Table S1 Comprehensive Summary of Data-Driven MDD Subtyping Studies.

Study (Author, Year)	Subtype/Cluster Name	Data Type/Method	Biological Basis (Results/Conclusion)	Clinical/Cognitive Features	Key Findings	Clinical Correlates (Etiology, Comorbidity, Course, Severity, Response)
Tao et al. 2025 [1]	**Cognitive Deficit (CD)**	Neuropsychological (CANTAB) + MRI	Reduced GMV in left fusiform gyrus and cerebellum.	Significant deficits in attention and memory.	Reductions in specific regions serve as potential neuroimaging biomarkers for the CD subtype.	**Comorbidity:** CD group associated with significantly fewer education years (13.40 ± 3.49, $p < 0.0001$).
Tang et al. 2025 [2]	**Subtype 1 (Genetic)**	Multi-omics + ALFF	High PRS for MDD; enrichment in synaptic regulation pathways.	Most severe HAMD symptoms and cognitive decline (WCST).	Subtype 1 is characterized by higher polygenic risk and synaptic pathway enrichment.	**Severity:** Significantly higher HAMD scores (21.1 ± 10.1) than other subtypes. **BMI:** S1 BMI (21.7 ± 3.6) lower than HC.
Tang et al. 2025 [2]	**Subtype 2 (Immuno-inflammatory)**	Multi-omics + ALFF	Elevated IL-1 beta; lower Epigenetic Inflammation Score (EIS).	Intermediate symptom severity; cognitive deficits linked to precuneus ALFF.	Homogeneous subtype potentially responsive to adjunctive anti-inflammatory drugs.	**Etiology:** Driven by immune-inflammatory dysregulation. **Response:** May be a candidate for targeted anti-inflammatory therapy.

Hagenberg et al. 2025 [3]	**High Immune-Related (HIRDS 1/2)**	Multi-omics (Plasma + RNA-seq)	High inflammation (CRP); altered PBMC gene expression (*NRCAM*).	Elevated depression severity (BDI-II and MADRS).	Identified 4 biological clusters corresponding to distinct clinical phenotypes.	**Etiology:** Strongly linked to immuno-metabolic dysregulation. **Comorbidity:** Increased BMI and sleep in HIRDS.
Tse et al. 2024 [4]	**Youth MDD Networks**	Meta-analysis of rs-FC (fMRI)	Dysconnectivity consistently involving DMN, DAN, and VAN.	Symptom severity associated with specific network alterations.	Models predicted diagnostic status (AUC of 0.731 for predicting diagnostic status and symptom severity (r = 0.14, p = 0.008).	**Etiology:** Largest pooled mega-analysis identifying robust youth connectivity biomarkers.
Kashiwagi et al. 2026 [5]	**Weak-Correlation (Subtype 2)**	rs-FC (resting-state fMRI)	Thalamo-somatomotor functional connectivity strength.	Standard depressive symptom profile.	Connectivity strength closer to the mean predicted superior antidepressant outcomes.	**Response:** Tended toward higher improvement rates (p = 0.085) than other subtypes.
Colombo et al. 2024 [6]	**High Treatment-Resistant (Cluster associated with higher treatment resistance (TRD))**	Structural MRI (CT, GMV, FA)	Widespread reduction in cortical thickness and white matter integrity.	BDI Anergy domain: fatigue and decreased appetite.	Structural signatures identified TRD subtypes with 68.4% accuracy.	**Response:** Significantly higher rates of treatment resistance compared to other clusters.

