitted, since the number of LV components was always 3 for each assembled OPLS-DA model during validation. The figure results preview the data presented in Table 1.

Table 1. For the full dataset, a list of identified relevant features for subset 1, consisting of the 16 most relevant features from the 61 total features in this subset; and for subset 2, consisting of the 14 most relevant features from the 48 total features in this subset.

Subset 1		Subset 2	
ID Number	Feature Names	ID Number	Feature Names
24	LPC 16:0 A1	24	LPC 16:0 A1
107	SM 33:1;O2	107	SM 33:1;O2
106	SM 32:1;O2	106	SM 32:1;O2
1	C30H58O3	1	C30H58O3
37	PC 30:0	37	PC 30:0
17	m/z 293.1779	17	m/z 293.1779
93	PC O-34:2 A	93	PC O-34:2 A
144	TG 50:3 A2	144	TG 50:3 A2
101	PC O-38:5 B	101	PC O-38:5 B
35	MG 16:0	35	MG 16:0
103	PI 40:3	103	PI 40:3
140	TG 50:2 A1	140	TG 50:2 A1
118	SM 40:2;O2	118	SM 40:2;O2
123	SM 42:2;O2	98	PC O-36:5
82	PC 38:6 A		
135	TG 49:1		

The smallest subsets of features (1 and 2) identified through the subsequent selection procedure (Figures 4–6), together with the corresponding feature coding lists, are presented in Table 1. Their more detailed identification can be found in Table S1 (Supplementary Materials).

3.3. Data Subset with Male Participants

The same approach of the selection procedure as in the previous Section 3.2 also enabled subgroup-specific reductions, such as for male and female subjects. Figure S7 (Supplementary Materials) shows various model performance metrics across increasing subsets of features for male participants. The features range from N (the number of LV components in the current OPLS-DA model) to all 183 features included in the data subsets. The validation plot in Figure S7 highlights a features sequence region between 11 and 60 features, with an extreme used performance value near the 130th feature. Features beyond this point, and within the identified region, are strong candidates for elimination using the previously described subsetting techniques. Notably, the first eight features show a sharp improvement in classification performance, based solely on their frequency positions in the general ranked list (see the aggregation methods for ranking presented in Section 3.2, Supplementary Materials).

Through feature subsetting, the smallest subset (subset 4) providing the best performance metrics was identified. The validation plot in Figure 7 compares subset 4 with an intermediate subset (subset 3), which contains a few additional features but shows worse average performance for all OPLS-DA model metrics. Subset 4, consisting of 12 relevant features, is presented alongside subset 3 in Table 2. Their more detailed identification can be found in Table S1 (Supplementary Materials).

Table 2. For the male subset of samples, the list of identified relevant features for subset 3, consisting of the 20 most relevant features from the 37 overall features in this subset; and for subset 4, consisting of the 12 most relevant features from the 29 total features in this subset.

Subset 3		Subset 4	
ID Number	Feature Names	ID Number	Feature Names
106	SM 32:1;O2	106	SM 32:1;O2
39	PC 32:1	39	PC 32:1
107	SM 33:1;O2	107	SM 33:1;O2
37	PC 30:0	37	PC 30:0
16	FA 16:1	16	FA 16:1
5	Cer 34:1;O2 B	93	PC O-34:2 A
93	PC O-34:2 A	35	MG 16:0
46	PC 34:0	99	PC O-38:4
35	MG 16:0	132	TG 48:2 A1
99	PC O-38:4	143	TG 50:3 A1
51	PC 34:2 A2	97	PC O-36:4
129	TG 48:1 A1	183	TG 60:3
59	PC 36:2 A1		
149	TG 51:2 A1		
118	SM 40:2;O2		
132	TG 48:2 A1		
143	TG 50:3 A1		
166	TG 54:5 A1		
97	PC O-36:4		
183	TG 60:3		

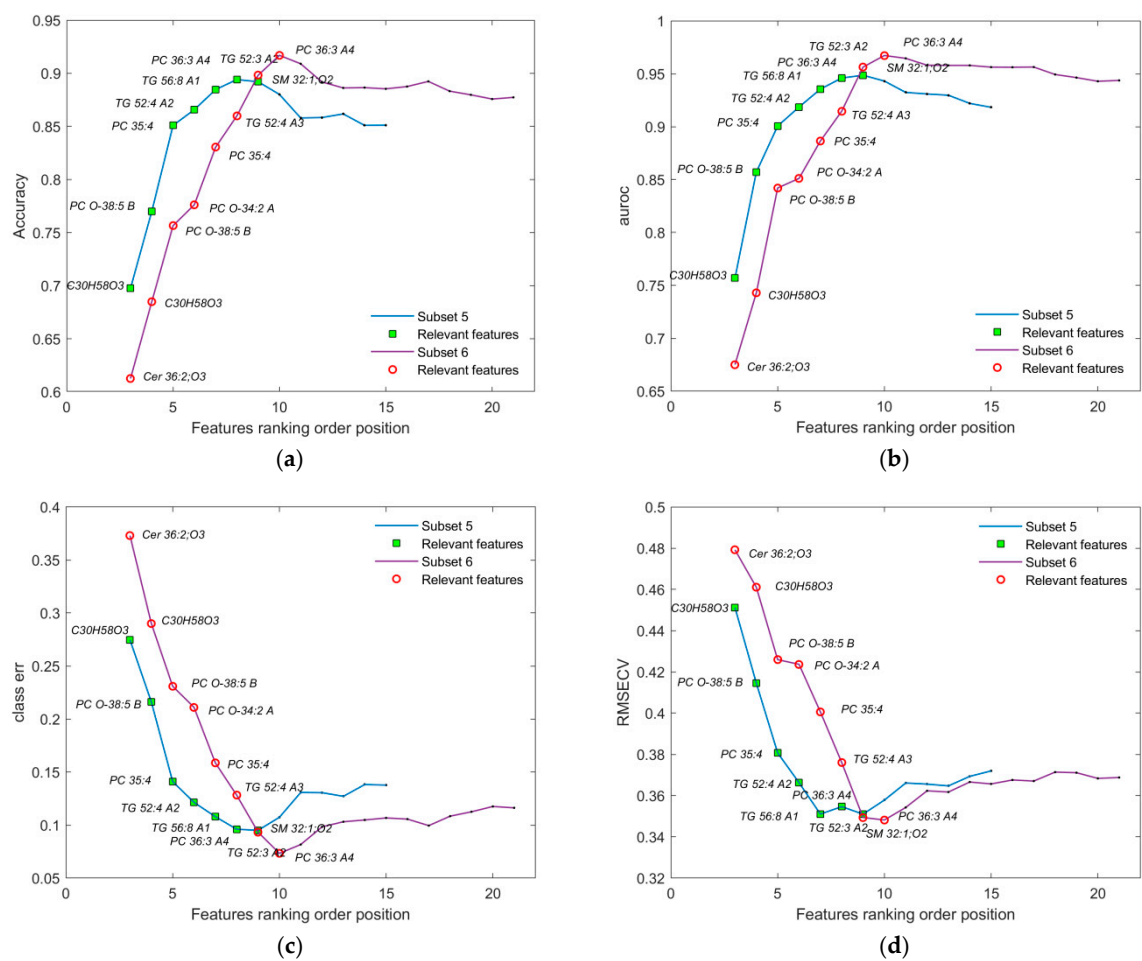


Figure 8. Performance metrics displaying the smallest subset of identified subset of features for partition of initial dataset containing only female subjects: (a) accuracy; (b) AUROC; (c) class error; and (d) RMSECV. Markers colored as green squares represent subset 5, consisting of the first 9 most relevant features out of the 15 total presented features on the plot. Markers colored as red circles represent subset 6, consisting of the first ten most relevant features from the 21 total presented features. The first two features were omitted, since the number of LV components was always 3 for each of the assembled OPLS-DA models during validation. Both subsets with corresponding feature coding lists are given in Table 3.

Table 3. For the female subset of samples, the list of identified relevant features for subset 5, consisting of the 9 most relevant features from the 15 total features in this subset; and for subset 6, consisting of the 10 most relevant features from the 21 total features in this subset.

Subset 5		Subset 6	
ID Number	Feature Names	ID Number	Feature Names
24	LPC 16:0 A1	24	LPC 16:0 A1
26	LPC 16:0 A3	26	LPC 16:0 A3
1	C30H58O3	8	Cer 36:2;O3
101	PC O-38:5 B	1	C30H58O3
57	PC 35:4	101	PC O-38:5 B
158	TG 52:4 A2	93	PC O-34:2 A
178	TG 56:8 A1	57	PC 35:4
66	PC 36:3 A4	159	TG 52:4 A3
155	TG 52:3 A2	106	SM 32:1;O2
		66	PC 36:3 A4

3.5. Comparison of Model Performance Metric Distribution

Model performance metrics (accuracy, AUROC, sensitivity, and specificity) were compared across 1000 independently generated CV OPLS-DA models using either all features or selected subsets (Tables 1–3; Supplementary Materials Figures S9–S11). Median values (dashed lines) for corresponding distribution plots revealed improved overall performance with feature selection.

During validation, each point in the plots (Figures 4–8) represents the mean of the corresponding metric distributions. Mean metric values were aggregated from 183 feature sequences (or fewer during subsetting) with 1000 outcomes. Wilcoxon rank-sum tests (1% significance level) indicated significant increases in accuracy, AUROC, sensitivity, and specificity (left-tailed), and significant decreases in RMSECV and class error (right-tailed) [71].

