Investigating peripheral metabolites and lipids as potential biomarkers for Major Depressive Disorder

Alexandra Lim Yun

Lady Margaret Hall, University of Oxford

Supervisors: Dr. Philip J. Burnet, Dr. Fay Probert

A thesis submitted for the degree of

Doctor of Philosophy,

Hilary 2024

Acknowledgements

This DPhil has been a most challenging yet fulfilling experience and would not have been possible without a few key people.

First, my family. Mum, Dad, Georgia, Philippa, Emilia. You have been invaluable to my growth and success as an individual overall, and were extremely supportive throughout the DPhil milestones, listening to my practice presentations, reading my work, and providing feedback. My family at Oxford– Daan, Tom, Tomas, Lorika, Max, Churl-Su, and Charlie, you were the best, most positive company and made life at LMH and Oxford in general fun and meaningful. I will really miss that time; however short it was.

Leia, Anastasia, Nina, Ruru, Miriam, Emma, Katie, Carla, Sarayu, whether near or far, you guys saw me through the lowest of lows and highest of highs, all of which happened to occur during my DPhil years. You are all beautiful inside and out, and I admire each of you tremendously. Core, lifelong friendships, indeed.

Finally, none of this would have been possible without my dear supervisor Phil, who played an integral role in giving me the positivity and motivation needed to challenge what I ever thought I was capable of. I am also extremely grateful for my secondary supervisor Fay Probert, who played a crucial role in helping me complete the thesis with her prowess in chemistry, coding, and statistics.

Hats off, to all of you.

Abstract:

Major Depressive Disorder (MDD) poses significant clinical challenges due to its complicated and elusive pathophysiology. This thesis seeks to fill research gaps in understanding MDD by investigating peripheral metabolites and lipids as potential biomarkers which could be used to improve diagnosis and prognosis of this debilitating mental health condition. Small molecules, such as amino acids, that are involved in energy homeostasis are of particular interest to investigate alongside lipids because mood disorders are strongly associated with metabolic dysfunction. A systematic review and meta-analysis was first conducted on articles reporting metabolomic analyses on amino acids and demonstrated significantly elevated citrate, alanine, and glutamate levels in peripheral biofluids from depressed people compared to healthy individuals. The investigation analyzed metabolomic and lipidomic data from UK twin populations, focusing on individuals assessed using the Hospital Anxiety and Depression Scale (HADS). Participants meeting inclusion criteria had at least one HADS score and <15% missing data. With 1,532 entries, the primary emphasis was on a two-year database, revealing insights into twin health, including 43 Normal, 35 Borderline Abnormal, and 35 Abnormal samples. This comprehensive study illuminated connections between lifestyle factors, inflammatory markers, lipidomics, and MDD severity. Positive associations between smoking and alcohol consumption underscored gender-specific implications and negative correlations between exercise and HADS suggested a protective relationship against depression, aligning with established benefits for mental well-being. Pyruvate and tyrosine were found to be positively associated with depression, while alanine and acetate were negatively associated with depression. However, none of these associations were statistically significant. Multiple regression study found positive relationships between IDLFC, SLDLL, LDLC, and HADS levels, while others exhibited significant negative associations. SLDLCE and SLDLFC were found to have the highest impact on HADS scores, highlighting the complex nature of lipid metabolism in depression. Furthermore, the inverse connection between cholesterol and cholesterol esters in small and medium LDL particles and depression severity shows that cholesterol may be a more important biomarker for major depressive disorder (MDD), regardless of lipoprotein size or density. The OPLS-DA model

effectively distinguished normal and abnormal mental health categories, highlighting the discriminative potential of metabolite biomarkers. Conspicuously, "AcAce" led with a VIP score of 2.24, followed by "BOHBut" (1.89) and "Ala" (1.75), emphasizing their impact. The lipid analysis illuminated complex shifts in cholesterol metabolism and lipid species, accentuating their crucial role in the pathophysiology of depression. "FreeC" stood out with a VIP score of 1.04, indicating its significant predictive influence, while "SM" closely followed with a score of 1.03. Additionally, "IDLFC" and "LLDLL" demonstrated remarkable importance, with VIP scores of 1.02 and 1.01, respectively, capturing essential data patterns. Comorbidity was found to influence lipid profiles, emphasizing the need to consider psychiatric conditions in research and clinical practice. Collectively, these findings contribute valuable insights into the multifaceted nature of MDD, proposing potential new biomarkers that may refine diagnostic accuracy and deepen our comprehension of this complex disorder.

Keywords: Major Depressive Disorder (MDD), Metabolomics, Lipids, OPLS-DA, Biomarkers, Lifestyle Factors

Contents

List of Figures

List of Tables

List of Abbreviations

Pyr= Pyruvate

Cit= Citrate

Ala= Alanine

Glc= Glucose

Gln= Glutamine

Gly= Glycine

His= Histidine

Phe= Phenylalanine

Tyr= Tyrosine

Ace= Acetate

AcAce= Acetoacetate

BOHBut=3-hydroxybutyrate

Crea= Creatinine

Alb= Albumin

Ile= Isoleucine

Leu= Leucine

Val= Valine

Lac= Lactate

Glol= Glycerol

Gp= Glycoprotein acetyls, mainly a1-acid glycoprotein

IDLC: total cholesterol in IDL (intermediate-density lipoprotein)

IDLFC: free cholesterol in IDL

LLDLL: Total lipids in large LDL (ow-density lipoprotein)

LLDLC: total cholesterol in large LDL

LLDLFC: free cholesterol in large LDL

MLDLL: total lipids in medium LDL

MLDLC: cholesterol in medium LDL

MLDLCE: cholesterol esters in medium LDL

MLDLFC: free cholesterol in medium LDL

SLDLL: total lipids in small LDL

SLDLC: total cholesterol in small LDL

SLDLCE: cholesterol esters in small LDL

SLDLFC: free cholesterol in small LDL

SLDLPL: phospholipids in small LDL

MLDLPL: phospholipids in medium LDL

LDLC: cholesterol in LDL

SM: sphingomyelin

DHA: docosahexaenoic acid

Free cholesterol

Smoking= At what age did you smoke your first whole cigarette?

Alcohol= For how long did you consume this maximum amount of alcohol each week?

Exercise = What is your leisure-time physical activity during the last 12 months?

BP_age = Age when high hypertension was first diagnosed

HC_Age = Age when high cholesterol was first diagnosed

Supplement=For how long (Years) have you taken this current supplement?

Chapter 1: Introduction

1.1 Major Depressive Disorder: A Complex Clinical Challenge

Major Depressive Disorder (MDD) is a formidable challenge in clinical practice due to the intricate and elusive nature of its molecular and cellular pathophysiology (Pitsillou et al., 2020; de Menezes et al., 2021). The realm of mental health, encompassing conditions such as depression and anxiety, stands as a global concern with profound ramifications for individuals and society as a whole (Kirmayer & Minas, 2023). The burden of these disorders extends far beyond emotional well-being, impacting physical health, interpersonal relationships, productivity, and even life expectancy (Firth et al., 2019: Kirmayer & Minas, 2023; Fazel et al., 2014). Understanding these disorders and their underlying mechanisms is essential, given their association with conditions like cardiovascular disease and early mortality in individuals with schizophrenia (Fazel et al., 2014; Lee et al., 2023). Suicide and disability stemming from depression further underline the urgency of this issue (Ross et al., 2023).

1.2 Exploring the pathogenesis of MDD and the role of metabolomics

MDD, with its complex pathogenesis involving factors like, the hypothalamic–pituitary–adrenal axis, genetics, metabolism, neurotrophic factors, and environmental influences, necessitates comprehensive research (Pitsillou et al., 2019: Belvederi et al., 2018; Juruena et al., 2017; Liu et al., 2015). In this context, metabolomics, a burgeoning -omics technology, holds promise for unravelling depression's enigmatic mechanisms (Brindle et al., 2002; German et al., 2004). However, the application of metabolomics in depression research is still evolving, yielding disparate results due to variations in experimental methods and

subjects (Liu et al., 2015). Various metabolites have been explored for their potential roles in MDD, but inconsistent findings persist (Liu et al., 2015; Zheng et al., 2017; Li Y. et al., 2020; Geng et al., 2020; Gui et al., 2018; Kawamura et al., 2018). Identifying pivotal metabolites and establishing causal relationships with depression remains an imperative area of investigation. Moreover, the potential significance of yet-undiscovered metabolites necessitates further exploration to refine our comprehension of depression's metabolic foundations. To enhance the clinical management of MDD, the integration of peripheral biomarkers with subjective symptom scoring presents a promising avenue. These biomarkers offer the potential for objective, cost-effective, efficient, and non-invasive diagnosis and monitoring (Domenici et al., 2010). While existing research has proposed biomarkers related to monoamine neurotransmission, immune-inflammation, neuroplasticity, and neuroendocrine function, the role of lipids cannot be understated. Lipids, pivotal for protein function by regulating transport and structural support, play a fundamental role in neuronal function. They influence membrane properties, vesicular processes, neurotransmission, and cell integrity and ultimately effect individual’s depression level (Gross et al., 2005; Tsui-Pierchala et al., 2002). Fig. 1.1 illustrates the core energy metabolism in organisms, highlighting the role of glucose, lipids, and proteins as primary energy sources. It showcases how certain amino acid breakdown products contribute to intermediate product synthesis and entry into the TCA cycle. The figure emphasizes mitochondria as the primary energy production site, synthesizing ATP crucial for life activities, aided by NADH from glycolysis and the citric acid cycle.

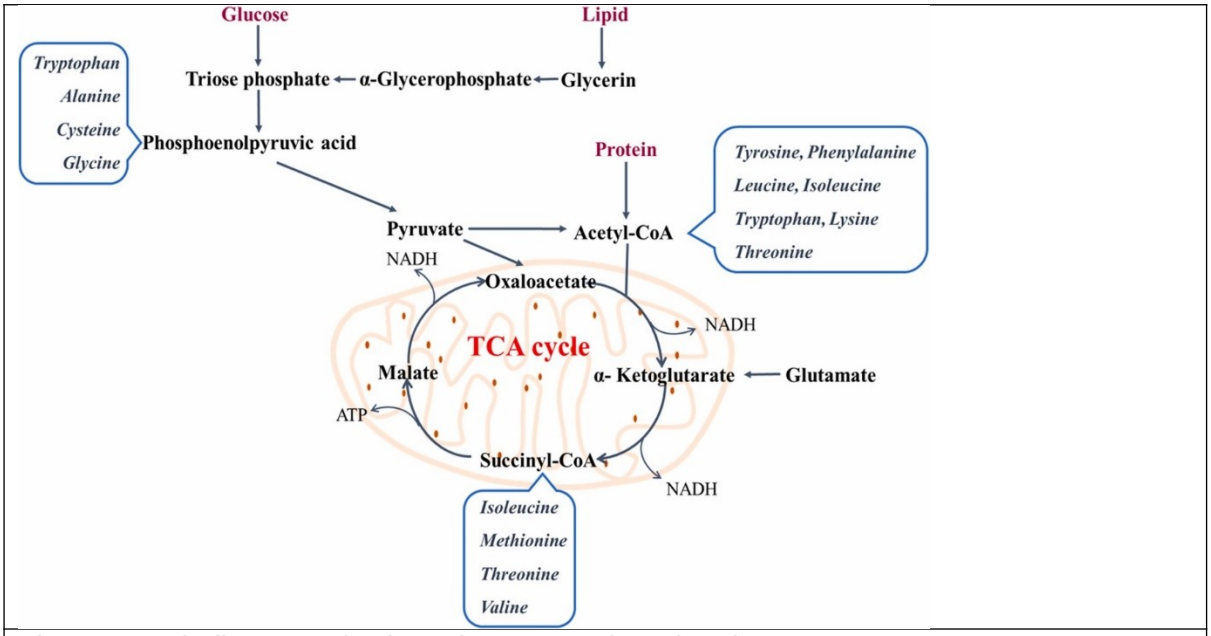


Fig 1.1 Metabolic Dynamics in Major Depressive Disorder
Source: Gu et al. (2021)

1.3 Role of Lipids in Depression

Lipids represent a compelling family of peripheral biomarkers for quantitative diagnosis, treatment response monitoring, and patient stratification. Their association with MDD may also provide insights into potential pharmacological or dietary interventions. Cholesterol and cholesterol-containing molecules have surfaced as potential links to MDD, with decreased total serum cholesterol frequently observed in depressed patients compared to healthy counterparts (Terao et al., 2000; Rabe-Jabłońska & Poprawska, 2000). Altered levels of high-density lipoprotein (HDL), low-density lipoprotein (LDL), and LDL/HDL ratios are also common among MDD patients (Gupta et al., 2013). Moreover, a lack of remission from depressive symptoms is associated with low cholesterol levels (Lehto et al., 2010).

Yet, challenges persist, with some studies challenging these findings, especially in elderly cohorts (Ergün et al., 2004). Variations in lipid levels, such as very low-density lipoprotein (VLDL), total cholesterol, and HDL, have been reported (Huang et al., 2005; Gupta et al., 2013). Particularly, a meta-analysis suggests an inverse association between LDL levels and

depression, especially in medically naïve samples (Persons & Fiedorowicz, 2016). Additionally, various other lipid species, including glycerolipids, sphingolipids, glycerophospholipids, and triglycerides, have been linked to MDD (Müller et al., 2015). Thus, to ascertain the precise nature of the cholesterol-depression link and address discrepancies, further research across diverse cohorts is essential. To clarify the precise link between genome, environment and depression and resolve discrepancies, the information is summarized in Fig 1.2.

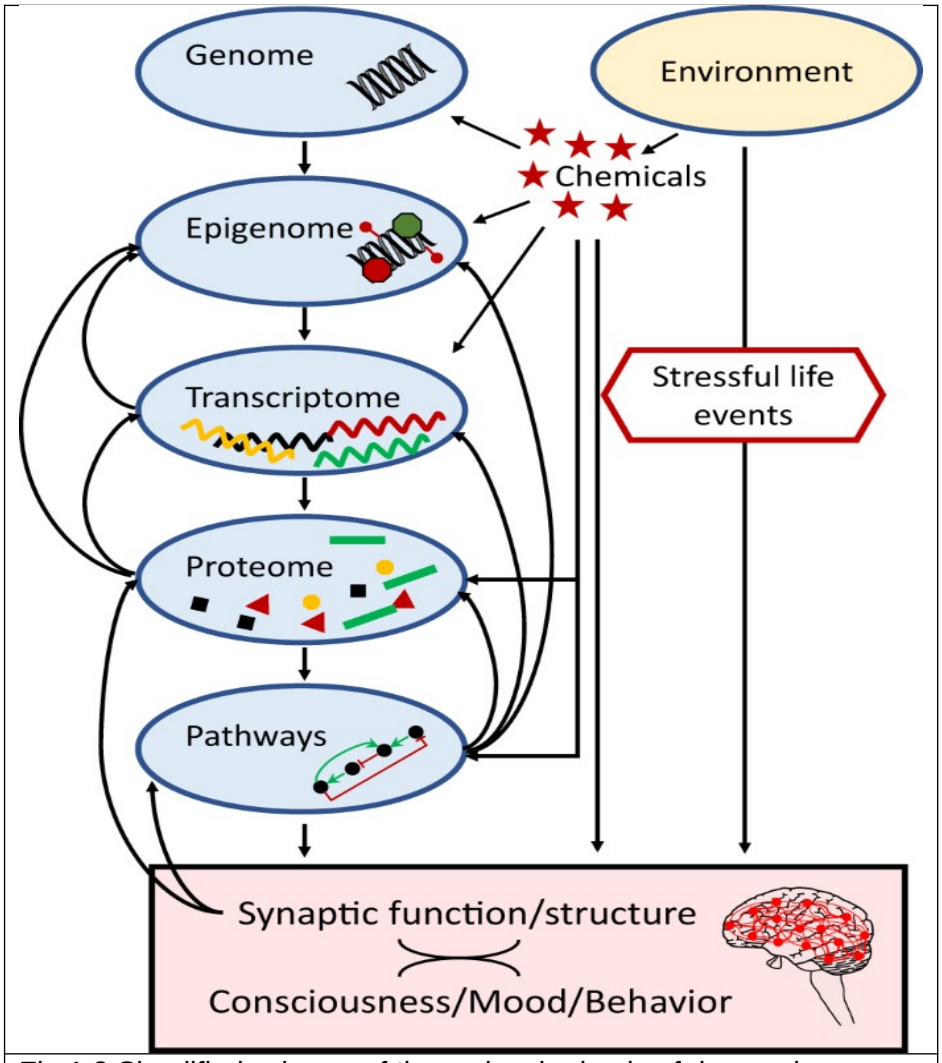


Fig 1.2 Simplified scheme of the molecular basis of depression.
Source: (Fries et al., 2023)

1.4 Lifestyle Dynamics Shaping MDD Severity

1.4.1 Relationship between depression and alcohol consumption

Research has shown a significant positive relationship between depression severity and alcohol consumption, with each preceding the other. Individuals who experience severe depression are more likely to consume alcohol to medicate their symptoms, but alcohol consumption itself can contribute to depressive symptoms, such as the positing of the tension-reduction theory posited by MacAndrew, where alcohol is related to reduction of tension in depressed patients (MacAndrew, 1982), and excessive alcohol consumption also disrupts jobs and relationships, which can lead to depression (Awaworyi Churchill et al., 2017).

1.4.2 Relationship between depression and exercise

People with MDD tend to be less active and fail to get the recommended 150 minutes of moderate to strenuous exercise per week (Marx et al., 2023) due to their mental health issues. Exercise targets neuroplasticity pathways in the brain that are impaired in depression, by improving the vascular delivery of neurotrophic factors and oxygen. Moreover, exercise has been shown to improve levels of self-confidence and self-sufficiency, while reducing suicide ideation, which is characteristic of severe depression. Exercise in a group context further increases social support that buffers depressive symptoms (Kandola et al., 2019).

1.4.3 Relationship between depression and smoking

There exists a plethora of literature showing the strong association between depression severity and smoking. Individuals who suffer from moderate to severe depression tend to smoke at a much higher frequency than those who are healthy. Some studies even suggest that depression can increase the likelihood of smoking initiation, so these individuals can pick up the habit even if they never smoked earlier on in their lives (Fluharty et al., 2016). This relationship is due partly to the self-medicating effects of nicotine, as individuals afflicted with depression turn to smoking to alleviate their symptoms of depression, anxiety, and low mood. This leads to a whole host of further health problems such as lung cancer, bronchitis, or emphysema (Fluharty et al., 2016).

1.4.4 Relationship between depression and BMI

There is a significant relationship between depression and BMI. Individuals who have a high BMI or a high waist circumference have been found to be at a greater risk of developing depression, with a Mendelian randomisation study even suggesting a causal relationship (Speed et al., 2019). This correlation can be attributed to quite a few factors such as chronic inflammation, hormonal imbalances, or poor diet, including over-consumption of processed and unhealthy foods that contribute not just to weight gain, but the depressive symptoms that made the individual turn to excess food and drink to medicate these symptoms in the first place. A key symptom of depression is loss or gain of weight in a relatively short period of time, and this accompanied by a potentially more negative body image only exacerbates depressive symptoms. That said, BMI does not distinguish between fat and non-fat mass, which is important from a physiological perspective since fat mass (adipose tissue) is what contains the inflammatory markers that contribute to negative depressive symptoms (Speed et al., 2019).

1.4.5 Relationship between depression and previous medication history or hormonal medication

It is common for individuals to seek out medication earlier to alleviate the depressive symptoms– feelings of sadness, hopelessness, and a loss of interest in activities that would normally interest them– and to find out which medication or which combination of medications works best for them. Some common depression medications include SSRIs and SNRIs among others, however long-term use of antidepressants is not associated with alleviated depressive symptoms, with studies showing common relapses despite long term (>52 weeks) (Saul H et al., 2022). Although there is plenty of research done on hormone-based intervention for depression, not a lot investigate the potential causal contribution of hormonal medication such as oral contraception on the severity of MDD, but the few that exist suggest a strong positive relationship between use of hormonal contraception and depression severity (Skovlund et al., 2016).

1.4.6 Relationship between depression and comorbidities

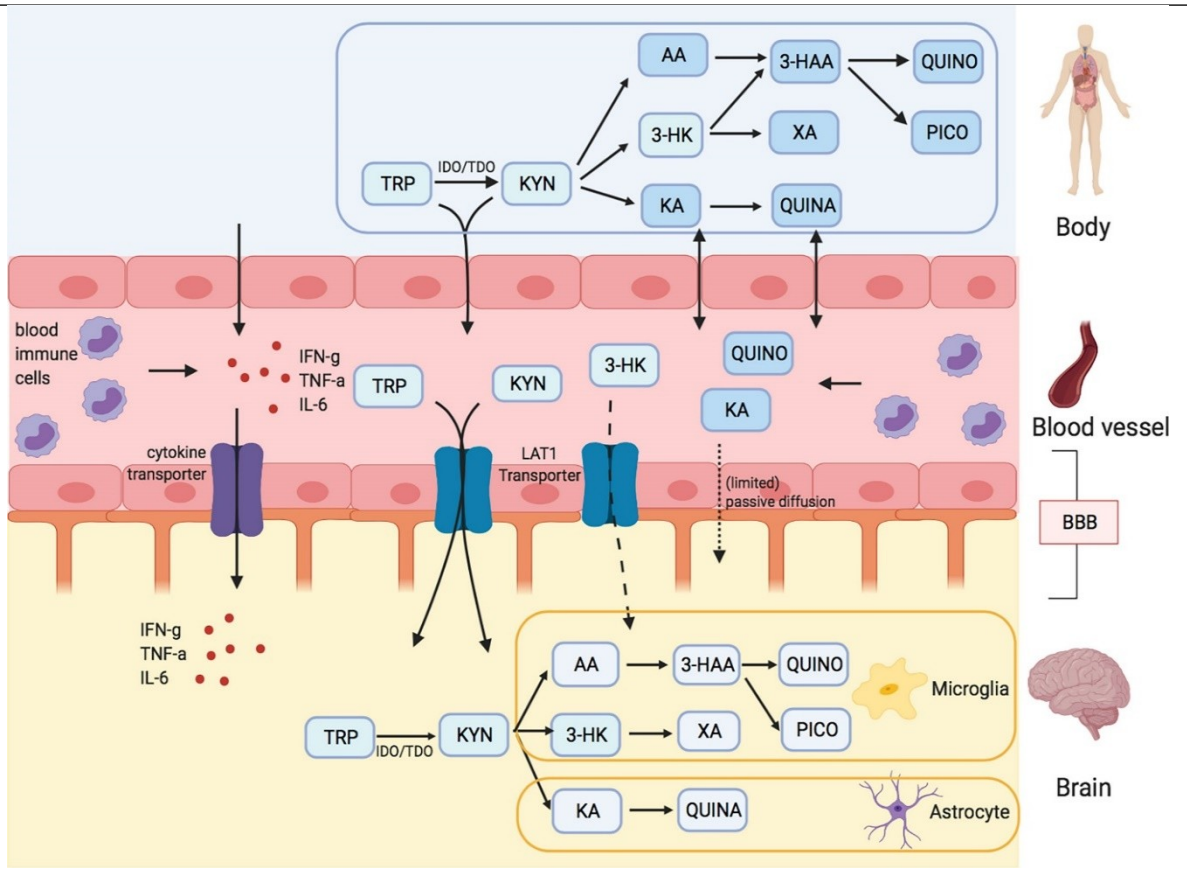
The full list of comorbidities involves the central nervous system such as Alzheimer's and dementia, cancer, cardiovascular disease, metabolic and neuroendocrine disease, autoimmune and gastrointestinal disease, respiratory disorders, and pain-related diseases. Comorbidities such as that with other mental illnesses, high blood pressure, and diabetes are also regarded as closely associated with major depressive disorder. Others include chronic pain, cardiovascular disease, and obesity. Depression can increase the likelihood of developing these debilitating chronic diseases, and individuals with these conditions may be more likely to experience depression (Gold et al., 2002). Additionally, this relationship can create a poorer cycle of outcomes for both depression and the associated comorbidities. One study pointed to a significant relationship between depression and the incidence of

Alzheimer's Disease and dementia, increased incidence and worsening of cardiovascular disease, diabetes, and obesity in younger women, increased incidence of metabolic syndrome, and has been strongly linked to worsening of cognitive symptoms in the existing state of major depressive disorder (Arnaud et al., 2022).

1.4.7 Relationship between depression and inflammatory markers

The association of specific inflammatory markers such as interleukins and neurotrophins with the severity of depression is somewhat well-established, due to their role in inflammatory processes like major depressive disorder. Interleukins and neurotrophins such as brain-derived neurotrophic factor (BDNF) and glial cell-derived neurotrophic factor (GDNF) are of specific interest here due to their known role in mood disorders, with IL-6, BDNF and GDNF all having been identified as significant biomarkers of major depressive disorder (Gadad et al., 2021). However, due to the scope of the variables available in the TwinsUK database, the role of interleukins was of specific interest for this investigation, especially since the relationship between depression severity and all known interleukins are not yet elucidated. IL-4, IL-10, IL-13, IL-19, and IL-33 are all considered anti-inflammatory interleukins, while IL-6, IL-1β and tumour necrosis factor (TNF)-α are established markers of inflammation in mood disorders like major depressive disorder. The metabolism of tryptophan (TRP) plays a crucial role in the synthesis of serotonin (5-HT) and melatonin. The initial and rate-limiting step in this pathway involves the conversion of TRP to kynurenine (KYN), catalysed by the enzymes indoleamine 2,3-dioxygenase (IDO) and tryptophan 2,3-dioxygenase (TDO) (refer to Fig 1.3). TDO, predominantly found in the liver and also present in the brain, is responsible for metabolizing 95% of total body TRP into KYN, with the liver contributing 90% to this process. In normal physiological conditions, TDO in the liver acts as the primary mechanism for utilizing diet-derived TRP, serving as the major source of KYN throughout the body. TDO functions as a housekeeping enzyme, channelling excess TRP

towards the Krebs cycle for energy production. Under stressful conditions, TDO is induced by glucocorticoids to meet increased energy demands, activated in response to psychophysiological stress through cortisol release. Additionally, TDO activity is influenced by various regulatory factors. It is inhibited by a reduction in nicotinamide, activated by heme, and stabilized by TRP. These complex regulatory mechanisms highlight the multifaceted role of TDO in the metabolic pathway, not only in normal physiological functions but also in responding to stress and maintaining balance in cellular energy dynamics.



Tryptophan (TRP) and kynurenine (KYN), and to a lesser degree 3-hydrox kynurenine (3-HK) are actively transported into the brain over LAT1 transporters. Downstream metabolites of the kynurenine pathway (KP), like quinolinic acid (QUINO) and kynurenic acid (KA), cannot make use of these transporters, but (probably limited) passive diffusion of these metabolites over the BBB is possible. Anthranilic acid and 3-hydroxy anthranilic acid (not shown in figure) may equally pass the blood brain barrier through passive diffusion, much like QUINO. In the brain, microglia are responsible for the production of metabolites 3-HK and QUINO, whereas astrocytes produce KA. Peripheral production of these KP metabolites is done by blood immune cells, such as blood monocytes (PBMC) and other organs, including liver and kidney. The gut microbiome, which plays a role in psychiatric illness through the gut-brain axis, also affects KP metabolization.

Fig 1.3. Kynurenine metabolites and the blood brain barrier (BBB)
Source: (Skorobogatov et al., 2921)

1.4.8 Relationship between depression and education or income level

Studies on both country and individual-level socioeconomic status and income level have been conducted- country-level income was found to have no association with depression, whereas on an individual level, lower education, fewer material assets (attributed to lower individual income) have been associated with higher depression levels (Rai et al., 2013).

1.5 Metabolites and their Implications in MDD

Despite improved understanding of the pathophysiology of MDD over the past few decades, this has not translated into the development of a biomarker panel that could improve the diagnosis, management, and prognosis of MDD. Metabolites range from peptides, amino acids, and coenzymes, to nucleotides that arise from cells and tissues, and which can be sampled from body fluids. Currently, depression is still mostly associated with the neurotransmitters serotonin and catecholamines (dopamine and norepinephrine) that are postulated to be altered in MDD patients, especially since antidepressants such as SSRIs target serotonergic receptors, and MAO inhibitors affect dopamine and norepinephrine turnover in the brain. However, there is still little understood on role of metabolites, of which there are a vast range and are intimately involved in the biochemical processes that produce these neurochemicals in the first place. The metabolites typically associated with MDD are those involved in neurotransmission, cell signalling, hormone activators, sleep controllers, and inflammatory mediators (Guerreiro Costa et al., 2022).

1.5.1 Relationship between glutamate and depression

By far the most profound metabolites associated with depression are those related to glutamate metabolism. Of emerging importance are the multipath links between dietary glutamate, endogenous glutamate, the gut, and brain. Glutamate is the primary excitatory neurotransmitter in the brain at which altered levels disrupt healthy neuronal transmission and function. Interestingly, dietary glutamate (in the form of monosodium glutamate) is hypothesised to have the potential to cause dysfunction in the blood-brain barrier alongside indirectly contributing to gut dysbiosis and has been shown to induce anhedonia and despair in rodent experiments (Onaolopo, 2021).

1.5.2 Relationship between kynurenine in depression

Kynurenine has been implicated in the pathophysiology of MDD in many studies, the consensus being that kynurenine is decreased in MDD patients compared to controls, alongside other metabolites involved in the kynurenine pathway such as tryptophan (Ryan et al., 2020). The hypothesis here is that kynurenine metabolism is preferentially switched to produce the neurotoxic quinolinic acid metabolite instead of the neuroprotective kynurenic acid metabolite, aggravating depressive symptoms.

1.5.3 Relationship between amino acids (proteins) and depression

Understanding the pathophysiology of MDD and predicting the therapy response may be possible through the study of amino acid metabolism (Duan & Xie, 2020). Serotonin and norepinephrine, for example, are well-known neurotransmitters with close ties to depression. In any event, amino acids serve as the precursors for the synthesis of neurotransmitters. Patients with MDD have been found to have low amounts of tryptophan and tyrosine, the amino acids that serve as these neurotransmitters' precursors (Nagasawa et al., 2012). Moreover, taking amino acid supplements, which can then be transformed into

neurotransmitters upon consumption, has been shown to lessen or treat depressive and other mental health issues (Lakhan & Vieira, 2008).

1.5.4 The role of metabolites in major depressive disorder

Research has linked metabolites, small molecules involved in body biochemical reactions, to depression incidence, suggesting their role in its development (Qiu et al., 2021). Disturbances in metabolite levels such as glutamate, aspartate, alanine, and others have been observed in depressed patients, but it's unclear if these changes cause or result from mood alterations (Erjavec et al., 2018). Various factors like lifestyle habits, medication use, and metabolic disorders can influence the relationship between metabolite levels and depression severity, necessitating further exploration. Amino acids like tryptophan and tyrosine, precursors of key neurotransmitters, are often depleted in individuals with depression (Nagasawa et al., 2012). Some studies suggest a significant role of certain amino acids in depression development (Kofler et al., 2019). Supplementation with amino acids transforming into neurotransmitters has shown potential in improving mental health (Lakhan & Vieira, 2008). Oxidative stress, mediated by antioxidant function, has been linked to depression, with deficiencies in vitamins A, E, and C suggested as potential contributors (Jimenez-Fernandez et al., 2015; Islam et al., 2020). Recent meta-analyses suggest elevated levels of specific metabolites like kynurenine and acylcarnitine in depression (Guerreiro Costa et al., 2022), although individual studies report a broader range of associated metabolites (Whip et al., 2022).

Large meta-analyses recently conducted by Guerreiro Costa et al. (2022) found that levels of kynurenine and acylcarnitine are elevated in MDD, while other individual studies have found a wider variety of associated metabolites.

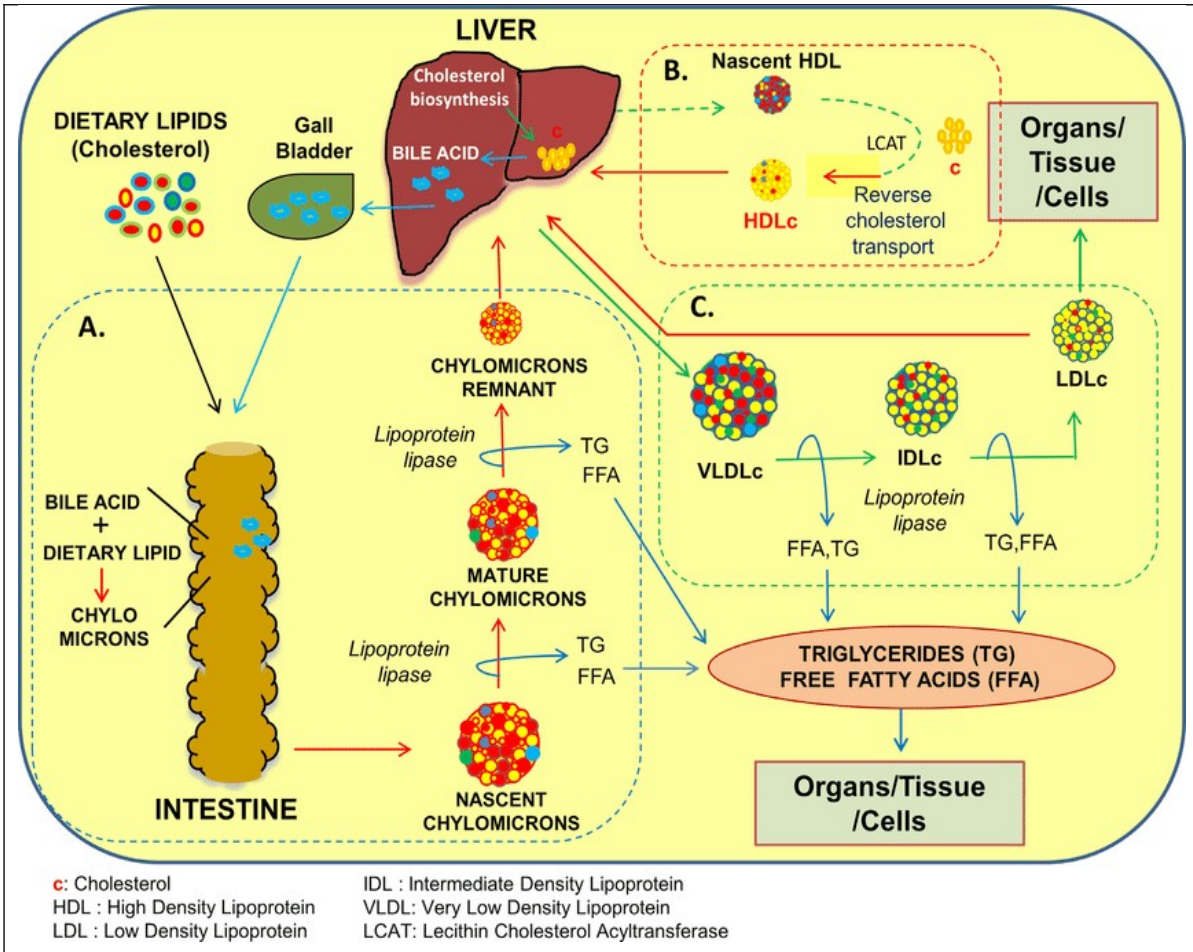
1.6 Lipid Metabolism's Impact on MDD

1.6.1 The importance of lipid research in understanding major depressive disorder

The role of lipids in major depressive disorder has been extensively studied, and higher levels of total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and triglycerides have been typically associated with a higher severity of depression (both objective and self-rated states) and cognitive impairment. Conversely, an increase in consumption of omega-3 fatty acids such as eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) is associated with decreasing levels of depression (Wani et al., 2019), potentially through an increase in grey matter in the dorsal and medial prefrontal cortex, as well as the anterior cingulate cortex, that are involved in the regulation of emotional behaviours.

1.6.2 Evidence for the role of lipid subclasses in major depressive disorder

Higher levels of DHA and EPA are known to be associated with lower levels of depression (Lin et al., (2010). Consuming certain fatty acids, like EPA, is thought to reduce inflammation by limiting the creation of pro-inflammatory molecules called eicosanoids and competing with arachidonic acid to prevent its conversion into these inflammatory compounds. Both DHA and EPA also prevent the release of proinflammatory molecules such as interferon-γ, TNF-α, IL-2, and IL-6, all of which are affected by eicosanoid synthesis from arachidonic acid (Liao et al., 2019), and of which are well known to be inflammatory biomarkers in major depressive disorder. A view of lipid metabolism has been summarized in Fig. 1.4.



Pictorial representation of cholesterol transport in the human body. (A) Transport of dietary lipid in the form of TG, FFA, and cholesterol from the intestine to various organs, tissues, and cells. Dietary lipids or cholesterol are converted into chylomicrons with the help of bile acid. Chylomicrons shed TG & FFA, and the remaining chylomicron remnant containing cholesterol reaches the liver. (B) Reverse cholesterol transport from organs, tissues, or cells and back to the liver. Reverse cholesterol transport is mediated by HDL to remove excess cholesterol from different organs/tissue, small "c " represents cholesterol (C) Transport of TG, FFA, and cholesterol from the liver to other organs, tissues, and cells. The liver releases VLDLc consisting of different lipid components i.e., TG, FFA, and cholesterol. VLDLc shed TG & FFA forming IDLc and LDLc, cholesterol from LDLc is either used up by organs/tissue/cells or being taken up by the liver.

Fig 1.4 Metabolism of dietary lipids
Source: (Mayengbam et al., 2021)

There are conflicting findings related to the role of lipoproteins in major depressive disorder– studies have reported that low levels of LDL and high levels of HDL contribute to higher depression severity, whilst other studies report that depression severity is associated with higher levels of lipids related more to metabolic syndrome, such as CHO and triglycerides (Jia et al., 2020). Another study showed higher, not lower, LDL levels correlated with greater depressive scores (Wagner et al., 2019). It was also unclear whether levels of these

lipoproteins or triglycerides occurred in tandem or contributed to depression as independent biomarkers i.e. a combination of higher levels of HDL and lower levels of LDL might indicate depression severity, or either higher levels of HDL or lower levels of LDL separately indicate higher depression severity.

1.6.3 TwinsUK Database: Lipid Subclasses and Depression Severity

The vast amount of data on various lipid subclasses contained within different sizes and classes of lipoproteins in the TwinsUK database allowed correlations between hitherto unexplored lipids and depression to be investigated. For instance, current and previous research do not state the relationship between depression severity and triglycerides in certain sizes or types of lipoproteins, for example cholesterol present in large low-density lipoprotein versus phospholipids present in medium low-density lipoproteins. There is currently plenty of literature on certain ratios of lipid such as total cholesterol to that or high or low-density lipoproteins (HDL/LDL) (Huang et al., 2003) but not when the two are separated.

Lipids in the TwinsUK database were identified as: cholesterol (9 types), lipoprotein subclasses (168 in total, including ratios of lipids in lipoproteins), glycerides and phospholipids (9 subtypes in total), fatty acids as classified by saturation (16 types), and apolipoproteins (9 total).

1.6.4 The relationship between lipoprotein subclasses, cholesterol, and depression

There is inconclusive and conflicting evidence on various lipoproteins which carry cholesterol, and their associations with depression severity. It would therefore be interesting to investigate the influence of such a wide spectrum of lipoprotein subclasses, instead of

simply looking at the two most common lipoprotein subclasses categorised by density– high-density versus low-density lipoprotein (HDL and LDL), alongside the typical lipid subclasses that are associated with depression, such as DHA or EPA. Lipoproteins can be split into 7 subclasses: chylomicrons, chylomicron remnants, intermediate-density lipoproteins (ILDL, also known as VLDL remnants), very low-density lipoproteins (VLDL), low-density lipoproteins (LDL) and high-density lipoproteins (HDL) (Feingold et al., 2021). Chylomicrons are large particles that are made in the intestine and are very rich in triglycerides (Gouni-Berthold et al., 2011). Chylomicron remnants typically carry more cholesterol as well as triglycerides and are similarly pro-atherogenic. VLDLs are produced in the liver and are more triglyceride than cholesterol-rich. IDLs are produced when triglycerides are removed from VLDLs by both fat and muscle, resulting in very cholesterol-rich particles and are thus pro-atherogenic. LDLs are derived from both VLDLs and IDLs and are the most cholesterol-rich of all the lipoprotein subclasses, carrying most of the cholesterol circulating in the body. Smaller LDL particles are more pro-atherogenic and inflammatory than larger LDL particles since they are more prone to oxidation, and also have lesser affinity for the LDL receptor to trigger endocytosis into cells. HDLs reverse cholesterol transport by moving cholesterol from peripheral tissues back to the liver to be eliminated, thus are anti-atherogenic. Additionally, these HDL particles are anti-inflammatory, anti-oxidant and anti-thrombotic, which enhance its anti-atherogenic nature (Feingold et al., 2021).

Current literature shows little, if any, relationship between chylomicron concentration and depression severity. Chylomicron remnants typically are rapidly removed from circulation by the liver, so they may be of lesser importance than the other lipoprotein subclasses, which, since are present longer in circulation, more accurately reflect their levels. A notable limitation here would be that levels of cholesterol, triglycerides or lipoproteins would differ depending on when these measurements were taken, since these vary depending on whether or not the person has eaten or fasted for a few hours, however ingested fatty acids

and cholesterol are poorly absorbed, and should not be considered a large confounding variable in this investigation (Parekh et al., 2017).

Some detailed and high-powered studies have found a significant correlation between levels of high-density lipoprotein and depression levels (Jia et al., 2020), however these did not reveal specific environmental factors which the TwinsUK database provides, that may interfere with the findings in multivariate analyses, and render them less reliable as lipid profile biomarkers for depression. One especially unique feature of the TwinsUK database that would help elucidate some particularly unique findings lay in the availability of not just lipoprotein levels, but the levels of triglycerides present in lipoproteins as categorised by their size and class– whether the lipoprotein was of high or low-density, and its size– small, medium or large. These permutations of triglycerides in lipoproteins were thus vast and interesting to compare–for example, triglycerides in small low-density lipoproteins versus triglycerides in large low-density lipoproteins versus triglycerides in medium low-density lipoproteins.

There is evidence to suggest that there may be a relationship between triglyceride levels carried in different-sized lipoproteins and depression severity. A meta-analysis published in the Journal of Clinical Psychiatry in 2017 found that higher levels of total cholesterol and low-density lipoprotein (LDL) cholesterol were associated with an increased risk of depression, while higher levels of high-density lipoprotein (HDL) cholesterol were associated with a reduced risk of depression. Triglycerides are another type of lipid that is carried in various lipoproteins in the blood, including VLDL and LDL particles. Some studies have suggested that higher levels of triglycerides carried in smaller, denser LDL particles may be associated with an increased risk of depression. In a study published in the Journal of Affective Disorders in 2019, researchers found that individuals with major depressive disorder had higher levels of triglycerides carried in small, dense LDL particles compared to healthy controls. They also found that there was a positive correlation between the severity

of depression and the amount of triglycerides carried in these small and denser LDL particles. Other studies have found an association between lipid ratios in men and not women, in a most recent meta-analysis involving 11,000 men and women, looking at the association between depression and high cholesterol (Han, 2022).