Xiao et al. 2024 [7]	**Archetypal vs. Atypical**	ALFF + Machine Learning	Frontal-posterior functional imbalance (Archetypal: Frontal ↑; Atypical: Frontal ↓).	Both show significant depressive and anxiety symptoms.	Subtype-specific rTMS targets (dmPFC, OCC) improved symptoms.	**Response:** classifiers allow for precise targeted rTMS therapy.
Habets et al. 2023 [8]	**Remission vs. Chronicity**	Integrated Multi-omics + Clinical	Proteomic pathways (IL-10 signaling, lipids).	Baseline symptom severity (IDS-SR) and Neuroticism.	Unimodal proteomic model (AUROC = 0.68) outperformed clinical predictions.	**Course:** Predicts remission vs. chronic persistence at 2-year follow-up.
Zhou et al. 2024 [9]	**Anxious Depression (AD)**	Anxiety Factor + Multimodal MRI	AD-specific abnormalities in Left ITG, STG, and SFGdor.	Agitation and restlessness.	Multimodal MRI features identify biological distinctions from Non-Anxious Depression.	**Severity:** AD has more complex and severe symptoms. **Course:** Prior literature suggests higher risk of recurrence and suicidality in anxious depression.
Liang et al. 2020 [10]	**hyperDMN vs. hypoDMN**	rs-FC within DMN	hyperDMN: increased prefrontal-PCC/PrC connectivity.	Comparable demographics and HAMD scores between groups.	Two robust neural subtypes identified and validated across multi-site datasets.	**Course:** No significant linear correlation was found between age and other clinical variables.
Feng et al. 2025 [11]	**PRG Subtypes A & B**	Transcriptomics + Machine Learning	DE-PRG signatures (*GZMA*, *TNF*); Subtype B upregulated.	Significant differences in immune infiltration (e.g., M2 macrophages).	Identified candidate immunological biomarkers (GZMA, AKR1C3, CD52) for MDD.	**Etiology:** Linked to pyroptosis-related inflammation.

Zhang et al. 2023 [12]	**Mitophagy Subtypes 1 & 2**	Transcriptomics + WGCNA	5 biomarkers (*MATR3*, *ACTL6A*, *FUS*, *BIRC2*, *RIPK1*).	Distinct immune cell patterns (Subtype 1: higher memory B cells).	Mitophagy-related gene signature distinguishes MDD and molecular subtypes.	**Etiology:** Linked to mitochondrial mechanisms and biological heterogeneity.
Ding et al. 2023 [13]	**Anhedonia Subgroup C**	Clinical (TEPS) + rs-fMRI	Aberrant connectivity between reward network and DMN.	Most profound impairment in consummatory pleasure (31.29 ± 4.77).	Anhedonia is non-uniform; Subgroup C has the most profound impairment.	**Etiology:** Subgroup C (lowest TEPS scores) is associated with the most profound impairment in both anticipatory and consummatory pleasure, consistent with features of melancholic depression.
Jing et al. 2023 [14]	**Subtype I (16%) vs. Subtype II (84%)**	rs-FC + Normative Modeling	Subtype I has larger extreme functional deviations in 10 networks.	Comparable baseline symptom severity (HAMD/HAMA).	Subtypes identified by characterising functional deviations from normative range.	**Course:** Subtype I illness duration (11.7 months); Subtype II longer (22.1 months).
Liu et al. 2016 [15]	**Melancholic (MelD)**	Blood Metabolomics	High biological homogeneity in 76 unique metabolites.	Pervasive anhedonia and weight loss.	Metabolomic biosignatures accurately classify MelD (83.84% accuracy).	**Severity:** Higher HAMD (24.57 ± 4.46) and CORE (11.24 ± 3.96) scores than others.

Grant et al. 2022 [16]	**Suicide Attempt (SA)**	Multi-omics + History	integrative genomic-metabolomic signature for attempt history.	SA group had earlier depression onset (65% vs 37% before age 18).	SNVs and metabolites identify signature of suicide attempts in MDD adults.	**Etiology:** Subgroup defined by history of suicide attempt.
Collins et al. 2022 [17]	**Affective vs Somatic**	STAR*D Clinical Data	IDS-30 component analysis identified somatic vs affective clusters.	Somatic cluster characterized by energy, sleep, and appetite symptoms.	Full IDS (30 items) better captures clinical heterogeneity than QIDS (12 items).	**Response:** Highlights importance of targeting biological correlates of somatic symptoms.
Yeung et al. 2024 [18]	**MDD Definitions**	Connectomics + ML	compared 6 MDD definitions (Ever, Recurrent, Narrow, etc.).	Defined by varying clinical criteria and childhood trauma exposure.	Classification accuracy varied across MDD phenotypes and improved in more homogeneous, childhood trauma-stratified subgroups.	**Course:** Childhood trauma stratification enhanced subgroup homogeneity and classification performance.
Ichikawa et al. 2020 [19]	**Melancholic Biotype**	rs-FC (MRI)	Development of melancholic MDD classifier.	Moderate-to-severe symptoms (BDI-II ≥ 17).	Classifier generalizable to independent cohorts and sites.	**Severity:** specifically targeted at patients with melancholic M.I.N.I. subtype.