Despite overall improvements, distribution plots showed slight negative skewness for models with selected features, suggesting marginally reduced performance in some cases. Skewness for the datasets including all observations ranged from -0.1 to -0.7 for the full and reduced feature sets, with stronger skewness in gender-partitioned data (e.g., AUROC for males: -2.6 ; specificity for females: -1.75 ; Table 4).

Table 4. Skewness for classification metrics presented for all distributions obtained from 1000 CV OPLS-DA models for each dataset (including all observations assigned as ds, and solely with males assigned as ds_male or females assigned as ds_female participants), obtained by using all and selected (subset 2, 4, or 6) lists of features (appended to the base name for the corresponding group of samples following snake_case convention) for their assembly.

Skewness	ds_All	ds_Selected	ds_Male_All	ds_Male_Selected	ds_Female_All	ds_Female_Selected
accuracy	−0.3355	−0.5339	−0.6120	−1.1246	−0.4323	−1.1702
AUROC	−0.6971	−0.7312	−1.2077	−2.5998	−0.5925	−1.2864
sensitivity	−0.4327	−0.4533	−0.7758	−1.0454	−0.4366	−1.1509
specificity	−0.1993	−0.8569	−0.5685	−1.1961	−0.4632	−1.7518

A plausible explanation for such behavior, at least for the former distribution corresponding to AUROC for male subjects presented in Supplementary Materials Figure S10b, could be that the distribution is truncated on the right side with an upper limit equal to 1; this is also relevant for all other distributions with given metrics where truncation was bonded within the $[0, 1]$ interval. In contrast, all other distributions with moderate observed skewness could not be fitted with a truncated normal distribution, since they result in the same average values for the corresponding metric distribution as the untruncated normal distribution.

This skewness may also result from outliers emerging under reduced feature sets, undetected in full-feature PCA, and from diminished predictive power in certain samples.

3.6. Performance of OPLS-DA Models with Selected Feature Subsets

From 1000 CV OPLS-DA models built for each dataset and feature subsets (Tables 1–3), the best model was determined based on the most frequent metric range on any given distribution (accuracy, sensitivity, and specificity). The optimal model, also exhibiting the best AUROC, was identified using multiple intersects [72]. The specific metric performance for these models is summarized in Table 5.

Table 5. The best values for accuracy, AUROC, sensitivity, and specificity obtained among 1000 generated CV OPLS-DA models are given for each processed dataset (including the complete dataset, assigned as ds, and with only males assigned as ds_male or females assigned as ds_female), with all and corresponding relevant feature subsets appended to the base name for the corresponding group of samples (following the snake_case convention) given in Tables 1–3.

Feature Subsets	Accuracy	AUROC	Sensitivity	Specificity	Number of Features
ds_all	0.8523	0.9401	0.8523	0.8523	183
ds_subset1	0.9091	0.9654	0.9318	0.8864	16
ds_subset2	0.8977	0.9551	0.9091	0.8864	14
ds_male_all	0.8625	0.9590	0.8636	0.8611	183
ds_male_subset3	0.9125	0.9867	0.9091	0.9167	20
ds_male_subset4	0.9375	0.9956	0.9318	0.9444	12
ds_female_all	0.8211	0.9314	0.8182	0.8235	183
ds_female_subset5	0.8947	0.9652	0.9091	0.8824	9
ds_female_subset6	0.9263	0.9857	0.9545	0.9020	10

The best OPLS-DA models rely on the relevant feature subsets (subsets 2, 4, and 6 in Tables 1–3) are visualized in Figures 9–14 with scores plots for the first two LV components (predictive and orthogonal) and loading plots, highlighting the most relevant features (VIP scores > 1). Permutation tests (using 500 iterations) were performed to address potential overfitting [73,74].

Model performance metrics in Figures S9–S11 (Supplementary Materials) align with Table 5, showing significant improvements for models with selected features compared to those obtained with all features. Comparison of Figure 2a vs. Figure 9a reveals pronounced sample shifts (“SCH” and “BD”) from the central coordinate line separating two classes along the LV1 predictive component, supporting this observation.

In the dataset with all observations, the best OPLS-DA model’s loading plot (Figure 9b) identified the following features based on the “greater than one rule”: “LPC 16:0 A1”, “SM 33:1;O2”, “SM 32:1;O2”, “C30H58O3”, “PC 30:0”, and “*m/z* 293.1779”. These features are ranked in increasing order, as shown in Table 1.

In the male-only dataset (Figure 11b), the OPLS-DA model using subset 4 (Table 2) identified five key features (VIP > 1): “SM 32:1;O2”, “PC 32:1”, “SM 33:1;O2”, “PC 30:0”, and “FA 16:1”.

For the female-only dataset (Figure 13b), three features were identified: “LPC 16:0 A1”, “LPC 16:0 A3”, and “C30H58O3”. The feature “Cer 36:2;O3,” despite ranking higher than “C30H58O3,” did not meet the VIP > 1 criterion for subset 6 (Table 3). All identified features appear in the top three positions of subset 5 for the same table.

A comparison of feature lists from OPLS-DA models across male, female, and all-observation datasets revealed limited overlap. Common features include “SM 32:1;O2,” “SM 33:1;O2,” and “PC 30:0” for males, and “LPC 16:0 A1” and “C30H58O3” for females, appearing in the full dataset as well. Each significant feature (VIP > 1) was specific to its gender group.

However, Table 5 shows that models with all observations perform worse than gender-specific models. This may result from interactions between features favoring one gender, causing confounding effects in combined datasets. These findings suggest that gender-based partitioning improves OPLS-DA model performance, like datasets containing SCH and C (control) groups [19].

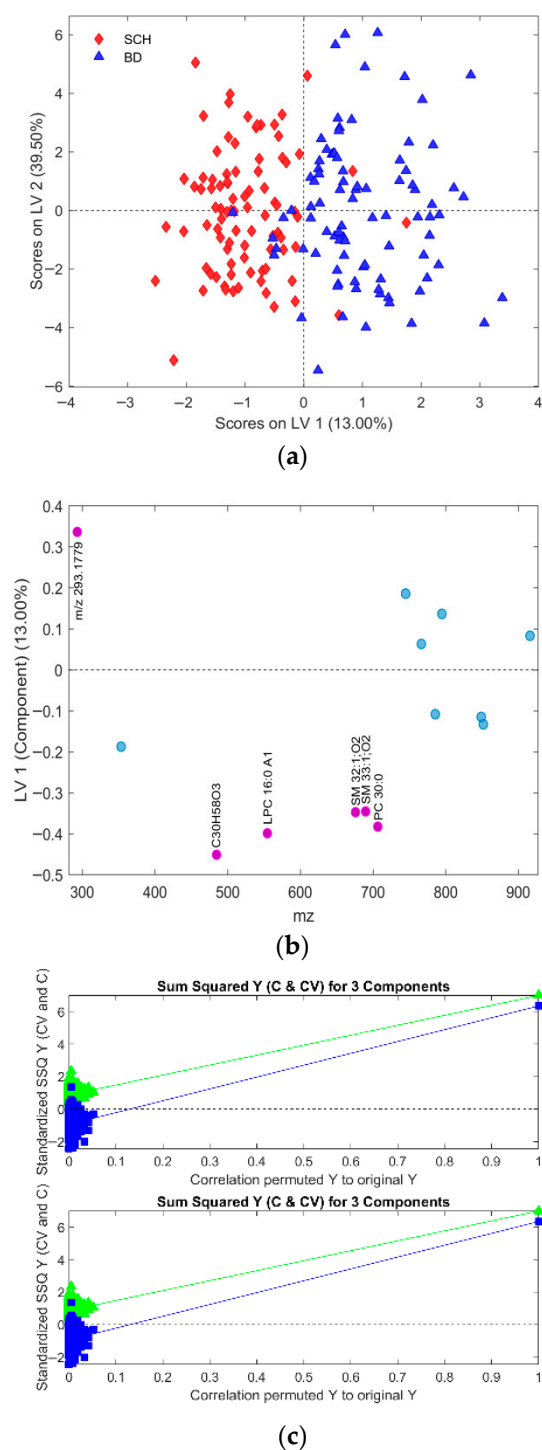


Figure 9. (a) Scores plot of the first two LV components of the OPLS-DA model using mean centering and unit variance scaling. The variance captured by the first two LV components, predictive LV1 and orthogonal LV2, is given in the figure axis labels. (b) Loading LV1 component of the PLSDA model for the set of features assigned as subset 2 (given in Table 1), with the relevant set of features including all initial samples with identified outliers excluded, where features with VIP scores greater than 1 are indicated as pink circles assigned with corresponding names. (c) Permutation test for the OPLS-DA model composed of three LV components and the set of selected features assigned as subset 2, performed with 500 iterations. Fractional y-block information captured by calibration (green) and cross-validation (blue) versus y-correlation is shown.