1.6.5 The relationship between glycerides, phospholipids and depression

Phospholipids contain a phosphorus group in place of an extra fatty acid group (like in triglycerides) and form an essential part of cell membranes to enhance their structure and function. There is not a lot of evidence to show any significant effect of phospholipids on depressive symptoms, however a 4-week trial executed by Wang et al (2021) that administered EPA-enriched phospholipids to mice was superior to administration of EPA-enriched ethyl esters in reducing neuroinflammation that contributes to depressive symptoms, via the hypothalamic-pituitary axis (Wang et al., 2021).

1.6.6 The relationship between apolipoproteins and depression

Apolipoproteins are not as well investigated as a lipid subclass as compared to lipoproteins or cholesterol, although they are needed for the transport of lipids by attaching to them. Apolipoprotein B is the main protein in cholesterol contained in low-density lipoprotein (LDL-C) while Apolipoprotein A is the main protein in cholesterol contained in high-density lipoprotein (HDL-C). Knowing this, it would be expected for these levels to correlate with levels of LDL-C and HDL-C respectively in the cohorts examined. Research done by Sevincok et al has shown a reduction in levels of apolipoprotein A in depressed patients compared to healthy controls (Sevincok et al., 2006), while apoliprotein B levels are positively predicted by depression levels (Sadeghi et al., 2011).

1.7 Research Gaps

Based on the above facts and figures, the following gaps in research can be inferred:

The earlier studies emphasise the complexity of the underlying mechanisms driving depression, which include elements such as heredity, metabolism, and neurotrophic factors. For more thorough research to better understand these mechanisms and how they affect the onset and progression of depression, it is necessary to concentrate on these individual factors.

While metabolomics is recognized as a promising approach to unravelling depression's mechanisms, the literature suggests that its application in depression research is still evolving and has yielded inconsistent findings. This highlights a research gap in identifying reliable metabolites associated with depression.

The previous research emphasizes the importance of lipids in neuronal function and their potential as peripheral biomarkers for depression. However, the precise role of different lipid species and their associations with depression remain areas that require further investigation.

The link between cholesterol and depression is discussed, but there are discrepancies in the findings, especially in elderly cohorts. This highlights the need for more research to clarify the relationship between cholesterol levels and depression across different age groups.

The preceding investigations mention the importance of conducting research across diverse cohorts to address discrepancies in findings. This suggests a research gap in terms of the need for more extensive and diverse studies to validate the associations between metabolites, lipids, and depression.

Prior studies further mention the use of machine learning algorithms to identify key metabolites and lipids associated with depression. Further research is needed to explore the full potential of machine learning like OPLS-DA in analysing complex metabolomics data for depression diagnosis and prognosis.

While the research focuses on metabolites and lipids as potential biomarkers, there is a research gap in terms of the need for a holistic approach that considers multiple factors, including lifestyle variables in understanding depression.

Addressing these research gaps could contribute to a more comprehensive understanding of depression's pathophysiology and improve the diagnosis and management of Major Depressive Disorder (MDD).

1.8 Research Questions

Based on the research gaps discussed above, the following research questions were identified:

What is the potential role of metabolites and lipids as peripheral biomarkers in the diagnosis and management of MDD, and how can this understanding contribute to more effective diagnostic and management strategies for MDD. The central research question that guides this investigation is: what are the most critical biomarkers for diagnosing and prognosing depression? This question arises from the pressing need to gain a deeper understanding of the role of lipids and metabolites in the underlying metabolic mechanisms of depression, and identifying reliable biomarkers that can significantly enhance our ability to diagnose and predict its onset and progression.

1.9 Research Objectives

To address these research questions, the current study outlines several key objectives:
The study will involve the measurement and comparison of lipids and metabolites in two groups: patients with depression (medicated and unmedicated) and a control group comprising individuals without depression. This comparative approach will help identify differences and patterns specific to the depressive condition.

Following the measurement and comparison phase, the study will aim to identify specific lipids or metabolites that demonstrate a strong association with depression. These identified molecules will be considered potential diagnostic and prognostic biomarkers.

1.10 OPLS-DA: A Novel Tool

To achieve these objectives, the study will employ robust statistical analyses, including multivariate analysis and machine learning algorithms such as Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA). These advanced analytical methods are well-suited to handle the complexity of metabolomics data and can help identify key metabolites and lipids that are associated with depression. By leveraging these techniques, the study aims to unravel the intricate relationships between specific molecules and the presence or severity of depression.

Chapter 2: Systematic review and meta-analysis of peripheral metabolites in major depressive disorder and psychosis as potential disorder-specific biomarkers

2.1 Introduction

2.1.1 The growing importance of metabolomics in psychiatric research

Psychiatric distress serves as an umbrella term encompassing various psychological conditions, ranging from subclinical symptoms to clinical diagnoses like depression, anxiety, stress, or posttraumatic stress disorder (Zhu et al., 2022). High levels of distress are often indicative of impaired mental health or common mental disorders such as depression and anxiety (Zhu et al., 2022). These conditions are known to have connections with metabolomics, reflecting physiological processes (Prince et al., 2023). To advance our understanding of common mental disorders, including Major Depressive Disorder (MDD), and to develop advanced clinical methods, it is essential to explore the role of metabolomics as potential biomarkers (Zhu et al., 2023). Prior research has demonstrated the significance of metabolomics in shedding light on mental health conditions such as schizophrenia, anxiety, and depression (Merritt et al., 2023; Humer et al., 2020; Prince et al., 2023; Caspani et al., 2021). However, these studies have not yet extensively explored symptom dimensions, which are vital for understanding the heterogeneity across psychiatric disorders (Pedrini et al., 2019). Despite the rapid and high-tech progress in the field, there remains limited research on specific metabolite profiles in neuropsychiatric diseases that can effectively discriminate between these conditions. Addressing this gap, this study focuses on a systematic review and meta-analysis of unique metabolites associated with depression and psychosis (schizophrenia) in adults. The ultimate goal is to identify metabolites that could potentially serve as biomarkers. Advanced medical technology has opened doors for

metabolomics in psychiatric research, offering a comprehensive view of metabolic status. This approach could lead to early intervention by detecting early disease stages (Kirwan, 2023; Jia et al., 2023). Major depressive disorder remains a debilitating condition, and its pathophysiology is not yet fully understood (Svensson et al., 2021). The lack of biological markers and clear treatment endpoints complicates both diagnosis and treatment evaluation. Metabolomics offers a promising approach to capture molecular functions that are perturbed both in distinct and overlapping dimensions of mental health. This approach aligns with the shift towards personalized medicine in the realm of mental health (Jia et al., 2023).

2.1.2 Existing literature and comorbidity between depression and psychosis

The paucity of literature on metabolites in depression is motivation for its advancement. Since the disease metabolome is in constant flux during various stages of depression, metabolomics can help uncover a biomarker specific to disease stage. Decreased levels of glutamate and its metabolites have been found in depression (Moriguchi et al., 2018). Digging deeper, literature also points out comorbidity between depression and psychosis, with suicidal tendencies, as well as glutamate dysfunction, also prevalent in psychosis (schizophrenia), most notably in first-episode psychosis (FEP) (Coentre et al., 2017), despite depression and psychosis (specifically psychotic mania) being on opposite sides of the mood spectrum, Thus, it would be interesting to go further and identify the metabolite profiles of depression within psychotic disorders too.

Overall, there are fewer systematic reviews on depression metabolites compared to schizophrenia, of which psychosis is characteristic, highlighting a need to further clarify peripheral markers that are present in depression. There is mixed literature on biomarkers for depressive subtypes– anxious, atypical, mixed, melancholic and psychotic depression, highlighting the need to distinguish between biomarkers for each, for a better diagnosis. It is

therefore crucial to develop a solid panel of biomarkers that distinguish it from other similar disorders and prevent inaccurate diagnoses, and also to see if metabolite levels correlate with severity of depression or schizophrenia, and if specific glutamatergic metabolites are implicated in either depression or schizophrenia. For example, brain-derived neurotrophic factor (BDNF) is reduced in both schizophrenia and depression (Angelucci et al., 2005), therefore the metabolite cannot be used to distinguish between the two.

Davison et al. (2018) examined 63 studies, including people with schizophrenia, schizophrenia spectrum disorders, or those at risk of acquiring the disease using the PANSS scale, in one of the most thorough systematic reviews on metabolites in schizophrenia. It was found that there is no single panel of biomarkers that define schizophrenia due to the lack of a universally accepted method for metabolomic analysis of several biochemical signatures, including elevated glutamate levels. In FEP specifically, glutamate was upregulated in psychosis, contributing to the negative psychotic symptoms (Davison et al 2018). In fact, glutamatergic metabolites are heavily implicated in all stages of schizophrenia (Merritt and Egerton, 2017). Clinical predictors of disorders are not fail-safe due to overlapping symptomology, such as that between psychotic disorders such as schizophrenia, and bipolar or autism (Davison et al, 2018). However, due to comorbid symptomology between depression and schizophrenia, there may be further overlap. Identifying the metabolites may also help untangle mechanisms underlying the pathophysiology of schizophrenia and depression, which must move from drawing associations to properly understanding the mechanisms of disease to develop more objective diagnoses and consequently novel drugs for treatment.

2.1.3 Spotlight on glutamate

The most ubiquitous neurotransmitter metabolite, glutamate, is required for neuroplasticity and long-term potentiation, and its dysfunction is well-known in various psychiatric diseases (Tsakapis and Travis, 2002). Focussing on metabolites here thus complements our aim of understanding how metabolites as biomarkers may help us develop our understanding of psychiatric disease, specifically in depression and psychosis, since glutamate and other metabolites are products of metabolism and contribute to other cycles in energy metabolism and is implicated in both disorders. Importantly, it would be interesting to steer focus from the established monoamine system, where monoamines such as dopamine, adrenaline and serotonin are decreased in many psychiatric diseases, to the glutamate system, which is a more nuanced way of understanding pathways specific to different psychiatric disorders, since its levels are not predictably high or low across the spectrum of psychiatric disorders.

Further, glutamate acts on many receptors throughout the brain, such as NMDA, AMPA and Glu receptors. The former two are heavily involved in long-term potentiation, with NMDAR signalling capable of promoting cell survival or neurotrophic functions, impaired in depression and psychosis. Even moderate signalling can activate cyclic AMP response-element binding protein (CREB)-mediated induction of survival genes, which promotes brain-derived neurotrophic factor (BDNF) expression in the brain, important since BDNF is significantly decreased in the brains of both depressed and psychotic individuals (Wang et al., 2018). Since glutamate acts on almost all major receptors in the brain, understanding its role, and possibly that of other metabolites it interacts with, and the metabolic pathways they participate in, will complement a whole-body concept of psychiatry, where the role of the brain is perceived to influence other major organs in the body, in this case being a specific phenotype that is characteristic of either depression or psychosis.

2.1.4 Aim of study

The aim of this review and meta-analysis is to, via a rigorous review and meta-analysis, investigate novel metabolite biomarkers found using metabolomics, to see if there is a biomarker found across most studies, which can be tested and validated for its robustness in further studies involving animals and humans.

2.2 Materials and methods

2.2.1 Protocol

This systematic review was conducted in accordance with PRISMA guidelines (Fig. 2.1) The protocol was registered with Prospero.

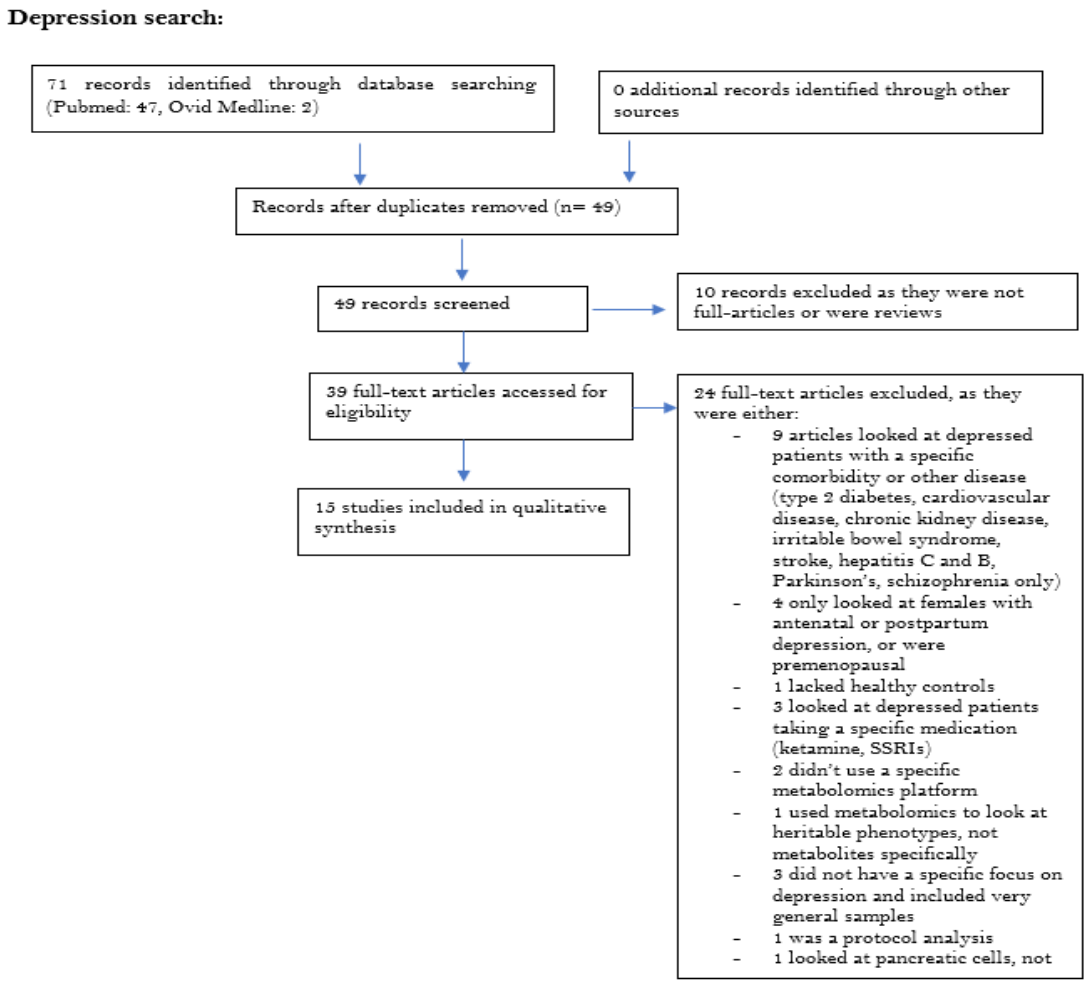


Fig. 2.1 The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart outlining the selection of studies included in the systematic review

with meta-analysis.

2.2.2 Selection Criteria, search strategy, and data extraction

The search was restricted to human subjects aged between 18-65 years. Subjects included both healthy and clinical samples. The three databases Pubmed, Ovid Medline and Google Scholar were searched for studies published from January 1st, 2004 to December 31st, 2022. Search terms for schizophrenia include metabolomics, metabolites, metabolome, schizophrenia, psychosis, first-episode psychosis, and those for depression were similar but psychosis-related terms were replaced by depression, major depressive disorder, biomarker and mood disorder. The inclusion and exclusion criteria were as follows: only independent, full-text studies were included, and these papers measured biofluid metabolite levels using specifically stated analytical platforms for metabolomic analysis. For the schizophrenia search, measurements were done in humans with schizophrenia, schizophrenia spectrum disorders or those at risk of developing schizophrenia. For depression, papers which looked at those already diagnosed with major-depressive disorder (MDD) and those who were at risk of developing depression were looked at. Neither the schizophrenia nor depression searches included samples that were already on medication to minimise bias. All papers would have compared the diseased or at-risk individuals with healthy controls, with the appropriate scales used for prior screening. For schizophrenia, the PANSS scale to measure positive and negative symptoms was used while for depression, the Beck's Depression Inventory and Hamilton Depression Rating Scale were the common scales used to assess extent of disease severity in recruited participants.

An a priori decision was made to only analyse glutamate/glutamine, the metabolites in the TCA cycle and associated amino acids. 4 or more studies that demonstrated changes of a metabolite in the same direction (either increased or decreased) were included in the

quantitative analysis. Means and standard deviations (SDs) were calculated using Microsoft Excel.

2.2.3 Statistical analysis

Data was extracted as standardised mean differences (SMDs) with 95% CIs and I^2 values used to measure heterogeneity between the studies in the meta-analysis. Pooled SMD was calculated using a random effects model in R.

2.3 Results

2.3.1 Study selection

The systematic review included a total of 282 studies, and the final output for the review was 40 studies, following screening by author AL. The final number of studies included for the meta-analysis was 17; 7 for the metabolite alanine, 5 for citrate and 5 for glutamate. It was found that there were significantly higher levels of citrate and alanine in depressed patients, and higher levels of glutamate in schizophrenic (psychotic) patients. Overall, there were more metabolites associated with depression than schizophrenia. Tables 1, 2 and 3 (See Appendix-1, 2 & 3) show the abnormal levels of molecules in depression and psychosis, with Table 3 showing the stratification of these metabolites.

Based on the criteria described, two metabolites including alanine and citrate were reported to be consistently increased in patients with MDD (Table 1, Appendix-1), with 6 out of 7 studies involving alanine showing a consistent increase in patients with MDD, derived mostly from urine. All 5 citrate studies included display a consistent increase in levels of citrate in patients with MDD, results derived more evenly from the biofluids urine and plasma.

Glutamate and glutamine levels were altered but only 2 studies were identified in total. The levels of α-ketoglutarate were increased in some studies but decreased in others, therefore it was not included in the quantitative analysis because results were inconsistent. Glycine and lactate were mostly reported to be elevated in depressed patients but did not meet the stringent criteria. Therefore, only alanine and citrate were further quantitatively analysed. In the case of psychosis, glutamate was reported to be consistently increased in patients with psychosis, although overall findings were not as significant as that for depression.

Here's a summary of whether each metabolite increased or decreased in the mentioned studies (Table 1, Appendix-1):

Increased Metabolites:

1. α-ketoglutarate: Increased in two studies (Pan et al., 2018, and Zheng et al., 2013).
2. Alanine: Increased in six studies (Tian et al., 2014; Tian et al., 2016; Liu et al., 2015; Chen et al., 2018; Zheng et al., 2016; Ding et al., 2014).
3. Ascorbic acid: Increased (Hashimoto et al., 2017).
4. Azelaic acid: Increased in two studies (Chen et al., 2018 and Zheng et al., 2016 for women).
5. Citrate: Increased in five studies (Tian et al., 2014; Tian et al., 2016; Liu et al., 2015; Setoyama et al., 2016; Zheng et al., 2013).
6. Dimethylamine: Increased (Tian et al., 2014).
7. Ethanolamine (EA): Increased (Ogawa et al., 2015).
8. GABA: Increased in two studies (Pan et al., 2018; Setoyama et al., 2016).
9. Glucose: Increased (Liu et al., 2015).
10. Glutamate: Increased (Martins-de-Souza, 2014).
11. Glycine: Increased in three studies (Liu et al., 2015; Chen et al., 2018; Ding et al., 2014).

12. Hippurate: Increased in one study (Tian et al., 2016).
13. Indoxyl sulphate: Increased (Zheng et al., 2016).
14. Lactate: Increased in three studies (Tian et al., 2014; Liu et al., 2015; Ding et al., 2014).
15. Lipids: Increased (Liu et al., 2015).
16. Palmitic, oleic, capric, dodecanoic acids: Increased (Zhou et al., 2018).
17. Phenylalanine: Increased in one study (Liu et al., 2015).
18. Sphingomyelin: Increased (Czysz et al., 2019).
19. Xanthurenic acid: Increased (Tian et al., 2014).

Decreased Metabolites:

1. α-ketoglutarate: Decreased in one study (Zheng et al., 2016 for men).
2. 3-hydroxybutyrate: Decreased in one study (Setoyama et al., 2016).
3. 3-hydroxyphenylacetate: Decreased in one study (Zheng et al., 2016 for women).
4. 4-HPA/4-HB: Decreased (Bhattacharyya et al., 2019).
5. Creatine: Decreased in one study (Tian et al., 2014).
6. Creatinine: Decreased in one study (Setoyama et al., 2016).
7. Glutamine: Decreased (Liu et al., 2015).
8. Taurine: Decreased in one study (Tian et al., 2014).
9. TMAO: Decreased in one study (Zheng et al., 2016 for women).
10. TMAO: Decreased in one study (Zheng et al., 2012).
11. Tryptophan: Decreased in two studies (Pan et al., 2018; Ryan et al., 2020).
12. Tyrosine: Decreased in two studies (Tian et al., 2016; Zheng et al., 2016 for men).
13. Valine: Decreased in one study (Liu et al., 2020).
14. Valine: Decreased in one study (Zheng et al., 2016 for men).
15. Valine: Increased in one study (Liu et al., 2015).
16. EPA, DHA: Decreased (Borsini et al., 2020).

Below is a concise overview of whether there was an increase or decrease in each metabolite as observed in the studies that compared schizophrenia or psychosis patients to control groups (Table 2, Appendix-2):

Increased Metabolites:

α-ketoglutarate: Increased in one study (Cai et al., 2012).

α-ketoisovaleric acid (valine metabolite): Increased (Tasic et al., 2017).

3-hydroxybutyrate: Increased in one study (Yang et al., 2011).

3-hydroxykynurenine (3-OHKY): Increased (Condray et al., 2011).

Aspartate: Increased (Yoshikawa et al., 2018).

Citrate: Increased in two studies (Cai et al., 2012; Yoshikawa et al., 2018).

Creatine: Increased in two studies (Koike et al., 2014; Fournier et al., 2014).

GABA: Increased (Tasic et al., 2017).

Glucose: Increased in one study (Huang et al., 2007).

Glutamate: Increased in six studies (Buckley, 2012; Yoshikawa et al., 2018; Fukushima et al., 2014; Oresic et al., 2011; Garip et al., 2019; Nagai et al., 2017).

Glycerate: Increased (Liu et al., 2015).

Glycerate, eicosenoic acid: Increased (Yang et al., 2011).

Glycine: Increased (Tasic et al., 2017).

Guanine: Increased (Tasic et al., 2017).

Homovanillic acid (HVA): Increased in two studies (Palsson et al., 2017; Ramirez-Bermudez et al., 2008).

IL-6, IL-10: Increased (Pedrini et al., 2012).

Lactate: Increased in one study (Fukushima et al., 2014).

Linoleic acid, eicosenoic acid, pentadecanoic acid: Increased (Suvitaival et al., 2016).

Mannitol: Increased (Tasic et al., 2017).

Nervonic acid (similar to oleic acid): Increased (Kageyama et al., 2017).

Palmitic, linoleic, stearic oleic acids: Increased (Yang et al., 2017).

p-aminobenzoic acid (PABA): Increased (Tasic et al., 2017).

Pantothenate: Increased (Tasic et al., 2017).

Pyroglutamic acid: Increased (Yoshikawa et al., 2018).

Triglycerides: Increased in two studies (Suvitaival et al., 2016; Buckley, 2012).

Decreased Metabolites:

α-ketoglutarate: Decreased in one study (Xuan et al., 2011).

α-tocopherol: Decreased (Liu et al., 2014).

γ-tocopherol: Decreased (Xuan et al., 2011).

2-hydroxybutyrate: Decreased (Kageyama et al., 2016).

3-hydroxybutyrate: Decreased in two studies (Fukushima et al., 2014; Yang et al., 2011).

Acetate: Decreased (Dietrich Mouszalska et al., 2012).

Arachidonic acid: Decreased in two studies (Fukushima et al., 2014; Yang et al., 2017).

Arginine: Decreased (Fournier et al., 2014).

Benzoic, nonanoic, perillic, betaine acids: Decreased (Koike et al., 2014).

Citrate: Decreased in two studies (Xuan et al., 2011; Liu et al., 2015).

Creatine: Decreased in one study (Kageyama et al., 2016).

D-serine (derivative of glycine): Decreased (Fukushima et al., 2014).

Glucose: Decreased in three studies (Xuan et al., 2011; Holmes et al., 2006; Liu et al., 2015).

Glutamine: Decreased (He at al., 2012).

Glutathione: Decreased (Mico et al., 2011).

Glutathione and its precursor γ-L-Glutamyl-cysteine: Decreased (Fukushima et al., 2014).

Histidine: Decreased (He at al., 2012).

Lactate: Decreased in three studies (Dietrich Mouszalska et al., 2012; Holmes et al., 2006; Huang et al., 2007).

Linoleic acid: Decreased (Fukushima et al., 2014).

Phosphatidylcholine: Decreased in two studies (Suvitaival et al., 2016; He at al., 2012).

Taurine: Decreased (Cai et al., 2012).

Threonine, tyrosine (derivative of phenylalanine): Decreased (Fukushima et al., 2014).

Tryptophan: Decreased in two studies (Xuan et al., 2011; Fournier et al., 2014).

In the context of schizophrenia, metabolite class analysis reveals notable alterations in various metabolic pathways (Table 3, Appendix-3). Fatty acids and lipoproteins like triglycerides, phosphatidylcholine, and cholesterol esters did not show consistent trends, with some studies reporting upregulation and others downregulation. Carbohydrate metabolism appears to exhibit significant changes, with glucose being consistently upregulated in multiple studies. Amino acid metabolism in schizophrenia is characterized by upregulation of glutamate and downregulation of N-acetyl aspartate (NAA), proline, phenylalanine, and other amino acids in different studies. Antioxidants like glutathione and taurine are commonly downregulated. Tryptophan metabolism displays varying trends, with tryptophan, serotonin, and kynurenine showing mixed regulation patterns. Dopamine metabolism, as indicated by homovanillic acid (HVA), is upregulated in schizophrenia, while other pathways like creatine, creatinine, pregnanediol, and 3-hydroxybutyrate exhibit diverse regulation patterns. In the tricarboxylic acid cycle, citrate and α-ketoglutarate are both upregulated and downregulated in different studies.

In the context of depression, metabolite class analysis highlights distinct metabolic alterations. Triglycerides are upregulated, and linoleic acid is consistently downregulated. Glucose is consistently upregulated in depression, whereas lactate displays mixed regulation patterns. Amino acid metabolism is characterized by downregulation of glutamine and upregulation of glutamate in some studies. Glycine, alanine, and dimethylamine are upregulated in depression, while other amino acids like phenylalanine, tyrosine, and valine show varying trends across different studies. Antioxidants like taurine and ascorbic acid

exhibit downregulation and upregulation, respectively. Tryptophan metabolism is marked by the downregulation of serotonin and kynurenine. Vitamin E metabolism alterations are not evident in depression. Dopamine metabolism, represented by HVA, is consistently downregulated. Other pathways like creatine, creatinine, hippurate, and 3-hydroxybutyrate also exhibit mixed regulation patterns in depression. In the tricarboxylic acid cycle, citrate and α-ketoglutarate show varying regulation patterns, with both upregulation and downregulation observed in different studies.

2.3.2 R analysis and forest plots

R was used to generate forest plots to visualise the overall effect difference, indicated by standardised mean difference (SMD). Although these three metabolites were implicated in both depression and psychosis in the systematic review, the relative comparison done in the review revealed higher levels of alanine, citrate for depression, and higher glutamate for schizophrenia. For alanine, 245 MDD patients and 260 healthy comparators were analysed (Table 1, Appendix-1). 6 studies reported elevated levels of alanine in patients with MDD, with a total of 245 patients, compared to 260 healthy controls (Table 2.1 & Fig.2.2). Across these studies, significantly higher levels of alanine were reported in depressed patients (SMD=0.79; 95% CI; 0.15-1.43; p< 0.01), with evidence of high heterogeneity (I2=91%). 5 studies reported significantly elevated levels of citrate in 269 depressed (MDD) compared to 307 healthy controls (SMD=1.58; 95% CI; 0.38-2.78; p<0.01, I2=97% Table 2.2, Fig. 2.3).

For the citrate level comparison, 5 studies were included that reported significantly elevated levels of glutamate in the plasma and urine of people with schizophrenia in 604 depressed patients compared to 590 healthy controls (SMD=2.70; 95% CI; 0.81-4.59; p<0.01), also with high heterogeneity (I2= 99%). Metabolite levels were found to be increased at any of the clinical subgroups, not just one. Since none of the studies reported levels of metabolites in

clinically stable depressed or psychotic patients, illness phase was eventually not used to describe data.

A significantly higher level of alanine was identified between patients with schizophrenia and comparators plot and tables given at the end of Appendix-3. The SMD mean range was 0.03-0.80. The forest plot shown in Figure 5 illustrates the SMD. The value of heterogeneity (I2=73%) was notably high, yet lower in comparison to the results for citrate in MDD and glutamate in psychosis. The weight was evenly distributed between the 6 studies. The results for glutamate observed a similar pattern. A total of 335 patients with schizophrenia patients and 280 comparators were analysed (Figure 6 after Appendix-3), (SMD = 1.67; 98% CI: -0.09-3.43; $p<0.01$). suggesting a larger effect size. The MDD mean covered a wide range of 0.044-49.59, and the forest plot reported a high sample heterogeneity of 97%. The weight was also nearly equally distributed among the 5 studies, with each study accounting for approximately 20%. The effect size was the largest out of the three compared metabolites, with a similarly high heterogeneity, as may be observed by the SMD and I2 values respectively.

2.3.3 Heterogeneity

Overall, the I^2 values, showing how heterogenous the studies are, were relatively high for the studies looking at all three metabolites, showing that the inconsistency observed may be due to something other than chance. A sub-analysis was done for alanine and glutamate, and the subsequent exclusion of the studies which contributed most to heterogeneity did reduce the heterogeneity without changing the significance of the findings. The tables showing this are in the appendix (Appendix1-3).

Table 2.1: Summary of studies and statistics showing significantly higher levels of plasma and urine alanine in depressed (MDD) compared to control subjects

Author and date	MDD (n)	Comparator (n)	MDD mean	Comparator (mean)	MDD SD	Comparator SD	Mean SD	Effect Size	CI (95%) Low	CI (95%) High
Mitani et al 2006	23	31	404.4	333.6	203.2	120.8	162.00	0.44	-0.11	0.99
Zheng et al 2013	82	82	0.77	0.9	0.53	0.57	0.55	-0.24	-0.54	0.07
Ding et al 2014	23	25	0.502	0.435	0.137	0.074	0.11	0.64	0.06	1.21
Tian et al 2014	21	21	0.00495	0.0037	0.00145	0.00085	0.00115	1.09	0.47	1.71
Tian et al 2016	25	33	97.2	53.18	16.38	11.94	14.16	3.11	2.58	3.00
Zheng et al 2016	23	20	0.011	-0.212	0.468	0.383	0.43	0.52	-0.09	1.14
Chen et al 2018	48	48	14.597	11.793	7.77	5.42	6.60	0.43	0.02	0.83

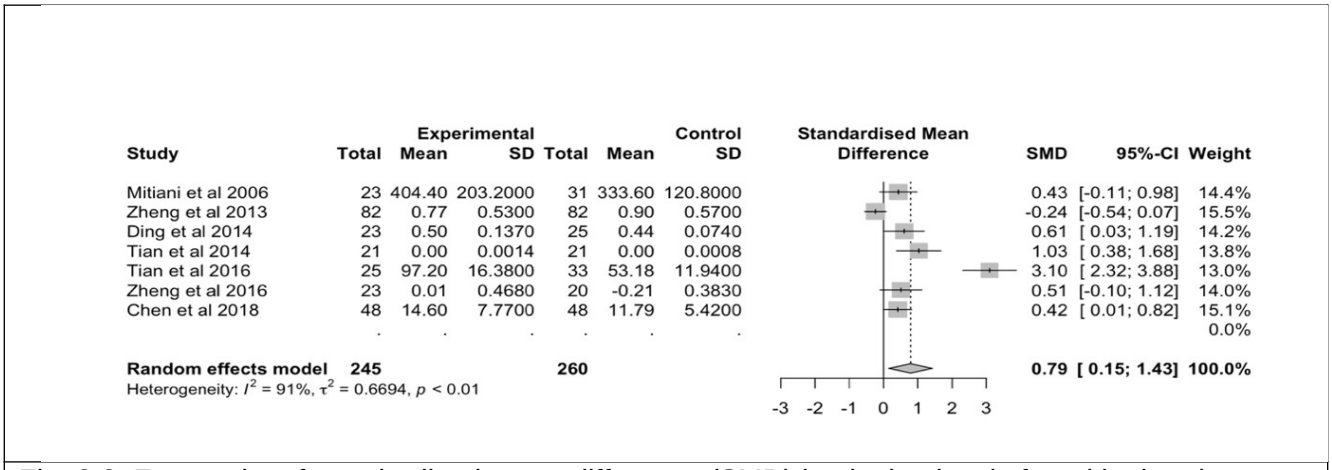


Fig. 2.2: Forest plot of standardised mean difference (SMD) in alanine levels found in the plasma or urine of patients with MDD and healthy controls. In the forest plots of standardised mean difference (SMD), the horizontal lines indicate 95% confidence intervals. The diamond at the bottom of the diagram represents the overall effect size between alanine levels between patients with MDD and healthy comparators. SMD = (mean of MDD group – mean of control group) / pooled standard deviation. I^2 at the bottom indicates the degree of heterogeneity, with values of 25%, 50% and 75% considered low, moderate, and high heterogeneity, respectively

Table 2.2 Summary of studies and statistics depicting significantly higher levels of citrate in depressed (MDD) patients than controls

Author and date	MDD (n)	Control (n)	MDD mean	Control (mean)	MDD SD	Control SD	Mean SD	Effect Size	CI (95%) Low	CI (95%) High
Zheng et al 2013	82	82	0.77	0.9	0.53	0.57	0.55	-0.24	-0.54	0.07
Tian et al 2014	21	21	0.00495	0.0037	0.00145	0.00085	0.00115	1.09	0.47	1.71
Zheng et al 2016	23	20	0.011	-0.212	0.468	0.383	0.43	0.52	-0.09	1.14
Chen et al 2018	48	48	14.597	11.793	7.77	5.42	6.60	0.43	0.02	0.83

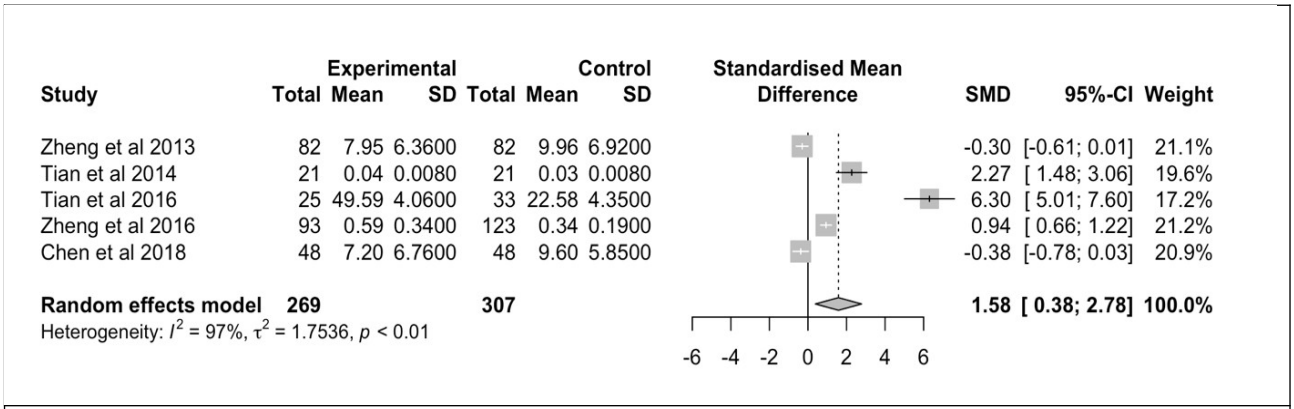


Fig 2.3 Forest plot of standardised mean difference (SMD) in plasma citrate levels of patients with MDD and healthy controls

Table 2.3: Summary of studies and statistics showing significantly higher levels of glutamate in the plasma and urine of people with schizophrenia compared to control subjects

Author and date	Schiz (n)	Comparator (n)	schiz (mean)	Comparator (mean)	schiz (SD)	Comparator (SD)	Mean SD	Effect Size	CI (95%) Low	CI (95%) High
Oresic et al 2011	75	75	5.06	4.46	1.89	1.77	1.83000	0.33	0.01	0.65
Fukushima et al 2014	25	27	71.3	35.5	12.3	3.2	7.75	4.62	4.06	5.17
Nagai et al 2017	19	16	3.26	1.42	1.74	1.26	1.50	1.23	0.54	1.91
Yoshikawa et al 2018	5	5	0.11	0.015	0.029	0.007	0.02	5.68	4.29	7.06
Garip et al 2019	41	35	187.5	125	40	61.6	50.80	1.23	0.77	1.69

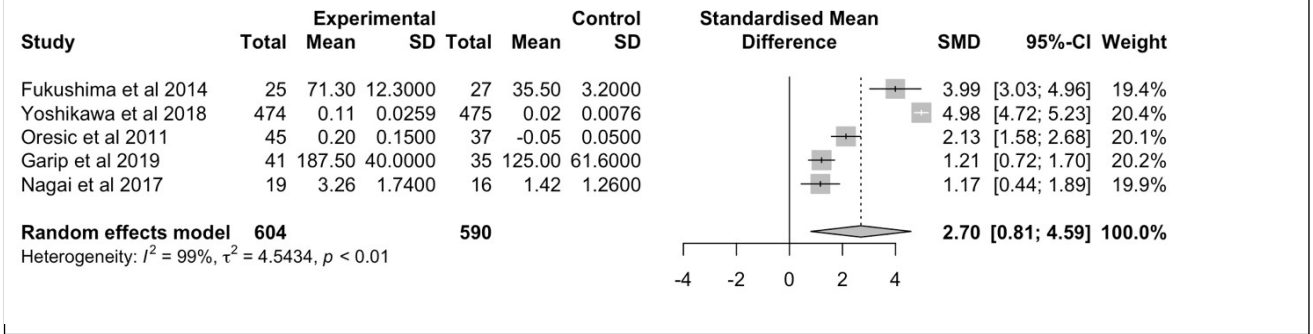


Fig.2.4 Forest plot of standardised mean difference (SMD) in glutamate levels found in the plasma

2.4 Discussion

2.4.1 Significantly elevated alanine levels associated with major depressive disorder

Alanine is important for the maintenance of muscle protein, but it can be also mobilized from muscle to liver cells where it is converted to glucose through the Cahill cycle (Felig, 1973). Alanine also modulates the secretion of insulin through the release of the incretin, glucagon-like peptide 1 (GLP-1). Based on these observations, it is reasonable to propose that mood dysfunction/depressive symptoms are associated with aberrant energy metabolism.

In the tricarboxylic acid (TCA) cycle, alanine contributes indirectly by providing substrates for the generation of intermediates like pyruvate and α-ketoglutarate, supporting energy production and maintaining cellular functions (Fig 2.5). It is proposed, therefore, that in depression, α-ketoglutarate availability is reduced because the progression of the TCA cycle is hindered. As a homeostatic response to replenish α-ketoglutarate, more glucose is generated from mobilized alanine to boost TCA cycle activity. This metabolite can also be produced from glutamate by glutamate dehydrogenase, which is highly expressed in the brain. The mobilization of glutamate into the TCA cycle may underlie reduced central levels of this amino acid in depression, which has been proposed by a recent meta-analysis. To support or refute these suppositions it is essential to examine the relationship between depression and metabolites of glucose oxidation (citrate, α-ketoglutarate, malate, alanine, glutamate other amino acids), and mediators of glucose homeostasis (alanine, insulin, glucagon, GLP-1, Gastric Inhibitory Peptide and other gut hormones).

MDD is a multifaceted psychiatric condition with intricate aetiology and pathophysiology, encompassing genetic, environmental, psychological, and biological factors (Marx et al., 2023). Among the biological factors, disruptions in amino acid metabolism, the building blocks of proteins and neurotransmitters, have emerged as potential contributors to MDD development and progression (Bortolato et al., 2017). Dysregulation of amino acid metabolism could perturb these processes, thereby influencing MDD pathogenesis (Gong et al., 12019). Alanine plays a crucial role in hepatic gluconeogenesis and participates in the alanine-glucose cycle, facilitating nitrogen and carbon exchange between skeletal muscle and the liver (Wu et al, 2019). Several studies have reported significantly elevated serum or plasma alanine levels in MDD patients compared to healthy controls (Islam et al., 2020; Li et al., 2021; Wu et al., 2022). The clinical implications of heightened alanine levels in MDD are multifaceted, potentially correlating with disease severity (Mitani et al., 2006) and response to treatment (Wu et al., 2022). However, alanine levels may be influenced by factors such as age (Islam et al., 2020), sex (Wu et al., 2022), body mass index (BMI) (Wu et al., 2022), medication use (Li et al., 2021), and comorbidities (Wu et al., 2022), indicating that it might not serve as a specific or consistent biomarker for MDD (Chandrasekaran & Swaminathan, 2015). In conclusion, while elevated alanine levels are intricately linked with MDD, further research is warranted to unravel the causal relationships and underlying mechanisms. These endeavors hold promise for advancing our understanding of the role of alanine metabolism in MDD pathophysiology and its potential utility as a biomarker or therapeutic target.