Zhang et al. 2020 [20]	**SSRI-Response Biotypes**	Clinical + tagSNPs (*CREB1*, *BDNF*)	Selected tagSNPs in CREB1 and BDNF improved prediction accuracy for SSRI resistance.	Non-Responsive (SSRI-R): young onset and frequent episodes.	Combined clinical and genetic variables improved prediction of SSRI response (PM9 accuracy = 77.2%).	**Response:** Young age at onset and episode frequency were associated with poorer SSRI response.
Yang et al. 2025 [21]	**Deteriorated IQ (DIQ)**	IQ Trajectories + MRI	Widespread structural (GMV) and functional (ALFF) abnormalities in DIQ.	Poor performance on logical memory and executive function tasks.	IQ decline is a trait marker for MDD cognitive heterogeneity; AUC = 0.8023.	**Severity:** DIQ patients exhibit significantly greater cognitive impairment.
Frässle et al. 2020 [22]	**REM, IMP, and CHR**	Generative embedding (fMRI)	Effective connectivity parameter estimates derived from DCM.	REM: fast remission; CHR: chronic (no improvement over 2 years).	Classifier discriminates CHR from REM with 79% balanced accuracy.	**Course:** Correctly identifies individuals at high risk for a chronic course.
Chang et al. 2019 [23]	**Antidepressant Response**	Deep Learning (ARPNet)	Integration of *BDNF*, *SLC6A4*, and MRI features.	Individual degree of antidepressant response.	ARPNet provides robust individual response prediction (Task 1 RMSE 3.30).	**Response:** predicts individual treatment response based on integrated markers.
Kessler et al. 2016 [24]	**Persistence/Severity**	Baseline Self-reports	Predictors focus on prognostic value over 10-12 years.	Predictive of MDD persistence, chronicity, and disability over time.	ML models provide prognostic subtypes (AUC 0.63-0.76) from survey data.	**Course:** Predicts high persistence (34.6-38.1%) and chronicity.
Chen et al. 2024 [25]	**Cluster A (MCI-like)**	MMSE domains + Plasma NfL	Higher NfL negatively correlated with MMSE scores.	Lower orientation, memory, and calculation MMSE scores.	Subgroup of older patients may progress to clinical dementia.	**Severity:** severe cognitive impairment observed in Cluster A.

Hack et al. 2023 [26]	**Cognitive Dyscontrol**	Behavioral Tests + Task fMRI	Reduced activation in right dLPFC and dACC.	Marked impairment in executive function and concentration.	Cognitive biotype represents a distinct subgroup; impairment persists regardless of symptom change.	**Response:** Was associated with poorer treatment outcomes (lower remission rates).
Pfarr et al. 2023 [27]	**Gyrification-based**	Cortical Gyrification (MRI)	Cluster differences in SFG, Middle Frontal Gyrus, and Cingulate.	Cluster 3: poorest cognitive performance (VF, DSST, CBBT).	Clinical diagnoses do not form biologically homogeneous groups in cortical folding.	**Etiology:** Cluster 3 shows significant association with prenatal-risk scores ($p = 0.026$).
Yeung et al. 2021 [28]	**Structural Natural Groupings**	Structural Metrics (CT, CV, subCV)	Two robust clusters identified in population structural data.	Significantly associated with general cognitive ability (g-factor).	Natural groupings identified in population data are **NOT associated with MDD status**.	**Severity:** MDD status was not significantly associated with these structural groupings ($p > 0.05$).
Tokuda et al. 2021 [29]	**Functional View-Based**	rs-FC + Multiple Clustering	Identification of transdiagnostic psychiatric subgroups.	Transdiagnostic features across HC, MDD, BD, and SCZ.	Distribution pattern of cluster labels was statistically consistent across KYO and UTO datasets.	**Etiology:** Unsupervised approach identifies transdiagnostic risk biotypes.