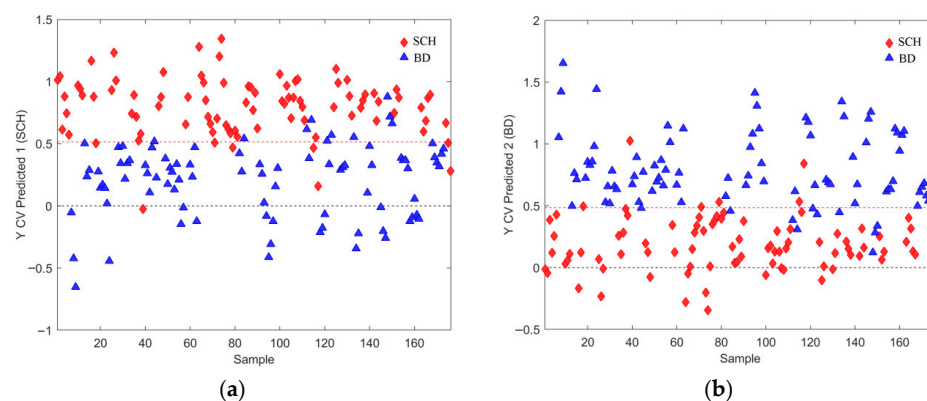


Figure 10. Cross-validated prediction results for the complete dataset and selected features from subset 2, given in Table 1: (a) class SCH (red filled diamonds), where the discrimination threshold was determined at 0.5150; (b) class BD (blue filled triangles), where the discrimination threshold was determined at 0.4850.

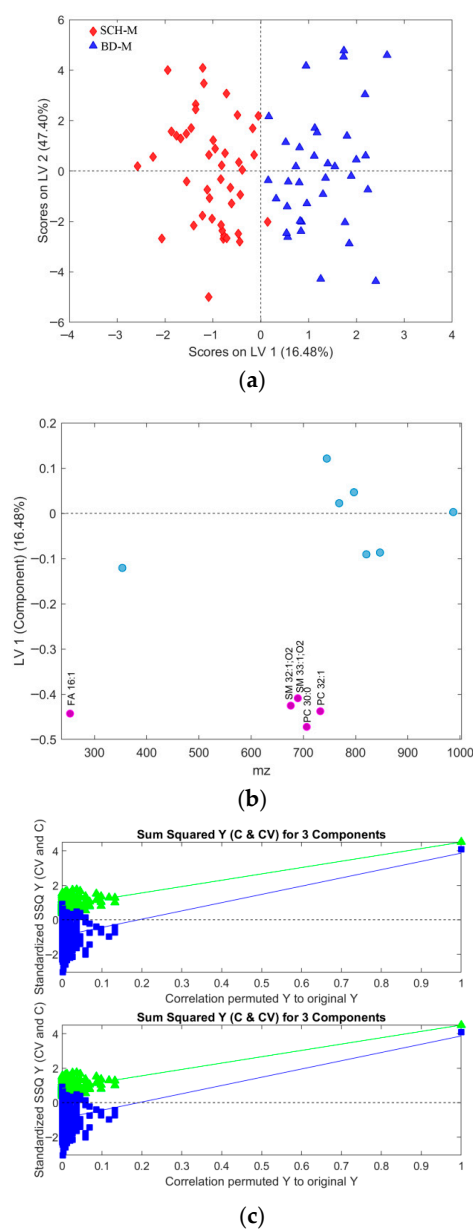


Figure 11. (a) Scores plot of the first two LV components of the OPLS-DA model using mean centering and unit variance scaling for only male subjects in the initial dataset. Variance captured by first two LV

components, predictive LV1 and orthogonal LV2, is given in the figure axis labels. **(b)** Loading LV1 component of the PLS-DA model for the set of features assigned as subset 4 (given in Table 2), with the relevant subset of features including only male subjects from the initial list of samples with identified outliers excluded, where features with VIP scores greater than 1 are indicated as pink circles assigned with corresponding names. **(c)** Permutation test for the OPLS-DA model composed of three LV components and the set of selected features assigned as subset 4, performed with 500 iterations. Fractional y-block information captured by calibration (green) and cross-validation (blue) versus y-correlation is shown.

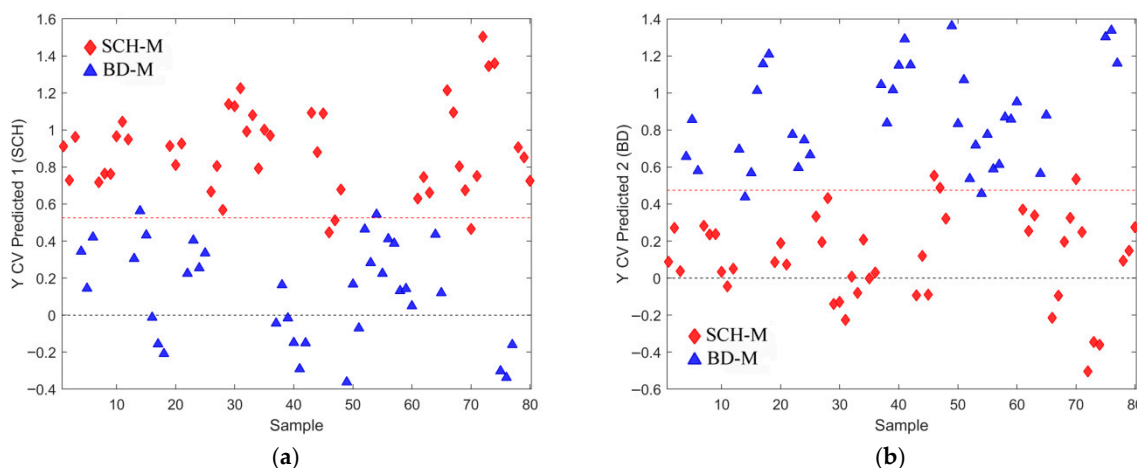


Figure 12. Cross-validated prediction results for male subjects only and selected features from subset 4, given in Table 2: **(a)** class SCH-M (red filled diamonds), where the discrimination threshold was determined at 0.5255 (red dashed line); **(b)** class BD-M (blue filled triangles), where the discrimination threshold was determined at 0.4745 (red dashed line).

On the other hand, a key limitation of this study is its relatively small dataset, which, as a consequence, prevents the partitioning of the independent subset needed for external validation with a feasible number of observations. Moreover, the dependency of classification tasks of the so-called wrapper method, which evaluates subsets of features based on the selected classifier performance, unavoidably leads to biased results, depending on the modeling algorithm on which they were evaluated [75]. To improve reliability in such circumstances, an additional modeling algorithm should be applied to confirm the selected features.

Besides the extensive application of CV used in all stages during the procedure of selection and subsequent validation of each subset of features in this work, Worley and Powers, based on their findings [58], also suggested using PCA models as a comparative practical indicator of OPLS-DA model reliability. Thus, if the procedure for feature selection followed in this study is acceptable, then we should expect to observe, from the assembled PCA models, more distinguishable group separation belonging to the recognized class of patients, in both partitions of datasets. This will also confirm the statistical significance of the class separation observed in the obtained OPLS-DA models. Initial PCA models for gender-specific lists of observations (Figure 1) showed no clear group separation, but models using the best subset of selected features (Figure 15) revealed distinct separation along PC2. The explained variance increased from 8.57% to 13.48% in males and from 12.42% to 15.93% in females, confirming the relevance of the selected features. Increased variance between patient groups also confirms the relevance of each selected feature subset, as independently validated by both OPLS-DA and PCA models. Furthermore, the PC2 loading plots (Figure 15b,d) closely matched those from OPLS-DA models (Figures 11b and 13b), reinforcing consistency across methods.

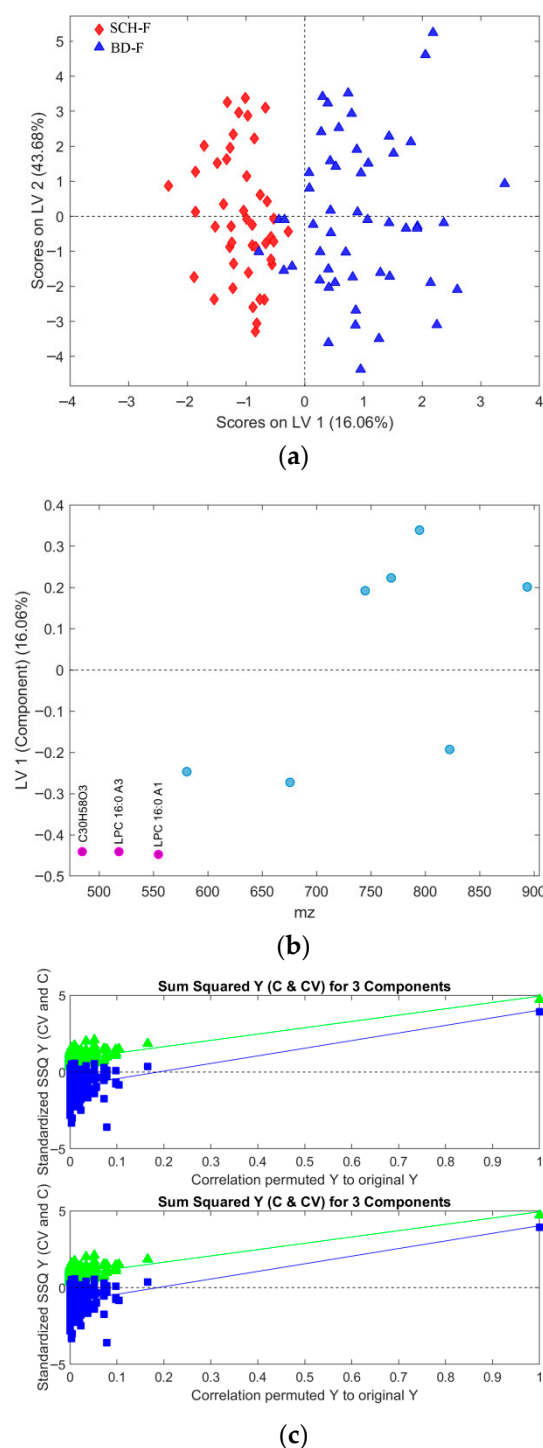


Figure 13. (a) Scores plot of first the two LV components of the OPLS-DA model using mean centering and unit variance scaling for only female subjects in the initial dataset. Variance captured by the first two LV components, predictive LV1 and orthogonal LV2, is given in the figure axis labels. (b) Loading LV1 component of the PLS-DA model for the set of features assigned as subset 6 (given in Table 3), with the relevant subset of features including only female subjects from the initial list of samples with identified outliers excluded, where features with VIP scores greater than 1 are indicated as pink circles assigned with corresponding names. (c) Permutation test for the OPLS-DA model composed of three LV components and the set of selected features assigned as subset 2, performed with 500 iterations. Fractional y-block information captured by calibration (green) and cross-validation (blue) versus y-correlation is shown.