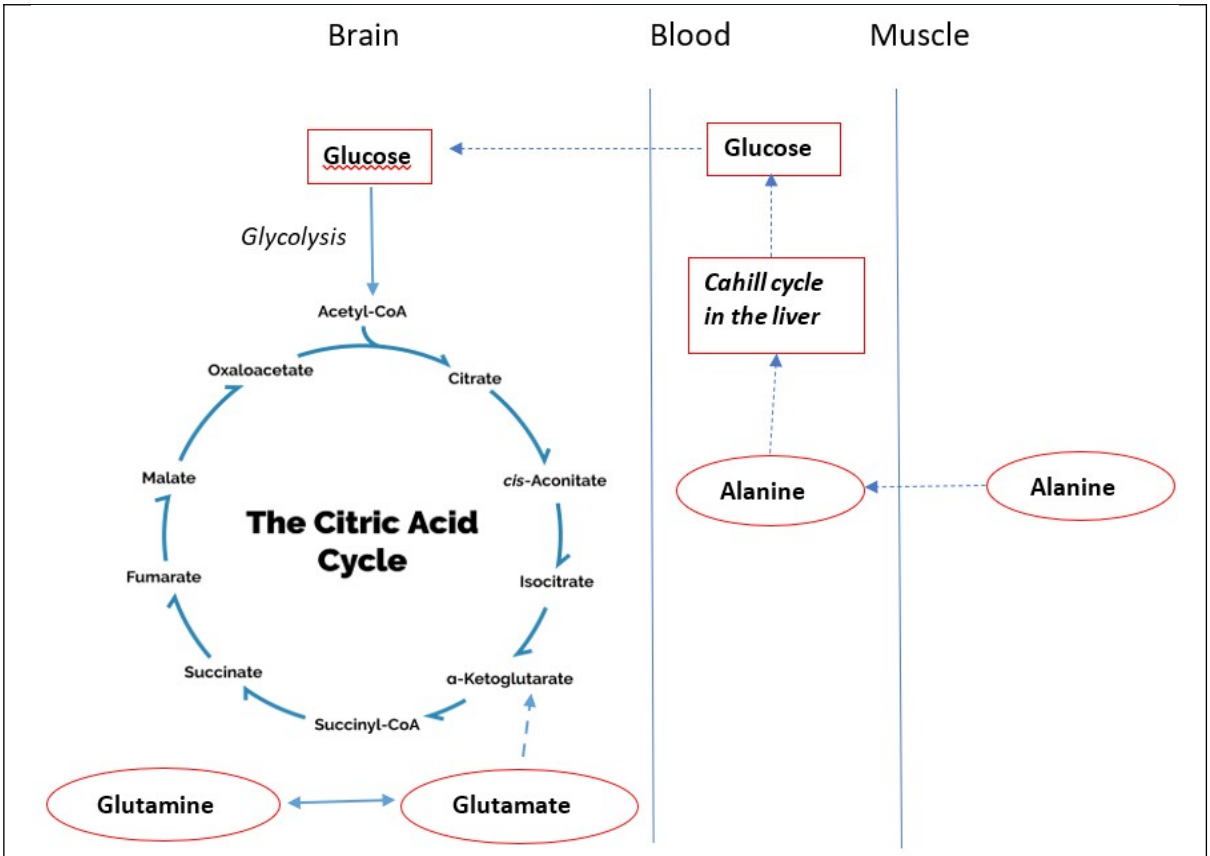


Fig 2.5 Mobilisation of alanine and the TCA cycle in a healthy individual

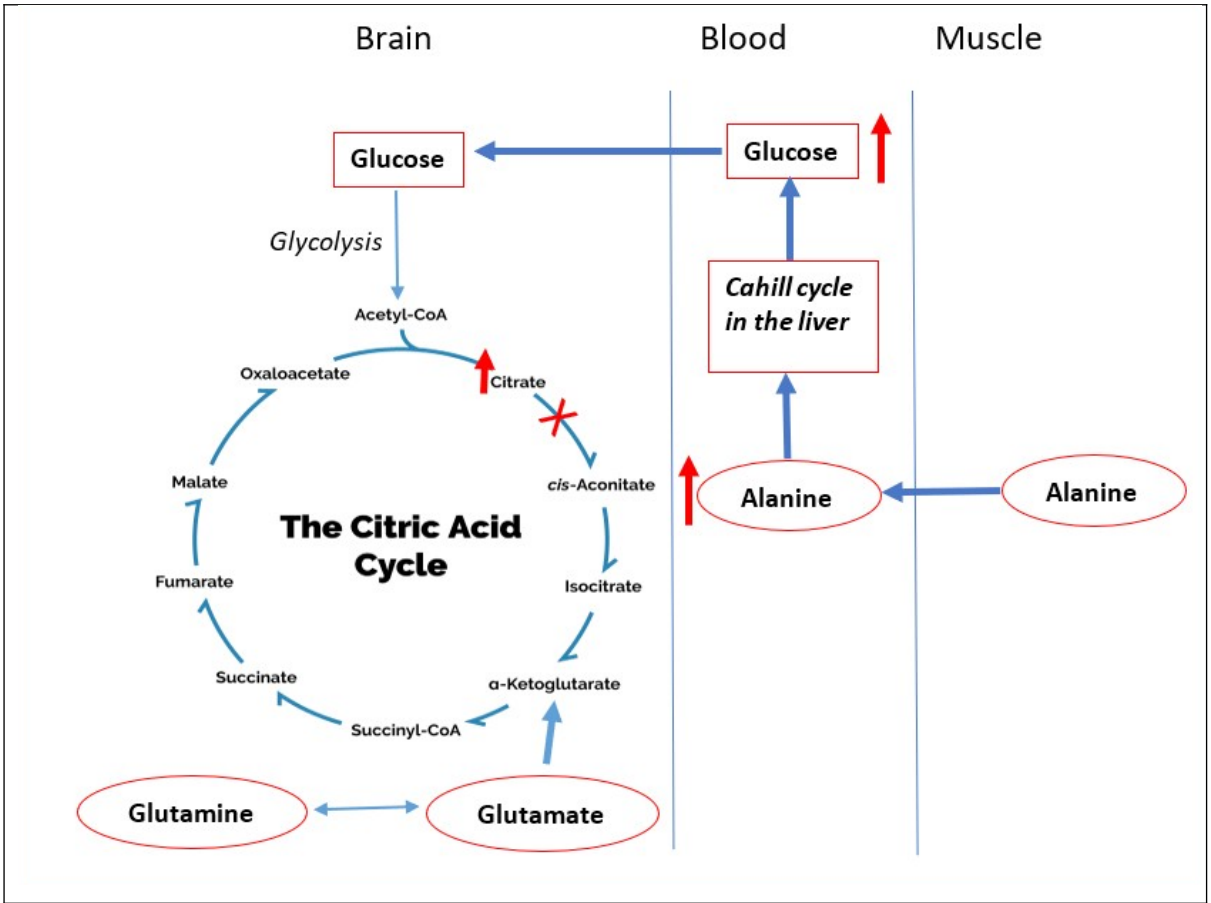


Fig. 2.6 Hypothesised TCA cycle disruption in a depressed individual

The Fig. 2.6 explores complicated metabolic pathways crucial for the conversion and transport of glucose, alanine, and glutamine among the brain, blood, and muscles. A vivid depiction of these processes is presented, highlighting key biochemical events such as glycolysis and the Citric Acid Cycle. Glycolysis emerges as a pivotal stage, signifying the breakdown of glucose into pyruvate. This intermediary compound holds the potential for further transformation into Acetyl-CoA, a vital molecule in energy production. The Citric Acid Cycle, another focal point, illuminates a sequence of reactions that yield Adenosine Triphosphate (ATP), Nicotinamide Adenine Dinucleotide (NADH), and FADH2 (Flavin Adenine Dinucleotide). These compounds are generated through the interaction of Acetyl-CoA with oxygen, culminating in a cascade of reactions. Furthermore, the cycle yields key intermediates like α-ketoglutarate and oxaloacetate, extending their utility in diverse cellular functions. Beyond the core metabolites, it navigates the roles of amino acids, specifically

glutamine and alanine. Glutamine and glutamate emerge as pivotal players in protein synthesis, neurotransmitter regulation within the brain, and their interconversion facilitated by enzymes like glutaminase and glutamine synthetase. Meanwhile, alanine assumes a crucial role in maintaining blood glucose levels during periods of fasting or exercise. Its transportation from muscles to the liver contributes to the Cahill cycle, a process ensuring glucose homeostasis. Therefore, it offers a thorough comprehension of the biochemical mechanisms that support physiological balance and cellular energy dynamics.

Compensatory hypothesis: a homeostatic response to the disruption in TCA cycle and glutamate. Considering citrate is a major product of the TCA cycle, and if the increase in alanine and citrate is related, can the increase in alanine be a result of the disruption of the TCA cycle in depression? A hypothesis on the role of alanine in depression is proposed (Figure 9). This proposal assumes that the elevation in alanine is a homeostatic response to an impeded TCA cycle. In depression, the TCA cycle activity is reduced, speculatively from a ‘block’ in the metabolism of citrate by the enzyme, aconitase. As a result, more alanine is needed to replenish the α-ketoglutarate level, therefore, alanine is increased to produce glucose through the Cahill cycle described in Figure 3. In this hypothesis, the increase in alanine is a compensatory mechanism to boost the TCA activity that was affected by the pathophysiology of depression whose mechanism is not precisely known.

Our current finding of increased glutamate in people with schizophrenia/psychosis is consistent with findings by Fukushima et al, (2014) and Yoshikawa et al (2013). Yang et al and Liemburg et al also found elevated glutamate levels in the prefrontal cortex in patients with schizophrenia and early-stage psychosis respectively (Yang et al., 2016; Liemburg et al., 2016).

Amino acid metabolites are important to look at as key neurotransmitters involved in hypothesised pathways of depression are themselves amino acids or derived from amino

acids, such as GABA, glutamate, dopamine, tryptophan, and noradrenaline (Ogawa et al., 2015). In schizophrenia, the metabolites which the studies seem to reach a consensus on are glutamate, N-acetyl aspartate (NAA), tyrosine and phenylalanine. Glutamate received the most overwhelming consensus, with 4 studies showing its upregulation. This complements the glutamate hypothesis of schizophrenia where there is hypofunction of glutamatergic function through NMDAR receptors, resulting in toxically elevated levels of glutamate (Fukushima et al., 2014). Yoshikawa et al (2013) found significant elevation of glutamate ($p=1.3x10^{-4}$) and additionally sequenced genes of the glycine cleavage system in a large cohort of schizophrenic patients prior to metabolomic analysis. Fukushima et al (2014) also found significantly higher levels of glutamate (p value=0.01). These findings complement in vivo studies using magnetic resonance spectroscopy which also found elevated glutamate levels in the medial prefrontal cortex (Yang et al., 2020). Liemburg et al also references high levels of glutamate in early-stage psychosis, making the differentiation between high and low levels of glutamate during the early and later stages of psychosis respectively (Liemburg et al., 2016).

Despite being informative, this study has its limitations. Firstly, these metabolites may well exist in different concentrations in circulating versus excreted amounts. It would thus be interesting to look into metabolites in just one specific biological tissue such as blood plasma or cerebrospinal fluid, and not urine. This further complements the whole-body concept of psychiatry, where the brain is seen to have a physiological effect on other parts and functions of the body.

Evidently, the heterogeneity in this meta-analysis was large, attributed to the various scales and instruments used across the studies. A sub-analysis done for alanine levels in depression, eliminating the studies that contributed the most to overall heterogeneity. The forest plots for this sub-analysis are in the appendix (Given after Appendix-3, Figures 5 and 6).

Fortunately, the significance level did not change for the alanine, showing that alanine was still significantly elevated in depressed patients and remains the most robust result in this study. Nevertheless, follow-up studies accounting for this large variance that contributed to the overall heterogeneity should be conducted, using a standard scale specific to the disease, alongside a standard metabolomics method. This will help clarify whether alanine is progressing depression itself, or whether it is a compensatory mechanism for the disease that is failing to work.

2.4.2 Alanine in depression:

The systematic review and meta-analysis revealed a novel link between alanine and depression. This increase has been observed in various depression subtypes and across genders. A higher plasma level of alanine was identified in both the MDD patients with ELS experience (ELS/MDD patients) and the MDD patients without it (non-ELS/MDD patients) (Ding et al., 2014), which suggests the increase in alanine can be present in different depression subtypes. Moreover, in a sex-specific urinary metabolite study, alanine was changed in both men and women MDD patients (Zheng et al., 2016), indicating alanine is not gender-specific. The association between alanine and depressive symptoms has also been confirmed in animal studies. For example, after inducing depressive-like behaviours in rats with chronic unpredictable mild stress, analysis of faecal metabolome revealed that changes in alanine abundance were related to the depressive-like behaviours (Jianguo et al., 2019).

Indeed, the alanine metabolism pathway has been implicated in the mechanism of existing antidepressants such as venlafaxine (Hu et al., 2018). Additionally, using hippocampal metabolomics, a study demonstrated that another herbal antidepressant treatment, Xiaoyaosan, can regulate alanine, aspartate, and glutamate metabolism (Liu et al., 2021).

Furthermore, the urinary alanine level was significantly reduced after 6-week treatment with Xiaoyaosan (Tian et al., 2016).

However, some studies identified no change or decreased alanine level in depression. For example, one study found no significant difference between patients with and without treatment-resistant depression (TRD) and control subjects in the serum concentrations of amino acids (e.g. serine, glycine, taurine, alanine, glutamate and aspartate) (Maes et al., 1998). Nevertheless, the study also showed that antidepressant treatment reduced the serum amino acid level (i.e. glutamate, taurine and aspartate) and increased the serum glutamine concentration. The study was not included because the article is not in the date range of the selection criteria (see Methods). Moreover, another study identified decreased alanine levels in depressed patients (Liu et al., 2015). However, this study was not included in the present meta-analysis because a later study by the same group (Tian et al., 2016) combined the data from Liu et al., (2015) with those from additional participants, to increase the power of their analysis. Therefore, the Liu et al., (2015) study was essentially updated which makes the original investigation redundant. Despite some contradicting evidence, alanine still demonstrated potential as a biomarker for both diagnosis and treatment of depression.

2.4.3 Citrate in depression:

The meta-analysis also identified increased citrate levels in depressed patients. Interestingly, all the five articles involved in the quantitative analysis studied urine (Table 1). Nevertheless, a study showed that blood citrate was found to be strongly associated with the severity of depression, especially suicidal ideation. This suggests that the chelating actions of citrate can cause a decrease in the blood coagulation capability to produce divalent Ca2+ and become a plausible pathway for suicide (Setoyama et al., 2016). The study was not included

in the meta-analysis because medicated patients were included, which did not meet the selection criteria.

Moreover, another study also identified increased citrate concentrations in the cerebrospinal fluid in patients with symptoms of depression compared to the non-psychiatric patients (Mellerup and Rafaelsen, 1981). Therefore, citrate levels in plasma and cerebrospinal fluid can also be associated with depression, despite the meta-analysis having only identified urine studies.

Moreover, the significance of citrate in depression has been supported by both new and traditional antidepressant studies. A significant shift of the citrate cycle in the hippocampus was observed in a mice metabolic study following ketamine administration (Weckmann, et al., 2014). Like alanine, after the treatment of Xiaoyaosan for 6 weeks, the citrate level was also decreased (Tian et al., 2016). Additionally, chronic administration of fluoxetine can divert the energy metabolism towards oxidative phosphorylation and the citric acid cycle (Filipović et al., 2017). It was suggested that citric acid can relieve the symptoms of depression (Hu et al., 2018). Overall, it seems that citrate also has the potential to be a biomarker for depression. This investigation is furthered in the chapters ensuing, to see if the same findings could be replicated in a larger population, with genetic factors controlled.

Following conclusions may be extracted from the above SLR and meta-analysis study:

The study encompassed 282 studies, finalizing 40 for review and 17 for meta-analysis. It revealed higher citrate and alanine levels in depressed patients and elevated glutamate in schizophrenic patients. There were more metabolites associated with depression than with schizophrenia.

The current systematic review and meta-analysis found significant elevation of alanine in plasma and urine from people with MDD in 7 studies, which is in-keeping with a previous

study showing that circulating alanine correlated with the severity of depression (Mitiani et al., 2006).

Various metabolites were consistently elevated in depressed patients, such as alanine and citrate, with different studies highlighting consistent increases in their levels, particularly derived from urine and plasma. The elevated concentration of citrate in depression (Table 2.1, Fig 2.2) also supports a metabolic dysfunction in this disorder as it is a key metabolite of the Tricarboxylic Acid (TCA) cycle (Fig 2.5).

Glutamate showed consistent elevation in schizophrenic patients, though findings were less significant compared to depression. In schizophrenia, diverse metabolic pathways like fatty acids, carbohydrate metabolism, amino acids, and antioxidants displayed varied trends, presenting inconsistencies in upregulation or downregulation across different studies.

Chapter 3: Methodology for TwinsUK Dataset Analysis

3.1 Data cleaning and exploratory analysis

Following acquisition of the database through the TwinsUK Request Portal, exploratory analysis was conducted as follows. Firstly, descriptive statistics were obtained for the entire TwinsUK database, to glean a general overview and breakdown of demographic variables. Next, participants who could provide answers to all 14 questions of the HADS questionnaire were isolated, during all three years that the scores were collected: 2002, 2017 and 2018. The next steps involved identifying the frequency of participants who could provide data for HADS scores and metabolite levels each year, ranging from 7 years before 2002 (since the first metabolite collection was in 1996) and 5 years before 2018 (since the last date of metabolite/lipid collection was in 2013). Participants with duplicate data were still included, but the patient ID was associated with the data that was closest to the date of HADS score collection.

3.2 Overcoming obstacles in the TwinsUK database

The biggest issue in navigating the database was in isolating participants who had sufficient data for all metabolites, lipids, and variables of interest, and in ensuring that this information lay close to the time at which they took the HADS questionnaire (explained below), which indicated how severe the participant's depression was, or if they were healthy. If the time of metabolite sample collection was much later or earlier than when information on a variable of interest was collected, it may be falsely assumed that levels of the metabolites or that of the variable was relevant to its impact on the severity of the participant's depression.

Missing data was thus one of the main obstacles in sorting through the database; it was difficult but important to isolate participants who could provide data for everything– metabolites, lipids, lifestyle variables, and HADS scores close enough to biological data and variable data. Participants may have provided alanine and citrate level data, but not that for smoking habit or frequency of alcohol consumption, and vice versa. Further, not all participants provided data at multiple time points, which would make for a less more robust analysis.

3.3 Inclusion and exclusion criteria

Only participants who could offer at least one HADS score and were missing <15% of metabolite and lipid data were included in this investigation, to preserve the most amount of data to power the study while ensuring reliability.

3.4 Formation of a test set for univariate and multivariate analysis

Depression phenotypes were determined by slotting the participants into their respective diagnostic groups (normal, borderline abnormal and abnormal), according to the HADS scores specified in the HADS questionnaire. The HADS questionnaire is split into two categories: HADS-D and HADS-A, to separate a patient's depression and anxiety-related score respectively. The addition of these two scores is the participant's total HADS score. A score of 0-7 in each category is considered healthy, 8-11 is borderline abnormal, and anything above 11 is considered abnormal (clinically depressed), out of a score of 21. A crucial aim in the data cleaning process was striking a balance between having enough participant data for a robust analysis and having the least amount of time elapsed between time of HADS score collection and time of blood sample collection, defined as the year cut-off.

Participants who could provide metabolite or lipid information within one year of providing a HADS score (in the year closest to the metabolite collection, in 2002 or 2017) were identified, before identifying how many abnormal, normal, and borderline participants comprised each group. A demographics table was drawn up to show the numbers in each year cut-off- 1, 2, 3, 4 and 5 years, alongside statistical tests to see if statistical differences were present between groups. This process was repeated for participants with lipid data.

Tables were constructed separately for metabolite data and lipid data, to compare the two. Variable data that was added not only had to lie within the specified year cut-off but was chosen to maximise the amount of data available that could match the participants in the test set group. Variables that could be attached from the raw database to each diagnostic group included: age, sex, total HADS score, depression score, anxiety score, BMI, alcohol consumption, smoking habit, glucose and insulin levels, inflammatory marker levels, previous medication history, prior use of hormones, and comorbidities.

3.5 Statistical analysis

3.5.1 Analyzing Diagnostic Group Disparities

All participant data was represented as mean± standard deviation. Participants were matched to variables using Microsoft Excel's Vlookup function. Categorical differences between groups such as gender, age, smoking was calculated using the chi-squared test.

3.5.2 Univariate analysis

Univariate analyses were executed in SPSS and R software environment to identify variables, metabolites and lipids that had significant associations with increasing severity of depression as defined by HADS scores. Pearson correlations were executed These correlations for metabolites, lipids or variables against depression severity were also visually represented in correlation plots. Confirmatory correlational analyses were executed using Spearman's correlation coefficient, with alpha set to 0.05, with accompanying p values to identify significantly correlated variables, metabolites, or lipids with HADS scores. Participants were also stratified by medication status. RStudio was also used to generate correlation plots to look at correlations between all variables and HADS scores.

Linear regression was executed to identify the degree to which a variable, metabolite or lipid affected the severity of a participant's depressive symptoms, as indicated by the HADS scores, using the lm function by fitting a model with HADS scores predicted by the independent variable (be it lifestyle variables, metabolite type or lipid type in the TwinsUK database). One-way ANOVA was conducted to identify significant differences between the three diagnostic groups (normal, borderline abnormal and abnormal), with Bonferroni to identify which of the diagnostic groups the difference lay between, visualised using bar graphs. ANOVA and Bonferroni calculations were calculated in SPSS.

3.5.3 Multivariate analysis

Multivariate analysis was conducted via the forward selection method in RStudio using the 'step.mode' function, to identify the most significant variables, metabolites or lipids that contributed to depression severity as indicated by HADS scores in the questionnaire. Four plots could be generated to visualise the relationship between the chosen independent variables and the dependent variable (HADS score): i) residuals versus fitted values, ii) a Q-

Q plot of standardised residuals, iii) fitted values against standardised residuals, and vi) residuals against leverage.

Significant associations were visually represented via: correlation plots formed in RStudio version 3.6 software, tables showing Pearson correlations and accompanying significance values, multiple regression, tables showing statistically significant differences for HADS against variables, and post-hoc Bonferroni tests to indicate significant differences in variables between the three diagnostic groups: normal, borderline abnormal and abnormal.

3.5.4 OPLS-DA

OPLS-DA was conducted in RStudio Version 4.1.2, using a code constructed by Dr. Fay Probert, based off its mechanism which matches two class sizes, before being shuffled and split into independent and training test sets, which build and predict models that decide the top important contributing variables. The aim of this method was to validate the model on an independent test set. First, the metabolite and lipid data was split into three separate Excel files, to align with the mechanism off which the code run, by splitting each of the three classes of interest: “Normal”, “Borderline Abnormal” and “Abnormal” so that two classes could be compared in each of the three excel files. The ‘vegan’, ‘ropls’, ‘scatterplot3d’ and ‘tcltk’ libraries were installed, and iteration was set to 10 for 10-fold cross validation.

OPLS-DA (orthogonal projection on latent structure discriminant analysis) was carried out to determine the directional relationship among these variables and to identify the most significant variables, which included variables pertaining to lifestyle as well as metabolites and lipids. After determining which factors were the most significant, structural equation modelling was employed to analyse the data and find out how strongly each variable was associated with the diagnosis of depression, as well as any possible grouping-based

directional relationships (which may be either positive or negative). Operating Characteristic (ROC) curves are graphical representations that show the performance of a binary classifier system at various discrimination thresholds (Gneiting & Walz, 2022). The graph compares the true positive rate (sensitivity) against the false positive rate (1 - specificity) at various threshold settings. It helps to see the trade-off between correctly detecting real positives and wrongly classifying false positives. ROC curves were then used to determine the discriminatory threshold of the variables that contributed most to differentiating between the three classes: Normal, Borderline and Abnormal.

Chapter 4: Investigating the relationship between demographic and lifestyle variables and severity of major depressive disorder in the TwinsUK database

4.1 Introduction

4.1.1 The importance of twin studies in understanding lifestyle variables and environmental influences on MDD

Twin studies have been especially helpful in elucidating relationships between potential environmental factors, and psychiatric disorders, over long periods of time. Since monozygotic (identical) twins share all their genes while dizygotic (fraternal) twins share half of their genes, twin studies can control for the heritability of depression. Polderman et al. (2015) reported a heritability of 49% for human traits, indicating that roughly half of phenotypic variation is due to genetics. Nevertheless, twin studies of psychiatric disorders, particularly depression, show varying results. Some suggest monozygotic twins are more prone to depression than dizygotic twins, while others indicate a three to four times higher risk for one twin if the other has depression (Flint, 2023; McGuffin et al., 1991). In their meta-analysis (Polderman et al., 2015) it was also found that genetic factors account for just one-third of variation in depression severity, complementing findings by Lombardi et al (2022) who showed that major depressive disorder (MDD) has a much broader set of causes that range from biological to environmental to psychological. Evans et al. (2021) found an uptick in MDD symptoms during the COVID-19 pandemic, which they believe may have been caused by increased alcohol consumption. Both direct and indirect effects of medications on the central nervous system (CNS), such as exhaustion, have been linked to an increase in

depression symptoms (Celano et al., 2022), too. The effects of antidepressant medication, as well as the pathophysiology and somatic symptoms of MDD, have been linked to pro-inflammatory cytokines, notably Interleukins like IL-6 (Ting et al., 2020).

Overall, there is no conclusive evidence to suggest that twins are more likely to be depressed than other individuals, although twins may face a unique set of challenges in their life such as confusion about their identity brought on by comparison to their co-twin, that could potentially lead to more psychological distress (Amani et al., 2021). The TwinsUK database is the largest adult registry comprising twins aged 16-98 and is predominantly middle-aged and female. The types of measures performed for the database are biochemical, self-reported, omics-related, and physical. Over 10,000 phenotypes have been collected in the 30 years of the database's existence, with measurements taken at multiple time points to assist longitudinal analysis, and these are continuously being updated for refreshed observations (Spector, 2021). For instance, the biological data available from the TwinsUK database includes metabolite and lipid concentrations in blood plasma, which can be coupled to donor demographic details ranging from alcohol consumption frequency, smoking habits, education level, to medication history. This provides, therefore, a rich source of information for a comprehensive investigation into the association of such variables with depression as well as other metabolites/lipids which may be involved in the pathophysiology of the disorder or are components of potential therapeutic pathways.

4.1.2 Rationale for investigation

The purpose of this investigation was to use data extracted from the TwinsUK database to confirm prior findings in the systematic review and meta-analysis (Chapter 2), which found an association between depression and elevated levels of the energy metabolites alanine and citrate, and to see if these findings could be replicated in a general population. The

potential of these metabolites to serve as biomarkers, to differentiate between depression and other mental illnesses such as psychosis, particularly in the early stages of the disorder, will greatly improve diagnosis and treatment. The present thesis is split into the following chapters: the current chapter will investigate the relationship between severity of depression and the variables available in the database, followed by the relationship between severity of depression and metabolites (Chapter 5), followed by the relationship between severity of depression and lipids (Chapter 6), and finally, the relationship between levels of metabolites and lipids with each other combined with the influence of lifestyle and demographic variables against depression severity (Chapter 7), increasing the scope of the study's rationale.

The first part of this chapter involved an exploratory analysis following thorough data cleaning, eliminate missing or redundant data, and verify assumptions on relationships between the variables of interest. This provided a robust basis upon which multilevel modelling could be performed, to see if variables available in the database such as medication, education level or alcohol consumption, play a role in the significance of elevated metabolite levels in a depressed versus a healthy person. Considering how environmental factors and stress can influence neuronal function epigenetically, understanding the influence of variables encompassing that of lifestyle and every day habit would prove to be invaluable information.

4.1.3 Measurement of depression

The HADS questionnaire, attached in the Appendix 4, was used to assess the severity of major depressive disorder for the TwinsUK cohort. This questionnaire is split into 14 questions, with a possible total score of 21. Results are grouped in the questionnaire according to three categories: 'Normal' represents scores of 0-7, 'Borderline Abnormal' represents scores between 8 and 10, and 'Abnormal' represents scores between 11-21.

4.2 Demographic and lifestyle variables that affect the severity of MDD

There are many lifestyle variables known to affect the severity of depression, and some which have an unclear influence on the disease, and MDD is significantly influenced by various demographic and lifestyle factors. Increased alcohol consumption often correlates with higher depression severity, engaging in more exercise tends to mitigate depressive symptoms, more smoking and higher BMI typically exacerbate the severity of depression. Additionally, long-term medication usage, especially with antidepressants, and hormonal medication may be associated with heightened symptom severity. Inflammatory markers like interleukins and neurotrophins also impact the intensity of MDD symptoms. Moreover, higher education levels, income, and the presence of certain comorbidities often coincide with a reduced severity of depression. Understanding these demographic and lifestyle variables is pivotal in assessing and addressing the severity of MDD.

For more detailed information relating to these variables please see Chapter 1, Section 1.4.

For full methodology, refer to Chapter 2. To summarise briefly, this chapter will:

- Provide descriptive statistics of the database
- Identify participants who had a HADS score, metabolite data, and lipid data
- Construct a test set of participants upon which univariate and multivariate analyses could be performed
- Use statistical analysis to identify differences between diagnostic groups, and to identify variables that are significantly associated with increasing severity of depression

4.3 Results

4.3.1 Isolation of a test set for univariate and multivariate analysis

Table 4.1 presents the statistical parameters for the TwinsUK cohort, a large population-based sample of twins from the United Kingdom. The cohort is categorized into three sections: Normal, Borderline abnormal, and Abnormal, based on the values of the variables measured within the cohort. To ensure data accuracy and relevance, we focused on a subset of this extensive TwinsUK database, specifically the data from five years before and after the benchmark year 2002. From this subset, we filtered for variables with missing data totalling less than 15%, resulting in a dataset of 1,532 records. Within this carefully selected dataset, our study investigated into metabolite and lipid analysis spanning a temporal range of one to five years. Commencing with a one-year cutoff, our analysis began with data extracted from 39 unique samples, setting the foundation for subsequent exploration. As we expanded our temporal purview to encompass a two-year cutoff, data from 138 unique samples enriched our understanding. Continuing along this trajectory, the three-year cutoff entailed the analysis of 287 unique samples, followed by a subsequent leap to a four-year cutoff, encapsulating data from 501 unique samples. Finally, the five-year cutoff, characterized by its expansive scope, involved a parallel engagement with 501 unique samples, culminating in a comprehensive analysis across all cutoff points. The total number of unique samples analyzed amounted to 1,466, post-exclusion of missing data, underscoring the depth and breadth of our metabolite and lipid analysis. Within this dataset, our primary focus was on a two-year database of metabolites and lipids, comprising 113 records in total. Among these, 43 were classified as Normal, 35 as Borderline Abnormal, and 35 as Abnormal, providing valuable insights into the health and well-being of this diverse twin population. The Table 4.1 includes 14 variables related to physical and mental health, such

as age, body mass index (BMI), glucose, insulin, smoking, alcohol, interleukin 10 (IL-10), interleukin 1 beta (IL-1B), exercise, blood pressure (BP) age, hypertension (HC) age, income, depression, and anxiety. The statistics includes the count, mean, standard deviation (SD), median, mode, maximum, minimum, and range of each variable for each section and for the total cohort. The normal group has the lowest mean values for BMI, glucose, insulin, smoking, alcohol, IL-10, IL-1B, BP age, HC age, income, depression, and anxiety compared to the other groups. The abnormal group has the highest mean values for BMI, glucose, insulin, smoking, alcohol, IL-10, IL-1B, BP age, HC age, income, depression, and anxiety compared to the other groups.

.

Table 4.1 Statistical parameters of TwinsUK cohort

Type	Variables	Statistical parameters							
		Count	Mean	SD	Median	Mode	Max.	Min.	Range
Normal	Age	43	57.39	8.81	59.16	41.64[a]	73.69	41.64	32.05
	BMI	43	26.91	6.04	26.50	19.19[a]	40.00	19.19	20.81
	Glucose	43	4.48	2.17	4.60	.00[a]	11.07	0.00	11.07
	Insulin	43	6.94	5.24	5.75	4.60	25.90	2.60	23.30
	Smoking	43	9.32	8.42	14.00	0.00	21.00	0.00	21.00
	Alcohol	43	7.93	10.24	5.00	0.00	41.00	0.00	41.00
	Interleukin10	43	3.59	9.74	0.00	0.00	36.20	0.00	36.20
	Interleukin1B	43	0.56	1.57	0.00	0.00	6.62	0.00	6.62
	Exercise	43	2.63	0.59	3.00	3.00	4.00	2.00	2.00
	BP_Age	43	15.26	24.80	0.00	0.00	70.00	0.00	70.00
	HC_Age	43	0.29	0.46	0.00	0.00	1.00	0.00	1.00
	Income	43	2.63	2.33	3.00	0.00	8.00	0.00	8.00
	Depression	43	1.93	1.53	2.00	2.00	5.00	0.00	5.00
	Anxiety	43	2.26	1.60	2.00	1.00	6.00	0.00	6.00
Borderline abnormal	Age	35	59.12	9.91	62.25	48.00[a]	74.80	41.63	33.17
	BMI	35	27.14	6.78	27.28	18.08[a]	38.53	18.08	20.45
	Glucose	35	3.12	2.24	4.17	0.00	5.95	0.00	5.95
	Insulin	35	7.27	4.66	6.05	2.20[a]	18.90	2.20	16.70
	Smoking	35	11.97	8.34	15.50	0.00	23.00	0.00	23.00
	Alcohol	35	10.26	10.67	5.00	0.00	35.00	0.00	35.00
	Interleukin10	35	1.12	4.23	0.00	0.00	21.99	0.00	21.99
	Interleukin1B	35	0.20	0.72	0.00	0.00	3.46	0.00	3.46
	Exercise	35	2.31	0.71	2.00	3.00	3.00	1.00	2.00
	BP_Age	35	16.88	25.41	0.00	0.00	68.00	0.00	68.00
	HC_Age	35	0.31	0.47	0.00	0.00	1.00	0.00	1.00
	Income	35	4.23	2.23	5.00	5.00	8.00	0.00	8.00
	Depression	35	4.20	1.98	4.00	2.00[a]	9.00	1.00	8.00
	Anxiety	35	4.77	1.80	5.00	5.00	8.00	1.00	7.00

Table 4.1 Statistical parameters of TwinsUK cohort (continued)

Type	Variables	Statistical parameters							
		Count	Mean	SD	Median	Mode	Max.	Min.	Range
Abnormal	Age	35	59.98	9.17	60.45	47.43[a]	73.69	41.63	32.06
	BMI	35	28.57	5.41	26.40	23.48[a]	39.46	23.48	15.98
	Glucose	35	3.32	2.17	4.20	0.00	7.63	0.00	7.63
	Insulin	35	8.16	5.33	6.40	2.30[a]	17.50	2.30	15.20
	Smoking	35	13.88	9.95	16.00	0.00	40.00	0.00	40.00
	Alcohol	35	13.40	13.74	8.00	0.00	45.00	0.00	45.00
	Interleukin10	35	0.54	3.20	0.00	0.00	18.91	0.00	18.91
	Interleukin1B	35	0.04	0.23	0.00	0.00	1.37	0.00	1.37
	Exercise	35	2.26	0.96	2.00	2.00[a]	5.00	0.00	5.00
	BP_Age	35	24.41	27.75	0.00	0.00	73.00	0.00	73.00
	HC_Age	35	0.26	0.45	0.00	0.00	1.00	0.00	1.00
	Income	35	3.67	2.39	5.00	5.00	8.00	0.00	8.00
	Depression	35	7.74	2.69	7.00	7.00	14.00	3.00	11.00
	Anxiety	35	8.20	2.91	8.00	7.00	17.00	3.00	14.00
Total	Age	113	58.73	9.25	59.89	66.01	74.80	41.63	33.17
	BMI	113	27.48	5.99	26.56	26.56	40.00	18.08	21.92
	Glucose	113	3.68	2.26	4.35	0.00	11.07	0.00	11.07
	Insulin	113	7.30	4.97	5.95	4.60	25.90	2.20	23.70
	Smoking	113	11.57	9.02	16.00	0.00	40.00	0.00	40.00
	Alcohol	113	10.35	11.67	5.00	0.00	45.00	0.00	45.00
	Interleukin10	113	1.88	6.78	0.00	0.00	36.20	0.00	36.20
	Interleukin1B	113	0.29	1.07	0.00	0.00	6.62	0.00	6.62
	Exercise	113	2.42	0.77	2.00	3.00	5.00	0.00	5.00
	BP_Age	113	18.44	25.89	0.00	0.00	73.00	0.00	73.00
	HC_Age	113	0.29	0.45	0.00	0.00	1.00	0.00	1.00
	Income	113	3.40	2.39	4.00	5.00	8.00	0.00	8.00
	Depression	113	4.43	3.18	4.00	2.00	14.00	0.00	14.00
	Anxiety	113	4.88	3.26	5.00	1.00[a]	17.00	0.00	17.00

a:Multiple modes exist. The smallest value is shown

The five-year cohort data set is related to metabolomics and HADS (Hospital Anxiety and Depression Scale) scores, which is organized based on different cutoff times ranging from one year to five years The five-year cohort data set has been separately sorted out into:
I) Metabolomics and II) Lipids based on Normal (0-7), Borderline Abnormal (8-10), and Abnormal (11-21):

I) Metabolomics

The following tables show subjects and associated demographic data ranging from 1 year to 5 years (Table 4.2 to Table 4.6) between the time of HADS score collection and the time of sample collection:

Table 4.2 Table showing samples with <15% missing metabolomics data and at least one HADS score: one year cut-off

Metabolites <15% and one year cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=39; NK=4)	21	8	6	
% Female	53.8%	20.5%	15.4%	No $\chi 2$ statistics is computed because sex is a constant.
Age (years, mean [SD])	65.22±5.93	67.35±8.00	71.26±1.41	ANOVA (F (21, 13) = [1.200], p = 0.375). Post hoc tests cannot be performed as at least one group has fewer than two cases.
Time between sample and HAD score (years, mean [SD])	-0.84±0.16	-0.78±0.27	-0.87±0.24	ANOVA not possible
Total HAD (median, IQR)	4.00, 4	9.50, 2	15.00, 6	$\chi 2$ test (24, N = 35) 70.00, **P<0.05**
Depression HAD (median, IQR)	2.00, 3	4.50, 2	7.50, 2	$\chi 2$ test (27, N = 39) = 56.096, **P<0.05**
Anxiety HAD (median, IQR)	1.00, 2	4.00, 2	8.00, 6	$\chi 2$ test (30, N = 39) = 54.68, **P<0.05**

Table 4.3 Table showing samples with <15% missing metabolomics data and at least one HADS score: two-year cut-off

Metabolites <15% and two years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
number of unique samples (Total=138; NK=11)	49	36	42	
% Female	36.1%	25.6%	30.8%	χ2 test (3, N = 138) = 1.864, P>0.05
Age (years, mean [SD])	56.73±9.18	59.07±9.77	58.71±9.60	ANOVA (F (79, 47) = [1.699], **p = 0.026**
Time between sample and HAD score (years, mean [SD])	-1.31±0.80	-1.09±1.04	-1.37±0.75	ANOVA not possible
Total HAD (median, IQR)	4.00, 5	9.00, 2	15.00, 5	χ2 test (50, N = 127) = 254.000, **P<0.05**
Depression HAD (median, IQR)	2.00, 2	4.00, 3	7.50, 3	χ2 test (36, N = 138) = 103.175, **P<0.05**
Anxiety HAD (median, IQR)	2.00, 2	5.00, 2	7.00, 4	No statistics are computed because Time between sample collection and HADS score is a constant

Table 4.4 Table showing samples with <15% missing metabolomics data and at least one HADS score: three-year cut-off

Metabolites <15% and three years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=287, NK=17)	128	50	92	χ2 test (69, N = 39) = 86.125, P>0.05
% Female	45.1%	16.7%	32.0%	χ2 test (3, N = 287) = 2.925, P>0.05
Age (years, mean [SD])	57.03±10.00	56.61±9.31	56.04±9.04	ANOVA (F (181, 66) = [1.642], **p = 0.005**).
Time between sample and HAD score (years, mean [SD])	-1.44±2.213	-1.78±1.973	-1.74±2.065	ANOVA not possible
Total HAD (median, IQR)	5.00, 3	9.00, 2	14.50, 7	χ2 test (58, N = 270) 540.000, P<0.05
Depression HAD (median, IQR)	2.00, 2	4.00, 2	7.0, 4	χ2 test (57, N = 287) = 250.907, P<0.05
Anxiety HAD (median, IQR)	2.00, 3	4.83, 2	8.00, 4	χ2 test (48, N = 287) = 245.617, P<0.05

Table 4.5 Table showing samples with <15% missing metabolomics data and at least one HADS score: four-year cut-off

Metabolites <15% and four years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=501; Missing data=34)	195	95	177	
% Female	38.6% (187)	18.4% (89)	36.3% (176)	χ2 test (3, N = 501) = 7.479, P>0.05 (.058)
Age (years, mean [SD])	57.61±8.36	58.28±8.46	56.01±8.06	ANOVA (F (313, 153) = [1.359], **p = 0.016**).
Time between sample and HAD score (years, mean [SD])	-2.28±2.68	-2.33±2.61	-2.17±2.78	χ2 test (57, N = 501) = 45.126, P>0.05 (0.872) ANOVA not possible
Total HAD (median, IQR)	4, 3	9, 2	15, 6	χ2 test (58, N = 467) = 934.0, P<0.05 (0.000)
Depression HAD (median, IQR)	2, 2	4, 2	7, 4	χ2 test (51, N = 501) = 419.712, P>0.05 (0.000)
Anxiety HAD (median, IQR)	2, 2	5, 2	8, 4	χ2 test (51, N = 501) = 432.924, P>0.05 (.000)

Table 4.6 Table showing samples with <15% missing metabolomics data and at least one HADS score: five-year cut-off

Metabolites <15% and five years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
number of unique samples (Total=501; Missing data=30)	188	116	167	
% Female	36.7% (180)	23.5% (115)	33.7% (165)	χ2 test (3, N = 501) = 6.123, P> (0.05) (0.11)
Age (years, mean [SD])	52.07±8.29	53.06±8.92	51.31±9.21	ANOVA (F (313, 157) = [0.985], p = 0.551)
Time between sample and HAD score (years, mean [SD])	1.72±4.23	1.32±4.30	1.46±4.30	χ2 test (51, N = 501) = 41.298, P> (0.05) (0.832) ANOVA not possible
Total HAD (median, IQR)	4, 4	9, 2	15, 5	χ2 test (56, N = 501) = 942.0**, P< (0.05)** (0.000)
Depression HAD (median, IQR)	2, 2	4, 2	7, 4	χ2 test (48, N = 501) = 427.077, **P< (0.05)** (0.000)
Anxiety HAD (median, IQR)	2, 3	5, 2	8, 4	χ2 test (48, N = 501) = 448.254, **P< (0.05)** (0.00)

LIPIDS

Below are the tables for Lipidomics which show participants and associated demographic data, ranging from 1 year to 5 years (Table 4.7 to Table 4.11) between time of HADS score collection and time of sample collection.