Wade et al. 2021 [30]	**Symptom Dimensions**	Random Forest + Volumetric MRI	Volumetric importance in right transverse temporal gyrus and left frontal pole.	Dimensions: Somatic (SoD), Core Mood/Anhedonia (CMA), and Insomnia.	Symptom dimension change is predicted more accurately than HDRS-17 total score ($R2$ up to 38-47%).	**Response:** Recovery along dimensions relates to unique patterns of volumetric change.
Watters et al. 2018 [31]	**Affective Dysfunction**	Circuitry + Behavior + Self-report	Hyper-reactivity in pregenual ACC (anger) and right amygdala (sadness).	Profile of less frequent use of emotion reappraisal strategies.	Unaffected relatives can be grouped into risk biotypes based on emotion profiles.	**Etiology:** Familial risk involves disruption to neural circuits for appraising emotions.

References

1. Tao Y, Yan Y, Wang M, Fan H, Dou Y, Zhao L, et al. Data-driven cognitive subtypes in major depressive disorder: Grey matter atrophy in the left fusiform gyrus and cerebellum. J Affect Disord. 2025; 391: 120068.
2. Tang L, Tang R, Zheng J, Zhao P, Zhu R, Tang Y, et al. Dissecting biological heterogeneity in major depressive disorder based on neuroimaging subtypes with multi-omics data. Transl Psychiatry. 2025; 15: 72.
3. Hagenberg J, Brückl TM, Erhart M, Kopf-Beck J, Ködel M, Rehawi G, et al. Dissecting depression symptoms: Multi-omics clustering uncovers immune-related subgroups and cell-type specific dysregulation. Brain Behav Immun. 2025; 123: 353-369.
4. Tse NY, Ratheesh A, Tian YE, Connolly CG, Davey CG, Ganesan S, et al. A mega-analysis of functional connectivity and network abnormalities in youth depression. Nat Ment Health. 2024; 2: 1169-1182.
5. Kashiwagi Y, Tokuda T, Takahara Y, Masaki Y, Sakai Y, Yoshimoto J, et al. Generalizable stratification based on thalamo-somatomotor functional connectivity predicts responses to antidepressants in patients with depression. Mol Psychiatry. 2026; 31: 270-281.
6. Colombo F, Calesella F, Bravi B, Fortaner-Uyà L, Monopoli C, Tassi E, et al. Multimodal brain-derived subtypes of major depressive disorder differentiate patients for anergic symptoms, immune-inflammatory markers, history of childhood trauma and treatment-resistance. Eur Neuropsychopharmacol. 2024; 85: 45-57.
7. Xiao Y, Womer FY, Dong S, Zhu R, Zhang R, Yang J, et al. A neuroimaging-based precision medicine framework for depression. Asian J Psychiatry. 2024; 91: 103803.
8. Habets PC, Thomas RM, Milaneschi Y, Jansen R, Pool R, Peyrot WJ, et al. Multimodal data integration advances longitudinal prediction of the naturalistic course of depression and reveals a multimodal signature of remission during 2-year follow-up. Biol Psychiatry. 2023; 94: 948-958.
9. Zhou E, Wang W, Ma S, Xie X, Kang L, Xu S, et al. Prediction of anxious depression using multimodal neuroimaging and machine learning. Neuroimage. 2024; 285: 120499.
10. Liang S, Deng W, Li X, Greenshaw AJ, Wang Q, Li M, et al. Biotypes of major depressive disorder: Neuroimaging evidence from resting-state default mode network patterns. NeuroImage Clin. 2020; 28: 102514
11. Feng L, Yuan J, Li L, Tang J. Identification of pyroptosis-related molecular subtypes and diagnostic model development in major depressive disorder. Mol Biotechnol. 2025; 67: 3281-3295.
12. Zhang J, Xie S, Xiao R, Yang D, Zhan Z, Li Y. Identification of mitophagy-related biomarkers and immune infiltration in major depressive disorder. BMC Genom. 2023; 24: 216.
13. Ding Y, Ou Y, Yan H, Liu F, Li H, Li P, et al. Uncovering the neural correlates of anhedonia subtypes in major depressive disorder: Implications for intervention strategies. Biomedicines. 2023; 11: 3138.
14. Jing R, Lin X, Ding Z, Chang S, Shi L, Liu L, et al. Heterogeneous brain dynamic functional connectivity patterns in first-episode drug-naive patients with major depressive disorder. Hum Brain Mapp. 2023; 44: 3112-3122.