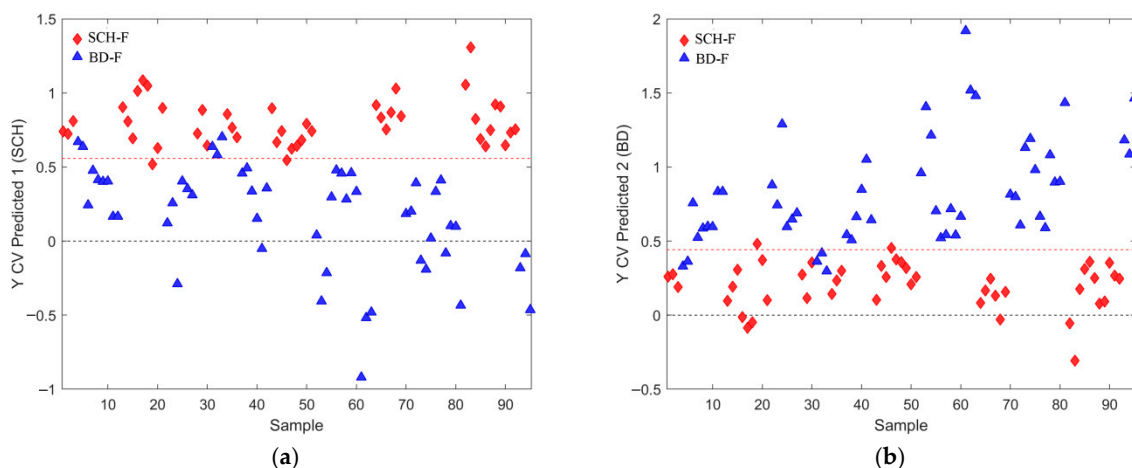


Figure 14. Cross-validated prediction results for female subjects only, and selected features from sub-set 6 given in Table 3: (a) class SCH-F (red filled diamonds), where the discrimination threshold was determined at 0.5581 (red dashed line); (b) class BD-F (blue filled triangles), where the discrimination threshold was determined at 0.4419 (red dashed line).

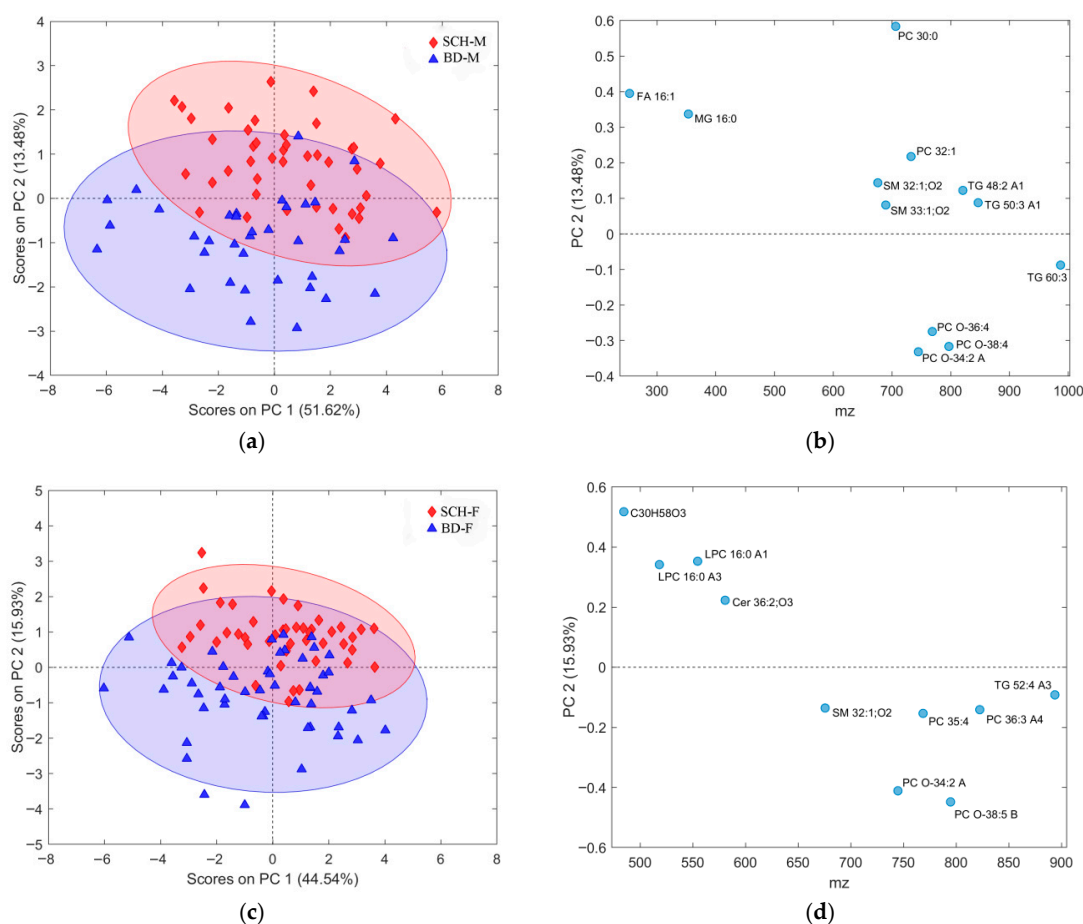


Figure 15. PCA model for the dataset partition with: (a) male subjects and the selected subset of features, using four PCA components with the explained 84.28% overall variability; (b) corresponding loading plot of the PC2 component of the dataset partition with male subjects and the selected subset of features; (c) female subjects and the selected subset of features, using four PCA components with the explained 78.51% overall variability; and (d) corresponding loading plot of the PC2 component of the dataset partition with female subjects and the selected subset of features. Each group of patients was assigned a different color (BD as blue full triangles and SCH as red full diamonds), where each feature in (b,d) was separately assigned a corresponding code name.

4. Discussion

Although PCA of gender-differentiated subgroups (SCH-M vs. BD-M and SCH-F vs. BD-F) did not reveal distinct separation patterns (Figure 1), the most relevant feature subset pairs—1 and 2 for all samples, 3 and 4 for male subgroups, and 5 and 6 for female subgroups—were robustly validated using optimal metrics from a 1000-fold cross-validated OPLS-DA model. Among these, subsets 2, 4, and 6 represent the smallest feature sets with the best metric performance for the overall, male, and female datasets, respectively (Tables 1–3; Figures 9–14). The OPLS-DA model applied to all samples (SCH vs. BD) using the minimal feature subset 2 identified six potential biomarkers: LPC 16:0, SM 33:1, SM 32:1, compound C30H58O3, PC 30:0, and an unidentified compound with m/z 293.1779, with VIP scores higher than 1 (Table 1, Figure 9b). Gender-differentiated models revealed five biomarkers in the male subgroups (SCH-M vs. BD-M): sphingomyelins SM 32:1 and SM 33:1, phosphatidylcholines PC 32:1 and PC 30:0, and FA 16:1 (Table 2, Figure 11b), and two biomarkers in the female subgroups (SCH-F vs. BD-F): two forms of lysophosphatidylcholine LPC 16:0 (A1 and A3, representing different adducts) and compound C30H58O3 (Table 3, Figure 13b), suggesting sex-specific lipidomic signatures that are also supported by the literature data [19,76–79]. Tabassum et al. (2023) highlighted that genetic factors, such as sex chromosomes and physiological differences like menopause and sex hormones, may contribute to gender-specific lipidomic profiles [79]. These findings suggest that lipid metabolism regulation differs between men and women, and blood lipidomics (serum and plasma) should be analyzed separately by gender, treating it as a confounding factor [76–79]. Actually, our results showed that in male patients (SCH-M vs. BD-M), three of five potential biomarkers (SM 32:1, SM 33:1, and PC 30:0) are common with those found in the overall SCH vs. BD comparison, while in female patients (SCH-F vs. BD-F), both of the identified biomarkers (two forms of LPC 16:0 and C30H58O3) are also present in the overall SCH vs. BD comparison (Tables 1–3, Figures 9, 11 and 13). Moreover, potential biomarkers appear to be distinct between female and male patients with no overlap.