Table 4.7 Table showing samples with <15% missing lipidomics data and at least one HADS score: one-year cut-off

Lipids<15% and five years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=39; NK=4)	21	8	6	
% Female	53.8%	20.5%	15.4%	No chi-square is possible with only one group i.e. Female
Age (years, mean [SD])	65.22±5.93	67.35±8.00	71.26±1.41	ANOVA (F (21, 13) = [1.200], p = 0.375).
Time between sample and HAD score (years, mean [SD])	-0.84±0.16	-0.78±0.27	1-0.87±0.24	ANOVA not possible
Total HAD (median, IQR)	4.00, 4	9.50, 2	15.00, 6	χ2 test (24, N = 35) 70.00, **P<0.05**
Depression HAD (median, IQR)	2.00, 3	4.50, 3	7.50, 2	χ2 test (27, N = 39) = 56.096, **P<0.05**
Anxiety HAD (median, IQR)	1.00, 2	4.00, 2	8.00, 6	χ2 test (30, N = 39) = 54.68, **P<0.05**

Table 4.8 Table showing samples with <15% missing lipidomics data and at least one HADS score: two-year cut-off

Lipids <15% and two year cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=124; NK=11)	43	35	35	
% Female	34.7%	28.1%	28.1%	χ2 test (3, N = 124) = 0.331, P>0.05
Age (years, mean [SD])	59.98±9.17	59.12±9.91	57.39±8.81	ANOVA (F (70, 42) = [1.845], **p = 0.017**).
Time between sample and HAD score (years, mean [SD])	-1.23±0.82	-1.07±1.05	-1.31±0.81	ANOVA not possible
Total HAD (median, IQR)	4.00, 5	9.00, 2	15.00, 5	χ2 test (50, N = 113) = 226, **P<0.05**
Depression HAD (median, IQR)	2.00, 2	4.00, 3	7.00, 3	χ2 test (36, N = 124) = 94.865, **P<0.05**
Anxiety HAD (median, IQR)	2.00, 2	5.00, 2	8.00, 5	χ2 test (42, N = 124) = 129.923, **P<0.05**

Table 4.9 Table showing samples with <15% missing lipidomics data and at least one HADS score: three-year cut-off

Lipids <15% and three year cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=280; NK=15)	128	46	91	
% Female	46.3%	15.7%	32.5%	$\chi 2$ test (3, N = 280) = 3.276, P>0.05
Age (years, mean [SD])	57.95±9.87	56.95±8.39	57.04±9.01	ANOVA (F (180, 84) = [1.938], **p = 0.000**).
Time between sample and HAD score (years, mean [SD])	-1.69±1.99	-1.94±1.81	-1.87±1.91	ANOVA not possible
Total HAD (median, IQR)	5.00, 2	9.00, 2	14.00, 6	$\chi 2$ test (60, N = 265) = 530.00, **P<0.05**
Depression HAD (median, IQR)	2.00, 2	4.00, 2	7.00, 4	$\chi 2$ test (57, N = 280) = 251.423, **P<0.05**
Anxiety HAD (median, IQR)	2.00, 3	5.00, 2	8.00, 4	$\chi 2$ test (48, N = 280) = 251.195, **P<0.05**

Table 4.10 Table showing samples with <15% missing lipidomics data and at least one HADS score: four-year cut-off

Lipids <15% and four year cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=409; Missing data=32)	160	80	137	
% Female	38.8% (153)	18.8% (74)	34.5% (136)	χ2 test (3, N = 409) = 6.926, P>0.05 (0.074)
Age (years, mean [SD])	58.07±8.30	58.86±8.21	55.98±7.44	ANOVA (F (266, 110) = [1.514], **p = 0.007**).
Time between sample and HAD score (years, mean [SD])	-2.26±2.68	-2.41±2.55	-2.26±2.74	χ2 test (57, N = 409) = 48.728, P>0.05 (0.774)
Total HAD (median, IQR)	4, 3	9, 2	15, 6	χ2 test (58, N = 377) 754.0, **P<0.05**
Depression HAD (median, IQR)	2, 2	3.50, 2	7, 3	χ2 test (51, N = 409) = 342.359, **P<0.05**)
Anxiety HAD (median, IQR)	2, 2	6, 2	8, 4	χ2 test (51, N = 409) = 364.355, **P<0.05**)

Table 4.11 Table showing samples with <15% missing lipidomics data and at least one HADS score: five-year cutoff

Lipids <15% and five year cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=400; Missing data=26)	148	93	133	
% Female	36.0% (140)	23.7% (92)	33.7% (131)	$\chi 2$ test (3, N = 400) = 6.385, P>0.05) (0.094)
Age (years, mean [SD])	52.91±8.18	53.66±8.17	51.94±9.27	ANOVA (F (272, 101) = [0.938], p = 0.660).
Time between sample and HAD score (years, mean [SD])	1.43±4.33	0.63±4.46	1.22±4.37	$\chi 2$ test (51, N = 400) = 42.663, P>0.05 (0.791)
Total HAD (median, IQR)	4, 3	9, 2	15, 5	$\chi 2$ test (56, N = 374) = 748.0, **P<0.05** (0.000)
Depression HAD (median, IQR)	2, 2	4, 2	7, 4	$\chi 2$ test (48, N = 400) = 328.445, **P<0.05** (0.000)
Anxiety HAD (median, IQR)	2.5, 3	5, 2	8, 4	$\chi 2$ test (45, N = 400) = 353.852, **P<0.05** (0.000)

4.4 Two-Year HADS Score and Metabolite Data Matching

It was agreed that a cohort that had two years between time of HADS score collection and time of sample collection would be an appropriate test set. There were enough participants (at least 30 in each diagnostic group) who were missing less than 15% of data, be it for metabolites or lipids, and could provide at least one HADS score, in either 2002 or 2017. Please refer to Table 4.12 for more details. This at once preserved the number of participants who could provide data, as well as ensuring the validity of the participants' depressive symptoms. Although the symptoms of major depressive disorder are typically shorter than persistent depressive disorder, of which symptoms can last for two years or longer, choosing the two-year cut-off seemed reasonable, tightening the scope of this investigation to focus on participants afflicted with major depressive disorder. 15 non-metabolite-related variables could be matched to the participants with <15% missing metabolite data and at least one HADS score to show: alcohol consumption, smoking frequency, glucose levels, insulin levels, BMI, exercise level, levels of interleukin 10, interleukin 1β, use of oral contraception, use of medication generally, history of high blood pressure, history of high cholesterol, education level, income level, and medication status.

Table 4.12: Table showing variables matched to participants with <15% missing metabolite data and at least one HADS score, in a two-year timespan between time of HADS score collection and blood sample collection

Metabolites <15% and two years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=127) Missing data=0.0% = 0	49 N*=0	36 N*=0	42 N*=0	$\chi 2$ test (2, N = 127) =2.000, P=0.3679
Depression HAD (median, IQR) N=127 Missing data=0% = 0	2.00, 2 N*=0	4.00, 3 N*=0	7.50, 3 N*=0	$\chi 2$ test (12, N = 138) = 58.698, **P<=0.001**
Anxiety HAD (median, IQR) N=127 Missing data=0% = 0	2.00, 2 N*=0	5.00, 2 N*=0	7.00, 4 N*=0	$\chi 2$ test (14, N = 127) = 65.261, **P =<0.001.**
Age (years, mean [SD]) N=127 Missing data=0% = 0	56.73±9.18 N*=0	59.07±9.77 N*=0	58.71±9.60 N*=0	ANOVA (F (1, 125) = [1.055], p = 0.306
BMI (median, IQR) N=47 Missing data=65.3% = 80	27.40, 5.05 N*=30	27.50, 6.16 N*=23	27.50, 9.08 N*=27	ANOVA (F 1, 45) = [2.16], p = 0.149
Glucose (median, IQR) N =118 Missing data=7.08% = 9	4.60,0.63 N*=4	4.40, 0.85 N*=1	4.62, 0.79 N*=4	ANOVA (F (1, 116) = [0.119], p = 0.73
Insulin (median, IQR) N=63 Missing data=50.39% = 64	6.60, 18.30 N*=24	4.74, 8.06 N*=20	7.35, 8.33 N*=20	ANOVA (F (1, 63) = [0.442], p = 0.509
Smoking (Q9_63; median, IQR) N=123 Missing data=3.14% = 4	10.00, 16 N*=2	15.00,18 N*=1	15.00, 18 N*=1	ANOVA (F (1, 121) = [3.622], p = 0.059
Exercise N=112 Missing data=11.81% = 15	40.2 % (45) N*=4	26.8% (30) N*=6	33.0% (37) N*=5	ANOVA (F (1, 110) = [5.418], p = **0.0218**
Alcohol N=120 Missing data=5.51% = 7	38.3% (46) N*=3	29.2% (35) N*=1	32.5% (39) N*=3	ANOVA (F (1, 118) = [4.979], **p = 0.0275**
Interleukin 10 N==126 Missing data=0.7% = 1	38.1% (48) N*=1	28.6% (36) N*=0	33.3% (42) N*=0	ANOVA (F (1, 124) = [4.433], **p = 0.0373**
Interleukin 1B N=126 Missing data=0.78% = 1	38.1% (48) N*=1	28.6% (36) N*=0	33.3% (42) N*=0	ANOVA (F (1, 124) = [4.519], p = **0.0355**
E1129_ BP, N=105 Missing data=17.32% =22	41.8% (44) N*=5	25.7% (27) N*=9	32.4% (34) N*=8	ANOVA (F (1, 103) = [2.597], p = 0.11
E1129_ HC, N=105 Missing data=17.32% =22	41.8% (44) N*=5	25.7% (27) N*=9	32.4% (34) N*=8	ANOVA (F (1, 103) = [0.42], p = 0.519
Income level, N=105 Missing data=17.32% =22	41.8% (44) N*=5	25.7% (27) N*=9	32.4% (34) N*=8	ANOVA (F (1, 103) = [2.53], p = 0.115
Years of taking oral contraceptive pill N=52 Missing data=59.05% = 75	34.6% (18) N*=31	26.9% (14) N*=22	38.5% (20) N*=22	ANOVA (F (1, 50) = [0.004], p = 0.951
Are they taking any medications/supplements Year N=113 Missing data=11.02% = 14	39.8% (45) N*=4	25.7% (29) N*=7	34.5% (39) N*=3	ANOVA (F (1, 111) = [4.783], **p = 0.0308**
N*= Number of missing observations				

4.5 Three-year metabolite cohort: differences between diagnostic groups for HADS scores and variables

In Table 4.13, which represents the three-year metabolite cohort, significant differences were observed between diagnostic groups (Normal, Borderline Abnormal, and Abnormal) in terms of depression HADS scores ($p < 0.001$) and anxiety HADS scores ($p < 0.001$). These findings indicate that there are substantial variations in HADS scores among the different diagnostic groups over the three-year time frame. "Smoking" is indeed significant ($p = 0.030$), indicating that there are significant differences in smoking status among the diagnostic groups (Normal, Borderline Abnormal, and Abnormal) in the three-year metabolite cohort. However, other variables such as age, BMI, glucose, insulin, exercise, alcohol consumption, interleukin levels, blood pressure, household income, oral contraceptive pill usage, and medication/supplement intake did not show significant differences among the diagnostic groups (all $p > 0.05$). This suggests that these variables did not vary significantly based on the HADS diagnostic groups over the three-year period.

Table 4.13: Table showing variables matched to participants with <15% missing metabolite data and at least one HADS score, in a three-year timespan between time of HADS score collection and blood sample collection

Metabolites <15% and three years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=270)	47.4% (128) N*=0	18.5% (50) N*=0	34.1% (92) N*=0	χ2 test (2, N = 287) =33.867, **P= <0.001**
Depression HAD (median, IQR) N=270	2.00, 2 N*=0	4.00, 3 N*=0	7.50, 3 N*=0	χ2 test (36, N = 270) = 222.02, **P<=0.001**
Anxiety HAD (median, IQR) N=270 Missing data=0%=0	2.00, 2 N*=0	5.00, 2 N*=0	7.00, 4 N*=0	χ2 test (36, N = 270) = 259.964, **P =<0.001**
Age (years, mean [SD]) N=270 Missing data=0%=0	56.73±9.18 N*=0 128	59.07±9.77 N*=0 50	58.71±9.60 N*=0 92	ANOVA (F 2, 267) = [0.290], p = 0.748
BMI (median, IQR) N=124 Missing data=54.07%=146	27.40, 5.05 N*=76	27.50, 6.16 N*=27	27.50, 9.08 N*=43	ANOVA (F 2, 121) = [.995], p = 0.373
Glucose (median, IQR) N=228 Missing data=15.55%=42	4.60,0.63 N*=23	4.40, 0.85 N*=7	4.62, 0.79 N*=12	ANOVA (F (2, 225) = [.275], p = 0.760
Insulin (median, IQR) N=109 Missing data=59.62%=161	6.60, 18.30 N*=82	4.74, 8.06 N*=24	7.35, 8.33 N*=55	ANOVA (F (2, 106) = [0.598], p = 0.552
Smoking (Q9_63; median, IQR) N=260 Missing data=3.70%=10	11, 16 N*=3	12,18 N*=1	15, 18 N*=6	ANOVA (F (2, 257) = [3.547], **p = 0.030**

Table 4.13 Table showing variables matched to participants with <15% (Continued)

Metabolites <15% and three years cut off	Normal (0-7)	Borderline Abnormal (8-10)	Abnormal (11-21)	Test Statistics
Q10_151__Exercise _E1129_ N=234 Missing data=13.33%=36	48.,1% (110) N*=18	18.8% (44) N*=6	33.1% (80) N*=12	ANOVA (F (2, 231) = [0.697], p = 0.449
Q2_44_Alcohol N=255 Missing data=5.55%=15	49.4% (126) N*=2	17.3% (44) N*=6	33.3% (85) N*=7	ANOVA (F (2, 252) = [0.609], p = 0.545
P013630_Interleukin10 N==262 Missing data=2.96%=8	48.1% (126) N*=2	17.9% (47) N*=3	34.0% (89) N*=3	ANOVA (F (2, 259) = [1.576], p = 0.209
P013632_Interleukin1B N=265 Missing data=1.85%=5	47.9% (127) N*=1	18.5% (49) N*=1	33.6% (89) N*=3	ANOVA (F (2, 262) = [0.502], p = 0.606
Q11A_288 _E1129_ BP, N=248 Missing data=8.14%=22	46.8% (116) N*=12	18.5% (46) N*=4	34.7% (86) N*=6	ANOVA (F (2, 245) = [0.266], p = 0.766
Q11A_291 _E1129_ HC, N=248 Missing data=8.51%=23	41.8% (116) N*=12	18.5.7% (46) N*=4	34.7% (86) N*=6	ANOVA (F (2, 245) = [1.311], p = 0.271
Q11A_75__E1129_Income, N=248 Missing data=8.14%=22	46.8% (116) N*=12	18.5% (46) N*=4	34.7% (86) N*=6	ANOVA (F (2, 245) = [0.334], p = 0.716
Q11B_259_Oral1 N=138 Missing data=48.88%=132	42.0% (58) N*=70	20.3 (28) N*=22	37.7% (52) N*=40	ANOVA (F (2, 135) = [0.777], p = 0.462
Hormone _Q11B_260_ \Oral contraceptive pill\In total how many years N=132 Missing data=48.88%=138	40.9% (54) N*=74	20.5% (27) N*=23	38.6% (51) N*=41	ANOVA (F (2, 129) = [1.833], p = 0.156
Q17D_62 _Are you currently taking any medications/supplements Year N=245 Missing data=9.25%=25	46.8% (117) N*=11	18.4% (45) N*=5	33.9% (83) N*=9	ANOVA (F (2,242) = [0.099], p = 0.905

4.6 Two-year lipid cohort

Considering the information presented (Table 4.14), we present descriptive statistics for several factors associated with HADs (Anxiety and Depression Scale) ratings for the 2-year lipid cohort. The variables include Age, BMI, Glucose, Insulin, Depression, Anxiety, Smoking, Alcohol, Interleukin10, Interleukin1B, Exercise, HC_Age (high cholesterol diagnosis age), BP_Age (hypertension diagnosis age), Income, Supplement, Contra1, and Contra2. The summaries include the count, mean, standard deviation (SD), median, mode, maximum (Max.), minimum (Min.), and range for each variable for normal, borderline abnormal and normal. Table 3.15 shows variables matched to participants with <15% missing lipid data and at least one HADS score, in a two-year timespan between time of HADS score collection and blood sample collection.

Table 4.14 Two-year lipid cohort

HADs Type	Variable	Count	Mean	SD	Median	Mode	Max.	Min.	Range
	Age	43	57.39	8.81	59.16	41.64[a]	73.69	41.64	32.05
	BMI	43	26.91	6.04	26.50	19.19[a]	40.00	19.19	20.82
	Glucose	43	4.48	2.17	4.60	.00[a]	11.07	0.00	11.07
	Insulin	43	6.94	5.24	5.75	4.60	25.90	2.60	23.30
	Depression	43	2	2	2	2	5	0	5
	Anxiety	43	2	2	2	1	6	0	6
	Smoking	43	9	8	14	0	21	0	21
	Alcohol	43	8	10	5	0	41	0	41
Normal	Interleukin10	43	3.59	9.74	0.00	0.00	36.20	0.00	36.20
	Interleukin1B	43	0.56	1.57	0.00	0.00	6.62	0.00	6.62
	Exercise	43	3	1	3	3	4	2	2
	HC_Age	43	15	25	0	0	70	0	70
	BP_Age	43	14	25	0	0	69	0	69
	Income	43	3	2	3	0	8	0	8
	Supplement	43	3	3	2	0	10	0	10
	Contra1	43	1	0	1	1	1	0	1
	Contra2	43	11	5	10	10	19	1	18

Table 4.14 Two-year lipid cohort (Continued)

HADs Type	Variable	Count	Mean	SD	Median	Mode	Max.	Min.	Range
	Age	35	59.12	9.91	62.25	48.00[a]	74.80	41.63	33.17
	BMI	35	27.13	6.78	27.28	18.08[a]	38.53	18.08	20.45
	Glucose	35	3.12	2.24	4.17	0.00	5.95	0.00	5.95
	Insulin	35	7.27	4.66	6.05	2.20[a]	18.90	2.20	16.70
	Depression	35	4	2	4	2[a]	9	1	8
	Anxiety	35	5	2	5	5	8	1	7
	Smoking	35	12	8	16	0	23	0	23
	Alcohol	35	10	11	5	0	35	0	35
Borderline abnormal	Interleukin10	35	1.12	4.23	0.00	0.00	21.99	0.00	21.99
	Interleukin1B	35	0.20	0.72	0.00	0.00	3.46	0.00	3.46
	Exercise	35	2	1	2	3	3	1	2
	HC_Age	35	17	25	0	0	68	0	68
	BP_Age	35	16	26	0	0	70	0	70
	Income	35	4	2	5	5	8	0	8
	Supplement	35	5	8	2	0	30	0	30
	Contra1	35	1	0	1	1	1	0	1
	Contra2	35	9	8	7	4[a]	34	0	34
	Age	35	59.98	9.17	60.45	47.43[a]	73.69	41.63	32.07
	BMI	35	28.57	5.41	26.40	23.48[a]	39.46	23.48	15.97
	Glucose	35	3.32	2.17	4.20	0.00	7.63	0.00	7.63
	Insulin	35	8.16	5.33	6.40	2.30[a]	17.50	2.30	15.20
	Depression	35	8	3	7	7	14	3	11
	Anxiety	35	8	3	8	7	17	3	14
	Smoking	35	14	10	17	0	40	0	40
	Alcohol	35	13	13	9	0	45	0	45
Abnormal	Interleukin10	35	0.54	3.20	0.00	0.00	18.91	0.00	18.91
	Interleukin1B	35	0.04	0.23	0.00	0.00	1.37	0.00	1.37
	Exercise	35	2	1	2	2[a]	5	0	5
	HC_Age	35	24	28	0	0	73	0	73
	BP_Age	35	15	27	0	0	71	0	71
	Income	35	4	2	5	5	8	0	8
	Supplement	35	6	8	2	0	30	0	30
	Contra1	35	1	1	1	1	1	0	1

Table 4.14 Two-year lipid cohort (Continued)

HADs Type	Variable	Count	Mean	SD	Median	Mode	Max.	Min.	Range
	Contra2	35	11	8	10	10	29	0	29
	Age	113	58.73	9.26	59.89	41.63[a]	74.80	41.63	33.17
	BMI	113	27.47	5.99	26.56	18.08[a]	40.00	18.08	21.93
	Glucose	113	3.68	2.26	4.35	0.00	11.07	0.00	11.07
	Insulin	113	7.30	4.97	5.95	4.60	25.90	2.20	23.70
	Depression	113	4	3	4	2	14	0	14
	Anxiety	113	5	3	5	1[a]	17	0	17
	Smoking	113	12	9	15	0	40	0	40
	Alcohol	113	10	11	5	0	45	0	45
Total	Interleukin10	113	1.88	6.78	0.00	0.00	36.20	0.00	36.20
	Interleukin1B	113	0.29	1.07	0.00	0.00	6.62	0.00	6.62
	Exercise	113	2	1	2	3	5	0	5
	HC_Age	113	18	26	0	0	73	0	73
	BP_Age	113	15	26	0	0	71	0	71
	Income	113	3	2	4	5	8	0	8
	Supplement	113	5	7	2	0	30	0	30
	Contra1	113	1	0	1	1	1	0	1
	Contra2	113	10	7	10	10	34	0	34

a. Multiple modes exist. The smallest value is shown

The two-year lipid cohort participants' demographic and HADS score-related information is shown in Table 4.15. There are substantial variations between the three groups in the median values of the Depression and Anxiety HADS scores in this cohort of 113 distinct samples ($p = 0.001$). Additionally, there are notable disparities in income ($p = 0.023$), exercise habits ($p = 0.015$), and glucose levels ($p = 0.032$). Age, BMI, insulin, alcohol use, exercise habits, medication/supplement use, and oral contraceptive pill use are some of the demographic and health-related factors that did not show any discernible variations between these groups.

Table 4.15 Table showing variables matched to participants with <15% missing lipid data and at least one HADS score, in a two-year timespan between time of HADS score collection and blood sample collection

Lipids <15% and two-year cut-off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=113) Missing data=0%=0	43	35	35	χ2 test (2, N = 113) =2.000, P=0.3679
Depression HAD (median, IQR) N=113 Missing data=0%=0	2, 2	4, 2.5	7, 2	χ2 test (24, N = 113) = 91.367, **P<=0.001**
Anxiety HAD (median, IQR) N=113 Missing data=0%=0	2, 2	5, 2	8, 4	χ2 test (26, N = 113) = 101.583, **P<=0.001**
Age (years, mean [SD]) N=113 Missing data=0%=0	57.4±8.81	59.1±9.91	60.0±9.17	ANOVA (F (2, 110) = [0.798], p = 453
BMI (median, IQR) N=37 Missing data=67.25%=76	26.9, 6.04 N*= 29	27.10, 6.78 N*= 23	28.60, 5.41 N*= 24	ANOVA (F (2, 34) = [0.255], p = 0.777
Glucose (median, IQR) N=73 Missing data=35.39%=40	4.48, 2.17 N*=9	3.12, 2.24 N*=4	3.32, 2.17 N*=27	ANOVA (F (2, 88) = [3.590], p = **0.032**
Insulin (median, IQR) N=38 Missing data=66.37%=75	6.94, 5.24 N*=25	7.27, 4.66 N*=23	8.16, 5.33 N*=27	ANOVA (F (2, 35) = [.161], p = 0.852
Smoking (Q9_63; median, IQR) N=109 Missing data=3.5%=4	37.6% (41) N*=2	31.2% (34) N*=1	31.2% (34) N*=1	ANOVA (F (2, 106) = [2.617], p = .078
Q2_44_Alcohol (median, IQR) N=109 Missing data=6.19%=7	38.5% (42) N*=1	31.2% (34) N*=1	30.3% (33) N*=2	ANOVA (F (2, 106) = [1.143], p = 0.323
Exercise (median, IQR) N=40 Missing data=64.60%=73	42.5% (17) N*=26	30.0% (12) N*=23	27.5% (11) N*=24	ANOVA (F (2, 37) = [4.754], p = **0.015**

Table 4.15 Table showing variables matched to participants with <15% missing lipid data (continued)

Lipids <15% and two-year cut-off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Q10_150 (median, IQR) N=100 Missing data=11.50%=13	40.0% (40) N*=3	29.0% (29) N*=6	31.0% (31) N*=4	ANOVA (F (2, 97) = [0.425], p = 0.655
Exercise N=100 Missing data=11.50%=13	40.0% (40) N*=3	29.0% (29) N*=6	31.0% (31) N*=4	ANOVA (F (2, 97) = [2.483], p = 0.089
P013630_Interleukin10 (median, IQR) N==113 Missing data=0.0%=0	38.1% (43) N*=0	31.0% (35) N*=0	31.0% (35) N*=0	ANOVA (F (2, 110) = [1.536], p = 0.220
P013632_Interleukin1B (median, IQR) N==113 Missing data=0.0%=0	38.1% (43) N*=0	31.0% (35) N*=0	31.0% (35) N*=0	ANOVA (F (2, 110) = [1.987], p = 0.142
Q11A_288 _E1129_ BP, (median, IQR) N=91 Missing data=19.4%=22	41.8% (38) N*=5	28.6% (26) N*=9	29.7% (27) N*=8	ANOVA (F (2, 88) = [1.052], p = 0.354
Q11A_291 _E1129_ HC, (median, IQR) N=91 Missing data=19.4%=22	41.8% (38) N*=5	28.6% (26) N*=9	29.7% (27) N*=8	ANOVA (F (2, 88) = [0.078], p = 0.925
Q11A_75__E1129_Income, (median, IQR) N=91 Missing data=19.4%=22	41.8% (38) N*=5	28.6% (26) N*=9	29.7% (27) N*=8	ANOVA (F (2, 88) = [3.931], p = **0.023**
Q11B_257 (median, IQR) N=91 Missing data=19.4%=22	38.6% (35) N*=8	30.5% (29) N*=6	32.6% (31) N*=8	ANOVA (F (2, 92) = [0.704], p = 0.487
Hormone _Q11B_260_ \Oral contraceptive pill\In total how many years (median, IQR) N=57 Missing data=49.55%=56	37.2% (16) N*=27	32.6% (14) N*=21	30.2% (13) N*=8	ANOVA (F (2, 40) = [0.365], p = 0.696
Q17D_62 (median, IQR)Are you currently taking any medications/supplements N=102 Missing data=9.34%=11	39.2% (40) N*=3	28.4% (29) N*=6	32.4% (33) N*=2	ANOVA (F (2, 99) = [2.626], p = 0.077

4.7 Discussion

In the analysis of the 2-year metabolite cohort, several crucial factors were identified. Depression and anxiety severity, as measured by the HADS scale, displayed significant differences among the normal, borderline abnormal, and abnormal groups, highlighting the clinical relevance of these psychological conditions. Exercise behaviour was notably related to overall health, with the normal group reporting higher exercise participation, suggesting a potential protective effect of physical activity (Livingston et al., 2020). Alcohol consumption patterns differed significantly among the groups, with the normal group exhibiting higher alcohol use prevalence. Immune responses, indicated by interleukin 10 and interleukin 1B levels, demonstrated group differences, suggesting potential implications for immunity and inflammation in these psychological conditions (Herrstedt et al., 2019). However, certain parameters, such as the use of oral contraceptives and medications/supplements, did not exhibit significant differences between groups.

The co-occurrence of depression and anxiety is a well-documented phenomenon with significant clinical implications. This dual presentation often leads to more severe and prolonged depressive disorders, impaired functioning, and challenging prognoses (Lenze et al., 2000; Murphy et al., 2017). Despite numerous collaborative studies, the shared causes of these conditions remain partially unknown, necessitating a comprehensive clinical evaluation to understand the underlying pathophysiological mechanisms and, importantly, environmental and lifestyle factors, as such patients tend to respond less effectively to traditional antidepressant treatments (Lenze et al., 2000; Murphy et al., 2017). The studies such as ones by Graham et al show that the relationship varies depending on the way in which alcohol consumption is measured (frequency, one-time consumption volume, and risk of dependence), more so than by gender or depression measure (Graham et al., 2007).

When examining the lipid 2-year cohort, similar significant differences were observed among the three groups in terms of depression and anxiety severity. Lower glucose levels in the abnormal group may indicate potential metabolic differences, and higher physical activity levels in the normal group suggest a protective role of exercise (Livingston et al., 2020). Moreover, disparities in income levels, like the metabolite cohort, could reflect broader health inequities.

In conclusion, the co-occurrence of depression and anxiety is a complex phenomenon with significant clinical implications. The results of this study demonstrate the importance of considering various factors, including exercise behaviour, alcohol consumption, immune responses, and income, in individuals with abnormal HAD scores. The potential benefits of exercise in preserving cognitive abilities and reducing depressive symptoms highlight the importance of incorporating physical activity interventions in the management of concurrent depression and anxiety (Livingston et al., 2020; Herrstedt et al., 2019). Overall, this part of the project highlights the importance of including certain demographic and lifestyle variables for analysis, so that deeper analysis in the ensuing chapters could proceed, since HADS scores correlated with a vast number of these variables. This investigation has succeeded in identifying the best data to use for the analysis. Although the elimination of participants was drastic, it was necessary, to include enough participants who possessed data for metabolites, lipids, variables and HADS scores. The complex interplay between mental health, lifestyle, and notably physical activity, provides valuable insights for further clinical practice and investigation.

Chapter 5: Correlation of energy metabolic profiling and depression in TwinsUK database

5.1 Introduction

Understanding MDD remains a challenge despite advances in pathophysiology knowledge. Focus on neurotransmitters like serotonin and catecholamines overlooks the broader role of metabolic products in various biochemical processes linked to MDD. Glutamate metabolism disruptions are strongly linked to depression. Dietary glutamate might influence blood-brain barrier function, gut health, and mood disturbances. Kynurenine's involvement in MDD centers on its shift toward generating neurotoxic rather than neuroprotective compounds, potentially worsening depressive symptoms. Amino acids, the precursors for neurotransmitters like serotonin and norepinephrine, are implicated in MDD. Lower tryptophan and tyrosine levels are observed in depressed individuals, and supplementing these amino acids could alleviate depressive symptoms. Metabolites, crucial in biochemical reactions, are linked to depression incidence. Disturbances in levels of amino acids and other compounds like kynurenine and acylcarnitine are observed in MDD. Lifestyle habits, medication use, and metabolic disorders impact these associations, requiring further exploration.

For detailed information relating to these variables please see Chapter one, Section 1.5.

5.1.1 The importance of the TwinsUK database in offering a fresh perspective on metabolomics and severity of MDD

Some metabolites associated with depression are well-established, however the TwinsUK database amassed a much wider range of metabolite data from thousands of participants, including protein metabolites such as amino acids, and metabolites from gluconeogenesis and glycolysis involved in biological pathways that are not directly linked to the brain. The main metabolites of interest were those involved in the Cahill cycle, the citric acid cycle or glucose metabolism generally, since levels of alanine and citrate, associated with these biological pathways, were elevated in patients with major depressive disorder from the systematic review and meta-analysis (Chapter 2). This finding is further supported by more recent research highlighting the link between major depressive disorder and neuroendocrine metabolism, showing the interaction between depressive pathology and impaired metabolism inherent in various metabolic diseases such as obesity, diabetes, and hypertension (Stuart & Baune, 2012). Characteristic traits of metabolic impairment such as dyslipidemia, hypertension, high C-Reactive Protein and insulin resistance have all been shown to be linked with increased risk of depression (Qiu et al., 2021). It is therefore important to examine whether these findings could be reproduced in a much larger population in the TwinsUK database and identify patterns and profiles of metabolites and factors that were previously not found in literature relating to their role in severity of major depressive disorder.

This chapter will investigate the relationship between depression severity and levels of certain metabolites, using statistical methods in R and SPSS such as linear regression, for both univariate and multivariate analysis.

5.2 Results

5.2.1 Identification of metabolomics included in the study

In our extensive metabolite investigation, we carefully examined 19 metabolites. The metabolites analysed are Pyr, Cit, Ala, Glc, Gln, Gly, His, Ile, Leu, Val, Phe, Tyr, Ace, AcAce, BOHBut, Crea, Alb, Gp, Lac, and Glol. Each metabolite, with its own chemical makeup and physiological significance, contributes to the rich web of metabolic interactions inside the biological system being studied. Table 5.1 shows the results of the correlation analysis conducted in SPSS. The assessment compares two sets of variables, "Medicated" and "Non-Medicated," and calculates the Pearson correlation coefficient, p-values, and 95% confidence intervals for each pair of variables. The table is organized into several columns. "Pairs" lists the pairs of variables being compared. "Pearson Correlation" displays the Pearson correlation coefficient, which measures the strength and direction of the linear relationship between the variables. "P-Value" indicates the statistical significance of the correlation. "95% Confidence Intervals" provides the lower and upper bounds of the 95% confidence interval for the correlation coefficient, giving an estimate of the range within which, the true correlation value is likely to fall.

The correlation between the variables "HADs" and "Gly" demonstrates the most significant disparity between the medicated and non-medicated groups (Table 5.1). With a value of r=-0.19, the value of the correlation coefficient between "HADs" and "Gly" in the Medicated category is negative and non-significant. However, the value of the correlation coefficient between "HADs" and "Gly" is positive and statistically significant; r = 0.324 p-value of 0.039 in the non-medicated category, which indicates that the correlation meets the criteria for statistical significance at the traditional level of 0.05. The table continues with similar information for all the pairs of variables analysed.

Table 5.1 Pearson correlations and significance values for HADS scores against each metabolite subclass for the two-year cohort, after stratifying for medication status

	Medicated				Non-Medicated			
Pairs	Pearson Correlation	P-Value	95% Confidence Intervals		Pearson Correlation	P-Value	95% Confidence Intervals	
			Lower	Upper			Lower	Upper
HADs - Pyr	0.114	0.382	-0.142	0.355	0.218	0.171	-0.096	0.493
HADs - Cit	0.002	0.99	-0.25	0.253	-0.042	0.796	-0.345	0.27
HADs - Ala	0.064	0.623	-0.191	0.311	0.249	0.116	-0.063	0.517
HADs - Glc	-0.084	0.52	-0.329	0.171	0.131	0.413	-0.184	0.422
HADs - Gln	0.014	0.918	-0.239	0.264	0.114	0.478	-0.201	0.407
HADs - Gly	-0.19	0.142	-0.422	0.065	0.324	**0.039**	0.018	0.574
HADs - His	-0.147	0.259	-0.384	0.109	-0.031	0.845	-0.336	0.279
HADs - Ile	0.102	0.436	-0.154	0.345	-0.191	0.231	-0.471	0.124
HADs - Leu	0.133	0.307	-0.123	0.372	-0.144	0.371	-0.432	0.172
HADs - Val	0.096	0.463	-0.16	0.339	-0.152	0.344	-0.439	0.164
HADs - Phe	0.079	0.545	-0.176	0.324	0.136	0.398	-0.18	0.425
HADs - Tyr	0.115	0.378	-0.141	0.356	0.238	0.134	-0.075	0.508
HADs - Ace	-0.058	0.657	-0.305	0.197	-0.295	0.061	-0.553	0.014
HADs - AcAce	-0.122	0.347	-0.363	0.133	-0.218	0.171	-0.493	0.096
HADs - BOHBut	0.097	0.456	-0.159	0.341	-0.282	0.074	-0.543	0.028
HADs - Crea	-0.076	0.561	-0.322	0.179	-0.019	0.907	-0.325	0.29
HADs - Alb	0.112	0.389	-0.144	0.354	-0.185	0.248	-0.466	0.13
HADs - Gp	0.217	0.093	-0.037	0.444	-0.132	0.409	-0.423	0.183
HADs - Lac	0.13	0.318	-0.126	0.37	-0.086	0.594	-0.383	0.228
HADs - Glol	0.11	0.398	-0.146	0.352	-0.111	0.489	-0.405	0.203

5.2.1 Correlation matrix for variables in the two-year metabolite cohort

Age and HADs had a positive but non-significant association, with a Pearson coefficient of 0.118. There was a positive and statistically significant relationship between age and all the other variables such as Interleukin10 (r = -0.191, p = 0.043), Interleukin1B (r = -0.204, p = 0.030), and Glucose (r = 0.221, p = 0.036), See Fig. 5.1. The variables ‘Depression’ and ‘Anxiety’ had stronger positive associations (r=0.754 and r=0.755, respectively) with HADs. There was a statistically significant inverse relationship between glucose and HADs (r = -0.221, p = 0.036). Among those with HADs, smoking was significantly positively associated (r = 0.211, p = 0.027). HADs was positively and significantly correlated with alcohol use (r = 0.194, p = 0.040). The inverse relationship between interleukin-10 and HADs was statistically significant (r = -0.191, p = 0.043), and likewise for interleukin1B (r = -0.204, p = 0.030). To a statistically significant degree (r = -0.207, p = 0.039), HADs was inversely

related to exercising. When comparing HADs and supplement taking, the latter showed a strong positive connection (r = 0.219, p = 0.027). For more details, please look at Fig 5.1.

Age and Insulin had positive associations (r = 0.598, p = 0.0001), while Insulin and Glucose had negative correlations (r = -0.639, p = 0.047). Cigarette smoking was positively associated with increasing age (r = 0.211, p = 0.027). Alcohol use increased with age in a statistically meaningful way (r = 0.194, p = 0.040). The level of interleukin-10 and interleukin-1B declined with age (r = -0.191, p = 0.043), (r = -0.204, p = 0.030). Inversely and significantly correlated with age (r = -0.27, p = 0.039) and marginally positively correlated with insulin (r = 0.320, p = 0.061) is exercise. There is a positive but non-significant connection between HC_Age (High Cholesterol) and both Age and Insulin (r = 0.156, p = 0.419; r = 0.126, p = 0.322). Aging (r = 0.219, p = 0.027), Insulin (r = 0.419, p = 0.010), and Mood (r = 0.251, p = 0.011) were all significantly positively correlated with supplement usage as well. For more details, refer to the Appendix-5.

5.2.2 Correlation plot for metabolites

The correlation coefficient measures the strength and direction of the linear relationship between two variables. A value close to 1 indicates a strong positive correlation, a value close to -1 indicates a strong negative correlation, and a value close to 0 suggests a weak or no linear relationship. In this case, the correlation between HADs and Pyr is 0.154, indicating a weak positive correlation (Fig. 5.1). However, the p-value of 0.103 suggests that this correlation may not be statistically significant. The correlation between HADs and Cit is -0.006, indicating a very weak negative correlation. Additionally, the high p-value of 0.952 suggests that there is likely no statistically significant correlation between HADs and Cit. Thus, we did find any significant relationship between HADs and all the metabolites.

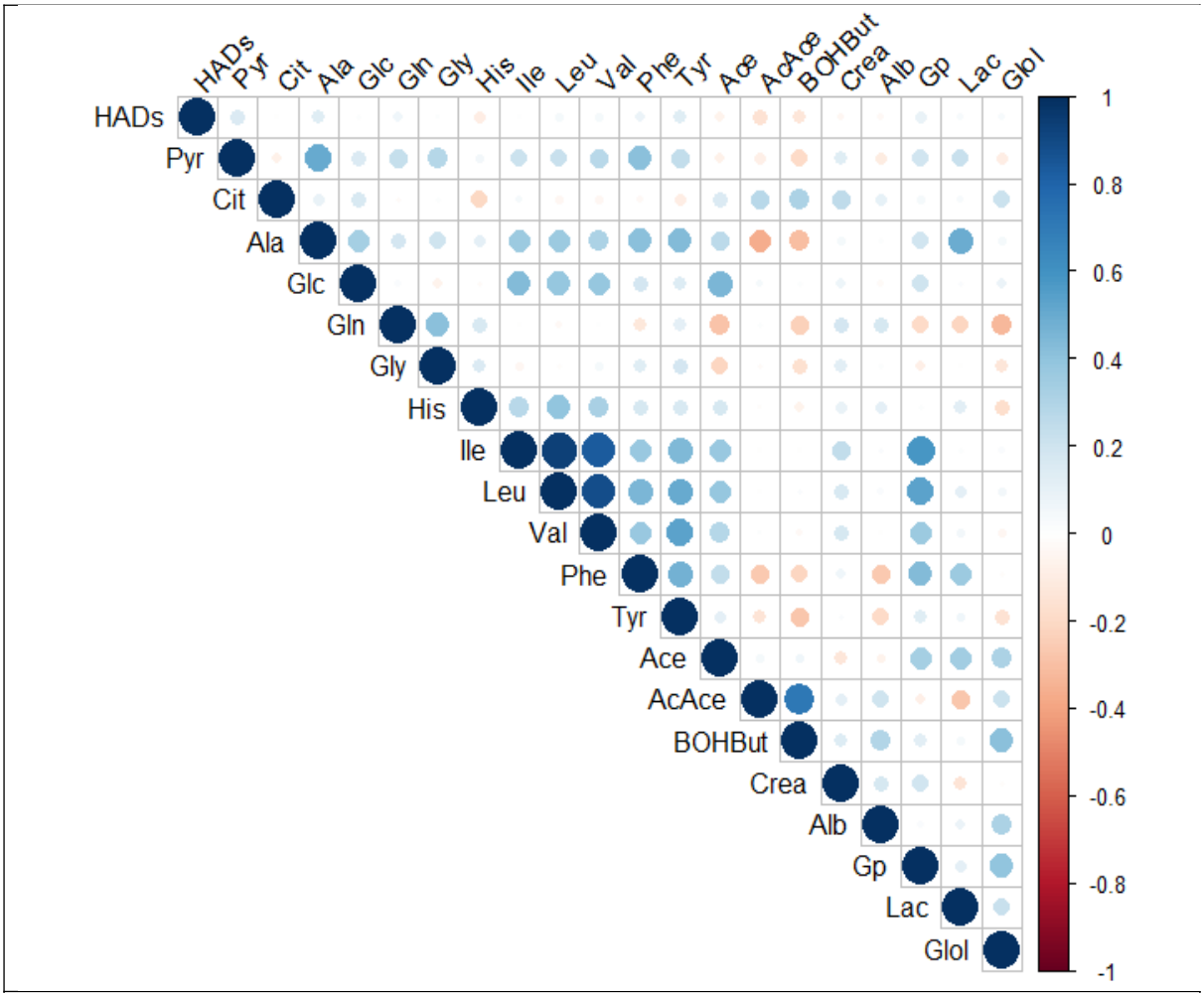


Fig 5.1 Correlation plot for metabolites in the two-year cohort

5.2.3 Identification of confounding variables based on medication status

In light of the information that was made available, the following table (Table 5.2) is a statistical summary of the variables, which includes the medication status, the number of observations (N), the mean, the standard deviation (StDev), the median, the maximum (Max.), the minimum (Min.), the interquartile range (IQR), the t-value, the degrees of freedom (DF), and the p-value. Based on these findings, it appeared that there was a statistically significant distinction between the "Medicated" and "Non-Medicated" groups for the variables age, glucose, smoking, exercise, age when hypertension was first diagnosed (BP_Age), and age when high cholesterol was first diagnosed (HC_Age). However, there was no significant difference between the groups for the variable’s alcohol, Interleukin10, Interleukin1B, or income.