15. Liu Y, Yieh L, Yang T, Drinkenburg W, Peeters P, Steckler T, et al. Metabolomic biosignature differentiates melancholic depressive patients from healthy controls. BMC Genom. 2016; 17: 669.
16. Grant CW, Wilton AR, Kaddurah-Daouk R, Skime M, Biernacka J, Mayes T, et al. Network science approach elucidates integrative genomic-metabolomic signature of antidepressant response and lifetime history of attempted suicide in adults with major depressive disorder. Front Pharmacol. 2022; 13: 984383.
17. Collins KA, Eng GK, Tural Ü, Irvin MK, Iosifescu DV, Stern ER. Affective and somatic symptom clusters in depression and their relationship to treatment outcomes in the STAR*D sample. J Affect Disord. 2022; 300: 469-473.
18. Yeung HW, Stolicyn A, Shen X, Adams MJ, Romaniuk L, Thng G, et al. Classification accuracy of structural and functional connectomes across different depressive phenotypes. Imaging Neurosci. 2024; 2: imag-2-00064.
19. Ichikawa N, Lisi G, Yahata N, Okada G, Takamura M, Hashimoto RI, et al. Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants. Sci Rep. 2020; 10: 3524.
20. Zhang H, Li X, Pang J, Zhao X, Cao S, Wang X, et al. Predicting SSRI-resistance: Clinical features and tagSNPs prediction models based on support vector machine. Front Psychiatry. 2020; 11: 493.
21. Yang X, Liao Q, Wang M, Yan Y, Fan H, Wei J, et al. Cognitive heterogeneity in major depressive disorder: Classification by IQ trajectory and multimodal neuroimaging profiles. BMC Psychiatry. 2025; 25: 754.
22. Frässle S, Marquand AF, Schmaal L, Dinga R, Veltman DJ, van der Wee NJA, et al. Predicting individual clinical trajectories of depression with generative embedding. NeuroImage Clin. 2020; 26: 102213.
23. Chang B, Choi Y, Jeon M, Lee J, Han KM, Kim A, et al. ARPNet: Antidepressant response prediction network for major depressive disorder. Genes. 2019; 10: 907.
24. Kessler RC, van Loo HM, Wardenaar KJ, Bossarte RM, Brenner LA, Cai T, et al. Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. Mol Psychiatry. 2016; 21: 1366-1371.
25. Chen CY, Chiu CC, Huang CY, Cheng YC, Huang MC, Kuo PH, et al. Cluster analysis dissecting cognitive deficits in older adults with major depressive disorder and the association with neurofilament light chain. BMC Geriatr. 2024; 24: 344.
26. Hack LM, Tozzi L, Zenteno S, Olmsted AM, Hilton R, Jubeir J, et al. A cognitive biotype of depression and symptoms, behavior measures, neural circuits, and differential treatment outcomes: A prespecified secondary analysis of a randomized clinical trial. JAMA Netw Open. 2023; 6: e2318411.
27. Pfarr JK, Meller T, Brosch K, Stein F, Thomas-Odenthal F, Evermann U, et al. Data-driven multivariate identification of gyrification patterns in a transdiagnostic patient cohort: A cluster analysis approach. NeuroImage. 2023; 281: 120349.
28. Yeung HW, Shen X, Stolicyn A, de Nooij L, Harris MA, Romaniuk L, et al. Spectral clustering based on structural magnetic resonance imaging and its relationship with major depressive disorder and cognitive ability. Eur J Neurosci. 2021; 54: 6281-6303.

29. Tokuda T, Yamashita O, Sakai Y, Yoshimoto J. Clustering of multiple psychiatric disorders using functional connectivity in the data-driven brain subnetwork. Front Psychiatry. 2021; 12: 683280.
30. Wade BSC, Hellemann G, Espinoza RT, Woods RP, Joshi SH, Redlich R, et al. Accounting for symptom heterogeneity can improve neuroimaging models of antidepressant response after electroconvulsive therapy. Hum Brain Mapp. 2021; 42: 5322-5333.
31. Watters AJ, Korgaonkar MS, Carpenter JS, Harris AWF, Gross JJ, Williams LM. Profiling risk for depressive disorder by circuit, behavior and self-report measures of emotion function. J Affect Disord. 2018; 227: 595-602.