Obviously, comparative analysis of lipid profiles between SCH and BD groups including all samples, as well as between their respective male and female subgroups, confirmed alterations in lipid metabolism, including three main lipid classes: sphingophospholipids (SPs), glycerophospholipids (GPs), and free fatty acids (FAs), which were also observed in our previous studies of the lipidomic profiles of SCH and BD patients compared to controls [19,20]. SPs and GPs are fundamental components of brain membranes. Sphingomyelins (SMs) are particularly important for transmembrane signaling, while GPs contribute significantly to the structural integrity and fluidity of neuronal membranes [80,81]. Our findings clearly indicate alterations in both GP (including PC 32:1, PC 30:0, and LPC 16:0) and SP (including SM 32:1 and SM 33:1) metabolic pathways, pointing to a disruption in lipid homeostasis. This imbalance likely affects membrane architecture and intracellular signaling mechanisms, which may underline the pathophysiology of SCH and BD. The potential biomarkers identified include lipid species from the classes of PC, SM, and LPC. According to the literature data, PC and SM are generally elevated in women, while LPC is higher in men, at least partly due to biological sex differences in the activities of phospholipases and the synthesis of SMs [76–79]. Since lipid species belonging to classes of PC, SM, and LPC could be influenced both by gender and psychiatric disorders, it could be crucial to perform gender-specific analysis when searching for disorder-specific biomarkers of the psychiatric disorders.

Only a few studies [18,82–84] have simultaneously analyzed the lipidomic profiles of plasma and serum in both BD and SCH patients, although in half of them [64,65], comparisons between patient cohorts and the control group were not presented as primary

outcomes. Costa et al. (2023) found that most differential lipids between BD and SCH were SPs (54.17%), followed by GLs and GPs (16.67% each), with sterol lipids and FA being the least common [82]. Their plasma lipidomics analysis suggests that focusing on these lipid classes could help distinguish between BD and SCH, which agrees with our results. On the other hand, however, Tkachev et al. (2023) found that changes in plasma lipidomic profiles associated with SCH and BD are largely similar [18]. One of the main findings was that alterations in ceramides (Cer) levels were characteristic of both BD and SCH patients. However, Cer (d32:1) and Cer (d38:1) levels were dysregulated in BD patients but not in SCH patients compared to healthy controls, and changes in serum Cer (d43:3) were specific to SCH [18].

Yu et al. (2024) used plasma lipidomics data from 7145 Finnish individuals to study the association of five psychiatric disorders, including SCH and BD, with plasma lipid profiles [84]. They identified 179 lipid species across 13 classes. In BD patients, several phosphatidylcholine (PC) lipid species were strongly correlated with disease, with some showing protective effects and others linked to higher risk. For SCH patients, 22 lipid species, mainly from the GP class, were associated with the diagnosis.

Tao et al. (2022) analyzed the lipidomic profiles of 112 SCH patients, 132 BD patients, 105 MDD patients, and 198 healthy controls [83]. They aimed to identify psychosis subtypes rather than differences from healthy controls. Using the UHPLC-MS method, 1164 lipid molecules were identified, with 10 key lipids distinguishing psychotic patients from controls. Seven lipids were upregulated (mostly derivatives of PCs and FA) and three were downregulated (derivatives of vitamin D3, an unsaturated FA, and gamma-butyric acid (GBA)). Patients were classified into two subtypes (Cluster 1 and Cluster 2) based on these lipids, and 66% of SCH and 48% of BD patients were included in Cluster 2, with lower global assessment scores and significant white brain matter alterations compared to Cluster 1 patients. This suggests that lipid biomarkers can identify transdiagnostic subtypes across psychiatric diseases.

Our analysis of lipid profiles between SCH and BD groups, especially considering the common potential biomarkers in gender-differentiated groups (SCH-M vs. BD-M and SCH-F vs. BD-F) with those found in the overall SCH vs. BD comparison (Tables 1–3, Figures 9, 11 and 13), suggests that diagnosis-specific alterations in lipid metabolism in SCH and BD may be closely linked to the modulation or upregulation of phospholipase A2 (PLA2) enzyme activity. PLA2 plays a catalytic role in hydrolyzing GPs, producing corresponding unsaturated fatty acids and lysophospholipids. Elevated PLA2 activity has previously been reported in SCH and is strongly associated with changes in neuronal function, which may contribute to affective and cognitive symptoms [80]. The lipidomic biomarkers identified in our study, particularly those consistently altered in both gender-differentiated groups (SCH-M vs. BD-M and SCH-F vs. BD-F), highlight the significance of enzymes regulating these metabolites.

In the overall comparison (SCH vs. BD), all lipid putative biomarkers (LPC 16:0, SM 33:1, SM 32:1, compound C30H58O3, and PC 30:0) showed higher abundance in SCH compared to BD, except for m/z 293.1779, which was decreased in SCH compared to BD patients (Table 1, Figure 9b). Similarly, our results of male patients (SCH-M vs. BD-M) confirmed that all potential biomarkers, SM 32:1, SM 33:1, PC 32:1, PC 30:0, and FA 16:1, were more abundant in SCH-M than BD-M (Table 2, Figure 11b). In addition, both of the identified biomarkers (two forms of LPC 16:0 and C30H58O3) in female patients (SCH-F vs. BD-F) had elevated levels in SCH-F compared to BD-F (Table 3, Figure 13b). To the best of our knowledge, the literature-available data for the comparison of the serum lipidomics profiles of SCH and BD patients are limited and inconsistent. Costa et al. (2023) found that most differential lipids belonging to SP, GL, and GP were downregulated in SCH compared

to BD and healthy controls [80], which is the opposite of our results. Tkachev et al. (2023) found no differences in choline and ethanolamine-containing GPs and TGs in BD patients, suggesting these changes might be more pronounced in SCH patients [18], which agrees with our results. Yu et al. (2024) confirmed the importance of PC lipid alterations in both BD and SCH, suggesting that their positive association might be more pronounced in BD [84], which is the opposite of our results.

Finally, the limitations of this study are very important to emphasize. The study limitations include a moderate sample size ($n = 61$ —SCH 30 and 31 BD), which may affect statistical power. Additionally, the results were obtained using a semi-quantitative method without full validation against reference standards. Although sex differences were considered, other confounding factors, such as dietary habits, lifestyle, comorbidities, duration of illness, and the use of first- and second-generation antipsychotics, may influence GP and SP content, and should be addressed in future studies incorporating rigorous control or stratification for these factors [85]. Therefore, in future research, larger multicenter studies are essential to validate these potential biomarkers across diverse populations.

5. Conclusions

This study demonstrates the utility of advanced lipidomic profiling in addressing one of the most persistent challenges in psychiatry, the differential diagnosis of schizophrenia (SCH) and bipolar disorder (BD). Using LC-HRMS combined with robust chemometric modeling, we identified distinct lipid signatures that reliably discriminate between these two disorders. The validated feature subsets, pairs 1 and 2, captured the most relevant metabolites differentiating SCH and BD patients, identifying five key biomarkers: LPC 16:0, SM 33:1, SM 32:1, compound C30H58O3, and PC 30:0. Gender-specific models further revealed five biomarkers in males (SM 32:1, SM 33:1, PC 32:1, PC 30:0, and FA 16:1) and two in females (LPC 16:0 and C30H58O3), highlighting potential sex-related metabolic differences. The potential biomarkers, primarily glycerophospholipids and sphingophospholipids, not only highlight metabolic differences between SCH and BD, but also reveal gender-specific patterns, underscoring the importance of personalized approaches in psychiatric research.

The elevated levels of these lipid classes in SCH patients suggest disruptions in membrane lipid homeostasis and signaling pathways, yet the precise mechanistic links between these lipids and disease pathophysiology remain unclear. Functional studies, including pathway analyses and experimental validation, are necessary to elucidate these relationships. The lack of validation in independent cohorts or functional assays represents a critical gap for clinical application, highlighting the need for future replication studies and targeted metabolomic assays to confirm biomarker reliability and assess their predictive value for treatment response and disease progression.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/metabo16010069/s1>, Figure S1. (a) Boxplot of log10-transformed data for the schizophrenia (SCH) and bipolar disorder (BD) classes of patients. (b) Boxplot of log10-transformed and mean centered data for the same classes. Figure S2. PCA model obtained for four consecutive batches. Figure S3. ROBPCA outlier map of the LC-MS dataset composed using: (a) four PC components for the whole dataset including all observations of both classes; (b) three PC components for the subset including only male participants of SCH and BD classes; and (c) four PC components for the subset including only female participants of SCH and BD classes. Figure S4. Results for: (a) skewness and (b) kurtosis of the log10 transformed raw data before and after the removal of identified outliers. The x-axis shows m/z (mass-to-charge ratio), while the y-axis displays related statistics for any variable in the dataset. Figure S5: Histogram of frequency counts obtained from: (a) the minimum of RMSECV; (b) the minimum of class error; and (c) the maximum of AUROC