Table 5.2: Significant variables stratifying for medication

Variable	Medication	N	Mean	StDev	Median	Max.	Min.	IQR	T-Value	DF	P-Value
Age	Medicated	66	60.0	9.7	60.9	73.7	41.6	15.1	2.78	106	**0.006**.
	Non-Medicated	47	55.3	8.3	55.6	74.2	41.6	13.9			
	Total	113									
Glucose	Medicated	59	5.0	1.5	4.6	12.9	3.6	1.0	2.04	88	**0.044**
	Non-Medicated	45	4.5	0.8	4.4	7.8	2.6	0.4			
	Total	104									
Smoking	Medicated	63	12.5	9.6	16.0	40.0	0.0	18.0	2.61	103	**0.01**
	Non-Medicated	46	8.0	8.3	5.5	22.0	0.0	16.0			
	Total	109									
Alcohol	Medicated	64	11.2	11.8	6.0	41.0	0.0	19.0	1.38	97	0.171
	Non-Medicated	43	8.2	10.4	4.0	45.0	0.0	9.0			
	Total	107									
Exercise	Medicated	59	2.3	0.9	2.0	5.0	0.0	1.0	-2.26	97	**0.026**.
	Non Medicated	42	2.6	0.5	3.0	4.0	2.0	1.0			
	Total	101									
Interleukin10	Medicated	65	2.5	8.4	0.0	40.5	0.0	0.0	1.35	99	0.179
	Non Medicated	47	0.8	4.2	0.0	27.8	0.0	0.0			
	Total	112									
Interleukin1B	Medicated	65	0.3	1.1	0.0	6.6	0.0	0.0	0.52	108	0.601
	Non Medicated	47	0.2	0.9	0.0	5.8	0.0	0.0			
	Total	112									
BP_Age	Medicated	57	26.3	27.3	31.0	73.0	0.0	50.0	4.23	90	**0.000**
	Non Medicated	36	6.6	17.6	0.0	62.0	0.0	0.0			
	Total	93									
HC_Age	Medicated	57	20.0	27.9	0.0	71.0	0.0	51.5	3.18	90	**0.002**
	Non Medicated	36	5.1	17.3	0.0	63.0	0.0	0.0			
	Total	93									
Income	Medicated	57	3.5	2.2	4.0	8.0	0.0	3.0	-0.4	65	0.688
	Non Medicated	36	3.7	2.6	5.0	8.0	0.0	5.5			
	Total	93									
Pyr	Medicated	66	0.0	0.0	0.0	0.1	0.0	0.0	1.2	74	0.236
	Non Medicated	47	0.0	0.0	0.0	0.1	0.0	0.0			
	Total	113									
Cit	Medicated	66	0.1	0.0	0.1	0.2	0.1	0.0	0.04	105	0.972
	Non Medicated	47	0.1	0.0	0.1	0.2	0.1	0.0			
	Total	113									
Ala	Medicated	66	0.4	0.1	0.4	0.6	0.3	0.1	0.77	92	0.444
	Non Medicated	47	0.4	0.1	0.4	0.6	0.3	0.1			
	Total	113									
Glc	Medicated	66	4.5	1.7	4.0	13.2	2.8	1.2	3.06	85	**0.003**
	Non Medicated	47	3.8	0.6	3.7	5.2	2.7	0.9			
	Total	113									

Table 5.2: Significant variables stratifying for medication (continued)

Variable	Medication	N	Mean	StDev	Median	Max.	Min.	IQR	T-Value	DF	P-Value
Gln	Medicated	66	0.4	0.1	0.4	0.5	0.2	0.1	0.52	100	0.602
	Non Medicated	47	0.4	0.1	0.4	0.5	0.2	0.1			
	Total	113									
Gly	Medicated	66	0.3	0.1	0.3	0.5	0.2	0.1	-0.94	94	0.349
	Non Medicated	47	0.3	0.1	0.3	0.4	0.2	0.1			
	Total	113									
His	Medicated	66	0.1	0.0	0.1	0.1	0.0	0.0	-0.43	101	0.671
	Non Medicated	47	0.1	0.0	0.1	0.1	0.0	0.0			
	Total	113									
Ile	Medicated	66	0.1	0.0	0.1	0.1	0.0	0.0	1.2	100	0.234
	Non Medicated	47	0.1	0.0	0.0	0.1	0.0	0.0			
	Total	113									
Leu	Medicated	66	0.1	0.0	0.1	0.1	0.0	0.0	0.54	106	0.594
	Non Medicated	47	0.1	0.0	0.1	0.1	0.1	0.0			
	Total	113									
Val	Medicated	66	0.2	0.0	0.2	0.2	0.1	0.1	1.69	106	0.094
	Non Medicated	47	0.2	0.0	0.1	0.3	0.1	0.0			
	Total	113									
Phe	Medicated	66	0.1	0.0	0.1	0.1	0.1	0.0	0.22	103	0.825
	Non Medicated	47	0.1	0.0	0.1	0.1	0.1	0.0			
	Total	113									
Tyr	Medicated	66	0.1	0.0	0.1	0.1	0.0	0.0	0.38	85	0.709
	Non Medicated	47	0.1	0.0	0.1	0.1	0.0	0.0			
	Total	113									
Ace	Medicated	66	0.1	0.0	0.1	0.1	0.0	0.0	2.84	110	**0.005**
	Non Medicated	47	0.1	0.0	0.1	0.1	0.0	0.0			
	Total	113									
AcAce	Medicated	66	0.1	0.0	0.0	0.1	0.0	0.0	-0.29	68	0.774
	Non Medicated	47	0.1	0.0	0.0	0.3	0.0	0.0			
	Total	113									
BOHBut	Medicated	66	0.1	0.0	0.1	0.2	0.1	0.1	-1.04	72	0.303
	Non Medicated	47	0.1	0.1	0.1	0.4	0.1	0.0			
	Total	113									
Crea	Medicated	66	0.1	0.0	0.1	0.1	0.0	0.0	-0.11	94	0.916
	Non Medicated	47	0.1	0.0	0.1	0.1	0.0	0.0			
	Total	113									
Alb	Medicated	66	0.1	0.0	0.1	0.1	0.1	0.0	-0.98	90	0.33
	Non Medicated	47	0.1	0.0	0.1	0.1	0.1	0.0			
	Total	113									
Gp	Medicated	66	1.3	0.2	1.2	2.0	0.9	0.3	1.23	99	0.222
	Non Medicated	47	1.2	0.2	1.1	1.8	0.9	0.2			
	Total	113									
Lac	Medicated	66	1.8	0.7	1.7	4.3	0.5	0.7	0.26	107	0.793
	Non Medicated	47	1.8	0.6	1.7	3.2	0.8	0.9			
	Total	113									
Glol	Medicated	66	0.1	0.0	0.1	0.1	0.0	0.0	1.44	106	0.154
	Non Medicated	47	0.1	0.0	0.1	0.1	0.0	0.0			

	Total	113

5.2.4 Multiple regression analysis for two-year metabolite cohort

The following output represents the results of a linear regression model where the dependent variable is HADs, and the independent variables are Smoking, Alcohol, Exercise, Interleukin1B, Supplement, Pyr, Ala, Tyr, and Ace. The R-squared value of 0.2912 indicates that the independent variables explain approximately 29.12% of the variance in the dependent variable HADs, or depression severity (Table 5.3). The F-statistic of 4.702 with its associated p-value is 3.088e-05, suggesting that the overall model is statistically significant. The estimated intercept is 2.916712, indicating the expected value of HADs when all independent variables are set to zero. The coefficient for Smoking was 0.012250, suggesting a positive relationship between smoking and HADs. However, it was not statistically significant ($p = 0.142128$). The coefficient for Alcohol was 0.016255, indicating a positive relationship between alcohol consumption and depression severity, which was statistically significant ($p = 0.019261$). The coefficient for Exercise was -0.375524, indicating a negative relationship between exercise and HADs and was also statistically significant ($p = 0.000379$). The coefficient for Interleukin1B was -0.203526, indicating a negative relationship between Interleukin1B and HADs. was statistically significant ($p = 0.005195$). The coefficients for Supplement, Pyr, Ala, Tyr, and Ace represented the relationships between these variables and HADs, all of them were statistically significant at $P<0.01$. The summary of correlation variables is given in Table 5.4

Table 5.3: Output for multiple linear regression analysis for 2-year metabolite cut-off group

Call: lm (formula=HADS~Smoking+Alcohol+Exercise+Interleukin1B+Supplement+Pyr+Ala+Tyr+Ace)

Coefficients	**Estimate**	**Std. Error**	**t-Value**	**Pr (>|t|)**
Intercept	2.916712	0.622028	4.6890	8.45E-06
Smoking	0.012250	0.008282	1.4790	0.142128
Alcohol	0.016255	0.006836	2.3780	0.019261 *
Exercise	-0.375524	0.102168	-3.6760	0.000379 ***
Interleukin1B	-0.203526	0.071271	-2.8560	0.005195 **
Supplement	0.016801	0.011165	1.5050	0.135461
Pyr	8.799149	5.842190	1.5060	0.135092
Ala	-2.101494	1.404237	-1.4970	0.137571
Tyr	12.423141	6.559915	1.8940	0.061058
Ace	-7.996821	4.134438	-1.9340	0.055833
Residual Std Error on 103 df:	0.7589			
Multiple R2	0.2912			
Adj. R2	0.2293			
F-Statistic: on 9 and 103 DF	4.702			
P-value	3.09E-05			

Table 5.4 Summary of correlation variables

Pair	Pearson Correlation	P-Value	95% Confidence Intervals	
			Lower	Upper
HADs - Age	0.118	0.215	-0.069	0.296
HADs - BMI	0.112	0.511	-0.22	0.42
HADs - Glucose	-0.221	**0.036**	-0.408	-0.015
HADs - Insulin	0.092	0.584	-0.235	0.4
HADs - Smoking	0.211	**0.027**	0.024	0.384
HADs - Alcohol	0.194	**0.04**	0.009	0.365
HADs - Interleukin10	-0.191	**0.043**	-0.363	-0.006
HADs - Interleukin1B	-0.204	**0.03**	-0.374	-0.02
HADs - Exercise	-0.207	**0.039**	-0.387	-0.011
HADs - BP_Age	0.144	0.174	-0.064	0.34
HADs - HC_Age	-0.025	0.815	-0.23	0.182
HADs - Income	0.201	0.057	-0.006	0.39
HADs - Supplement	0.219	**0.027**	0.025	0.396
HADs - Medicated	0.147	0.141	-0.049	0.332
HADs - Depression	0.754	**<0.01**	0.662	0.824
HADs - Anxiety	0.755	**<0.01**	0.663	0.825
HADs - Pyr	0.154	0.103	-0.032	0.329
HADs - Cit	-0.006	0.952	-0.19	0.179
HADs - Ala	0.138	0.146	-0.048	0.314
HADs - Glc	0.016	0.866	-0.169	0.2
HADs - Gln	0.066	0.486	-0.12	0.248
HADs - Gly	0.011	0.906	-0.174	0.196
HADs - His	-0.096	0.313	-0.276	0.09
HADs - Ile	0.002	0.983	-0.183	0.187
HADs - Leu	0.046	0.631	-0.14	0.228
HADs - Val	0.043	0.652	-0.143	0.226
HADs - Phe	0.084	0.376	-0.102	0.265
HADs - Tyr	0.135	0.153	-0.051	0.312
HADs - Ace	-0.067	0.484	-0.248	0.12
HADs - AcAce	-0.153	0.105	-0.329	0.032
HADs - BOHBut	-0.131	0.167	-0.308	0.055
HADs - Crea	-0.041	0.663	-0.224	0.144
HADs - Alb	-0.034	0.721	-0.217	0.152
HADs - Gp	0.096	0.311	-0.09	0.276
HADs - Lac	0.038	0.692	-0.148	0.221
HADs - Glol	0.032	0.734	-0.153	0.216

Pyruvate and Tyrosine were found to be positively associated with depression while Alanine and Acetate were negatively associated with depression, although none of these associations were significant. The linearity and homoscedasticity assumptions of a linear regression model were tested by plotting the residuals against the fitted values (Fig 5.3). This is useful for illustrating how the residuals (the gaps between the observed and predicted values) relate to the fitted values (the predicted values). The assumption of linearity was fulfilled since the points lay on a horizontal line centered at zero. Homoscedasticity (constant variance) is suggested by the fact that the residual distribution remains constant over the predicted value range (-1.5 to 2.0). The normality of a dataset can be visually evaluated with a Normal Q-Q (Quantile-Quartile) plot (Fig 5.3). It does this by contrasting the data's observed quantiles with the expected quantiles based on a normal distribution. As the data is normally distributed, the points on the graph cluster in a straight line. One way to visualize which observations matter more in a regression study is via a Residuals vs. Leverage plot (Fig 5.2). It aids in the identification of outlier observations that significantly alter the calculated regression coefficients. Examining points with high leverage values and/or big standardized residuals in the Residuals vs Leverage plot reveals influential observations. These data points may have a substantial effect on the final regression outcomes; more analysis is warranted to ascertain their relevance to the model. It is clear that observations 45, 52, and 105 are outliers, however their presence has no appreciable effect on the results of the regression analysis.

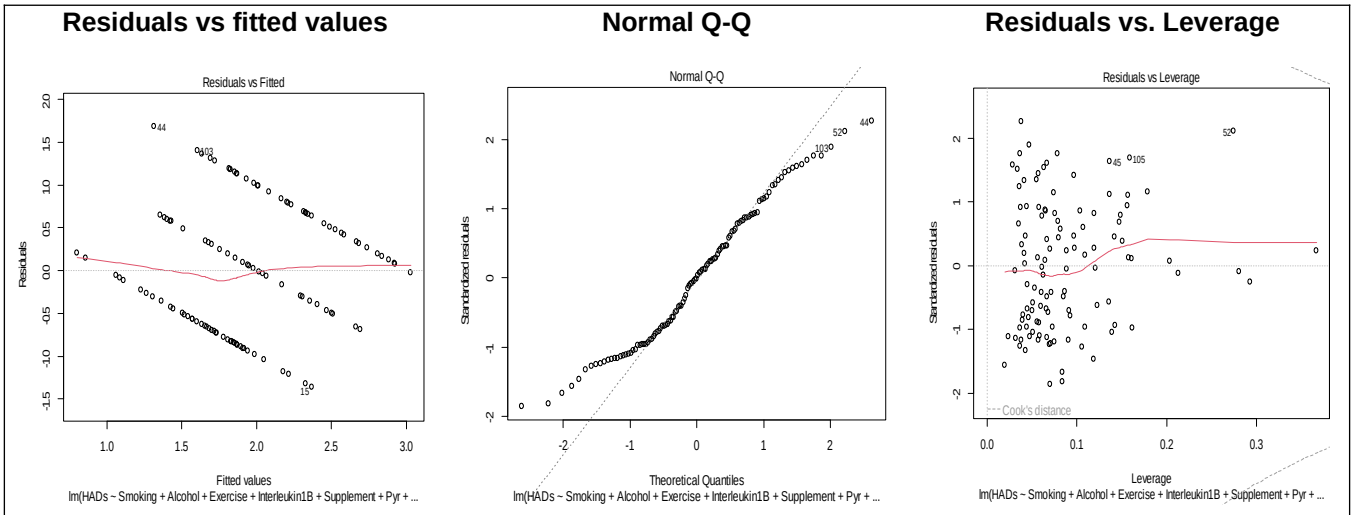


Fig 5.2 Normality charts of multiple regression for two-year metabolite cohort

5.2.5 Background and aims of using OPLS-DA

The application of Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) in this investigation emerges with a resolute aim: the contribution of specific variables, metabolites and lipids to the differentiation between depression subtypes based on Hospital Anxiety and Depression Scale (HADS) scores (Kong et al., 2023; Zhang et al., 2019). Central to the OPLS is its proficiency in dissecting the systematic variations encompassed within a set of input variables (X) into two distinct components. The group variation is separated into two parts: one with a linear relationship to the response variables (Y) and one that is orthogonal to or perpendicular to Y and has no influence on it (Frazier et al., 2023). This unique split, articulated by Vajargah et al. (2014) and Eriksson et al. (2006), finds its inception in the pre-processing stage. This phase involves the centralization of means and the scaling of variance within the input data. The reasoning behind this is based on the likelihood that specific model parameters could be affected by fluctuations in the magnitudes of the variables (Simić et al., 2023). As a result, the OPLS model excels at separating out systematic changes in X that have no association with the response data Y, enriching the subsequent insights derived from it (Kong et al., 2023). It is crucial to look at the VIP scores (Zhang et al., 2023) for significant contributing variables after confirming the model's predictive capability and to evaluate the AUC curve (Rozali et al., 2023) for information on model accuracy and classification performance. VIP scores and AUC values give a thorough understanding of variable relevance and the model's capacity for class discrimination.

5.2.5.1 Metabolite-OPLSDA

The metabolite-OPLSDA model was employed to analyze a dataset comprising 78 samples characterized by 19 predictor variables and one response variable. The model's performance is summarized below:

The cumulative R-squared ($R^2X_{(Cum)}$) value for predictor variables (X) was 0.332, indicating that the model explained approximately 33.2% of the variance in predictor variables (Table 5.5). The cumulative R-squared ($R^2Y_{(Cum)}$) value for the response variable (Y) was 0.094, suggesting that the model explained approximately 9.4% of the variance in the response variable. The cumulative Q-squared ($Q^2_{(Cum)}$) value in cross-validation was -0.269. A negative Q^2 value suggests that the model's predictive ability is poor. RMSEE: The Root Mean Squared Error of Estimation (RMSEE) was 0.483, representing the average error of the model's predictions on the training data. The model utilized one predictive component (pre). The model utilized one orthogonal component (ort). The cumulative predictive R-squared value (pR^2Y) was 1.05. The cumulative Q-squared value (pQ^2) was 0.75, representing the percentage of variance in the response variable explained by the model's predictive component(s). The dataset was subjected to standard scaling of predictors and response(s) to ensure consistency in data analysis. Overall, while the model demonstrated relatively good explanatory power for predictor variables (R^2X(cum)), the predictive performance (Q^2(cum)) was poor, as indicated by the negative Q^2 value. The overall framework of the model is illustrated in Fig 5.3

Table 5.5 OPLS-DA summary statistics

R^2X(cum)	R^2Y(cum)	Q^2(cum)	RMSEE	pre	ort	pR^2Y	pQ^2
0.332	0.094	-0.269	0.483	1	1	1.05	0.75

78 samples x 19 variables and 1 response
standard scaling of predictors and response(s)

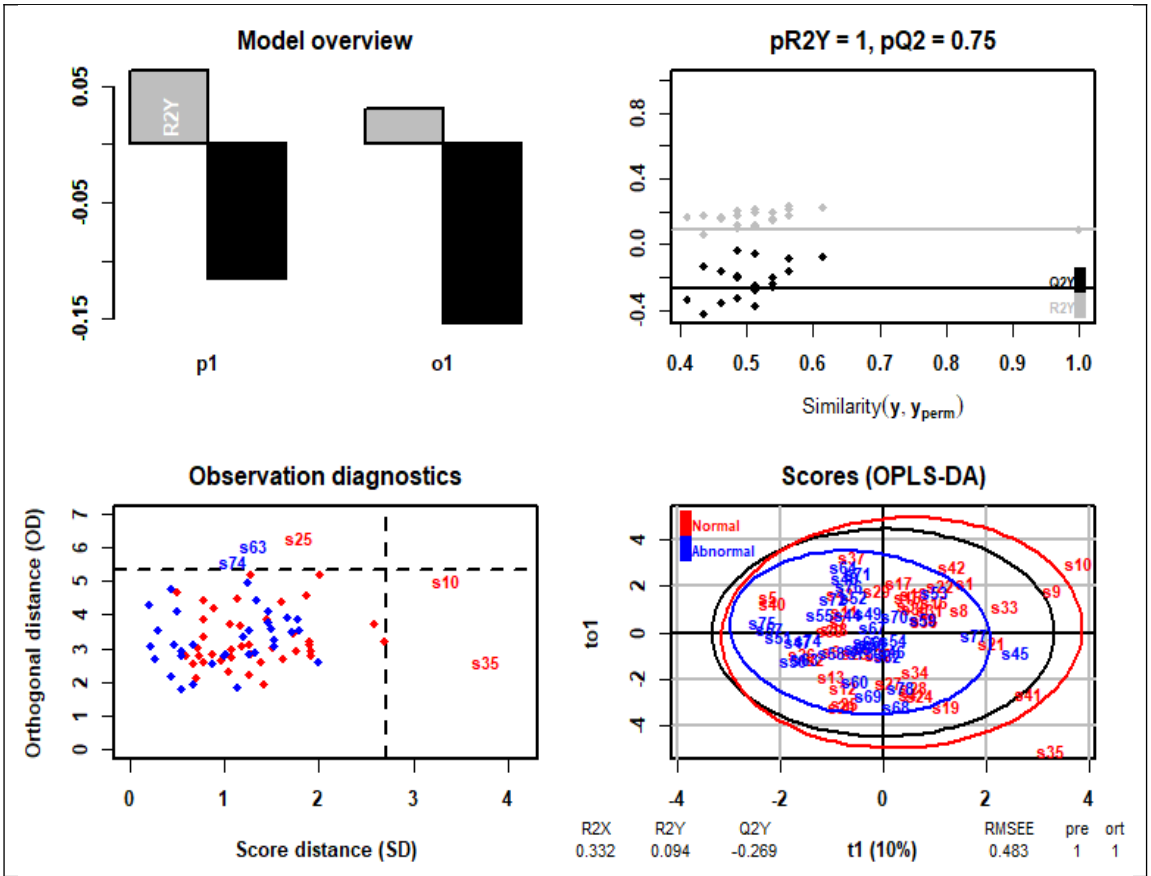


Fig 5.3 OPLS-DA Model overview, similarity index, observation diagnostics and scores for pattern analysis

5.2.5.2 Significance Spectrum for Normal versus Abnormal

Table 5.6, titled "Significance across Spectrum of Metrics," provides a comprehensive comparison between the model and random chance across various performance metrics. The model demonstrates an accuracy (Acc) of 0.42, indicating that it correctly predicts outcomes for approximately 42% of cases, with a relatively low standard deviation of 0.06, implying consistency in its predictive accuracy. Specificity (Spec) and Sensitivity (Sens) stand at 0.44 and 0.43, respectively, indicating the model's ability to correctly identify negative and positive instances, with consistent performance indicated by their standard deviations. The model exhibits a notably higher cumulative R-squared for predictor variables (R2X) at 0.49, suggesting that it effectively explains approximately 49% of the variance in the predictor space, with a remarkably low standard deviation of 0.01, underlining its robustness in capturing variance. However, the cumulative R-squared for the response variable (R2Y) is comparatively lower at 0.18, albeit with a modest standard deviation of 0.03, indicating a lesser degree of explained variance in the response space. For more details see at Fig. 5.4.

Table 5.6 Significance across Spectrum of Metrics

	Acc	SD	Spec	SD	Sens	SD	R2X	SD	R2Y	SD	Q2	SD
model	0.42	0.06	0.44	0.08	0.43	0.08	0.49	0.01	0.18	0.03	-0.67	0.18
random	0.50	0.07	0.49	0.13	0.51	0.14	0.30	0.03	0.22	0.05	-0.25	0.14

Accuracy	Q^2 Cum

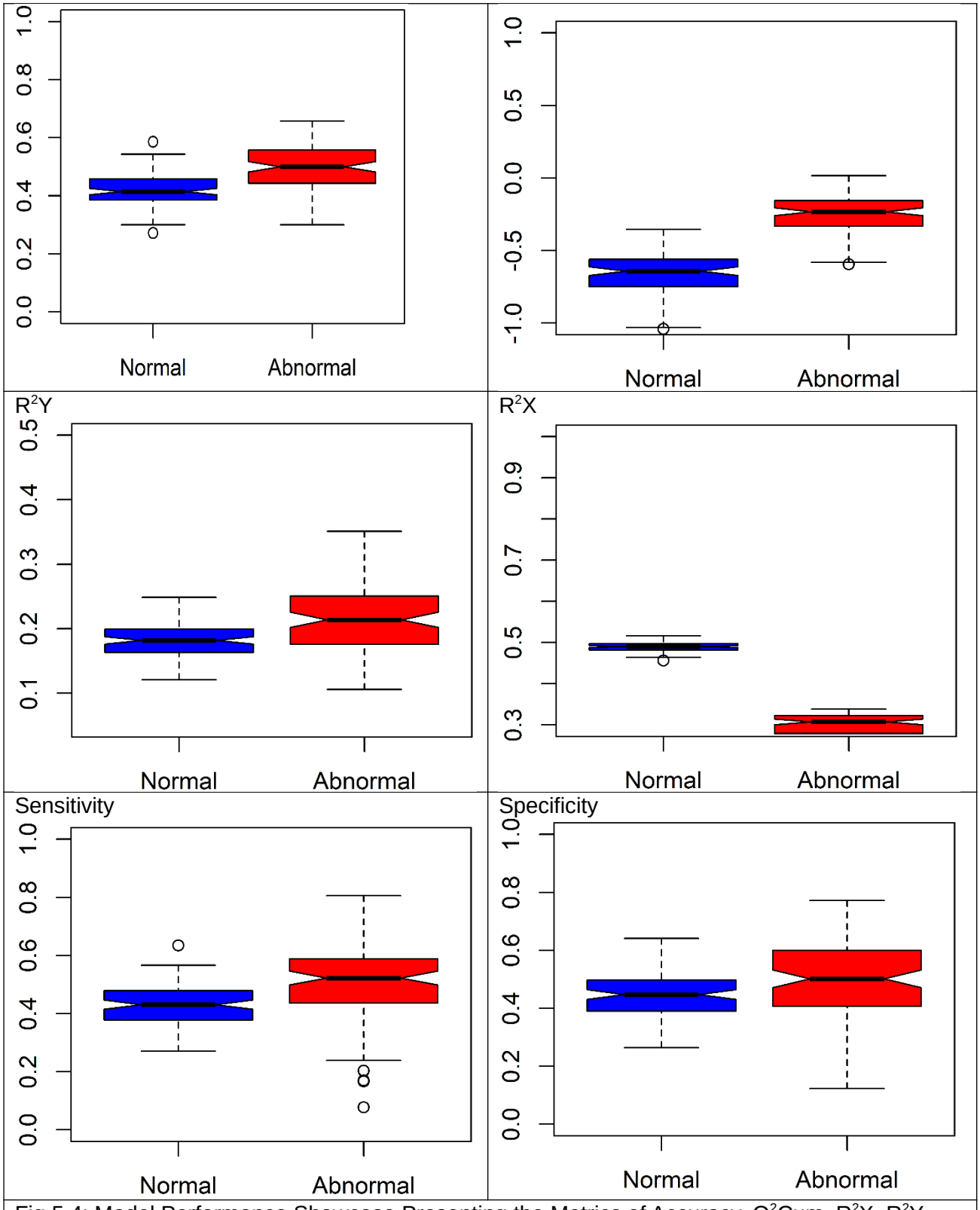


Fig 5.4: Model Performance Showcase Presenting the Metrics of Accuracy, Q^2Cum, R^2X, R^2Y, Sensitivity, Specificity and Component Distribution

5.2.5.3 Model Performance Indices

Table 5.7 presents the VIP scores as analytical guides and significance for distinguishing between Normal and Abnormal states. The table specifically outlines model performance indices derived from two distinct tests: the t-test and the ks-test. For the t-test, all performance metrics including accuracy (Acc), specificity (Spec), and sensitivity (Sens) achieve perfect scores of 1.00. Overall, Table 7.3 provides a comprehensive overview of model performance indices, showcasing the efficacy of the model in distinguishing between Normal and Abnormal states, as well as its predictive reliability under different analytical approaches.

Table 5.7: Model Performance Indices

	Acc	Spec	Sens	R2X	R2Y	Q2
t-test	1.00	1.00	1.00	0.00	1.00	1.00
ks-test	1.00	0.44	0.61	0.00	0.91	1.00

5.2.5.4 VIP Scores as Analytical Guides and Significance for Normal versus Abnormal

In section 5.2.7.4, VIP scores serve as analytical guides and indicators of significance for distinguishing between Normal and Abnormal states. VIP scores, standing for Variable Importance in Projection, provide valuable insights into the contribution of each variable to the discrimination process. The Fig. 5.5 presents VIP scores for various variables, with higher scores indicating greater importance in separating Normal from Abnormal states. At the top of the list, "AcAce" boasts the highest VIP score of 2.238771, signifying its significant contribution to the discrimination process. Following closely are "BOHBut" and "Ala" with VIP scores of 1.88677 and 1.748646, respectively, indicating their substantial importance in distinguishing between the two states. Additionally, "Phe," "His," and "Tyr" exhibit notable VIP scores above 1.2, underscoring their relevance in the discrimination process. As the list progresses, the VIP scores gradually decrease, with variables such as "Glc," "Cit," "Glol," "Val," and "Leu" showing relatively lower scores. Overall, the VIP scores serve as crucial analytical guides, aiding in the identification of key variables essential for distinguishing between Normal and Abnormal states.

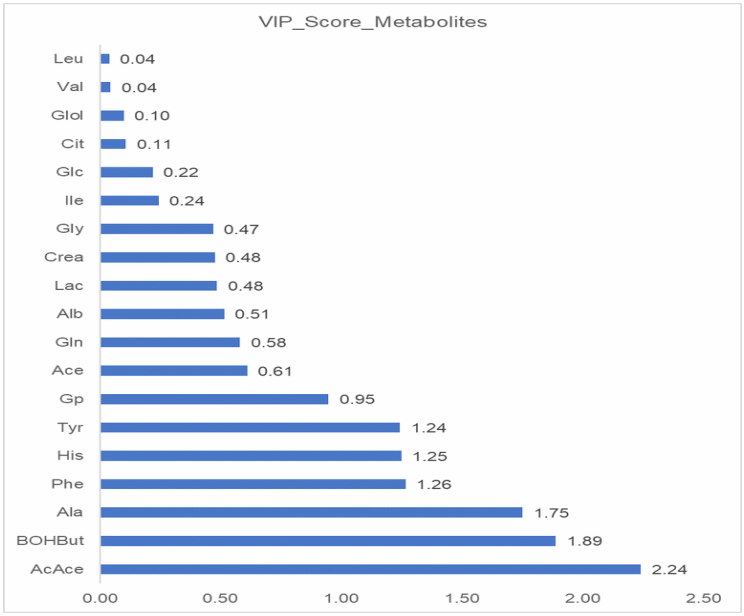


Fig 5.5 Significant Contributors through VIP Scores for Normal versus Abnormal

5.2.5.5 Inflection point for Normal vs Abnormal

In section 5.2.7.5, an inflection point is identified for distinguishing between Normal and Abnormal states. The inflection point signifies a crucial threshold where the significance of the variable becomes pronounced in the discrimination process. Among the variables listed, "Tyr" stands out with a VIP score of 1.241259, indicating its pivotal role in the discrimination process (Fig 5.6). This VIP score serves as the inflection point, suggesting that beyond this value, variables may exhibit increasingly significant contributions to distinguishing between Normal and Abnormal states. Preceding "Tyr," variables such as "AcAce," "BOHBut," "Ala," "Phe," and "His" exhibit relatively higher VIP scores, indicating their importance in the discrimination process. Conversely, variables with lower VIP scores, such as "Val" and "Leu," contribute less significantly to the discrimination between the two states. Identifying the inflection point allows for a deeper understanding of the critical variables driving the discrimination process. In this case, "Tyr" serves as a pivotal marker where the significance of variables transitions, potentially guiding further analysis and exploration into the underlying mechanisms distinguishing between Normal and Abnormal states.

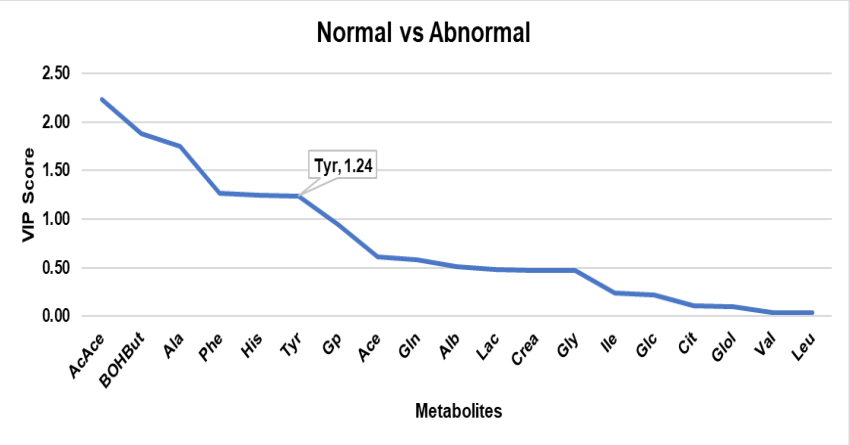


Fig 5.6 Inflection point for Normal vs Borderline abnormal

5.2.5.6 AUC Curve in Distinguishing Normal from Abnormal

In section 7.3.5, the AUC (Area Under the Curve) curve is utilized as a metric to assess the discriminatory capability of various variables in distinguishing between Normal and Abnormal states. The AUC scores for each variable provide insights into their effectiveness in classification, with higher AUC values indicating better discriminatory performance.

Among the examined variables, "Pyr" demonstrates a relatively high AUC score of 0.6192691, suggesting strong discriminatory ability in distinguishing between Normal and Abnormal states (Fig 5.7). Similarly, "Tyr" also exhibits a notable AUC score of 0.6106312, indicating effective discrimination. However, "Ala," "Phe," and "AcAce" show slightly lower but still significant AUC scores of 0.5847176, 0.5568106, and 0.5734219, respectively, suggesting moderate discriminatory power. On the other hand, "Ace" demonstrates a lower AUC score of 0.4916944, indicating relatively weaker discriminatory capability. These findings underscore the variable-specific differences in discriminatory ability and highlight the importance of considering individual variables' contributions when evaluating the effectiveness of classification models in distinguishing between Normal and Abnormal states.

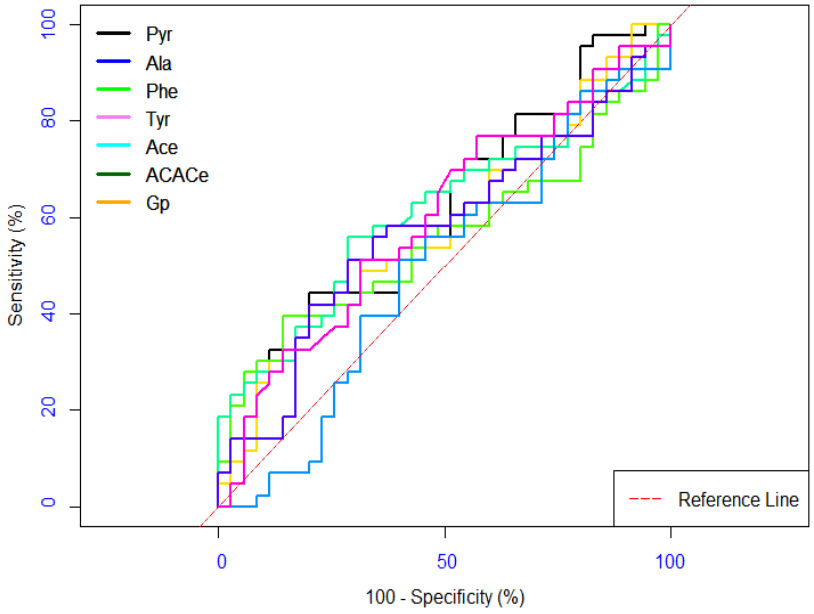


Fig 5.7 Significance of AUC in Distinguishing Normal from Borderline abnormal

5.3 Evaluation and Implications

5.3.1 Correlational analyses

Significant positive correlations were found for quite a few variables: smoking, alcohol, and supplement usage. Age had correlation that was not significant, but this finding of worsening depressive symptoms with age, despite increased supplement usage, highlights the lack of efficacy of antidepressants they may have been taking to mitigate these symptoms.

The glucose finding matches previous research showing how impaired glucose metabolism is linked to depression symptoms specifically (Bouwman et al., 2010), but the significant negative correlation indicates low, rather than high, glucose levels associated with higher depression scores. This is contrary to research linking diabetes (high blood glucose) to higher depression rates, potentially due to oxidative stress and decreased insulin sensitivity (Peng et al., 2016). Hyperglycemia may contribute to depression by inducing inflammation and reductions in neurotrophic function, which may lead to decreased plasticity of neuronal networks and, ultimately, depression (Dean and Keshavan, 2017; Lustman et al., 2000; Lima Giacobbo et al., 2019).

It is possible that these low levels reflect a fluctuation in blood sugar levels, rather than chronically low blood glucose. This fluctuation has been shown to disrupt hormone balance, which in turn affects mood regulation in depressed patients, and furthermore affects levels of inflammation in the body, which would aggravate depressive symptoms (Holt et al., 2014; Mantantzis et al., 2019). Pal (2021) also evidenced how fluctuating glucose levels might also affect glutamate levels which have consistently been shown to be dysregulated in mood disorders such as MDD, disrupting connectivity between regions of the brain that involve emotion and self-perception (Pal, 2021). It is possible that this cohort have also had higher

glucose levels but the time at which samples were taken, their glucose levels may have happened to be low. That said, the relationship between impaired glucose levels and risk of depression is still inconclusive in literature, with a large meta-analysis by Nouwel et al (2011) reinforcing this ambivalence, where risk of depression was not increased in people with impaired glucose metabolism or normal glucose metabolism (OR 0.96, 95% CI 0.85-1.08) (Nouwel et al., 2011).

Both interleukins IL-10 and IL-1β were significantly negatively associated with HADS scores in the two-year cut-off cohort. Literature has already established the importance of inflammatory markers IL-6 and IL-8 in MDD, but the roles of IL-10 and IL-1β are not as well-known. IL-10 is anti-inflammatory and its administration has been shown to reduced learned helplessness (LH), which is characteristic in depressed patients (Mesquita et al., 2008; Keaton et al., 2021). IL-1β as an inflammatory biomarker should have revealed the opposite, with its higher levels correlating with higher depression scores instead (Ng et al., 2018), and interestingly, this inverse relationship was shown once again in the multiple regression analysis afterwards. This could be due to IL-1β playing a lesser role in the inflammatory pathophysiology of depression compared to interleukins such as IL-6, IL-1 and TNF- α (Ting et al., 2020), although it has also been shown that significance between depression levels and IL-6 is lost when, interestingly, the variables smoking, exercise and BMI are controlled for, two of which (exercise and smoking) were significantly correlated in the cohort for this present investigation. Low levels of IL-1β could also be a new biomarker of interest for MDD, aside from higher levels of other interleukins, especially since it was extremely significantly inversely associated in the multiple regression too.

Exercise as inversely associated with depression severity was not a surprising finding. Physical activity can boost levels of BDNF and IGF-1 by improving vascular delivery of these neurotrophic or growth factors and thus rectifying impaired neuroplasticity pathways in the brain that are characteristic of MDD (Kandola et al., 2019). This relationship between

exercise and depression severity became more pronounced during the multiple regression analyses, showing a very significant relationship between exercise levels and depression severity all round, underlining the antidepressant effects of physical activity, via its cardiorespiratory benefits and ability to boost cognitive function. Recchi et al. (2022) in a meta-analysis looking at 21 randomised controlled trials found that there was no difference between exercise and antidepressant efficacy in reducing depressive symptoms, suggesting exercise may be a powerful adjuvant or alternative to antidepressant medication (Rechhi et al., 2022). However, a potential confounder may be that those who were not active might have had a health condition that they were afflicted with at a young age, for example a cardiovascular condition like carditis or congenital heart disease, or had an eye disorder or disability, which meant that they could not be active at a young age. This sedentary lifestyle may continue into adulthood, thus increasing the likelihood of developing depression that only increases with time due to an inability to initiate and maintain an active lifestyle.

Alcohol was significantly positively correlated with increasing depression severity, and this significance, like exercise, was upheld in the multiple regression analysis. Although the extent to which alcohol plays a part in aggravating depression severity is generally unknown, research does display an obvious bidirectional relationship: alcohol can be used to numb depressive symptoms, but alcohol consumption itself can exacerbate depressive symptoms, such as the positing of the tension-reduction theory posited by MacAndrew (1982), mentioned earlier in this chapter, or disrupt a normal work-life balance and relationships, which further exacerbates depressive symptoms. The main limitation here lies in the measurement of alcohol consumption, since just one question (chosen as it could be matched to the most participants in the cohort) did not show volume of alcohol consumption each time, but rather the frequency of maximum alcohol consumption per week. It is also possible that some people may have grossly underestimated or overestimated their frequency of consumption.

Smoking as an activity associated with increased oxidative stress in the body exacerbates depressive symptoms of MDD, explaining its positive association with the disease. Lower levels of exercise are usually found to be correlated with higher depression severity- a sedentary lifestyle exacerbates depressive symptoms in a meta-analysis by Huang et al (2020) (RR = 1.10, 95% CI 1.03–1.19, $I^2 = 60.6\%$, $P < 0.01$). However, this same positive association was not found the multiple linear regression results, showing that it may not be as important a variable to consider compared to exercise or alcohol. That said, smoking's detrimental health effects overall and contribution to higher inflammatory markers does not erase its significance in exacerbating depressive symptoms.

There was only a significant difference found for income level between the Normal and Borderline groups, which makes sense despite the lack of a finding for an association between income level overall and depression severity as indicated by HADs scores. More so than country-level income, individual income was seen to be more significantly correlated with depression rates, in a study that also showed how being female increased vulnerability to depressive symptoms (Rai et al., 2018), complementing findings in this cohort which comprises females.

The only other significant correlation that came up for the metabolite analysis for the two-year cohort was for glycine, showed to be positively correlated only in the unmedicated fraction of the cohort following stratification for medication. On the one hand, this finding is expected since glycine is one of the four main amino acid neurotransmitters alongside glutamate, γ-aminobutyric acid, and it is known that an imbalance of these excitatory or inhibitory neurotransmitters exacerbates depressive symptoms. On the other hand, glycine being the only significantly associated metabolite out of the four puts forward its potential as a unique MDD biomarker. Complementing this notion is new research that has shown how knockout of the gene for glycine receptor GPR158 resulted in mice that were resilient to chronic stress (Laboute et al., 2023) by regulating neuronal excitability in cortical neurons. Contrasting this is research showing that glycine is associated with favourable plasma lipid

profiles. This finding of acetate and glycine most strongly associated with depression severity more so than other metabolites implicated in the TCA cycle warrants further research into their potential as MDD biomarkers and the importance of the TCA cycle in understanding MDD pathophysiology, as referenced in the systematic review and meta-analysis prior to this investigation. Diagnostic group differences agreed with correlational analyses findings, with no significant differences between the three diagnostic groups.

5.3.2 Stratification for medication

Diagnosis of high blood pressure and high cholesterol being higher in the medicated group indicates that the medicated subgroup were being treated for high blood pressure or high cholesterol levels, which must be taken into account when interpreting the other significant findings. Acetate and glucose levels were higher in the medicated subgroup of the cohort, alongside smoking frequency and age. Glucose being higher in the medicated group was not surprising, since statins (for high cholesterol levels) and medications to lower blood pressure such as beta-blockers have the ability to increase blood glucose levels by increasing insulin resistance; Abbasi et al demonstrates a drastic increase in insulin resistance by a median of 8% ($p=0.01$), as well as that for insulin secretion by a median of 9% ($p=0.01$) (Abbasi et al., 2021). Acetate is an oft-overlooked metabolite that is keenly involved in the TCA cycle (coupling with CoA to form acetyl-CoA), and its higher levels in the medicated cohort parallel the glucose finding, as each glucose molecule produces two acetyl-CoA molecules that can enter the TCA cycle.