in all OPLS-DA models applied on 1000 subsets of randomly selected triplicates for each sample of each of the 5 folds. The maximum number of LV components in all models was 10. The applied CV method for the presented results was “venetian blinds”. Figure S6. Permutation test performed with 500 iterations for the best OPLS-DA model depicted in Figure 4, selected among 1000 CV models with a complete set of observations with (a) three LV components and (b) five LV components. The sum of squares of responses for the calibrated models are presented in green, while the CV models are presented in blue. Figure S7. Performance metrics displayed in relation to the increasing sequence of sets of ranked features for partition of the initial dataset containing male subjects only: (a) accuracy; (b) AUROC; (c) class error; and (d) RMSECV. Additionally, markers colored as green squares (subset 3) and red circles (subset 4) represent the smallest sets of features identified through the selection procedure. The first two features were omitted, since the number of LV components was always 3 for each of the assembled OPLS-DA models during validation. Both subsets with corresponding feature coding lists are presented in Table 2. Figure S8. Performance metrics displayed in relation to the increasing sequence of sets of ranked features for partition of the initial dataset containing female subjects only: (a) accuracy; (b) AUROC; (c) class error; and (d) RMSECV. Additionally, markers colored as green squares (subset 5) and red circles (subset 6) represent the smallest sets of features identified through the selection procedure. The first two features were omitted, since the number of LV components was always 3 for each of the assembled OPLS-DA models during validation. Both subsets with corresponding feature coding lists are presented in Table 3. Figure S9. Comparison of the performance metric distribution of 1000 assembled CV OPLS-DA models of the complete dataset of all included features (given in blue), and with the subset of relevant selected features assigned as subset 2 in Table 1 (given in red), for: (a) accuracy; (b) AUROC; (c) sensitivity; and (d) specificity. At the same time, for each presented distribution, the median for the given metric distribution was assigned a dashed line in a corresponding color. Histograms in the distributions of each metric are given with equal bin width. Figure S10. Comparison of the performance metric distribution of 1000 assembled CV OPLS-DA models of the subset of data including male subjects only, with all included features (given in blue) and with the subset of relevant selected features assigned as subset 4 in Table 2 (given in red), for: (a) accuracy; (b) AUROC; (c) sensitivity; and (d) specificity. At the same time, for each presented distribution, the median for the given metric distribution was assigned a dashed line in a corresponding color. Histograms in the distributions of metrics are given with equal bin width.; Figure S11. Comparison of the performance metric distribution of 1000 assembled CV OPLS-DA models of the subset of data including female subjects only, with all included features (given in blue), and with the subset of relevant selected features assigned as subset 6 in Table 3 (given in red), for: (a) accuracy; (b) AUROC; (c) sensitivity; and (d) specificity. At the same time, for each presented distribution, the median for the given metric distribution was assigned a dashed line in a corresponding color. Histograms in the distributions of metrics are given with equal bin width. Table S1: Identification of relevant features for subsets 1–6 (Tables 1–3) relating to the lipids found differently in all, male, and female schizophrenia and bipolar disorder patients.

Author Contributions: Conceptualization, Z.M., M.B.J., M.T., and B.M.; methodology, Z.M., M.B.J., V.T., and B.M.; software, Z.M. and M.B.J.; validation, Z.M., M.B.J., G.K., and B.M.; formal analysis, M.R.J., M.B.J., Z.M., D.S., and S.L.; investigation, Z.M., M.T., N.A., A.G., Č.M., M.P., Z.A., S.R., and M.B.J.; resources, V.T., M.T., and B.M.; data curation, Z.A., S.R., A.G., Č.M., M.P., M.R.J., and G.K.; writing—original draft preparation, Z.M., M.B.J., M.T., and N.A.; writing—review and editing, N.A., Z.M., M.B.J., and B.M.; visualization, S.L., D.S., and N.A.; supervision, B.M.; project administration, G.K.; and funding acquisition, M.T., V.T., and B.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Organization for the Prohibition of Chemical Weapons (OPCW) (Grant Number: L/ICA/ICB-108/21) and the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia (Contract Numbers 451-03-136/2025-03/200168, 451-03-136/2025-03/200026, 451-03-136/2025-03/200051, 451-03-136/2025-03/200015, and 451-03-136/2025-03/200110).

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Special Hospital for Psychiatric Diseases “Kovin” (protocol code: 01-3342/1-1, date of approval: 7 November 2018), the University of Belgrade–Faculty of Chemistry (protocol code: 1-4/21, date of approval: 19 April 2021), and the Blood Transfusion Institute of Serbia (protocol code: 1226, date of approval: 15 March 2021).

Informed Consent Statement: Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: The lipidomics data presented in this study are unavailable due to privacy or ethical restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Janoutová, J.; Janáčková, P.; Serý, O.; Zeman, T.; Ambroz, P.; Kovalová, M.; Varechova, K.; Hosák, L.; Jirik, V.; Janout, V. Epidemiology and risk factors of schizophrenia. *Neuro Endocrinol. Lett.* **2016**, *37*, 1–8.
2. Charlson, F.J.; Ferrari, A.J.; Santomauro, D.F.; Diminic, S.; Stockings, E.; Scott, J.G.; McGrath, J.; Whiteford, A.H. Global epidemiology and burden of schizophrenia: Findings from the global burden of disease study 2016. *Schizophr. Bull.* **2018**, *44*, 1195–1203. [\[CrossRef\]](#)
3. Rowland, T.A.; Marwaha, S. Epidemiology and risk factors for bipolar disorder. *Ther. Adv. Psychopharmacol.* **2018**, *8*, 251–269. [\[CrossRef\]](#) [\[PubMed\]](#)
4. WHO. Schizophrenia. Available online: <https://www.who.int/news-room/fact-sheets/detail/schizophrenia> (accessed on 6 October 2025).
5. WHO. Mental Disorders. Available online: <https://www.who.int/news-room/fact-sheets/detail/mental-disorders> (accessed on 30 September 2025).
6. Cloutier, M.; Greene, M.; Guerin, A.; Touya, M.; Wu, E. The economic burden of bipolar I disorder in the United States in 2015. *J. Affect. Disord.* **2018**, *226*, 45–51. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Barbosa, W.B.; Costa, J.O.; de Lemos, L.L.P.; Gomes, R.M.; de Oliveira, H.N.; Ruas, C.M.; Acurcio, F.A.; Barbui, C.; Bennie, M.; Godman, B.; et al. Costs in the treatment of schizophrenia in adults receiving atypical antipsychotics: An 11-year cohort in Brazil. *Appl. Health Econ. Health Policy* **2018**, *16*, 697–709. [\[CrossRef\]](#)
8. Kadakia, A.; Catillon, M.; Fan, Q.; Williams, G.R.; Marden, J.R.; Anderson, A.; Kirson, N.; Dembek, C. The economic burden of schizophrenia in the United States. *J. Clin. Psychiatry* **2022**, *83*, 5377. [\[CrossRef\]](#)
9. Alda, M.; Manchia, M. Personalized management of bipolar disorder. *Neurosci. Lett.* **2018**, *669*, 3–9. [\[CrossRef\]](#)
10. Bzdok, D.; Meyer-Lindenberg, A. Machine learning for precision psychiatry: Opportunities and challenges. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2018**, *3*, 223–230. [\[CrossRef\]](#)
11. Holmes, E.; Tsang, T.M.; Huang, J.T.; Leweke, F.M.; Koethe, D.; Gerth, C.W.; Nolden, B.M.; Gross, S.; Schreiber, D.; Nicholson, J.K.; et al. Metabolic profiling of CSF: Evidence that early intervention may impact on disease progression and outcome in schizophrenia. *PLoS Med.* **2006**, *3*, e327. [\[CrossRef\]](#)
12. Ribeiro, H.C.; Klassen, A.; Pedrini, M.; Carvalho, M.S.; Rizzo, L.B.; Noto, M.N.; Zeni-Graiff, M.; Sethi, S.; Fonseca, F.A.H.; Tasic, L.; et al. A preliminary study of bipolar disorder type I by mass spectrometry-based serum lipidomics. *Psychiatry Res.* **2017**, *258*, 268–273. [\[CrossRef\]](#)
13. Tasic, L.; Pontes, J.G.M.; Carvalho, M.S.; Cruz, G.; Dal Mas, C.; Sethi, S.; Pedrini, M.; Rizzo, L.B.; Zeni-Graiff, M.; Asevedo, E.; et al. Metabolomics and lipidomics analyses by ¹H nuclear magnetic resonance of schizophrenia patient serum reveal potential peripheral biomarkers for diagnosis. *Schizophr. Res.* **2017**, *185*, 182–189. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Yang, J.; Chen, T.; Sun, L.; Zhao, Z.; Qi, X.; Zhou, K.; Cao, Y.; Wang, X.; Qiu, Y.; Su, M.; et al. Potential metabolite markers of schizophrenia. *Mol. Psychiatry* **2013**, *18*, 67–78. [\[CrossRef\]](#)
15. Chen, J.; Liu, Z.; Fan, S.; Yang, D.; Zheng, P.; Shao, W.; Qi, Z.; Xu, X.; Li, Q.; Mu, J.; et al. Combined application of NMR- and GC-MS-based metabolomics yields a superior urinary biomarker panel for bipolar disorder. *Sci. Rep.* **2014**, *4*, 5855. [\[CrossRef\]](#)
16. Lan, M.J.; McLoughlin, G.A.; Griffin, J.L.; Tsang, T.M.; Huang, J.T.; Yuan, P.; Manji, H.; Holmes, E.; Bahn, S. Metabonomic analysis identifies molecular changes associated with the pathophysiology and drug treatment of bipolar disorder. *Mol. Psychiatry* **2009**, *14*, 269–279. [\[CrossRef\]](#)
17. Sethi, S.; Pedrini, M.; Rizzo, L.B.; Zeni-Graiff, M.; Mas, C.D.; Cassinelli, A.C.; Noto, M.N.; Asevedo, E.; Cordeiro, Q.; Pontes, J.G.M.; et al. ¹H-NMR, ¹H-NMR T₂-edited, and 2D-NMR in bipolar disorder metabolic profiling. *Int. J. Bipolar Disord.* **2017**, *5*, 23. [\[CrossRef\]](#)

18. Tkachev, A.; Stekolshchikova, E.; Vanyushkina, A.; Zhang, H.; Morozova, A.; Zozulya, S.; Kurochkin, I.; Anikanov, N.; Egorova, A.; Yushina, E.; et al. Lipid alteration signature in the blood plasma of individuals with schizophrenia, depression, and bipolar disorder. *JAMA Psychiatry* **2023**, *80*, 250–259. [[CrossRef](#)] [[PubMed](#)]
19. Marković, S.; Jadranin, M.; Miladinović, Z.; Gavrilović, A.; Avramović, N.; Takić, M.; Tasic, L.; Tešević, V.; Mandić, B. LC-HRMS lipidomic fingerprints in Serbian cohort of schizophrenia patients. *Int. J. Mol. Sci.* **2024**, *25*, 10266. [[CrossRef](#)]
20. Jadranin, M.; Avramović, N.; Miladinović, Z.; Gavrilović, A.; Tasic, L.; Tešević, V.; Mandić, B. Untargeted lipidomics study of bipolar disorder patients in Serbia. *Int. J. Mol. Sci.* **2023**, *24*, 16025. [[CrossRef](#)]
21. Yan, L.; Zhou, J.; Wang, D.; Si, D.; Liu, Y.; Zhong, L.; Yin, Y. Unbiased lipidomic profiling reveals metabolomic changes during the onset and antipsychotics treatment of schizophrenia disease. *Metabolomics* **2018**, *14*, 80. [[CrossRef](#)]
22. Wang, F.; Guo, L.; Zhang, T.; Cui, Z.; Wang, J.; Zhang, C.; Xue, F.; Zhou, C.; Li, B.; Tan, Q.; et al. Alterations in plasma lipidomic profiles in adult patients with schizophrenia and major depressive disorder. *Medicina* **2022**, *58*, 1509. [[CrossRef](#)] [[PubMed](#)]
23. Shi, M.; Du, X.; Jia, Y.; Zhang, Y.; Jia, Q.; Zhang, X.; Zhu, Z. The identification of novel schizophrenia-related metabolites using untargeted lipidomics. *Cereb. Cortex.* **2024**, *34*, bhae160. [[CrossRef](#)]
24. Brunkhorst-Kanaan, N.; Klatt-Schreiner, K.; Hackel, J.; Schröter, K.; Trautmann, S.; Hahnefeld, L.; Wicker, S.; Reif, A.; Thomas, D.; Geisslinger, G.; et al. Targeted lipidomics reveal derangement of ceramides in major depression and bipolar disorder. *Metab. Clin. Exp.* **2019**, *95*, 65–76. [[CrossRef](#)]
25. Guo, L.; Zhang, T.; Li, R.; Cui, Z.Q.; Du, J.; Yang, J.B.; Xue, F.; Chen, Y.H.; Tan, Q.R.; Peng, Z.W. Alterations in the plasma lipidome of adult women with bipolar disorder: A mass spectrometry-based lipidomics research. *Front. Psychiatry* **2022**, *13*, 802710. [[CrossRef](#)]
26. Tomasik, J.; Harrison, S.J.; Rustogi, N.; Olmert, T.; Barton-Owen, G.; Han, S.Y.S.; Cooper, J.D.; Eljasz, P.; Farrag, L.P.; Friend, L.V.; et al. Metabolomic Biomarker Signatures for Bipolar and Unipolar Depression. *JAMA Psychiatry* **2024**, *81*, 101–106. [[CrossRef](#)] [[PubMed](#)]
27. Hashimoto, M.; Katakura, M.; Tanabe, Y.; Al Mamun, A.; Inoue, T.; Hossain, S.; Arita, M.; Shido, O. n-3 fatty acids effectively improve the reference memory-related learning ability associated with increased brain docosahexaenoic acid-derived docosanoids in aged rats. *Biochim. Biophys. Acta* **2015**, *1851*, 203–209. [[CrossRef](#)] [[PubMed](#)]
28. McDougall, M.; Choi, J.; Magnusson, K.; Truong, L.; Tanguay, R.; Traber, M.G. Chronic vitamin E deficiency impairs cognitive function in adult zebrafish via dysregulation of brain lipids and energy metabolism. *Free Radic. Biol. Med.* **2017**, *112*, 308–317. [[CrossRef](#)]
29. Hussain, G.; Wang, J.; Rasul, A.; Anwar, H.; Imran, A.; Qasim, M.; Zafar, S.; Kamran, S.K.S.; Razzaq, A.; Aziz, N.; et al. Role of cholesterol and sphingolipids in brain development and neurological diseases. *Lipids Health Dis.* **2019**, *18*, 26. [[CrossRef](#)]
30. Araque, A.; Castillo, P.E.; Manzoni, O.J.; Tonini, R. Synaptic functions of endocannabinoid signaling in health and disease. *Neuropharmacology* **2017**, *124*, 13–24. [[CrossRef](#)] [[PubMed](#)]
31. Egawa, J.; Pearn, M.L.; Lemkuil, B.P.; Patel, P.M.; Head, B.P. Membrane lipid rafts and neurobiology: Age-related changes in membrane lipids and loss of neuronal function. *J. Physiol.* **2016**, *594*, 4565–4579. [[CrossRef](#)]
32. Wu, L.; Zhang, X.; Zhao, L. Human ApoE isoforms differentially modulate brain glucose and ketone body metabolism: Implications for Alzheimer's disease risk reduction and early intervention. *J. Neurosci.* **2018**, *38*, 6665–6681. [[CrossRef](#)]
33. Wu, S.; Panganiban, K.J.; Lee, J.; Li, D.; Smith, E.C.C.; Maksyutynska, K.; Humber, B.; Ahmed, T.; Agarwal, S.M.; Ward, K.; et al. Peripheral lipid signatures, metabolic dysfunction, and pathophysiology in schizophrenia spectrum disorders. *Metabolites* **2024**, *14*, 475. [[CrossRef](#)]
34. Tkachev, A.; Stekolshchikova, E.; Anikanov, N.; Zozulya, S.; Barkhatova, A.; Klyushnik, T.; Petrova, D. Shorter chain triglycerides are negatively associated with symptom improvement in schizophrenia. *Biomolecules* **2021**, *11*, 720. [[CrossRef](#)]
35. Hsu, M.C.; Huang, Y.S.; Ouyang, W.C. Beneficial effects of omega-3 fatty acid supplementation in schizophrenia: Possible mechanisms. *Lipids Health Dis.* **2020**, *19*, 159. [[CrossRef](#)]
36. Psara, E.; Papadopoulou, S.K.; Mentzelou, M.; Voulgaridou, G.; Vorvolakos, T.; Apostolou, T.; Giaginis, C. Omega-3 fatty acids for the treatment of bipolar disorder symptoms: A narrative review of the current clinical evidence. *Mar. Drugs.* **2025**, *23*, 84. [[CrossRef](#)]
37. O'Brien, K.A.; Atkinson, R.A.; Richardson, L.; Koulman, A.; Murray, A.J.; Harridge, S.D.R.; Martin, D.S.; Levett, D.Z.H.; Mitchell, K.; Mythen, M.G.; et al. Metabolomic and lipidomic plasma profile changes in human participants ascending to Everest Base Camp. *Sci. Rep.* **2019**, *9*, 2297. [[CrossRef](#)]
38. MATLAB, version 9.7.0.1737446 (R2019b) Update 9; The MathWorks Inc.: Natick, MA, USA, 2022. Available online: <https://www.mathworks.com> (accessed on 1 January 2020).
39. Albanese, D.; Visintainer, R.; Merler, S.; Riccadonna, S.; Jurman, G.; Furlanello, C. mlpy: Machine Learning Python. *arXiv* **2012**, arXiv:1202.6548. [[CrossRef](#)]
40. Smith, C.A.; Want, E.J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **2006**, *78*, 779–787. [[CrossRef](#)]