Exercise levels found to be significantly lower in the medicated subgroup could be a compensatory result of taking medication– taking statins to lower blood pressure or cholesterol may reduce the patient’s desire to exercise more to reduce cholesterol levels

5.3.3 Multiple regression and multivariate analysis

Lower levels of acetate complement the overall low levels of glucose found in the cohort during Pearson correlation, despite the opposite being found for both metabolites in the medicated subgroup, and this was a more significant finding than that for tyrosine, positioning it as a potential MDD biomarker. This short-chain fatty acid (SCFA) was the most significantly associated metabolite with a coefficient of -10.4156, and significance for glycine was lost. This sparks interest in a relatively overlooked metabolite in investigating fatty acids implicated in depression pathophysiology– research focussing on acetate is sparse since humans typically have lower plasma acetate levels compared to other mammalian species (Moffett et al., 2020), since modern diets lack fermentable dietary fibre to produce SCFAs. Acetate must be converted to acetyl-CoA, which is needed for energy derivation and lipogenesis in all human cells. Acetate, alongside other SCFAs (short-chain fatty acids) such as butyrate, also has been hypothesised to affect depression pathophysiology by direct stimulation of central nervous system receptors, immune mediation, and epigenetic regulation (Caspani et al., 2019). Acetate supplementation has even been shown by Huang et al (2021) to reduce depressive-like symptoms in animal studies, utilising the forced swim and tail suspension tests, and even upregulated BDNF levels and increasing dendritic connections and spinal density in pyramidal neurons (Hang et al., 2021). Although this study was limited to just this part of the mice brain, this is a valuable finding that reinforces the importance of acetate in balancing aberrant energy metabolism in the brains of depressed patients. Acetyl-L-carnitine which is produced from acetyl-CoA reacting with L-carnitine, and thus requires acetate, has also been shown to be at much lower levels in chronically depressed patients (Zagorski, 2018). Ironically, with current research sparking interest in the relationship between the gut microbiota and neurological health, acetate metabolism and its role in depression pathophysiology may be of renewed importance.

The lack of any significant associations in the 3-year cohort could have been due to a range of reasons– one year is ample time for someone such as a family member or friend to help the patient find professional help or assist with diagnosing the patient and reduce his/her depressive symptoms. Participants may have found a more suitable setting in which they could recover quickly than they would have if they were attempting to recover in another environment, for example in unsupportive social settings at home or school. If consistent with medication, a year could drastically reduce depressive symptoms at least temporarily.

5.3.4 Metabolite-OPLSDA model

The Metabolite-OPLSDA model was employed to analyze a dataset comprising 78 samples characterized by 19 predictor variables and one response variable. While the model demonstrated relatively good explanatory power for predictor variables ($R^2X(cum)$), its predictive performance ($Q^2(cum)$) was poor, as indicated by the negative Q^2 value. Despite explaining a significant portion of variance in predictor variables ($R^2X(cum) = 0.332$), the model's ability to predict responses was limited, as indicated by the low cumulative Q-squared value in cross-validation ($Q^2(cum) = -0.269$). Nonetheless, the model consistently outperforms random chance in distinguishing between Normal and Abnormal states, with VIP scores serving as valuable guides to identify key variables essential for discrimination. "AcAce," "BOHBut," and "Ala" emerge as top contributors, while "Tyr" acts as the inflection point, guiding further analysis. The AUC curve analysis further underscores the discriminatory capability of various variables, with "Pyr" and "Tyr" showing notable discriminatory ability, and "Ace" exhibiting relatively weaker performance.

5.3.4 Limitations

A major limitation was the restriction of this investigation to the two-year cohort, since this also reduced the number of variables that could be matched to the participants. Therefore,

other important confounding factors would not have been taken into consideration, such as certain comorbidities like dementia, health issues or sociodemographic variables.

An important note of consideration is the relevance of disease course, and not just the onset of the disease itself. For example, this cohort was mostly above middle age, which would mean a higher incidence of age-related disease such as cardiovascular, lung and brain complications such as pulmonary (of which data was not collected in the database anyway), due to the accumulation of certain health habits and deterioration of health overall. The HADS questionnaire as the only means of assessing the cohorts' depression severity also does not offer comprehensive insight into a patient's depression severity, and the scope of research should be widened to incorporate more methods such as brain scans that can complement the assessment of the neurocognitive profiles of participants.

Conclusively, exercise may be a powerful addition to the therapeutic options for MDD, alongside reduction for alcohol consumption. MDD is linked to low IL-1β levels, may indicate immune dysfunction, but further investigation is needed, especially in a more defined depressed group while levels of pyruvate, alanine, tyrosine, and acetate should not be overlooked due to their strong associations with depression severity overall.

Chapter 6: Investigating the role of lipids in major depressive disorder in the TwinsUK database

6.1 Introduction

The TwinsUK database presents an opportunity to explore new networks between various lipid subclasses and depression. Contradictory findings on the association between lipoprotein subclasses, cholesterol, and depression necessitate broader investigations beyond HDL and LDL. Limited evidence hints at the potential impact of EPA-enriched phospholipids in reducing neuroinflammation tied to depressive states. Furthermore, the varying trends observed in the relationship between apolipoproteins and depression emphasize the pressing need for more comprehensive research to clarify their role in depressive conditions. For detailed information relating to these variables please see Chapter one, Section 1.6.

Exploratory analysis was conducted in the same way as done for the metabolites in the previous chapter, using the master file which comprised all information on participants who could provide data for lipids with less than 15% of this data missing, alongside at least one HADS score (taken in 2002, 2017 or both years). Participants were only chosen if they possessed a HADS score taken within two years of offering a blood sample.

6.2 Results

6.2.1 Identification of lipids included in the analysis

Out of the 208 lipids selected in this analysis (Please see appendix-8 for details), there were 19 lipid subclasses in total that were included in the final analysis, as shown in the list below:

IDLC: total cholesterol in IDL (intermediate-density lipoprotein)

IDLFC: free cholesterol in IDL

LLDLL: Total lipids in large LDL (ow-density lipoprotein)

LLDLC: total cholesterol in large LDL

LLDLFC: free cholesterol in large LDL

MLDLL: total lipids in medium LDL

MLDLC: cholesterol in medium LDL

MLDLCE: cholesterol esters in medium LDL

MLDLFC: free cholesterol in medium LDL

SLDLL: total lipids in small LDL

SLDLC: total cholesterol in small LDL

SLDLCE: cholesterol esters in small LDL

SLDLFC: free cholesterol in small LDL

SLDLPL: phospholipids in small LDL

MLDLPL: phospholipids in medium LDL

LDLC: cholesterol in LDL

SM: sphingomyelin

DHA: docosahexaenoic acid

Free cholesterol

6.2.2 Univariate analysis

Just as was done for the metabolite analysis, Table 6.1 displays data for various lipid variables based on different categories of lipids (of which data there was <15% missing), with a two-year cut-off between time of answering the questionnaire and time of sample collection. The categories in the tables are Normal, Borderline Abnormal, and Abnormal. The summaries include the mean and standard deviation (SD) for each variable. For example, there is a significant difference in IDLC levels among the three categories (Normal, Borderline Abnormal, and Abnormal), as indicated by the F-statistic of 3.296. The p-value of 0.041 suggests that the observed difference is statistically significant at a probability level of 0.05. It is observed that there is a significant difference in all lipid variable levels between the three categories, as indicated by the F-statistic and the associated p-values.

Table 6.1: Table showing lipids significantly associated with HADS scores, and statistical tests to show significance between the three diagnostic groups

Lipids <15% and two years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
Number of unique samples (Total=113) Missing data=0%=0	43	35	35	χ2 test (2, N = 113) =1.133, P=0.568
IDLC (mean, SD) N=113 Missing data=0%=0	0.672±0.22	0.618±0.171	0.562±0.157	ANOVA (F (2, 110) = [3.296], p = **0.041**).
IDLFC (mean, SD) N=113 Missing data=0%=0	0.187±0.06	0.171±0.046	0.155±0.051	ANOVA (F (2, 110) = [3.394], p = **0.037**).
LLDLL (mean, SD) N=113 Missing data=0%=0	1.216±0.392	1.112±0.307	1.025±0.295	ANOVA (F (2, 110) = [3.100], p = **0.049**).
LLDLC (mean, SD) N=113 Missing data=0%=0	0.804±0.288	0.728±0.225	0.662±0.226	ANOVA (F (2, 110) = [3.135], p = **0.047**).
LLDLFC (mean, SD) N=113 Missing data=0%=0	0.231±0.067	0.213±0.053	0.196±0.056	ANOVA (F (2, 110) = [3.443], p = **0.035**).
MLDLL (mean, SD) N=113 Missing data=0%=0	0.686±0.234	0.621±0.184	0.57±0.176	ANOVA (F (2, 110) = [3.251], p = **0.042**).
MLDLC (mean, SD) N=113 Missing data=0%=0	0.443±0.176	0.393±0.137	0.352±0.141	ANOVA (F (2, 110) = [3.364], p = **0.038**).
MLDLCE (mean, SD) N=113 Missing data=0%=0	0.307±0.144	0.267±0.111	0.233±0.118	ANOVA (F (2, 110) = [3.316], p = **0.040**).
MLDLFC (mean, SD) N=113 Missing data=0%=0	0.136±0.032	0.126±0.026	0.119±0.024	ANOVA (F (2, 110) = [3.469], p = **0.035**).

Table 6.1: Table showing lipids significantly associated with HADS scores…(continued)

Lipids <15% and two years cut off	**Normal (0-7)**	**Borderline Abnormal (8-10)**	**Abnormal (11-21)**	**Test Statistics**
SLDLL (mean, SD) N=113 Missing data=0%=0	0.441±0.143	0.401±0.113	0.369±0.108	ANOVA (F (2, 110) = [3.276], p = **0.042**).
SLDLC (mean, SD) N=113 Missing data=0%=0	0.27±0.105	0.24±0.082	0.215±0.086	ANOVA (F (2, 110) = [3.474], p = **0.034**).
SLDLCE, (mean, SD) N=113 Missing data=0%=0	0.185±0.086	0.161±0.067	0.14±0.072	ANOVA (F (2, 110) = [3.416], p = **0.036**).
SLDLFC, (mean, SD) N=113 Missing data=0%=0	0.085±0.02	0.079±0.016	0.075±0.015	ANOVA (F (2, 110) = [3.501], p = **0.034**).
SLDLPL, (mean, SD) N=113 Missing data=0%=0	0.141±0.03	0.133±0.026	0.127±0.022	ANOVA (F (2, 110) = [2.760], p = 0.068).
MLDLPL (mean, SD) N=113 Missing data=0%=0	0.194±0.047	0.179±0.042	0.172±0.033	ANOVA (F (2, 110) = [2.830], p = 0.063).
LDLC (mean, SD) N=113 Missing data=0%=0	1.517±0.569	1.362±0.443	1.229±0.452	ANOVA (F (2, 110) = [3.273], p = **0.042**).
Free cholesterol (mean, SD) N=113 Missing data=0%=0	1.266±0.309	1.184±0.244	1.114±0.247	ANOVA (F (2, 110) = [3.068], p = **0.050**).
SM (mean, SD) N=113 Missing data=0%=0	0.447±0.082	0.42±0.069	0.404±0.065	ANOVA (F (2, 110) = [3.375], p = **0.038**).
DHA (mean, SD) N=113 Missing data=0%=0	0.125±0.052	0.104±0.038	0.098±0.045	ANOVA (F (2, 110) = [3.634], p = **0.030**).

Table 6.2 below presents the correlation between HADs scores and each lipid type, along with the associated p-values at 95% confidence intervals. The negative correlation coefficients suggest that higher HADs scores are associated with lower levels of the respective lipids, and vice versa. Both the medicated and non-medicated groups' interactions between HADs scores and lipid variables are displayed here, along with their respective correlation coefficients and p-values. There was not a single type of lipid that was identified as having any kind of substantial association with HADs when looking at the non-medicated category. In the medicated cohort, significant negative correlations were observed between HADS scores and several lipid subclasses, indicating potential associations between lipid metabolism and depression severity in individuals receiving medication. Conspicuously, HADS scores exhibited negative correlations with IDLC ($r = -0.269$, $p = 0.036$), IDLFC ($r = -0.285$, $p = 0.026$), LLDLC ($r = -0.254$, $p = 0.049$), LLDLFC ($r = -0.273$, $p = 0.034$), MLDLL ($r = -0.252$, $p = 0.050$), MLDLC ($r = -0.271$, $p = 0.035$), MLDLCE ($r = -0.278$, $p = 0.030$), SLDLC ($r = -0.274$, $p = 0.033$), SLDLCE ($r = -0.284$, $p = 0.027$), LDLC ($r = -0.263$, $p = 0.041$), and DHA ($r = -0.426$, $p = 0.001$). Among these, the strongest negative correlation was observed between HADS scores and DHA, highlighting a potentially robust relationship between depression severity and docosahexaenoic acid levels.

Table 6.2: Pearson correlations and significance values for HADS scores against each lipid subclass and significant associations between the lipids with each other in the two-year cohort, after stratifying for medication status

	Medicated				Non-Medicated			
Pairs	**Pearson Correlation**	**P-Value**	**95% Confidence Intervals**		**Pearson Correlation**	**P-Value**	**95% Confidence Intervals**	
			Lower	**Upper**			**Lower**	**Upper**
HADs - IDLC	-0.269	***0.036***	-0.488	-0.019	-0.222	0.163	-0.496	0.092
HADs - IDLFC	-0.285	***0.026***	-0.501	-0.035	-0.201	0.207	-0.479	0.113
HADs - LLDLL	-0.246	0.056	-0.469	0.006	-0.233	0.143	-0.504	0.081
HADs - LLDLC	-0.254	***0.049***	-0.475	-0.002	-0.227	0.153	-0.500	0.087
HADs - LLDLFC	-0.273	***0.034***	-0.491	-0.022	-0.219	0.168	-0.494	0.095
HADs - MLDLL	-0.252	***0.050***	-0.474	0.000	-0.233	0.142	-0.505	0.080
HADs - MLDLC	-0.271	***0.035***	-0.489	-0.021	-0.220	0.168	-0.494	0.094
HADs - MLDLCE	-0.278	***0.030***	-0.495	-0.028	-0.210	0.187	-0.486	0.104
HADs - MLDLFC	-0.235	0.069	-0.459	0.018	-0.260	0.101	-0.525	0.052
HADs - SLDLL	-0.246	0.056	-0.469	0.006	-0.232	0.145	-0.504	0.082
HADs - SLDLC	-0.274	***0.033***	-0.492	-0.023	-0.217	0.172	-0.492	0.097
HADs - SLDLCE	-0.284	***0.027***	-0.500	-0.034	-0.205	0.199	-0.482	0.110
HADs - SLDLFC	-0.218	0.091	-0.446	0.036	-0.267	0.092	-0.531	0.044
HADs - SLDLPL	-0.164	0.207	-0.399	0.092	-0.272	0.086	-0.535	0.039
HADs - MLDLPL	-0.177	0.173	-0.410	0.078	-0.274	0.083	-0.536	0.037
HADs - LDLC	-0.263	***.041***	-0.483	-0.012	-0.223	0.161	-0.497	0.091
HADs - FreeC	-0.251	0.051	-0.473	0.001	-0.210	0.187	-0.486	0.104
HADs - SM	-0.246	0.056	-0.469	0.006	-0.229	0.150	-0.501	0.085
HADs - DHA	-0.426	***0.001***	-0.612	-0.195	-0.038	0.811	-0.342	0.272

Bold and italic are significant at P<0.05.

6.2.3 Correlation matrix for variables in the two-year lipid cohort

These results show the correlation coefficients and p-values for the associations between HADs scores and various factors, such as age, BMI, glucose, insulin, depression, anxiety, smoking, alcohol, interleukin levels, exercise, health conditions, income, and supplement usage. The correlation coefficients indicate the strength and direction of the relationship, while the p-values provide information on the statistical significance of the correlations. With a 0.000 p-value, the positive correlation of 0.754 for depression is statistically significant. This shows that larger levels of depression are associated with higher HADs scores. Concerning anxiety, a strong association (r = 0.755) exists (p = 0.000). This data strongly suggests that increased anxiety is associated with higher HADs scores. A correlation between smoking and unfavourable health outcomes of 0.216 at a significance level of 0.024 is found. This demonstrates that higher concentrations of HADs are associated with a higher rate of smoking. The correlation between alcohol consumption and health issues is 0.198 (p = 0.037). This demonstrates that higher HADs scores are associated with heavier alcohol use. Interleukin-10 is inversely correlated with HADs (-0.191), with a p-value of 0.043. Lower levels of Interleukin10 are associated with higher HADs scores, as shown here. The p-value for the Interleukin1B correlation is 0.030, indicating a -0.204 negative association. Higher HADs scores seem to be associated with lower Interleukin1B levels. The p-value for the workout correlation is 0.039 (or -0.207). This suggests that people with higher HADs scores engage in less physical exercise. (For more details on this analysis, look at Appendix-5).

6.2.4 Correlation plot for lipids

The correlation coefficients indicate the strength and direction of the linear relationship between variables. A negative correlation means that as one variable increases, the other tends to decrease, while a positive correlation means that both variables tend to increase or decrease together. The significance levels indicate whether the observed correlations are likely to be statistically significant or occurred by chance. The correlation coefficient between HADs and IDLC is -0.238 (Fig 6.1). This indicates a moderate negative correlation between the two variables. As the "HADs" score increases, the "IDLC" value tends to decrease, and vice versa. The p-value associated with this correlation is 0.011, which is below the typical significance level of 0.05. This suggests that the correlation is statistically significant, indicating that the observed relationship between "HADs" and "IDLC" is unlikely to have occurred by chance alone. We discovered a statistically significant connection between lipid factors and the HADs score. For more information look at Appendix-6.

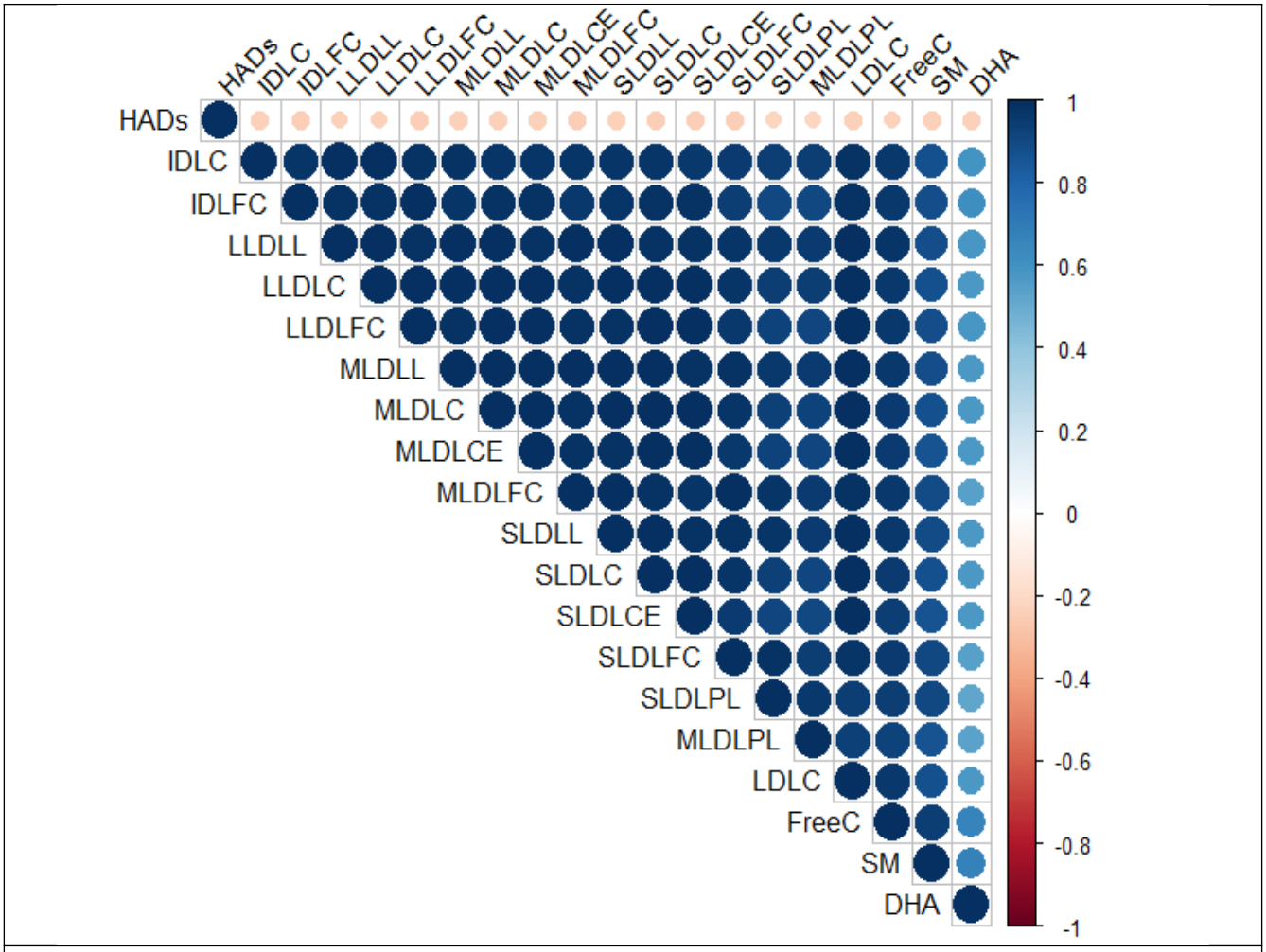


Fig 6.1 Correlation plot for lipids in the two-year cohort

6.2.5 Identification of confounding variables based on medication

Regarding the Medicated and Non-Medicated groups, the t-test outcomes reveal a statistically significant age difference between the Medicated and Non-Medicated groups, as shown in Table 6.3. The average age of those who take medicine is significantly older than the average age of people who do not take medication. When the Medicated and Non-Medicated groups are compared using a t-test, the results show that there is a statistically significant difference in depression; however, there is no such difference in anxiety. Based on the significance of the t-test, we could isolate confounding variables such as age, depression, smoking, exercise, age at the time of the initial diagnosis of hypertension (BP_age), age at the time of the initial diagnosis of high cholesterol (HC_age).

Table 6.3: Significant variables stratifying for medication in the two-year lipid cohort

Variable	Medicated	N	Mean	StDev	Median	Max.	Min.	IQR	T-Value	DF	P-Value
Age	Medicated	61	60.49	9.45	61.85	73.69	41.63	14.35	-2.51	94	***0.0140***
	Non_Medicated	41	56.13	8.01	55.79	74.22	41.64	11.31			
	Total	102									
Depression	Medicated	61	5	3.286	4	14	0	5	-2.32	92	***0.0230***
	Non_Medicated	41	3.561	2.924	3	12	0	4.5			
	Total	102									
Anxiety	Medicated	61	5.41	3.432	5	14	1	5.5	-1.62	92	0.1080
	Non_Medicated	41	4.366	3.015	4	17	0	4			
	Total	102									
Smoking	Medicated	58	13.59	9.2	17	40	0	19	-2.77	88	***0.0070***
	Non_Medicated	40	8.6	8.42	11.5	22	0	16			
	Total	98									
Alcohol	Medicated	61	11.77	11.36	8	41	0	17.5	-1.45	86	0.1500
	Non_Medicated	40	8.53	10.73	4.5	45	0	9			
	Total	102									
Interleukin10	Medicated	61	1.972	7.112	0	36.2	0	0	-0.19	91	0.8510
	Non_Medicated	41	1.72	6.47	0	30.88	0	0			
	Total	102									
Interleukin1B	Medicated	61	0.309	1.173	0	6.62	0	0	-0.06	92	0.9500
	Non_Medicated	41	0.295	1.045	0	5.82	0	0			
	Total	102									
Exercise	Medicated	54	2.315	0.865	2	5	0	1	1.97	89	***0.0500***
	Non_Medicated	38	2.6053	0.5472	3	4	2	1			
	Total	92									

Table 6.3: Significant variables stratifying for medication in the two-year lipid cohort… (continued)

Variable	Medicated	N	Mean	StDev	Median	Max.	Min.	IQR	T-Value	DF	P-Value
HC_Age	Medicated	52	26.54	27.97	0	15.5	73	50	-4.21	79	***0.0000***
	Non_Medicated	30	5.9	16.39	0	0	62	0			
	Total	82									
BP_Age	Medicated	52	20.27	28.46	0	0	71	53.5	-3.29	79	***0.0010***
	Non_Medicated	30	4.17	15.86	0	0	63	0			
	Total	82									
Income	Medicated	52	3.385	2.233	0	4	8	3	0.26	52	0.7960
	Non_Medicated	30	3.533	2.636	0	5	8	6			
	Total	82									
Contra1	Medicated	54	0.6111	0.4921	1	1	0	1	0.61	64	0.5430
	Non_Medicated	31	0.6774	0.4752	1	1	0	1			
	Total	85									
Contra2	Medicated	25	8.48	6.76	0	8	29	6	0.98	30	0.3350
	Non_Medicated	14	10.5	5.83	3	10	21	8.25			
	Total	39									
IDLC	Medicated	61	0.5833	0.1854	0.5497	1.104	0.2277	0.2594	2.52	83	***0.0140***
	Non_Medicated	41	0.6798	0.1921	0.6901	1.294	0.2949	0.2356			
	Total	102									
IDLFC	Medicated	61	0.16058	0.05458	0.151	0.318	0.0583	0.07485	2.57	88	***0.0120***
	Non_Medicated	41	0.18824	0.05248	0.1939	0.3346	0.0735	0.06015			
	Total	102									
LLDLL	Medicated	61	1.0597	0.3402	1.018	1.976	0.4778	0.4587	2.4	85	***0.0190***
	Non_Medicated	41	1.2257	0.3439	1.225	2.367	0.5254	0.3812			
	Total	102									
LLDLC	Medicated	61	0.6854	0.2518	0.6491	1.35	0.2289	0.3453	2.55	85	***0.0130***
	Non_Medicated	41	0.8155	0.254	0.8103	1.625	0.2953	0.2927			
	Total	102									
LLDLFC	Medicated	61	0.20193	0.06055	0.1937	0.3678	0.0948	0.0878	2.65	87	***0.0100***
	Non_Medicated	41	0.23373	0.0588	0.236	0.4005	0.1096	0.0741			
	Total	102									

Table 6.3: Significant variables stratifying for medication in the two-year lipid cohort… (continued)

Variable	Medicated	N	Mean	StDev	Median	Max.	Min.	IQR	T-Value	DF	*P-Value*
MLDLL	Medicated	61	0.5893	0.203	0.5743	1.136	0.2597	0.28	2.5	84	***0.0150***
	Non_Medicated	41	0.6929	0.2072	0.6843	1.384	0.272	0.2237			
	Total	102									
MLDLC	Medicated	61	0.3662	0.1547	0.3517	0.7788	0.0975	0.2387	2.66	85	***0.0090***
	Non_Medicated	41	0.4498	0.1562	0.4635	0.933	0.1285	0.199			
	Total	102									
MLDLCE	Medicated	61	0.2445	0.1269	0.2314	0.5832	0.0191	0.1918	2.65	85	***0.0100***
	Non_Medicated	41	0.3126	0.1279	0.3172	0.7033	0.0483	0.1677			
	Total	102									
MLDLFC	Medicated	61	0.12171	0.02835	0.1181	0.1956	0.0785	0.03935	2.69	85	***0.0090***
	Non_Medicated	41	0.1372	0.02864	0.1342	0.2297	0.0802	0.0332			
	Total	102									
SLDLL	Medicated	61	0.3803	0.1247	0.3783	0.7157	0.1795	0.1782	2.6	85	***0.0110***
	Non_Medicated	41	0.4462	0.1262	0.4335	0.8697	0.1879	0.14			
	Total	102									
SLDLC	Medicated	61	0.223	0.0936	0.2183	0.4753	0.0621	0.1427	2.74	86	***0.0070***
	Non_Medicated	41	0.2747	0.0931	0.2817	0.5593	0.0841	0.1173			
	Total	102									
SLDLCE	Medicated	61	0.14706	0.07694	0.1356	0.3543	0.0116	0.11725	2.7	86	***0.0080***
	Non_Medicated	41	0.1888	0.0761	0.187	0.4163	0.0318	0.1022			
	Total	102									
SLDLFC	Medicated	61	0.0759	0.01733	0.074	0.121	0.0504	0.025	2.83	85	***0.0060***
	Non_Medicated	41	0.08587	0.01756	0.0835	0.143	0.0522	0.02085			
	Total	102									
SLDLPL	Medicated	61	0.12915	0.02648	0.1244	0.2062	0.0921	0.0388	2.52	85	***0.0140***

Non_Medicated	41	0.14273	0.02682	0.1398	0.2404	0.0872	0.02575
Total	102						

Table 6.3: Significant variables stratifying for medication in the two-year lipid cohort… (continued)

Variable	Medicated	N	Mean	StDev	Median	Max.	Min.	IQR	T-Value	DF	*P-Value*
MLDLPL	Medicated	61	0.1746	0.0417	0.1643	0.2987	0.1	0.05895	2.49	85	***0.0150***
	Non_Medicated	41	0.19572	0.04218	0.1925	0.3532	0.1134	0.0433			
	Total	102									
LDLC	Medicated	61	1.2745	0.4993	1.237	2.584	0.3885	0.7392	2.62	85	***0.0100***
	Non_Medicated	41	1.54	0.5027	1.558	3.117	0.5079	0.6115			
	Total	102									
FreeC	Medicated	61	1.1372	0.2709	1.093	1.889	0.6336	0.3953	2.42	84	***0.0180***
	Non_Medicated	41	1.2718	0.2786	1.286	2.192	0.6057	0.3005			
	Total	102									
SM	Medicated	61	0.41205	0.07404	0.4109	0.5945	0.2824	0.10595	2.14	84	***0.0350***
	Non_Medicated	41	0.4445	0.0757	0.4347	0.661	0.3	0.0734			
	Total	102									
DHA	Medicated	61	0.10262	0.04678	0.0973	0.2404	0.0267	0.0639	1.6	84	0.1140
	Non_Medicated	41	0.11799	0.04813	0.1118	0.243	0.0175	0.0585			
		102									

Interestingly, stratifying for medication revealed that only the medicated portion of participants had significant findings for certain lipid levels– there were significant negative correlations between the medicated group in the two-year lipid cohort for cholesterol and cholesterol esters in small low-density lipoproteins, cholesterol and free cholesterol in intermediate-density lipoproteins, cholesterol and free cholesterol in large low-density lipoproteins, cholesterol and cholesterol esters in medium low-density lipoproteins, and SM. These same findings overlap with the overall two-year findings. Significant differences between the medicated and unmedicated participants were found for high blood pressure and high cholesterol diagnosis, and age of onset for smoking cigarettes (medicated cohort started a little later). No significant difference between the two groups was found for antidepressant or oral contraception use.

6.2.6 Multiple regression analysis for two-year lipid cohort

Multiple regression was conducted to control for confounding factors that may interfere with the relationship observed between increasing depression severity and lipid levels. Here, all the predictor variables were added, and through the forward selection method selected the significant variables. All other non-significant variables were excluded from the model. The multiple regression model was found significant at $P<0.0001$.

Multiple regression analysis revealed that more than one third of variation could be explained by the predictor variables, or the variables that had the strongest associations with HADS scores for this cohort (Table 6.4). The predictors accounted for 34.39% of the overall variance in the HADS scores. IDLFC, SLDLL, and LDLC revealed substantial positive relationship with HADS score, in contrast to the other lipid subclasses, which all possessed strong negative associations with HADS score. Smoking and consuming supplements have a positive association with HADs while interleukin-1B and exercise have negative associations with HADs. Thus, to alleviate symptoms of depression, increasing levels of interleukin-1B and engaging in greater physical activity would be beneficial.

The values of SLDLCE and SLDLFC's coefficients– 9.309e+02 and -9.882e+02 respectively, show SLDLCE and SLDLFC having the greatest impact on HADS scores. This suggests that a decrease in HADs of around -9.000e+02 units would be anticipated to occur in response to an increase of one unit in each of these variables. In the same vein, an increase of one unit in smoking is anticipated to result in an increase of 1.088e-02 in the HADS. Increasing exercise by one unit would decrease HADS score by -1.756e-01. The situation was the same regarding interleukin-1β.

The residual standard error is 0.7237, which represents the average deviation of the observed values from the predicted values. It is a measure of the model's accuracy in predicting the dependent variable. The multiple R-squared value is 0.3439, which indicates that approximately 34.39% of the variance in the dependent variable is explained by the independent variables included in the regression model. The adjusted R-squared value is 0.2425, which considers the number of independent variables and the sample size. It is a more conservative measure of the model's goodness of fit, penalizing the inclusion of unnecessary variables. The F-statistic is 3.39 with 15 and 97 degrees of freedom. It tests the overall significance of the regression model by comparing the variability explained by the model to the variability not explained. The p-value associated with the F-statistic is 0.0001321, which is below the conventional significance level of 0.05. This indicates that the regression model is statistically significant, and at least one of the independent variables is significantly related to the dependent variable.

We have the standard error, t-value, and p-value for each coefficient. The statistical significance of the estimated coefficient is shown by its p-value, while the t-value is the ratio of the estimated coefficient to its standard error. Estimated impacts of IDLFC and LLDLL on the dependent variable are shown as coefficients in the regression model. If we leave all other factors constant, a rise in the dependent variable is associated with an increase in IDLFC (because the coefficient for IDLFC is positive). However, the correlation for LLDLL is negative, indicating that, everything else being equal, a rise in LLDLL is correlated with a drop in the dependent variable. The value of LLDLC's coefficient is -6.214e+02. This means that a change of -6.214e+02 (or -621.4) units in the dependent variable is predicted for every 1 unit rise in LLDLC. The statistical significance of this coefficient can be inferred from the standard error, t-value, and p-value. Here, the coefficient is statistically significant (p-value 0.05), indicating a strong association between LLDLC and the outcome variable. MLDLC, SLDLCE, SLDLFC, SLDLPL, FreeC, and DHA all behave in the same manner.

Smoking has a coefficient of 1.988e-02. That's a predicted rise of 0.01988 units in the dependant variable for every unit increase in smoking. There is a significant relationship between smoking and the outcome variable, as indicated by the statistical significance of the coefficient (p value 0.05). Interleukin1B's coefficient is estimated to be -2.163e-01. This implies that a one-unit rise in Interleukin1B should lead to a 0.2163-unit fall in the dependant variable. The p-value for this coefficient indicates that there is a statistically significant association between Interleukin1B and the dependent variable "HADs" (p-value 0.05). In case of exercise the value of coefficient -1.765e-01. That is a drop of 0.1765 units in the dependant variable for every 1 unit increase in exercise. The coefficient for the supplementary variable is also 1.595e-02. This suggests that we can anticipate a 0.01595-unit rise in the dependent variable for every 1-unit increase in Supplement. If the p-value is less than 0.10, then the coefficient is not statistically significant; otherwise, it would be. There appears to be a strong correlation between supplement and the outcome variable.

Table 6.4: Output for multiple linear regression analysis for 2-year lipid cut-off group

```
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.461e+00  1.410e+00   3.874 0.000194 ***
IDLFC          1.005e+02  4.121e+01   2.439 0.016557 *
LLDLL         -5.854e+01  1.834e+01  -3.191 0.001908 **
LLDLC         -6.214e+02  2.655e+02  -2.341 0.021279 *
MLDLC         -6.509e+02  2.627e+02  -2.477 0.014965 *
SLDLL          1.974e+02  6.198e+01   3.185 0.001948 **
SLDLCE        -9.309e+02  2.898e+02  -3.213 0.001785 **
SLDLFC        -9.882e+02  2.930e+02  -3.373 0.001069 **
SLDLPL        -1.056e+02  6.112e+01  -1.727 0.087317 .
LDLC           6.786e+02  2.654e+02   2.557 0.012115 *
FreeC         -6.304e+00  3.088e+00  -2.042 0.043885 *
DHA           -4.093e+00  2.452e+00  -1.670 0.098236 .
Smoking        1.988e-02  8.274e-03   2.403 0.018181 *
Interleukin1B -2.163e-01  7.279e-02  -2.972 0.003732 **
Exercise_151  -1.765e-01  9.658e-02  -1.828 0.070617 .
Supplement     1.595e-02  1.059e-02   1.506 0.135260
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7237 on 97 degrees of freedom
Multiple R-squared:  0.3439,    Adjusted R-squared:  0.2425
F-statistic:  3.39 on 15 and 97 DF,  p-value: 0.0001321
```

A close match between residuals and fitted values (Fig 6.2) verify the regression assumptions to be correct. The regression assumptions were further verified as we observed a straight line of normal Q-Q plot about 45-degree orientation.

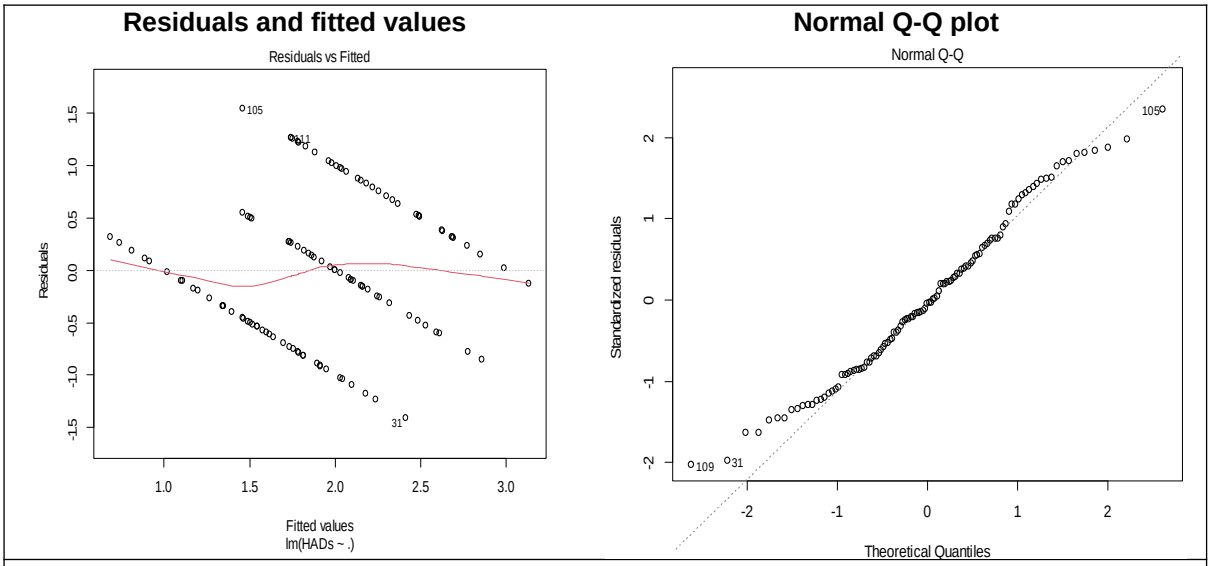


Fig 6.2 Evaluation of multiple regression through i) residuals and fitted values), and normal Q-Q plot

6.2.7 Lipid Correlations in Three-Year Cohort

Following the incorporation of an additional year between time of HADs score and time of blood sample collection (a three-year cohort), nothing was found to be significant between lipid subclasses and depression severity. The number of subclasses that could be included in this analysis was also reduced from 19 to 12. Although significance between each of the lipid subclasses was retained, nothing was found to be significant for any lipid subclass either between the three diagnostic groups or when correlational analysis was performed to look at the associations between HADs scores and lipid levels.

6.3 OPLS-DA Model for Lipids

The OPLS-DA model analyzed a dataset consisting of 78 samples characterized by 19 predictor variables (X) and one response variable. The response variable was utilized for classification purposes, with the "Class" representing the normal and abnormal categories. To ensure consistency, both predictor and response variables were standardized through standard scaling before conducting the OPLS-DA analysis.

The summary statistics of the OPLS-DA model are presented in Table 6.5:

The cumulative R-squared $R^2X(cum)$ value for the predictor variables (X) was 0.975, indicating that the model explained approximately 97.5% of the variance in the predictor variables. This suggests a high level of variance captured by the model in the predictor space. The cumulative R-squared $R^2Y(cum)$ value for the response variable (Y) was 0.09, indicating that the model explained approximately 9% of the variance in the response variable. The cumulative Q-squared $Q^2(cum)$ value in cross-validation was 0.041, indicating the model's predictive power. This value suggests that the model could make accurate predictions of the "Class" labels, differentiating normal vs abnormal classes. The RMSEE was 0.484, representing the average error of the model's predictions on the training data. A smaller RMSEE indicates better predictive performance, so a value of 0.484 suggests reasonable predictive accuracy. The model utilized one predictive component, indicating its focus on separating classes based on predictive variables. The model utilized one orthogonal component, indicating its ability to capture variations unrelated to class separation. In this case, one orthogonal component was deemed sufficient to account for such variations. The cumulative predictive R-squared pR^2Y value (pR^2Y) was 0.30, representing the cross-validated predictive capacity of the model. This value suggests that the model's predictive component(s) could explain approximately 30% of the variance in the response variable. The cumulative Q-squared pQ^2value (pQ2) was 0.1, representing the percentage of variance in the response variable that can be explained by the model's

predictive component(s). Overall, the OPLS-DA model demonstrated high cumulative R-squared for predictor variables, modest Q^2(cum) in cross-validation, and reasonable predictive accuracy, as indicated by RMSEE. The overall framework of the model is illustrated in Fig 6.3

Table 6.5 OPLS-DA summary statistics

R^2X(cum)	R^2Y(cum)	Q^2(cum)	RMSEE	pre	ort	pR^2Y	pQ^2
0.975	0.09	0.041	0.484	1	1	0.30	0.1

78 samples x 19 variables and 1 response
standard scaling of predictors and response(s)

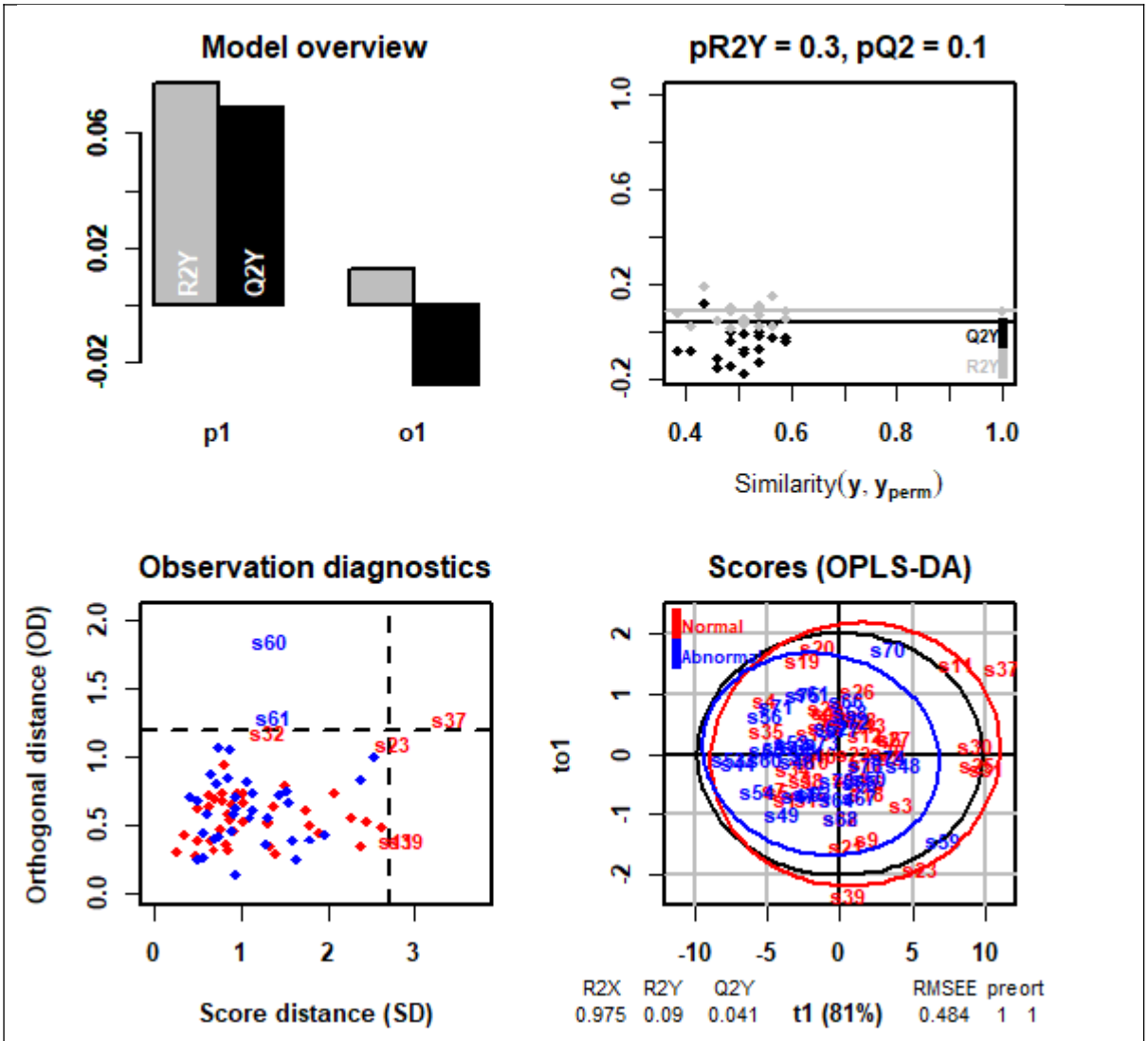


Fig 6.3 OPLS-DA Model overview, similarity index, observation diagnostics and scores for pattern analysis

6.3.1 Comparative Analysis of Model Performance

The Table 6.6 presents the significance of various metrics across different models, including the mean values (Acc, Spec, Sens, R^2X, R^2Y, Q^2) and their respective standard deviations (SD). The mean accuracy for the model is 0.521 with a standard deviation of 0.041, while for the random model, it's 0.495 with a larger standard deviation of 0.070. This suggests that the model generally performs better in terms of accuracy compared to random chance. The mean specificity for the model is 0.516 with a standard deviation of 0.064, while for the random model, it's 0.473 with a higher standard deviation of 0.197. This indicates that the model demonstrates higher specificity on average compared to the random model. The mean sensitivity for the model is 0.550 with a standard deviation of 0.070, whereas for the random model, it's 0.524 with a standard deviation of 0.206. This suggests that the model exhibits slightly higher sensitivity on average compared to the random model. The mean cumulative R-squared R^2X for X variables in the model is 0.989 with a very low standard deviation of 0.001, indicating a high level of explained variance in the predictor variables. The mean cumulative R-squared R^2Y for the response variable (Y) in the model is 0.146 with a standard deviation of 0.025. In comparison, for the random model, the mean R^2Y is 0.063 with a slightly lower standard deviation of 0.028. This implies that the model explains more variance in the response variable compared to the random model. The mean cumulative Q-squared $\mathbf{Q^2}$for the model is 0.136 with a standard deviation of 0.064, while for the random model, it's 0.091 with a similar standard deviation of 0.060. This indicates that the model has a higher predictive ability on average compared to random chance. Overall, the model demonstrates superior performance across various metrics compared to random chance, as evidenced by higher mean values and generally lower standard deviations, indicating more consistent performance. For more comparative in-depth details, refer to Fig 6.4.

Table 6.6 Significance across Spectrum of Metrics

	Acc	SD	Spec	SD	Sens	SD	R^2X	SD	R^2Y	SD	Q^2	SD
model	0.521	0.041	0.516	0.064	0.550	0.070	0.989	0.001	0.146	0.025	0.136	0.064
random	0.495	0.070	0.473	0.197	0.524	0.206	0.966	0.007	0.063	0.028	0.091	0.060

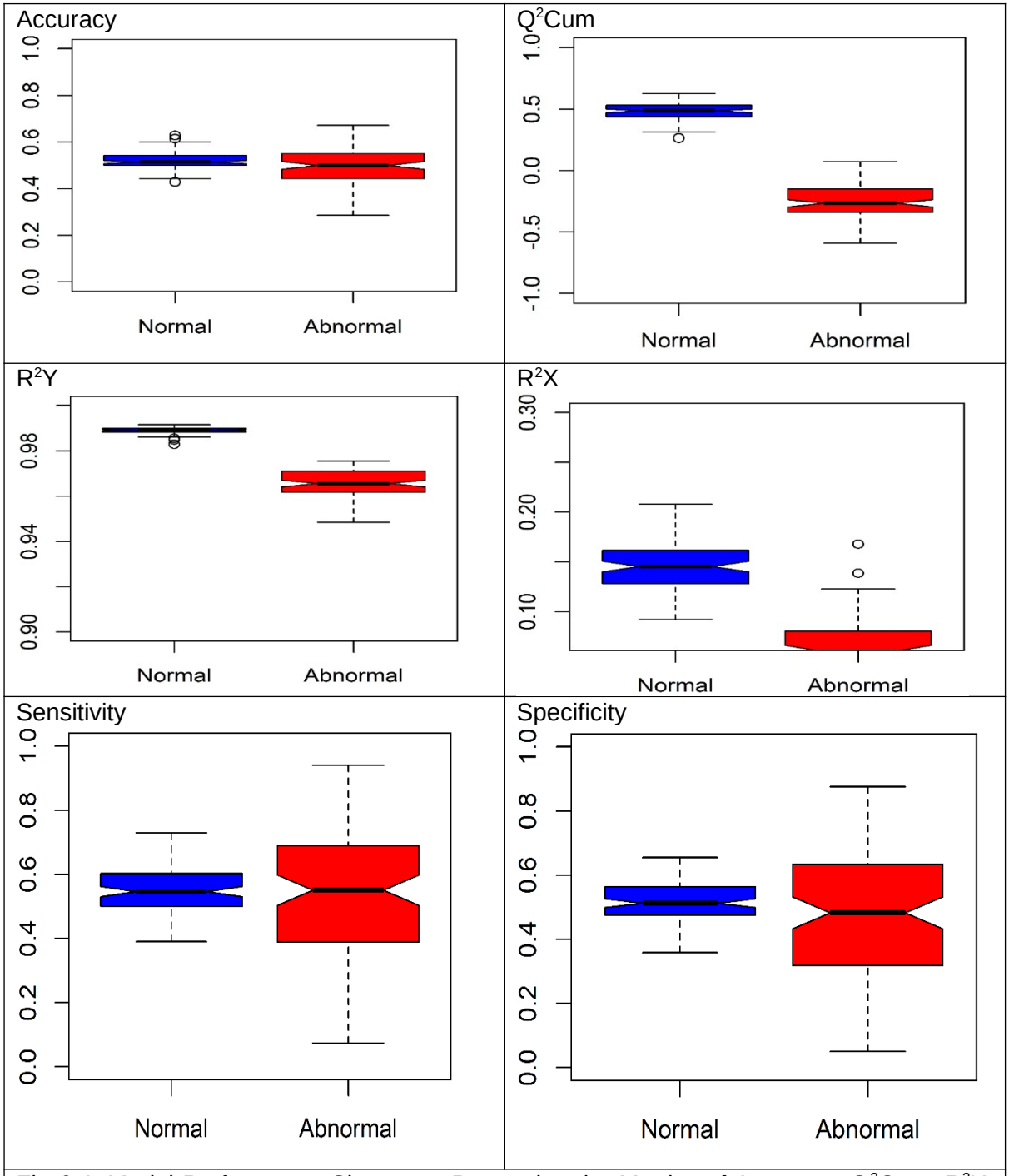


Fig 6.4: Model Performance Showcase Presenting the Metrics of Accuracy, Q^2Cum, R^2X,

R^2Y, Sensitivity, Specificity and Component Distribution

6.3.2 Model Performance Significance Spectrum for Normal versus Abnormal

The Table 6.8 presents the performance indices of the OPLS-DA model for lipids, focusing on accuracy (Acc), specificity (Spec), sensitivity (Sens), cumulative R-squared for predictor variables (R^2X), cumulative R-squared for response variable (R^2Y), and cross-validated predictive power (Q^2). For the t-test, the accuracy (Acc) was 0.000919, indicating a very low misclassification rate. Specificity (Spec) was 0.020556, suggesting that the model correctly identified the proportion of true negatives among all actual negatives. Sensitivity (Sens) was 0.093682, indicating the proportion of true positives identified correctly by the model. The cross-validated predictive power (Q^2) was 1, suggesting perfect prediction capability in cross-validation. For more details look at table 6.7.

Table 6.7: Model Performance Indices

	Acc	Spec	Sens	R^2X	R^2Y	Q^2
t-test	0.000919	0.020556	0.093682	7.76E-61	3.31E-56	1
ks-test	0.000682	1.86E-05	0.000223	3.72E-44	3.98E-35	1

6.3.3 VIP Scores as Analytical Guides and Significance for Normal versus Abnormal

Understanding the importance of variables in data analysis and predictive modeling is fundamental for constructing accurate models. This study investigates variable importance using VIP scores, which quantify each variable's impact on predictive performance. VIP scores, commonly utilized in multivariate data analysis methods like OPLS-DA, assess the contribution of predictor variables to explaining variance in the response variable. Higher VIP scores indicate greater influence on predictive capability, while lower scores suggest lesser importance. These scores aid researchers in identifying relevant variables for prediction, facilitating feature selection, model interpretation, and prioritization for further investigation.

The VIP scores presented in the table represent the relative importance of each variable in predicting the target outcome. A VIP score greater than 1 indicates a significant impact on the model's performance, while scores closer to 1 suggest moderate importance, and scores below 1 signify relatively lesser influence. Each variable in the dataset contributes uniquely to the predictive power of the model, with its importance reflected in the VIP score. "FreeC" emerges as a highly influential variable with a VIP score of 1.04273, indicating its significant impact on prediction accuracy (Fig. 6.5). Similarly, "SM" follows closely with a VIP score of 1.03259, highlighting its substantial contribution to the model. "IDLFC" and "LLDLL" also demonstrate noteworthy importance with VIP scores of 1.0155 and 1.00512, respectively, suggesting their relevance in capturing essential patterns in the data. While variables such as "MLDLC" and "SLDLC" exhibit slightly lower VIP scores, they still contribute significantly to the predictive performance of the model. Each variable's unique contribution underscores the complexity of the predictive process, emphasizing the importance of comprehensive variable analysis in model development and interpretation.

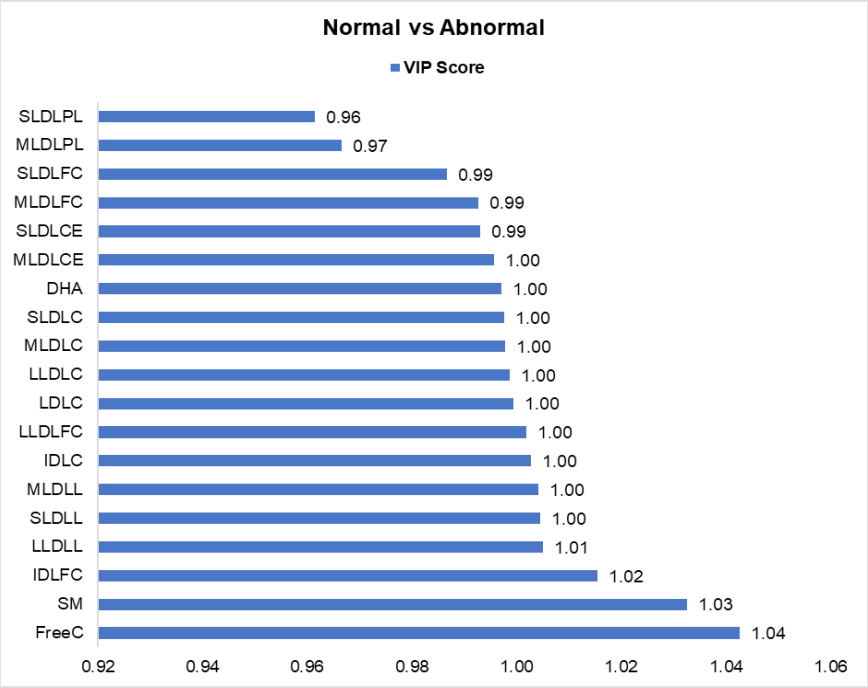


Fig 6.5 Significant Contributors through VIP Scores for Normal versus Abnormal

6.3.4 Inflection point for Normal vs Abnormal

At the inflection point for distinguishing between Normal and Abnormal states, denoted by LLDLL with a VIP score of 1.01, a critical transition in predictive importance is observed (Fig. 6.6). Preceding this inflection point, variables such as FreeC, SM, and IDLFC exhibit VIP scores above 1.00, indicating their substantial impact on the predictive model. However, at the inflection point, LLDLL marks a pivotal shift, where the predictive significance of subsequent variables begins to diminish gradually. Variables following LLDLL, including SLDLL, MLDLL, IDLC, LLDLFC, LDLC, LLDLC, MLDLC, SLDLC, and DHA, exhibit VIP scores of 1.00, suggesting a relatively uniform, albeit diminishing, contribution to the model. Beyond this inflection point, variables such as MLDLCE, SLDLCE, MLDLFC, SLDLFC, MLDLPL, and SLDLPL demonstrate VIP scores below 1.00, indicating a decreasing influence on the predictive distinction between Normal and Abnormal states. This analysis emphasises the importance of LLDLL as a pivotal threshold in the predictive framework, delineating the transition in variable importance and providing valuable insights for model interpretation and refinement.

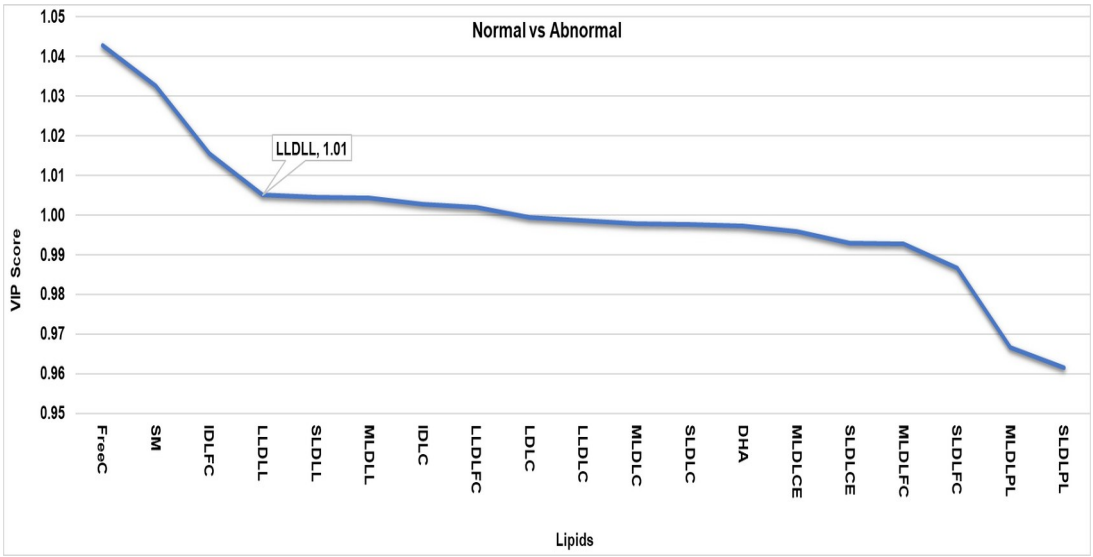


Fig 6.6 Inflection point for Normal vs abnormal

6.3.5 AUC Curve in Distinguishing Normal from Abnormal

In the context of discriminating between different states (Normal from Abnormal), several variables were examined, each with its corresponding AUC score, representing their discriminatory ability. Particularly, "IDLFC," "LLDLL," and "SLDLC" all demonstrated robust discriminatory performance with AUC scores of 0.64 each, indicating their strong ability to distinguish between states observed (Fig. 6.7). Similarly, "IDLC" and "FreeC" exhibited AUC scores of 0.63, suggesting considerable discriminatory power in their respective contexts. These findings highlight the variable-specific differences in discriminatory power, providing value.

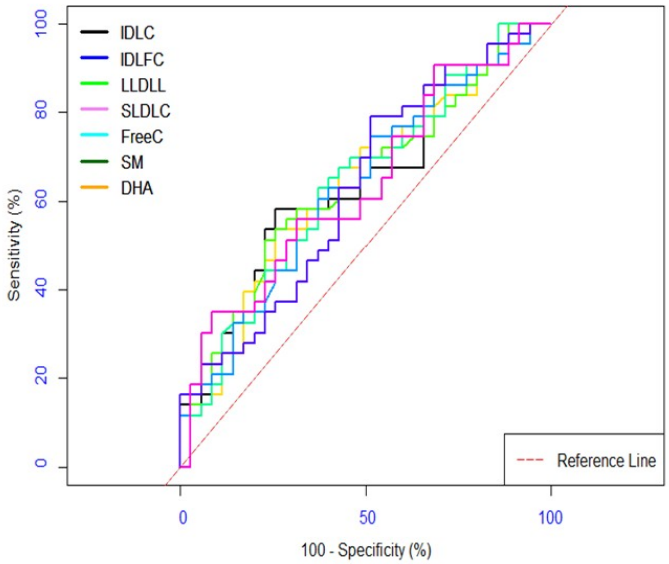


Fig 6.7 AUC values for Normal vs Abnormal
(AUC for IDLC: 0.63, AUC for IDLFC: 0.64, AUC for LLDLL: 0.64, AUC for SLDLC: 0.64, AUC for FreeC: 0.63, AUC for SM: 0.629, AUC for DHA: 0.635).

6.4 Evaluation and Implications

6.4.1 Correlational analyses

The finding for a significant inverse correlation between exercise and depression severity highlights the significance of physical activity in mitigating major depressive disorder, (with accompanying explanations in the preceding chapter), and the finding of glucose levels and both interleukins 1β and 10 negatively correlated with depression severity matched the relationship in the metabolite cohort as well.

The main difference in findings for significant variables when comparing the metabolite cohort lay in that for exercise– though found to be extremely significantly inversely associated with depression severity, its significance was taken away during multiple linear regression, while smoking was maintained as the most positively associated lifestyle variable with increasing depression, alongside levels of interleukin 1β. Smoking, of which relationship with depression is belaboured already in the metabolite chapter, is found to be a frequent habit in depressed female patients, even when controlling for BMI and alcohol consumption, showing its importance as a risk factor of depression., may induce oxidative stress by generating free radicals, and oxidised proteins and lipids, enhancing depressive symptoms (Pasco et al., 2008). Low levels of the inflammatory cytokine IL-1β found to be associated with depression does not fit the traditional view of its relationship with depression since higher levels of IL-1β are implicated particularly in an older depressed cohort (Thomas et al., 2005), matching the demographic of this TwinsUK cohort, however interestingly, a review by Farooq and his co-workers has revealed the importance of this interleukin in stimulating serotonergic transport in the MAPK pathway (Farooq et al., 2016), so findings in this area are still conflicted. As mentioned in the previous chapter, these conflicting findings may point to low levels of IL-1β being a potential new biomarker for major depressive disorder. Lower

physical activity associated with higher depression in the analyses, despite lack of significance in the multiple linear regression analysis, reinforces its importance as a lifestyle variable that should still be integrated into treatment approach. Like the metabolite group, there was a significant difference found for income level between the Normal and Borderline groups; Hinata et al (2021) found income ($p<-0.0001$) and education (OR=0.79, $p=0.0007$) to be independently negatively associated with higher depression levels, with income showing a more robust association (Hinata et al., 2021).

Regarding lipid correlation analysis, it showed significant inverse correlations for all lipid subclasses against HADS scores and there were even significant differences between normal and abnormal groups for every lipid except for SLDLPL, and MLDLPL, see Table 6.1. Cholesterol and cholesterol esters in small and medium LDL particles found to have an inverse relationship with depression severity may indicate that cholesterol is a more influential biomarker for MDD than lipoprotein size or density, since regardless of the latter, the negative correlation between cholesterol levels and higher depression levels was maintained. Low cholesterol has been reported to alter serotonergic levels in depressed patients, thus contributing to higher depression. Cholesterol as one of the three main brain lipids (alongside phospholipids and fatty acids such as polyunsaturated ones like DHA), is also necessary for myelin sheaths to enhance neurotransmission, so its lower levels will alter GABAergic or glutaminergic neurotransmission networks that are implicated in depression.

Countering this finding is a more recent comprehensive overview of studies looking into the influence of cholesterol on depression severity found this claim inconsistent (Zhang et al., 2022). The significant inverse relationship found between LDL levels with depression severity does not match established literature on higher depression severity associated with higher LDL levels (Parekh et al., 2017), except for the phospholipids that were carried in these LDL particles.

However, our findings were strongly supported by the work of Wagner et al. (2019), who made similar conclusions– they discovered that higher total cholesterol (ρ = -0.233, P = 0.010), LDL cholesterol (ρ = 0.235, P = 0.010), and LDL/HDL ratio (ρ = 0.199, P = 0.029) were associated with a more significant decrease in HAM-D score between study inclusion and direct follow-up.

It is not yet entirely clear what the exact pathophysiological process is that connects cholesterol and depression– studies have shown that patients who suffer from major depression (MD) have low HDL-C concentrations as well as higher ratios of total cholesterol to high-density lipoprotein cholesterol (TC/HDL-C) and low-density lipoprotein cholesterol to high-density lipoprotein cholesterol (LDL-C/ HDL-C) (Shibata et al., 1999; Maes et al., 1997; Borgherini et al., 2002; Ross et al., 2007). The findings from most of the investigations point to a possible connection between high cholesterol levels and feelings of depression, and there have been studies showing a lack of remission from depressive symptoms in patients with low LDL cholesterol (Parthasarathy, 2017). General discrepant findings could be explained by a wide variety of study designs and changes in methodology, such as a restricted age range, a small sample size, and heterogeneous study populations (Beydoun et al., 2015; Dortland et al., 2009).

Cholesterol subtypes contained in low and intermediate-density lipoproteins cropping up as the most significantly negatively correlated to depression severity across all analysis methods reveals an important new finding, since previous research emphasises the role of LDL rather than the types of cholesterol, for example cholesterol esters, contained within these lipoproteins. Literature on cholesterol esters specifically is scant, with some studies showing that they are positively associated with depression severity (Mamalakis et al., 2006), although the same paper showed the greater importance of fatty acid concentration in determining depression severity. Cholesterol esters are less polar than free cholesterol and are thus the preferred form of transport in the plasma (Christie, 2023), so it could be

hypothesised that cholesterol esters more so than free cholesterol are better peripheral biomarkers, although these esters are usually rich in arachidonic acid, an inflammatory omega-6 fatty acid which is still needed in healthy amounts to maintain neuronal function and myelin construction. Since peripheral cholesterol can also alter CNS cholesterol concentration as it can cross the blood-brain barrier (BBB) by altering its own structure (Demirkan et al., 2013), cholesterol's importance in developing a lipid biomarker signature for more reliable MDD diagnosis must not be understated.

Significant differences in IDLC, IDLFC, LLDLL, LLDLC, LLDLFC, MLDLL, MLDLC, MLDLCE, MLDLFC, SLDLC, SLDLL, DLDLFC, SLDLCE, LDLC and sphingomyelin between the three diagnostic groups points to these lipid subclasses possessing the potential to be ideal stratification biomarkers for MDD, more so than others, confirming the correlational analyses results. Cholesterol and its variations (free cholesterol and cholesterol esters) consistently show up as significantly inversely correlated with higher depression scores, suggesting its greater importance a biomarker more so than other lipid subclasses such as free fatty acids and phospholipids. Sphingomyelin (a phospholipid) inversely associated with depression is in line with Demirkan's (2013) finding, being implicated in depression severity alongside phosphatidylcholine. A crucial component of brain lipid composition, sphingomyelin is needed for effective cell membrane structure and intercellular signalling (Paul et al., 2023) which when disrupted will interfere with healthy brain circuitry such as that of the kynurenine pathway or cortico-striatal circuit, contributing to elevated depressive symptoms.

6.4.2 Stratification for medication

Stratifying for medication in this cohort revealed no significant difference for taking of either antidepressants or contraception between the medicated and unmedicated, meaning that the relationship between levels of lipids and depression severity was not confounded by

these variables. Instead, these participants were likely medicated for other things such as high cholesterol and high blood pressure, evidenced by the significant difference in those who were afflicted with either of the two conditions.

6.4.3 Multiple regression analysis

Some studies showed that associations found between lipids and depression severity were attenuated to an extent, after adjusting for confounding variables such as BMI, smoking, age and education (van Reedt; Dortland et al., 2009), suggesting that lifestyle factors may play a larger role than an adverse lipid profile in patients afflicted with major depressive disorder, however countering this is the multivariate and multiple regression analyses done in this investigation which showed the very opposite. Opposing the metabolite cohort findings is that for this lipid cohort, where low lipid levels are significantly inversely related to higher depression severity, and outweigh the significance of certain variables, namely: high smoking levels and low IL-1β levels. The latter was the only non-metabolite/lipid-related variable found also to be significantly inversely related to depression severity in the metabolite cohort, underlining its potential as a biomarker.

Multiple regression results further support the correlational cholesterol findings, revealing the overwhelming contribution of cholesterol subtypes more so than other variables, showing the extent to which cholesterol subtypes, specifically cholesterol and cholesterol esters contained in small LDL particles (SLDLFC and SLDLCE, also found significantly associated in the correlational analysis), contribute to increased depression severity even when controlling for potential confounding factors like exercise and use of medication.

DHA being the most significantly negatively associated lipid subtype after cholesterol is in line with literature positing that increased DHA levels are associated with decreasing levels of depression (Wani et al., 2019), potentially due to an increase in grey matter in regions that

regulate depression. Formation of neurotransmitters and prostaglandins is also affected by DHA levels, essential for healthy mood regulation that mitigates depressive symptoms (Liao et al., 2016). As Liao (2016) notes, DHA also prevent the release of proinflammatory molecules which are affected by eicosanoid synthesis, enhancing its important implication in MDD diagnosis.

In a similar vein to the metabolite analysis, the addition of an extra year not adding any extra significant observations for the lipid cohorts could be due to the one year adding ample recovery time, drastically reducing depressive symptoms, during which a patient may have changed his/her environment to be more conducive to mitigating depressive symptoms. Considering the nature of this TwinsUK cohort specifically, in which all the patients were at least middle-aged, it may be noted that patients over a certain age may not share the same pathological processes as a younger group of patients with major depressive disorder.

6.4.4 Conclusions

The systematic review and meta-analysis encompassed 282 studies, finalizing 40 for review and 17 for meta-analysis revealed higher citrate and alanine levels in depressed patients and elevated glutamate in schizophrenic patients. There were more metabolites associated with depression than with schizophrenia. A significant elevation of alanine in plasma and urine was found from people with MDD in 7 studies, which is in-keeping with a previous study showing that circulating alanine correlated with the severity of depression. The elevated concentration of citrate in depression identified with systematic review and meta-analysis supports a metabolic dysfunction in this disorder as it is a key metabolite of the Tricarboxylic Acid (TCA) cycle. Similarly, glutamate showed consistent elevation in schizophrenic patients, though findings were less significant compared to depression.

Smoking exhibited a positive association with depression among those with HADs ($r = 0.211$, $p = 0.027$), as did alcohol consumption ($r = 0.194$, $p = 0.040$). Smoking is associated with higher levels of anxiety and depression symptoms, especially among men. This suggests that smoking may serve as a coping mechanism for stress or contribute to the development of these symptoms. Understanding the gender differences in this relationship is important for tailored interventions.

HADs scores displayed a positive and significant correlation with alcohol use. Higher alcohol consumption was associated with elevated levels of anxiety and depression symptoms. Consistent with prior research, excessive alcohol use is linked to an increased risk of anxiety and depression.

Interleukin-10 and Interleukin-1B levels inversely correlated with depression severity ($r = -0.191$, $p = 0.043$ and $r = -0.204$, $p = 0.030$, respectively), as did exercise ($r = -0.207$, $p = 0.039$). Elevated levels of the anti-inflammatory cytokine IL-10 are linked to reduced symptoms of anxiety and depression. This indicates that IL-10 may play a crucial role in modulating inflammatory responses in the brain, potentially offering a target for therapeutic interventions in mood disorders. HADs scores exhibited an inverse relationship with exercise participation. Regular physical activity was linked to lower levels of anxiety and depression symptoms. The findings align with existing evidence supporting exercise as a beneficial strategy for managing mental health, attributed to the release of endorphins and stress reduction. These conclusions emphasize the influential role of lifestyle factors, such as alcohol use and exercise, in shaping anxiety and depression symptoms. They underscore the importance of promoting healthy behaviours, like regular physical activity, to enhance mental well-being while discouraging excessive alcohol consumption, which may exacerbate these conditions. Supplement-taking showed a strong positive connection with depression severity ($r = 0.219$, $p = 0.027$).

The OPLS-DA model for mental health classification identified cholesterol subtypes, particularly those within small LDL particles, as significant contributors to increased depression severity, even after adjusting for exercise and medication use. FreeC and SM emerged as a crucial lipid subtypes negatively associated with depression severity, aligning with existing literature, and IL-1β maintained significance as a potential biomarker for MDD.

The OPLS-DA model analyzed a dataset of 78 samples with 19 predictor variables and one response variable, demonstrating high explanatory power for predictor variables ($R^2X(cum) = 0.975$) but relatively low explanatory power for the response variable ($R^2Y(cum) = 0.09$). The model showed modest predictive power ($Q^2(cum) = 0.041$) and reasonable predictive accuracy (RMSEE = 0.484), with one predictive and one orthogonal component utilized in the analysis.

Variable Importance in Projection (VIP) scores provided valuable insights into the contribution of each variable to the model's performance, with certain variables such as "FreeC," "SM," "IDLFC," and "LLDLL" demonstrating significant impact on prediction accuracy. The inflection point marked by LLDLL delineated a critical transition in predictive importance, emphasizing the importance of comprehensive variable analysis in model development and interpretation.

In the context of discriminating between Normal and Abnormal states, several variables including FreeC, SM, and IDLFC demonstrated robust discriminatory performance, highlighting variable-specific differences in discriminatory power.

Lipidomics studies revealed negative correlations between various lipid species and anxiety and depression symptoms, suggesting that higher levels of certain lipids may be associated with lower levels of mood disturbances. Alterations in cholesterol metabolism are commonly

observed in individuals with depression, potentially contributing to increased cardiovascular risk.

Comorbidity between depressive and anxiety disorders can influence lipid profiles, indicating distinct lipid profiles in individuals with multiple psychiatric conditions. Low levels of various lipid species have been associated with depression, emphasizing the complexity of lipid involvement in mood disorders. These findings contribute to understanding the relationships between lifestyle factors, inflammatory markers, metabolic factors, lipid profiles, and the severity of Major Depressive Disorder, alongside the identification of potential new biomarkers. Lipidomics studies reveal that various lipid species have negative correlations with anxiety and depression symptoms. This implies that higher levels of certain types of lipids may be associated with lower levels of mood disturbances. Comorbidity between depressive and anxiety disorders can influence lipid profiles, suggesting that individuals with multiple psychiatric conditions may exhibit distinct lipid profiles. This highlights the importance of considering comorbidity in research and clinical practice. Beyond cholesterol, low levels of a range of lipid species, such as glycerolipids, sphingolipids, glycerophospholipids, and triglycerides, have been associated with depression. This complexity underscores the need for comprehensive research to better understand the role of lipids in mood disorders.

These findings contribute significantly to our comprehension of the tangled interconnections among lifestyle factors, inflammatory markers, metabolic factors, and lipid profiles concerning the severity of Major Depressive Disorder (MDD). Most importantly, at the inflection point represented by LLDLL, there is a notable shift in predictive significance, with subsequent variables demonstrating diminishing contributions to the model. This phenomenon, observed in variables such as SLDLL, MLDLL, IDLC, LLDLFC, LDLC, LLDLC, MLDLC, SLDLC, and DHA, underscores the necessity of considering the collective impact of

multiple biomarkers in predictive modeling. It emphasizes the requirement for comprehensive approaches to accurately capture the complexity of depression pathophysiology, ensuring more effective diagnostic and therapeutic strategies tailored to individual needs.

Moreover, the identification of potential new biomarkers, such as reduced levels of FreeC, SM, and IDLFC, alongside parameters like heart rate variability, metabolic flexibility, and VO2 max, sheds light on the intricate interplay between physiological and psychological aspects in depression. These emerging biomarkers not only provide valuable insights into the underlying mechanisms of MDD but also hold promise for the development of a blood marker signature specific to the disorder. Such a signature could significantly enhance diagnostic precision and pave the way for personalized treatment interventions, thereby improving clinical outcomes and patient care in the realm of depression management.

7. Discussion

Major Depressive Disorder is a prevalent mental health ailment that impacts a considerable portion of the population, estimated at around 300 million individuals (Winter et al., 2018). Common symptoms of depression encompass the loss of interest in once-enjoyed activities, alterations in appetite, disruptions in sleep patterns, persistent feelings of sadness, and a sense of worthlessness (Wallace & Milev, 2017). When left untreated or poorly managed, depression could lead to severe consequences, including thoughts of suicide or even suicide attempts (Xu et al., 2010).

However, despite its widespread prevalence and impact, the precise underlying causes of depression remain elusive, and there are still no universally accepted biomarkers for its diagnosis. This situation underscores the importance of investigating and pinpointing the primary contributing factors to depression, as this will enhance diagnostic methods and the effectiveness of treatments for this intricate mental health disorder (Horato et al., 2022). With this background in mind, the central objective of this study was to precisely classify individuals into three distinct groups: Normal, Borderline abnormal, and Abnormal. To achieve this, the study employed OPLS-DA, a technique explored by Vajargah et al., 2014, and Gabrielsson et al., 2006, who addressed the challenges related to multicollinearity in multivariate data and emphasized OPLS-DA's proficiency in eliminating irrelevant systematic changes within the X matrix, or predictor variables. Sadeghi et al., (2011) and Thacker et al., (2007) validated and provided interpretation for the model's performance. The utilization of R^2 as a measure of goodness of fit, Q^2 as an indicator of predictive capacity, and the application of CV-ANOVA showcased adherence to established practices. This rigorous validation enhanced the credibility and reliability of the model in capturing variations and making precise predictions.

The discussion on Variable Importance in Projection (VIP) scores was enriched by insights from Banoei et al., (2023), who explained the significance of VIP scores in identifying variables contributing to group differentiation. By providing guidelines for assessing variable importance using VIP scores, these references supported the assertion that psychological indicators, lipid profiles, inflammatory markers, fatty acids, and HDL subtypes played a crucial role in explaining observed variations. The findings from this analysis have far-reaching implications for both research and practical applications. The validated OPLS-DA model's potential to differentiate between the two groups opened doors to diagnostic, prognostic, and even therapeutic avenues. The identified influential variables can guide future studies that delve into the biological, psychological, and lifestyle factors contributing to the observed distinctions.

The OPLS-DA model for lipids analyzed a dataset of 78 samples with 19 predictor variables and one response variable, demonstrating high explanatory power for predictor variables ($R^2X(cum) = 0.975$) but relatively low explanatory power for the response variable ($R^2Y(cum) = 0.09$). The model showed modest predictive power ($Q^2(cum) = 0.041$) and reasonable predictive accuracy (RMSEE = 0.484), with one predictive and one orthogonal component utilized in the analysis. These findings aligned with previous research showcasing the efficacy of OPLS-DA in multivariate data analysis (Chong & Jun, 2005; Westerhuis et al., 2008). The incorporation of orthogonal components in the model further underscored its capacity to account for unrelated variance, enhancing its reliability in practical applications (Trygg & Wold, 2002). Variable Importance in Projection (VIP) scores provided valuable insights into the contribution of each variable to the model's performance, with certain variables such as "FreeC," "SM," "IDLFC," and "LLDLL" demonstrating significant impact on prediction accuracy. These values underscored their discriminatory power and significant roles in achieving high classification accuracy and robust model performance (Fawcett, 2006; Hand, 2009). Multivariate analysis revealed a drastic drop on accuracy from around 80% to 50% once the variables Depression and Anxiety, characteristic clinical references

from the HADS scale, were excluded. This means that these two variables are important in diagnosing major depressive disorder, and looking at just metabolites, lipids or lifestyle variables cannot replace clinical questionnaires in diagnosing patients. That said, the striking findings for certain lipids as being significant contributing factors to depression severity for all comparisons of the three classes show that lipid level measurements, alongside dietary guidelines involving increased consumption of these lipids, would do well as adjuncts to current clinical diagnostic measurements for major depressive disorder.

Strikingly, and as proven also in previous univariate and multilinear correlations, lipid profiles have a profound relationship with depression severity, as proven by the VIP and AUC scores across the board for all class comparisons. There is scant research on the relationship between low lipid levels and higher depression severity, so this newer finding underlines the importance of one's lipid profile in ascertaining the severity of his/her illness. Although previous research has shown low levels of DHA and EPA linked to higher depression severity, there has not been enough research showing the relationship between other lipid subclasses and depression severity. This has implications for dietary and supplement guidelines to aid the therapeutic response of major depressive disorder and should be used as adjuncts to current antidepressant treatment.

The literature puts forward a hypothesis: low cholesterol levels may impact serotonin sensitivity by reducing the density of 5-HT receptors (Scanlon et al., 2001). In this study, variables like "FreeC," "SM," "IDLFC," and "LLDLL", or mostly low-density lipoproteins, could potentially influence serotonin function and thus increase the severity of depressive symptoms. Future research could investigate the specific relationships between these lipid variables and serotonin receptors in the context of depression. Cholesterol-related changes in inflammatory profiles, including the release of pro-inflammatory cytokines like IL-6 and TNF, are also discussed by Kurano et al., (2011) and Torres et al., (2014), and may overlap

with reduced serotonin function in contributing to the observed relationship between low cholesterol and high depression severity.

This study identifies variables like SM (sphingomyelin), and low levels of SM are in line with literature stating that sphingomyelin is required for the synthesis of ceramide, of which an imbalance of concentration in the brain can lead to premature cell death and brain degeneration, which could explain the worsened depression symptoms and altered brain structure in afflicted individuals.

Literature has mentioned that individuals with depression often have abnormal lipid profiles, including lower total serum cholesterol, lower HDL, and higher LDL/HDL ratios (Terao et al., 2000; Horsten et al., 1997; Rabe-Jaboska et al., 2000). In this study, we've found that certain lipid-related variables, such as FreeC, SM, IDLFC, LLDLL, SLDLL, MLDLL and DHA is associated with depression severity. These variables align with the patterns mentioned in the literature, indicating that lipid profiles indeed play a role in depression severity. The literature highlights the connection between higher BMI, depression symptoms, and higher LDL levels (Guardiola et al., 2015). In this study, we could not find association of BMI other lipid variables may be due to a reasonable number of missing data of BMI.

The serotonin (5-HT) system, a primary focus of antidepressant treatment, may be at the heart of the opposite relationship between serum cholesterol and suicidality due to depression (Scanlon et al., 2001). Cholesterol has been demonstrated to play a function in regulating serotonin sensitivity, with studies linking low cholesterol content in cell membranes to a lower density of 5-HT receptors (Scanlon et al., 2001; Heron et al., 1980; Sun et al., 1980). Cholesterol's effect on 5-HT1A receptor sensitivity in the pathophysiology and treatment of depression was further proven in a rat model of chronic mild stress (CMS; Sun et al., 1980). This effect was most pronounced in the prefrontal brain. However, the serotonin hypothesis is insufficient because antidepressants that target the serotonergic

system do not work for many people (de Sousa et al., 1980; Stewart et al., 2009; Maes, 2008). Cholesterol-related changes to the inflammatory profile are another possible explanation. Atherosclerosis immunological upregulation is aided by high LDL and low HDL. Animal models and in vitro studies show that high cholesterol concentrations enhance the release of pro-inflammatory cytokines such as interleukin 6 (IL-6) and tumour necrosis factor alpha (TNF) (Kurano et al., 2011; Torres et al., 2014). Increased levels of these cytokines have been linked to depression and have been shown to decrease with the use of antidepressant medication (Torres et al., 2014; Liu et al., 2012). Interleukin 1 (IL-1), another cytokine elevated in depressive individuals (Lin et al., 2003; Ovaskainen et al., 2000), has also been linked to oxidised LDL therapy. Importantly, patients with depression frequently have abnormal lipid profiles, including lower total serum cholesterol, lower HDL, and higher LDL/HDL ratios (Terao et al., 2000; Horsten et al., 1997; Rabe-Jaboska et al., 2000).

Depressive symptoms that do not improve after treatment have been linked to reduced total serum cholesterol and LDL cholesterol levels (Lehto et al., 2010; FREITAS et al., 2002). Changes in HDL and LDL apolipoprotein levels, as well as those of its constituent apolipoproteins, have been associated with major depressive disorder. Compared to healthy controls, those with depression had higher levels of LDL and its apolipoprotein B (apoB), but lower levels of HDL and apoA (Sadeghi et al., 2010)- please link this to findings- because for example in one class comparison (normal vs abnormal), it was low levels of LDL, not high, that are linked to higher depression etc. Yet other research contradicts these conclusions, arguing that there is no causal link between low blood cholesterol and depression in senior cohorts (Uguz et al., 2004). There is still much uncertainty around the link between cholesterol and depression, with conflicting findings across age groups. Serum levels of very low-density lipoprotein (VLDL) have been shown to rise without concomitant increases in total cholesterol or high-density lipoprotein (HDL) (Huang et al., 2005). The need for monitoring total cholesterol levels has been highlighted by a meta-analysis showing an inverse connection between total cholesterol and depression, especially in medically naive

groups (Shin et al., 2008). Another meta-analysis (Persons et al., 2016) discovered an inverse correlation between circulating LDL levels and the prevalence of depression. However, additional large-scale studies with several cohorts are necessary to better understand the cholesterol-depression connection.

The association between low-density lipoprotein (LDL) cholesterol and depression may be moderated by factors such as body mass index (BMI). Higher body mass index (BMI) has been connected to depressed symptoms (Guardiola et al., 2015), which is in turn associated with higher LDL levels. In fact, it's not just depression and bipolar illness; other mental health issues, like generalised anxiety disorder (GAD), have also been linked to abnormal cholesterol levels. Sevincok et al. (2001) and Roohafza and colleagues (2010) found that GAD is associated with increased total serum cholesterol and LDL and decreased HDL. Comorbid MDD and GAD patients have unique lipid profiles compared to patients with either condition alone (Roohafza et al., 2010; Sevincok et al., 2001). This provides support for the idea that a healthy range for cholesterol levels may exist, and that these levels may play a role in warding off both depression and anxiety.