41. Hubert, M.; Rousseeuw, P.J.; Vanden Branden, K. ROBPCA: A new approach to robust Principal Component Analysis. *Technometrics* **2005**, *47*, 64–79. [\[CrossRef\]](#)
42. Hubert, M.; Rousseeuw, P.; Verdonck, T. Robust PCA for skewed data and its outlier map. *Comput. Stat. Data Anal.* **2009**, *53*, 2264–2274. [\[CrossRef\]](#)
43. Verboven, S.; Hubert, M. MATLAB library LIBRA. *WIREs Comput. Stat.* **2010**, *2*, 509–515. [\[CrossRef\]](#)
44. *PLS_Toolbox*, 8.9.1; Eigenvector Research, Inc.: Manson, WA, USA, 2021. Available online: <http://www.eigenvector.com> (accessed on 1 January 2021).
45. Simić, K.; Todorović, N.; Trifunović, S.; Miladinović, Z.; Gavrilović, A.; Jovanović, S.; Avramović, N.; Gođevac, D.; Vujisić, L.; Tešević, V.; et al. NMR metabolomics in serum fingerprinting of schizophrenia patients in a Serbian cohort. *Metabolites* **2022**, *12*, 707. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Simić, K.; Miladinović, Z.; Todorović, N.; Trifunović, S.; Avramović, N.; Gavrilović, A.; Jovanović, S.; Gođevac, D.; Vujisić, L.; Tešević, V.; et al. Metabolomic profiling of bipolar disorder by 1H-NMR in Serbian patients. *Metabolites* **2023**, *13*, 607. [\[CrossRef\]](#)
47. MATLAB Help Center. Fill Missing Entries—MATLAB Fillmissing. Available online: <https://www.mathworks.com/help/matlab/ref/fillmissing.html> (accessed on 11 February 2024).
48. Xia, J.; Broadhurst, D.I.; Wilson, M.; Wishart, D.S. Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics* **2013**, *9*, 280–299. [\[CrossRef\]](#)
49. Ejigu, B.A.; Valkenborg, D.; Baggerman, G.; Vanaerschot, M.; Witters, E.; Dujardin, J.-C.; Burzykowski, T.; Berg, M. Evaluation of normalization methods to pave the way towards large-scale LC-MS-based metabolomics profiling experiments. *OMICS J. Integr. Biol.* **2013**, *17*, 473–485. [\[CrossRef\]](#)
50. Tharwat, A. Classification assessment methods. *Appl. Comput. Inform.* **2021**, *17*, 168–192. [\[CrossRef\]](#)
51. Eigenvector Research Documentation Wiki. Plsda. Available online: <https://wiki.eigenvector.com/index.php?title=Plsda> (accessed on 28 September 2022).
52. Ballabio, D.; Grisoni, F.; Todeschini, R. Multivariate comparison of classification performance measures. *Chemom. Intell. Lab. Syst.* **2018**, *174*, 33–44. [\[CrossRef\]](#)
53. Barker, M.; Rayens, W. Partial least squares for discrimination. *J. Chemom.* **2003**, *17*, 166–173. [\[CrossRef\]](#)
54. Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* **2002**, *16*, 119–128. [\[CrossRef\]](#)
55. Szymańska, E.; Saccenti, E.; Smilde, A.K.; Westerhuis, J.A. Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* **2012**, *8*, 3–16. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Lee, L.C.; Liong, C.-Y.; Jemain, A.A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps. *Analyst* **2018**, *143*, 3526–3539. [\[CrossRef\]](#)
57. Haenlein, M.; Kaplan, A.M. A beginner's guide to partial least squares analysis. *Underst. Stat.* **2004**, *3*, 283–297. [\[CrossRef\]](#)
58. Worley, B.; Powers, R. PCA as a practical indicator of OPLS-DA model reliability. *Curr. Metabolomics* **2016**, *4*, 97–103. [\[CrossRef\]](#)
59. Pérez, N.F.; Ferré, J.; Boqué, R. Calculation of the reliability of classification in discriminant partial least-squares binary classification. *Chemom. Intell. Lab. Syst.* **2009**, *95*, 122–128. [\[CrossRef\]](#)
60. Eigenvector Research Documentation Wiki. Faq Is Prediction Probability and Threshold Calculated for PLSDA. Available online: https://wiki.eigenvector.com/index.php?title=FaQ_is_prediction_probability_and_threshold_calculated_for_PLSDA (accessed on 10 December 2024).
61. Eigenvector Research Documentation Wiki. Crossval. Available online: <https://www.wiki.eigenvector.com/index.php?title=Crossval> (accessed on 12 February 2025).
62. Morán-Fernández, L.; Bolón-Canedo, V. Finding a needle in a haystack: Insights on feature selection for classification tasks. *J. Intell. Inf. Syst.* **2024**, *62*, 459–483. [\[CrossRef\]](#)
63. Yun, Y.-H.; Li, H.-D.; Deng, B.-C.; Cao, D.-S. An overview of variable selection methods in multivariate analysis of near-infrared spectra. *Trends Anal. Chem.* **2019**, *113*, 102–115. [\[CrossRef\]](#)
64. Chong, I.-G.; Jun, C.-H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **2005**, *78*, 103–112. [\[CrossRef\]](#)
65. Gosselin, R.; Rodrigue, D.; Duchesne, C. A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications. *Chemom. Intell. Lab. Syst.* **2010**, *100*, 12–21. [\[CrossRef\]](#)
66. Mehmood, T.; Liland, K.H.; Snipen, L.; Sæbø, S. A review of variable selection methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* **2012**, *118*, 62–69. [\[CrossRef\]](#)
67. Jurman, G.; Merler, S.; Barla, A.; Paoli, S.; Galea, A.; Furlanello, C. Algebraic stability indicators for ranked lists in molecular profiling. *Bioinformatics* **2008**, *24*, 258–264. [\[CrossRef\]](#)
68. Leardi, R.; Seasholtz, M.B.; Pell, R.J. Variable selection for multivariate calibration using a genetic algorithm: Prediction of additive concentrations in polymer films from Fourier transform-infrared spectral data. *Anal. Chim. Acta* **2002**, *461*, 189–200. [\[CrossRef\]](#)
69. Fernández Pierna, J.A.; Abbas, O.; Baeten, V.; Dardenne, P. A Backward Variable Selection method for PLS regression (BVSPLS). *Anal. Chim. Acta* **2009**, *642*, 89–93. [\[CrossRef\]](#)

70. Leardi, R. Application of genetic algorithm–PLS for feature selection in spectral data sets. *J. Chemom.* **2000**, *14*, 643–655. [\[CrossRef\]](#)
71. Gibbons, J.D.; Chakraborti, S. *Nonparametric Statistical Inference*, 4th ed.; Marcel Dekker: New York, NY, USA, 2003; p. 289.
72. MATLAB Help Center. MINTERSECT—Multiple Set Intersection. Available online: <https://www.mathworks.com/matlabcentral/fileexchange/6144-mintersect-multiple-set-intersection>
73. Leardi, R. Application of a genetic algorithm to feature selection under full validation conditions and to outlier detection. *J. Chemom.* **1994**, *8*, 65–79. [\[CrossRef\]](#)
74. Lindgren, F.; Hansen, B.; Karcher, W.; Sjöström, M.; Eriksson, L. Model validation by permutation tests: Applications to variable selection. *J. Chemom.* **1996**, *10*, 521–532. [\[CrossRef\]](#)
75. Jović, A.; Brkić, K.; Bogunović, N. A review of feature selection methods with applications. In Proceedings of the 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 25–29 May 2015; pp. 1200–1205. [\[CrossRef\]](#)
76. Barupal, D.L.; Zhang, Y.; Fan, S.; Hazen, W.H.; Tang, W.; Cajka, T.; Irvin, M.R.; Arnett, D.K.; Kind, T.; Kaddurah-Daouk, R.; et al. The circulating lipidome is largely defined by sex descriptors in the GOLDN, GeneBank and the ADNI studies. *bioRxiv* **2019**. bioRxiv:731448. [\[CrossRef\]](#)
77. Ishikawa, M.; Maekawa, K.; Saito, K.; Senoo, Y.; Urata, M.; Murayama, M.; Tajima, Y.; Kumagai, Y.; Saito, Y. Plasma and serum lipidomics of healthy white adults shows characteristic profiles by subjects' gender and age. *PLoS ONE* **2014**, *9*, e91806. [\[CrossRef\]](#)
78. Slade, E.; Irvin, M.R.; Xie, K.; Arnett, D.K.; Claas, S.A.; Kind, T.; Fardo, D.W.; Graf, G.A. Age and sex are associated with the plasma lipidome: Findings from the GOLDN study. *Lipids Health Dis.* **2021**, *20*, 30. [\[CrossRef\]](#)
79. Tabassum, R.; Widén, E.; Ripatti, S. Effect of biological sex on human circulating lipidome: An overview of the literature. *Atherosclerosis* **2023**, *384*, 117274. [\[CrossRef\]](#)
80. Knowles, E.E.M.; Meikle, P.J.; Huynh, K.; Göring, H.H.; Olvera, R.L.; Mathias, S.R.; Duggirala, R.; Almasy, L.; Blangero, J.; Curran, J.E.; et al. Serum phosphatidylinositol as a biomarker for bipolar disorder liability. *Bipolar Disord.* **2017**, *19*, 107–115. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Bernal-Vega, S.; García-Juárez, M.; Camacho-Morales, A. Contribution of ceramides metabolism in psychiatric disorders. *J. Neurochem.* **2023**, *164*, 708–724. [\[CrossRef\]](#)
82. Costa, A.C.; Riça, L.B.; van de Bilt, M.; Zandonadi, F.S.; Gattaz, W.F.; Talib, L.L.; Sussulini, A. Application of lipidomics in psychiatry: Plasma-based potential biomarkers in schizophrenia and bipolar disorder. *Metabolites* **2023**, *13*, 600. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Tao, S.; Zhang, Y.; Wang, Q.; Qiao, C.; Deng, W.; Liang, S.; Wei, J.; Wei, W.; Yu, H.; Li, X.; et al. Identifying transdiagnostic biological subtypes across schizophrenia, bipolar disorder, and major depressive disorder based on lipidomics profiles. *Front. Cell Dev. Biol.* **2022**, *10*, 969575. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Yu, L.; Long, Q.; Zhang, Y.; Liu, Y.; Guo, Z.; Cao, X.; Qin, F.; Xu, Y.; Qian, Q.; Gao, B.; et al. Bidirectional Mendelian randomization analysis of plasma lipidome and psychiatric disorders. *J. Affect. Disord.* **2024**, *379*, 871–883. [\[CrossRef\]](#) [\[PubMed\]](#)
85. Liu, Y.; Song, X.; Liu, X.; Pu, J.; Gui, S.; Xu, S.; Tian, L.; Zhong, X.; Zhao, L.; Wang, H.; et al. Alteration of lipids and amino acids in plasma distinguish schizophrenia patients from controls: A targeted metabolomics study. *Psychiatry Clin. Neurosci.* **2021**, *75*, 138–144. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.