Cholesterol supplements as a treatment for major depressive disorder are not likely to gain traction in the medical community. While aggressive cholesterol therapy has been shown to be useful in the treatment of cardiovascular disease (CVD), its use in the treatment of depression through diet alone remains controversial. In addition to conventional antidepressants, statins (drugs that inhibit cholesterol synthesis) have been demonstrated to have positive effects in the treatment of depression (Ghanizadeh & Hedayati, 2013; Haghighi et al., 2014). The correlation between depression and cholesterol levels is complex, thus studies focusing on the functions of other lipid species in depression are also warranted. Nevertheless, lipid supplements and altered dietary guidelines should be considered mainstay adjuncts to current antidepressant treatment.

The potential of free cholesterol (FreeC), sphingomyelin (SM), intermediate density lipoprotein free cholesterol (IDLFC), large low-density lipoprotein lipids (LLDLL), small dense low-density lipoprotein lipids (SLDLL), medium low-density lipoprotein lipids (MLDLL), and docosahexaenoic acid (DHA) in expressing differences between normal and abnormal groups in depression research is supported by various studies. FreeC and SM have been associated with membrane structure and neuronal signaling, as indicated by previous research (Walther et al., 2018; Buschiazzo et al., 2017). Alterations in IDLFC levels may reflect dysregulated lipid metabolism, contributing to depression pathology (Rhee et al., 2017). Furthermore, studies have highlighted the significance of LLDLL, SLDLL, and MLDLL in influencing lipid homeostasis and cardiovascular risk factors, which are interconnected with depression (Lemaitre et al., 2018; van Himbergen et al., 2012). Additionally, DHA, an omega-3 fatty acid, plays a crucial role in brain function and has been linked to mood disorders, suggesting its relevance as a biomarker in depression research (Dyall, 2015; Grosso et al., 2014). These findings collectively underscore the potential of these lipid subclasses as biomarkers for distinguishing between normal and abnormal states in the context of depression, providing valuable insights for future diagnostic and therapeutic strategies.

Inflammatory and Immunological Factors: Cholesterol levels are known to influence the inflammatory and immunological responses in the body (Kurano et al., 2011; Torres et al., 2014). Depression has been linked to increased levels of pro-inflammatory cytokines like interleukin 6 (IL-6) and tumor necrosis factor alpha (TNF) (Torres et al., 2014; Liu et al., 2012). These inflammatory factors may contribute to the development and severity of depressive symptoms. The presence of specific cholesterol subtypes like FreeC, SM, IDLFC, LLDLL, and SLDLL might reflect this inflammatory aspect of depression.

Relationship Between Cholesterol and Brain Function: Cholesterol plays a role in brain function, including neurotransmitter regulation. The serotonin (5-HT) system, which is implicated in depression, has been linked to cholesterol sensitivity (Scanlon et al., 2001;

Heron et al., 1980; Sun et al., 1980). Changes in cholesterol levels, including FreeC, SM, IDLFC, LLDLL, and SLDLL, could potentially affect neurotransmitter function, which might be relevant to depressive symptoms.

Diagnostic Potential: Since FreeC, SM, IDLFC, LLDLL, and SLDLL are identified as potential biomarkers for distinguishing between normal and abnormal states in depression research, their measurement holds significant diagnostic potential. These lipid subclasses offer insights into lipid profile abnormalities associated with depression pathology, providing valuable indicators for diagnostic assessments (Sadeghi et al., 2010). Abnormalities in cholesterol subfractions, such as FreeC, SM, and various LDL subclasses, may serve as diagnostic markers or indicators of depressive disorders, offering clinicians additional tools for accurate diagnosis and treatment planning. By incorporating lipid profiling alongside clinical assessments, healthcare practitioners can enhance diagnostic accuracy and develop personalized treatment strategies tailored to individual lipid profiles, thereby improving patient outcomes in depression management. Further research exploring the diagnostic utility of these lipid biomarkers in diverse populations and clinical settings is warranted to validate their efficacy and utility in depression diagnosis.

It's important to note that the exact mechanisms underlying the relationship between FreeC, SM, IDLFC, LLDLL, and SLDLL and depression may require further investigation, and the relationship between cholesterol and depression is complex with conflicting findings in different studies (Uguz et al., 2004; Shin et al., 2008; Persons et al., 2016). However, the presence of FreeC, SM, IDLFC, LLDLL, SLDLL, MLDLL, and DHA in the model's discriminatory power suggests its potential importance in distinguishing between normal and abnormal groups in the context of depression. These lipid subclasses may reflect underlying physiological processes or metabolic dysregulations associated with depressive disorders, contributing to their discriminatory ability in diagnostic assessments. Further research exploring the intricate interactions between lipid metabolism and depression pathophysiology

is warranted to elucidate the underlying mechanisms and validate the diagnostic utility of these lipid biomarkers in clinical practice.

The findings from the analysis of metabolite chapter in this study reveal several noteworthy relationships, shedding light on the complex interplay between various factors, and anxiety and depression symptoms (HADs). Although no metabolite was found to be significantly correlated with HADS scores, one significant correlation observed in our study is the positive coefficient ($r = 0.211$) between smoking and HADs, indicating that smokers tend to exhibit higher levels of anxiety and depression symptoms than non-smokers. This finding aligns with previous research suggesting that smoking may serve as a coping mechanism for stress and anxiety or that the adverse health effects of smoking contribute to depressive and anxious states (Bonnet et al., 2005). Interestingly, Bonnet et al. found gender differences in the relationship, with anxiety and depression being correlated with smoking in men but only depression being linked to smoking in women.

The study revealed significant associations between the Hospital Anxiety and Depression Scale (HADs) scores and two important lifestyle factors: alcohol use and exercise. HADs scores showed a positive and significant correlation with alcohol use ($r = 0.194$, $p = 0.040$ for metabolites; $r= 0.198$, $p = 0.037$ for lipids), indicating that individuals who reported higher alcohol consumption tended to have elevated levels of anxiety and depression symptoms. This finding is in line with previous research highlighting the link between alcohol consumption and mental health issues. Excessive alcohol use has been associated with an increased risk of anxiety and depression (Boden & Fergusson, 2011; Swendsen et al., 2009). It's important to note that while some individuals may turn to alcohol as a means of coping with stress or emotional distress, alcohol can ultimately exacerbate these mental health conditions over time. Conversely, the study found an inverse relationship between

HADs scores and exercise (r = -0.21, p = 0.038 for metabolites; r -0.207, p=0.039 for lipids). In other words, individuals who engaged in regular physical activity tended to have lower levels of anxiety and depression symptoms. This result aligns with a substantial body of research demonstrating the mental health benefits of exercise (Schuch et al., 2018; Stubbs et al., 2017). Physical activity has been shown to release endorphins, reduce stress hormones, and improve overall mood, making it an effective strategy for managing anxiety and depression symptoms.

Interleukin-10 (IL-10) exhibited a negative coefficient (r = -0.191, P<0.05), indicating that elevated levels of this anti-inflammatory cytokine are associated with reduced symptoms of anxiety and depression. These findings suggest a potential role for IL-10 in mitigating the risk of depression by modulating inflammatory responses. Interleukin-10 (IL-10) is a cytokine that acts as an anti-inflammatory agent. It can suppress the production of proinflammatory cytokines, which are involved in inflammation and tissue damage. IL-10 also has various effects on the brain, such as regulating neurogenesis, protecting neurons, and modulating memory and synaptic plasticity in the hippocampus (Kenis & Maes, 2002). Therefore, IL-10 may have a beneficial role in preventing neuronal injury and enhancing brain recovery. IL-10 has been shown to prevent the morphological changes and activation of glial cells, which are the support cells of the brain. Glial cell activation can lead to the release of proinflammatory cytokines, enzymes, free radicals, and neurotoxic substances, which can cause neuroinflammation and neuronal death (Ogłodek, 2022; Wang et al., 2022). IL-10 can thus control the activity of microglia, which are the immune cells of the brain, and stop them from producing harmful substances and triggering neuroinflammation.

The role of interleukin-1 beta (IL-1β), a prominent proinflammatory cytokine, in neuroinflammatory processes has been well-established in previous studies (Basu et al., 2004; Mittli et al., 2023). Moreover, both IL-1α and IL-1β exhibit expression across a wide array of tissues and immune cells, instigating pro-inflammatory transcription pathways by

activating the shared receptor complex IL-1R1:IL-1R3 (Wang et al., 2010; Korherr et al., 1997).

We have identified a statistically significant inverse correlation between interleukin-1B (IL-1β) and HADs ($r = -0.204$, $p = 0.030$). This implies that elevated levels of IL-1β correspond to lower scores on the HADs scale, and conversely, lower IL-1β levels are associated with higher scores. Nevertheless, intriguingly, our findings align with a study conducted by Ovaskainen et al., which examined IL-1β levels in males and concluded that reduced levels of this cytokine and increased levels of the interleukin-1 receptor antagonist (IL-1RA) were more prevalent in individuals with depression. This may hint at IL-1β levels being a unique biomarker in MDD individuals, its significance in the OPLS-DA analysis confirmed this finding with VIP score >1. The mention of a VIP score >1 suggests that the finding has been confirmed using Variable Importance in Projection (VIP) analysis. VIP is a statistical method employed in multivariate analysis, such as in the context of metabolomics or other high-dimensional data. In this case, a VIP score greater than 1 indicates that the variable (possibly IL-1β or others mentioned) is deemed important in explaining the observed variation in the data. It suggests that the identified biomarkers or variables are influential and contribute significantly to the model or analysis, reinforcing their relevance in the context of the research.

Furthermore, the analysis revealed a negative coefficient for glucose ($r = -0.221$, $P<0.05$), suggesting that higher blood glucose levels are associated with a reduction in anxiety and depression symptoms. This observation may be attributed to the pivotal role of glucose in providing essential energy to the brain and body (Smith & Johnson, 2020; Kujawski et al., 2018).

7.1 Future hypotheses for further Research

Based on the findings presented, several hypotheses can be formulated for future research:

Impact of Smoking and Alcohol on Depression:

Hypothesis: Gender differences exist in the association between smoking/alcohol consumption and depression, with smoking being more strongly correlated with depression in men compared to women. This relationship may be mediated by coping mechanisms for stress or other psychological factors.

Inflammatory Markers and Exercise in Depression:

Hypothesis: Regular exercise modulates inflammatory responses, as indicated by the inverse correlation between exercise participation and interleukin levels. Further investigation may reveal the underlying mechanisms by which exercise influences mood through inflammatory pathways.

Supplement-Taking and Depression Connection:

Hypothesis: Certain types of supplements may exacerbate depressive symptoms, possibly through biochemical interactions or underlying health conditions. Investigating specific types and dosages of supplements and their association with depression severity could provide understandings into potential risk factors.

OPLS-DA Model for Mental Health Classification:

Hypothesis: Lipid variables identified by the OPLS-DA model may serve as biomarkers for distinguishing between normal and abnormal mental health classes. Further validation studies are needed to confirm the discriminatory ability of these lipid profiles across diverse populations.

Lipidomics and Mood Disorders:

Hypothesis: Specific lipid species play a role in the pathophysiology of mood disorders, with some lipids exhibiting negative correlations with anxiety and depression symptoms.

Investigating the biological mechanisms underlying these associations could inform targeted therapeutic interventions.

Cholesterol and Depression:

Hypothesis: Dysregulation of cholesterol metabolism contributes to the pathophysiology of depression, possibly through alterations in neurotransmitter function or inflammatory processes. Longitudinal studies exploring the temporal relationship between changes in cholesterol levels and depressive symptoms are warranted.

Comorbidity and Lipid Profiles:

Hypothesis: Comorbidity between depressive and anxiety disorders influences lipid profiles, suggesting a distinct metabolic profile in individuals with multiple psychiatric conditions. Exploring the interplay between psychiatric comorbidity and lipid metabolism could provide visions into shared biological pathways.

Diverse Lipid Species in Depression:

Hypothesis: Beyond cholesterol, various lipid species are associated with depression, indicating a complex interplay between lipid metabolism and mood regulation. Targeted lipidomic analyses may identify novel biomarkers for depression diagnosis and treatment.

These hypotheses highlight the importance of further investigating the complex relationships between lifestyle factors, inflammatory markers, lipid profiles, and depression severity. By elucidating the underlying mechanisms and identifying potential biomarkers, future research can contribute to the development of personalized diagnostic and therapeutic strategies for Major Depressive Disorder.

7.2 Future Research Directions and Implications:

Longitudinal Studies: Future research should focus on longitudinal studies to better understand the causal relationships between lifestyle factors, inflammatory markers, metabolic factors, and depression severity. Longitudinal designs can help determine whether changes in these factors precede or follow changes in depression symptoms.

Pathways Studies: Pathways studies are needed to elucidate the underlying biological mechanisms linking lipid metabolism, inflammatory pathways, and mood disorders. Investigating the molecular pathways involved can provide insights into potential therapeutic targets for depression.

Intervention Studies: Interventional studies exploring the effects of lifestyle modifications, such as exercise programs or dietary interventions, on lipid profiles and depression severity are warranted. These studies can help establish evidence-based interventions for managing depression and improving overall mental well-being.

Biomarker Validation: Further validation of potential biomarkers identified in this research, such as specific lipid species and inflammatory markers, is essential. Large-scale validation studies across diverse populations can confirm the reliability and clinical utility of these biomarkers for diagnosing and monitoring depression.

Gender-Specific Research: Given the observed gender differences in the relationship between smoking, alcohol consumption, and depression, future research should investigate gender-specific factors contributing to depression risk. Understanding these differences can inform tailored interventions for men and women.

Psychiatric Comorbidity: Research focusing on the impact of psychiatric comorbidity on lipid profiles and depression severity is crucial. Examining lipid profiles in individuals with multiple psychiatric conditions can provide insights into shared biological pathways and potential treatment implications.

Integration of Omics Approaches: Integrating lipidomics with other omics approaches, such as genomics and proteomics, can offer a comprehensive understanding of the molecular basis of depression. Multi-omics studies may uncover novel biomarkers and therapeutic targets for personalized treatment strategies.

Thus, future research should adopt a multidisciplinary approach, including longitudinal studies, mechanistic investigations, intervention trials, biomarker validation, gender-specific research, psychiatric comorbidity studies, and integration of omics approaches. By addressing these research gaps, we can advance our understanding of the complex interplay between lipid/metabolite metabolism, inflammation, and depression, leading to improved diagnostic tools and personalized treatment strategies for individuals with depression.

8 References

Abbasi, F. et al. (2021) ‘Statins are associated with increased insulin resistance and secretion’, Arteriosclerosis, Thrombosis, and Vascular Biology, 41(11), pp. 2786–2797. doi:10.1161/atvbaha.121.316159.

Agargün, M. Y., Algün, E., Şekeroğlu, R., Kara, H., & Tarakçioğlu, M. (1998). Low cholesterol level in patients with panic disorder: the association with major depression. Journal of affective disorders, 50(1), 29-32.

Agargun, M. Y., Dulger, H., Inci, R., Kara, H., Ozer, O. A., Sekeroglu, M. R., & Besiroglu, L. (2004). Serum Lipid Concentrations in Obsessive—Compulsive Disorder Patients with and without Panic Attacks. The Canadian Journal of Psychiatry, 49(11), 776-778.

Alpaydin, E. (2020). Introduction to machine learning. MIT press.

Amani, M. (2021) ‘Comparison of Self-Differentiation and Identity Statuses in Twins and Nontwins’, in A. Shariatipour (ed.) Twin research and human genetics. 3rd edn. Cambridge, Cambridge: Cambridge University Press, pp. 176–183.

Andalo, D. (2016) ‘Antidepressants associated with increased risk of suicidal thoughts in healthy adults’, The Pharmaceutical Journal [Preprint]. doi:10.1211/pj.2016.20201834.

Angelucci, F., Brene, S. and Mathé, A., 2005. BDNF in schizophrenia, depression and corresponding animal models. Nature, [online] 10, pp.354-352. Available at: <https://www.nature.com/articles/4001637> [Accessed 9 April 2021].

Anjum, S., Qusar, M. S., Shahriar, M., Islam, S. M. A., Bhuiyan, M. A., & Islam, M. R. (2020). Altered serum interleukin-7 and interleukin-10 are associated with drug-free major depressive disorder. Therapeutic advances in psychopharmacology, 10, 2045125320916655.

Ariyo, A. A., Haan, M., Tangen, C. M., Rutledge, J. C., Cushman, M., Dobs, A., & Furberg, C. D. (2000). Depressive symptoms and risks of coronary heart disease and mortality in elderly Americans. Circulation, 102(15), 1773-1779.

Arnaud, A.M. et al. (2022) “Impact of major depressive disorder on comorbidities,” The Journal of Clinical Psychiatry, 83(6). Available at: https://doi.org/10.4088/jcp.21r14328.

Awaworyi Churchill, S. and Farrell, L. (2017) ‘Alcohol and depression: Evidence from the 2014 health survey for england’, Drug and Alcohol Dependence, 180, pp. 86–92. doi:10.1016/j.drugalcdep.2017.08.006.

Bajwa, W. K., Asnis, G. M., Sanderson, W. C., Irfan, A., & Van Praag, H. M. (1992). High cholesterol levels in patients with panic disorder. The American journal of psychiatry, 149(3), 376-378.

Banoei, M. M., Rafiepoor, H., Zendehdel, K., Seyyedsalehi, M. S., Nahvijou, A., Allameh, F., & Amanpour, S. (2023). Unraveling complex relationships between COVID-19 risk factors using machine learning based models for predicting mortality of hospitalized patients and identification of high-risk group: a large retrospective study. Frontiers in Medicine, 10, 1170331.

Basu, A., Krady, J. K., & Levison, S. W. (2004). Interleukin-1: a master regulator of neuroinflammation. Journal of neuroscience research, 78(2), 151-156.

Belvederi Murri, M., Cervigni, R., Masotti, M., & Respino, M. (2018). The hypothalamic–renal axis in major depression: the role of vasopressin and the risk of hyponatremia. International journal of molecular sciences, 19(11), 3579 .

Beydoun, M. A., Beydoun, H. A., Dore, G. A., Fanelli-Kuczmarski, M. T., Evans, M. K., & Zonderman, A. B. (2015). Total serum cholesterol, atherogenic indices and their longitudinal association with depressive symptoms among US adults. Translational Psychiatry, 5(3), e518-e518.

Bhattacharyya, S., Ahmed, A., Arnold, M., Liu, D., Luo, C., Zhu, H., Mahmoudiandehkordi, S., Neavin, D., Louie, G., Dunlop, B., Frye, M., Wang, L., Weinshilboum, R., Krishnan, R., Rush, A. and Kaddurah-Daouk, R., 2019. Metabolomic signature of exposure and response to citalopram/escitalopram in depressed outpatients. Translational Psychiatry, 9(1).

Bjelland, I., Tell, G., Vollset, S., Konstantinova, S. and Ueland, P., 2009. Choline in anxiety and depression: the Hordaland Health Study. The American Journal of Clinical Nutrition, 90(4), pp.1056-1060.

Bjerkeset, O., Romundstad, P., Evans, J., & Gunnell, D. (2008). Association of adult body mass index and height with anxiety, depression, and suicide in the general population: the HUNT study. American journal of epidemiology, 167(2), 193-202.

Boden, J. M., & Fergusson, D. M. (2011). Alcohol and depression. Addiction, 106(5), 906-914.

Bonnet, F., Irving, K., Terra, J. L., Nony, P., Berthezène, F., & Moulin, P. (2005). Anxiety and depression are associated with unhealthy lifestyle in patients at risk of cardiovascular disease. Atherosclerosis, 178(2), 339-344.

Borgherini, G., Dorz, S., Conforti, D., Scarso, C., & Magni, G. (2002). Serum cholesterol and psychological distress in hospitalized depressed patients. Acta Psychiatrica Scandinavica, 105(2), 149-152.

Borsini, A. et al. (2021) “Omega-3 polyunsaturated fatty acids protect against inflammation through production of LOX and CYP450 lipid mediators: Relevance for major depression and for human hippocampal neurogenesis,” Molecular Psychiatry, 26(11), pp. 6773–6788. Available at: https://doi.org/10.1038/s41380-021-01160-8.

Bortolato, B., Hyphantis, T. N., Valpione, S., Perini, G., Maes, M., Morris, G., ... & Carvalho, A. F. (2017). Depression in cancer: the many biobehavioral pathways driving tumor progression. Cancer treatment reviews, 52, 58-70.

Bouwman, V., Adriaanse, M. C., van’t Riet, E., Snoek, F. J., Dekker, J. M., & Nijpels, G. (2010). Depression, anxiety and glucose metabolism in the general dutch population: the new Hoorn study. PloS one, 5(4), e9971.

Brindle, J. T., Antti, H., Holmes, E., Tranter, G., Nicholson, J. K., Bethell, H. W. L., et al. (2002). Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using H-1-NMR-based metabonomics. Nat. Med. 8, 1439–1444. doi: 10.1038/nm1202-802

Brown, M. S., & Goldstein, J. L. (1986). A receptor-mediated pathway for cholesterol homeostasis. Science, 232(4746), 34-47.

Buschiazzo, J., Ríos, G. L., Canizo, J. R., Antollini, S. S., & Alberio, R. H. (2017). Free cholesterol and cholesterol esters in bovine oocytes: Implications in survival and membrane raft organization after cryopreservation. PloS one, 12(7), e0180451. https://doi.org/10.1371/journal.pone.0180451

Cai, H., Li, H., Yan, X., Sun, B., Zhang, Q., Yan, M., Zhang, W., Jiang, P., Zhu, R., Liu, Y., Fang, P., Xu, P., Yuan, H., Zhang, X., Hu, L., Yang, W. and Ye, H., 2012. Metabolomic Analysis of Biochemical Changes in the Plasma and Urine of

First-Episode Neuroleptic-Naïve Schizophrenia Patients after Treatment with Risperidone. Journal of Proteome Research, 11(8), pp.4338-4350.

Carter, J. V., Pan, J., Rai, S. N., & Galandiuk, S. (2016). ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. Surgery, 159(6), 1638-1645.

Caspani, G., Turecki, G., Lam, R. W., Milev, R. V., Frey, B. N., MacQueen, G. M., ... & Swann, J. R. (2021). Metabolomic signatures associated with depression and predictors of antidepressant response in humans: A CAN-BIND-1 report. Communications biology, 4(1), 903.

Castagné, V., Rougemont, M., Cuenod, M. and Do, K., 2004. Low brain glutathione and ascorbic acid associated with dopamine uptake inhibition during rat's development induce long-term cognitive deficit: relevance to schizophrenia. Neurobiology of Disease, 15(1), pp.93-105.

Castrén, E., & Rantamäki, T. (2010). Role of brain-derived neurotrophic factor in the aetiology of depression: implications for pharmacological treatment. CNS drugs, 24, 1-7.

Celano, C. M., Freudenreich, O., Fernandez-Robles, C., Stern, T. A., Caro, M. A., & Huffman, J. C. (2022). Depressogenic effects of medications: a review. Dialogues in clinical neuroscience.

Chad A Bousman, T., 2013. Metabolomics of Psychotic Disorders. Journal of Postgenomics Drug & Biomarker Development, 03(01).

Chen, J., Bai, S., Li, W., Zhou, C., Zheng, P., Fang, L., Wang, H., Liu, Y. and Xie, P., 2018. Urinary biomarker panel for diagnosing patients with depression and anxiety disorders. Translational Psychiatry, 8(1).

Chen, J., Vargas-Bustamante, A., Mortensen, K., & Ortega, A. N. (2015). Racial and ethnic disparities in health care access and utilization under the Affordable Care Act. Medical Care, 53(4), 327-334.

Chong, I. G., & Jun, C. H. (2005). Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems, 78(1-2), 103-112.

Christie, W. (2023) Sterols: 1. cholesterol and cholesterol esters, Cholesterol and Cholesterol Esters - structure, occurrence, biochemistry and function. Available

at: https://lipidmaps.org/resources/lipidweb/lipidweb_html/lipids/simple/cholest/index.htm (Accessed: 30 May 2023).

Coentre, R., Talina, M., Góis, C. and Figueira, M., 2017. Depressive symptoms and suicidal behavior after first-episode psychosis: A comprehensive systematic review. Psychiatry Research, 253, pp.240-248.

Cohn, J. S., Kamili, A., Wat, E., Chung, R. W., Tandy, S., & Dietary, F. O. S. (2009). Dietary phospholipids, hepatic lipid metabolism and cardiovascular disease. Current Opinion in Lipidology, 20(2), 131-136.

Condray, R., Dougherty, G., Keshavan, M., Reddy, R., Haas, G., Montrose, D., Matson, W., McEvoy, J., Kaddurah-Daouk, R. and Yao, J., 2011. 3-Hydroxykynurenine and clinical symptoms in first-episode neuroleptic-naive patients with schizophrenia. The International Journal of Neuropsychopharmacology, 14(06), pp.756-767.

Courtet, P., 2018. Suicidal Risk During the First Episode of Psychosis. L'Encéphale, 44(6S), pp.39-43.

Czysz, A., South, C., Gadad, B., Arning, E., Soyombo, A., Bottiglieri, T. and Trivedi, M., 2019. Can targeted metabolomics predict depression recovery? Results from the CO-MED trial. Translational Psychiatry, 9(1).

Davison, J., O'Gorman, A., Brennan, L. and Cotter, D., 2018. A systematic review of metabolite biomarkers of schizophrenia. Schizophrenia Research, 195, pp.32-50.

De la Roca-Chiapas, J. M., Hernández-González, M., Candelario, M., Villafaña Mde, L., Hernández, E., Solorio, S., ... & Jasso, J. (2013). Association between depression and higher glucose levels in middle-aged Mexican patients with diabetes. Rev Invest Clin, 65(3), 209-13.

de Menezes Galvao, A. C., Almeida, R. N., de Sousa Jr, G. M., Leocadio-Miguel, M. A., Palhano-Fontes, F., de Araujo, D. B., ... & Galvao-Coelho, N. L. (2021). Pathophysiology of major depression by clinical stages. Frontiers in Psychology, 12, 641779.

de Sousa, R. T., V Zanetti, M., R Brunoni, A., & Machado-Vieira, R. (2015). Challenging treatment-resistant major depressive disorder: a roadmap for improved therapeutics. Current neuropharmacology, 13(5), 616-635.

Dean, J., & Keshavan, M. (2017). The neurobiology of depression: An integrated view. Asian journal of psychiatry, 27, 101-111.

Demirkan, A. et al. (2013) 'Plasma phosphatidylcholine and sphingomyelin concentrations are associated with depression and anxiety symptoms in a Dutch family-based lipidomics study', Journal of Psychiatric Research, 47(3), pp. 357–362. doi:10.1016/j.jpsychires.2012.11.001.

Demirkan, A., Isaacs, A., Ugocsai, P., Liebisch, G., Struchalin, M., Rudan, I., ... & van Duijn, C. M. (2013). Plasma phosphatidylcholine and sphingomyelin concentrations are associated with depression and anxiety symptoms in a Dutch family-based lipidomics study. Journal of psychiatric research, 47(3), 357-362.

Demirkan, A., Isaacs, A., Ugocsai, P., Liebisch, G., Struchalin, M., Rudan, I., Wilson, J., Pramstaller, P., Gyllensten, U., Campbell, H., Schmitz, G., Oostra, B. and van Duijn, C., 2013. Plasma phosphatidylcholine and sphingomyelin concentrations are associated with depression and anxiety symptoms in a Dutch family-based lipidomics study. Journal of Psychiatric Research, 47(3), pp.357-362.

Deussing, J., 2006. Animal models of depression. Drug Discovery Today: Disease Models, 3(4), pp.375-383.

Dietrich-Muszalska, A., Głowacki, R., Olas, B., Bald, E. and Wachowicz, B., 2012. P-1225 - The oxidative stress may be induced by the elevated homocysteine in schizophrenic patients. European Psychiatry, 27, p.1.

Ding, X., Yang, S., Li, W., Liu, Y., Li, Z., Zhang, Y., Li, L. and Liu, S., 2014. The Potential Biomarker Panels for Identification of Major Depressive Disorder (MDD) Patients with and without Early Life Stress (ELS) by Metabonomic Analysis. PLoS ONE, 9(5), p.e97479.

Domenici, E., Willé, D. R., Tozzi, F., Prokopenko, I., Miller, S., McKeown, A., ... & Muglia, P. (2010). Plasma protein biomarkers for depression and schizophrenia by multi analyte profiling of case-control collections. PLoS one, 5(2), e9166.

Duan, J., & Xie, P. (2020). The potential for metabolomics in the study and treatment of major depressive disorder and related conditions. Expert Review of Proteomics, 17(4), 309-322.

Dyall, S. C. (2015). Long-chain omega-3 fatty acids and the brain: a review of the independent and shared effects of EPA, DPA and DHA. Frontiers in aging neuroscience, 7, 52.

Eggers, A., 2013. A serotonin hypothesis of schizophrenia. Medical Hypotheses, 80(6), pp.791-794.

Ehrenreich, M. J. (2006). A case of the re-emergence of panic and anxiety symptoms after initiation of a high-protein, very low carbohydrate diet. Psychosomatics, 47(2), 178-179.

Enko, D., Brandmayr, W., Halwachs-Baumann, G., Schnedl, W. J., Meinitzer, A., & Kriegshäuser, G. (2018). Prospective plasma lipid profiling in individuals with and without depression. Lipids in health and disease, 17(1), 1-6.

Ergün, U. Ö., Uguz, S., Bozdemir, N., Güzel, R., Burgut, R., Saatçi, E., & Akpınar, E. (2004). The relationship between cholesterol levels and depression in the elderly. International journal of geriatric psychiatry, 19(3), 291-296.

Eriksson, L., Trygg, J., & Wold, S. (2008). CV-ANOVA for significance testing of PLS and OPLS® models. Journal of Chemometrics: A Journal of the Chemometrics Society, 22(11-12), 594-600.

Erjavec, G. N., Konjevod, M., Perkovic, M. N., Strac, D. S., Tudor, L., Barbas, C., ... & Pivac, N. (2018). Short overview on metabolomic approach and redox changes in psychiatric disorders. Redox biology, 14, 178-186.

Evans, S., Alkan, E., Bhangoo, J. K., Tenenbaum, H., & Ng-Knight, T. (2021). Effects of the COVID-19 lockdown on mental health, wellbeing, sleep, and alcohol use in a UK student sample. Psychiatry research, 298, 113819.

F. Guerreiro Costa, L.N. et al. (2022) ‘Metabolomics of major depressive disorder: A systematic review of Clinical Studies’, Cureus [Preprint]. doi:10.7759/cureus.23009.

Farooq, R.K. et al. (2016) ‘Role of inflammatory cytokines in depression: Focus on interleukin-1β’, Biomedical Reports, 6(1), pp. 15–20. doi:10.3892/br.2016.807.

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861-874.

Fazel, S., Wolf, A., Palm, C., & Lichtenstein, P. (2014). Violent crime, suicide, and premature mortality in patients with schizophrenia and related disorders: a 38-year total population study in Sweden. The Lancet Psychiatry, 1(1), 44-54

Felig, P., 1973. The glucose-alanine cycle. Metabolism, 22(2), pp.179-207.

Firth, J., Siddiqi, N., Koyanagi, A. I., Siskind, D., Rosenbaum, S., Galletly, C., ... & Stubbs, B. (2019). The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. The Lancet Psychiatry, 6(8), 675-712.

Flint, J. (2023). The genetic basis of major depressive disorder. Molecular Psychiatry, 1-12.

Fluharty, M. et al. (2016) “The Association of Cigarette Smoking with depression and anxiety: A systematic review,” Nicotine & Tobacco Research, 19(1), pp. 3–13. Available at: https://doi.org/10.1093/ntr/ntw140.

Forrest, A., Coto, C. and Siegel, S., 2014. Animal Models of Psychosis: Current State and Future Directions. Current Behavioral Neuroscience Reports, 1(2), pp.100-116.

Fournier, M., Ferrari, C., Baumann, P., Polari, A., Monin, A., Bellier-Teichmann, T., Wulff, J., Pappan, K., Cuenod, M., Conus, P. and Do, K., 2014. Impaired Metabolic Reactivity to Oxidative Stress in Early Psychosis Patients. Schizophrenia Bulletin, 40(5), pp.973-983.

Frazier, C. J., Gokool, V. A., Holness, H. K., Mills, D. K., & Furton, K. G. (2023). Multivariate regression modelling for gender prediction using volatile organic compounds from hand odor profiles via HS-SPME-GC-MS. Plos one, 18(7), e0286452.

Freeman, A. et al. (2016) “The role of socio-economic status in depression: Results from the Courage (Aging Survey in Europe),” BMC Public Health, 16(1). Available at: https://doi.org/10.1186/s12889-016-3638-0.

FREITAS, J. A. D., Lima, L. M. P., Ranieri, J. L., OLIVIERI Jr, C., Fragoso, H. J., & Chinzon, D. (2002). Efficacy, safety and tolerability of rabeprazole in treatment of acid-peptic diseases. Arquivos de Gastroenterologia, 39, 60-65.

Fries, G. R., Saldana, V. A., Finnstein, J., & Rein, T. (2023). Molecular pathways of major depressive disorder converge on the synapse. Molecular Psychiatry, 28(1), 284-297.

Fukushima, T., Iizuka, H., Yokota, A., Suzuki, T., Ohno, C., Kono, Y., Nishikiori, M., Seki, A., Ichiba, H., Watanabe, Y., Hongo, S., Utsunomiya, M., Nakatani, M., Sadamoto, K. and Yoshio, T., 2014. Quantitative Analyses of Schizophrenia-Associated Metabolites in Serum: Serum D-Lactate Levels Are Negatively Correlated with Gamma-Glutamylcysteine in Medicated Schizophrenia Patients. PLoS ONE, 9(7), p.e101652.

Gadad, B.S. et al. (2021) “Altered levels of interleukins and neurotrophic growth factors in mood disorders and suicidality: An analysis from periphery to central nervous system,” Translational Psychiatry, 11(1). Available at: https://doi.org/10.1038/s41398-021-01452-1.

Garip, B. and Kayir, H., 2019. Alteration in NMDAR-related amino acids in first episode psychosis. Synapse, 73(11).

Geng, C., Guo, Y., Wang, C., Liao, D., Han, W., Zhang, J., et al. (2020). Systematic impacts of chronic unpredictable mild stress on metabolomics in rats. Sci. Rep. 10, 700. doi: 10.1038/s41598-020-57566-x

German, J. B., Bauman, D. E., Burrin, D. G., Failla, M. L., Freake, H. C., King, J. C., et al. (2004). Metabolomics in the opening decade of the 21st century: building the roads to individualized health. J. Nutr. 134, 2729–2732. doi: 10.1093/jn/134.10.2729

Ghanizadeh, A., & Hedayati, A. (2013). Augmentation of fluoxetine with lovastatin for treating major depressive disorder, a randomized double-blind placebo controlled-clinical trial. Depression and anxiety, 30(11), 1084-1088.

Gneiting, T., & Walz, E. M. (2022). Receiver operating characteristic (ROC) movies, universal ROC (UROC) curves, and coefficient of predictive ability (CPA). Machine Learning, 111(8), 2769-2797. Receiver Gold, S. et al. (2020) ‘Comorbid depression in medical diseases’, Nature Reviews Disease Primers, 6(1). doi:10.1038/s41572-020-0211-z.

Gong, X., Huang, C., Yang, X., Mao, Q., Zeng, L., Zheng, P., ... & Xie, P. (2019). Proteomic analysis of the intestine reveals SNARE-mediated immunoregulatory and amino acid absorption perturbations in a rat model of depression. Life sciences, 234, 116778.

Gouni-Berthold, J. and Krone, W. (2011) “Lipids, lipoproteins, and atherogenesis,” Hyperlipidaemia, pp. 11–24. Available at: https://doi.org/10.1093/med/9780199543502.003.0002.

Graham, K. et al. (2007) “Does the association between alcohol consumption and depression depend on how they are measured?,” Alcoholism: Clinical and Experimental Research, 31(1), pp. 78–88. Available at: https://doi.org/10.1111/j.1530-0277.2006.00274.x.

Gross, R. W., Jenkins, C. M., Yang, J., Mancuso, D. J., & Han, X. (2005). Functional lipidomics: the roles of specialized lipids and lipid–protein interactions in modulating neuronal function. Prostaglandins & other lipid mediators, 77(1-4), 52-64.

Grosso, G., Pajak, A., Marventano, S., Castellano, S., Galvano, F., Bucolo, C., Drago, F., & Caraci, F. (2014). Role of omega-3 fatty acids in the treatment of depressive disorders: a comprehensive meta-analysis of randomized clinical trials. PloS one, 9(5), e96905. https://doi.org/10.1371/journal.pone.0096905

Gu, X., Ke, S., Wang, Q., Zhuang, T., Xia, C., Xu, Y., ... & Zhou, M. (2021). Energy metabolism in major depressive disorder: Recent advances from omics technologies and imaging. Biomedicine & Pharmacotherapy, 141, 111869.

Guardiola, M., Solà, R., Vallvé, J. C., Girona, J., Godàs, G., Heras, M., ... & Ribalta, J. (2015). Body mass index correlates with atherogenic lipoprotein profile even in nonobese, normoglycemic, and normolipidemic healthy men. Journal of clinical lipidology, 9(6), 824-831.

Gupta, A., Jadhav, A. A., Petkar, S. B., & Dubey, V. (2013). Study of Lipid Derangement in Pyschiatric Disorder.

Haghighi, M., Khodakarami, S., Jahangard, L., Ahmadpanah, M., Bajoghli, H., Holsboer-Trachsler, E., & Brand, S. (2014). In a randomized, double-blind clinical trial, adjuvant atorvastatin improved symptoms of depression and blood lipid values in patients suffering from severe major depressive disorder. Journal of psychiatric research, 58, 109-114.

Han, A.L. (2022) “Association between lipid ratio and depression: A cross-sectional study,” Scientific Reports, 12(1). Available at: https://doi.org/10.1038/s41598-022-10350-5.

Hand, D. J. (2009). Measuring classifier performance: a coherent alternative to the area under the ROC curve. Machine learning, 77(1), 103-123.

Harris, W. S., Connor, W. E., Alam, N., & Illingworth, D. R. (1988). Reduction of postprandial triglyceridemia in humans by dietary n-3 fatty acids. Journal of lipid research, 29(11), 1451-1460.

Hashimoto, K., Ishima, T., Sato, Y., Bruno, D., Nierenberg, J., Marmar, C., Zetterberg, H., Blennow, K. and Pomara, N., 2017. Increased levels of ascorbic acid in the

cerebrospinal fluid of cognitively intact elderly patients with major depression: a preliminary study. Scientific Reports, 7(1).

Hawton, K., Casañas i Comabella, C., Haw, C. and Saunders, K., 2013. Risk factors for suicide in individuals with depression: A systematic review. Journal of Affective Disorders, 147(1-3), pp.17-28.

He, P., Liu, X., Wen, J., & Zhang, Y. (2021). Major clinical advances of depression: Now and future. In E3S Web of Conferences (Vol. 292, p. 03102). EDP Sciences.

He, Y., Yu, Z., Giegling, I., Xie, L., Hartmann, A., Prehn, C., Adamski, J., Kahn, R., Li, Y., Illig, T., Wang-Sattler, R. and Rujescu, D., 2012. Schizophrenia shows a unique metabolomics signature in plasma. Translational Psychiatry, 2(8), pp.e149-e149.

Heron, D. S., Shinitzky, M., Hershkowitz, M., & Samuel, D. (1980). Lipid fluidity markedly modulates the binding of serotonin to mouse brain membranes. Proceedings of the National Academy of Sciences, 77(12), 7463-7467.

Herrstedt, A., Bay, M. L., Simonsen, C., Sundberg, A., Egeland, C., Thorsen-Streit, S., ... & Hojman, P. (2019). Exercise-mediated improvement of depression in patients with gastro-esophageal junction cancer is linked to kynurenine metabolism. Acta Oncologica, 58(5), 579-587.

Hinata, A. et al. (2021) ‘Education, household income, and depressive symptoms in middle-aged and older Japanese adults’, BMC Public Health, 21(1). doi:10.1186/s12889-021-12168-8.

Ho, R. C., Niti, M., Kua, E. H., & Ng, T. P. (2008). Body mass index, waist circumference, waist–hip ratio and depressive symptoms in Chinese elderly: a population-based study. International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences, 23(4), 401-408.

Holmes, E., Tsang, T., Huang, J., Leweke, F., Koethe, D., Gerth, C., Nolden, B., Gross, S., Schreiber, D., Nicholson, J. and Bahn, S., 2006. Metabolic Profiling of CSF: Evidence That Early Intervention May Impact on Disease Progression and Outcome in Schizophrenia. PLoS Medicine, 3(8), p.e327.

Holt, R., de Groot, M. and Golden, S. (2014) ‘Diabetes and Depression’, Current Diabetes Reports, 14(6). doi:10.1007/s11892-014-0491-3.

Horato, N., Quagliato, L. A., & Nardi, A. E. (2022). The relationship between emotional regulation and hemispheric lateralization in depression: a systematic review and a meta-analysis. Translational Psychiatry, 12(1), 162.

Horsten, M., Wamala, S. P., Vingerhoets, A. D., & Orth-Gomer, K. (1997). Depressive symptoms, social support, and lipid profile in healthy middle-aged women. Psychosomatic medicine, 59(5), 521-528.

Huang, J., Leweke, F., Tsang, T., Koethe, D., Kranaster, L., Gerth, C., Gross, S., Schreiber, D., Ruhrmann, S., Schultze-Lutter, F., Klosterkötter, J., Holmes, E. and Bahn, S., 2007. CSF Metabolic and Proteomic Profiles in Patients Prodromal for Psychosis. PLoS ONE, 2(8), p.e756.

Huang, T. L. (2005). Serum lipid profiles in major depression with clinical subtypes, suicide attempts and episodes. Journal of affective disorders, 86(1), 75-79.

Huang, T.-L. et al. (2003) “Correlation between serum lipid, lipoprotein concentrations and anxious state, Depressive State or major depressive disorder,” Psychiatry Research, 118(2), pp. 147–153. Available at: https://doi.org/10.1016/s0165-1781(03)00071-4.

Huang, W. et al. (2021) ‘Acetate supplementation produces antidepressant-like effect via enhanced histone acetylation’, Journal of Affective Disorders, 281, pp. 51–60. doi:10.1016/j.jad.2020.11.121.

Humer, E., Pieh, C., & Probst, T. (2020). Metabolomic biomarkers in anxiety disorders. International Journal of Molecular Sciences, 21(13), 4784.

Humer, E., Probst, T and Pieh C., 2020. Metabolomics in psychiatric disorders. Metabolites, 10(72), pp1-20.

Islam, M. R., Ali, S., Karmoker, J. R., Kadir, M. F., Ahmed, M. U., Nahar, Z., ... & Islam, M. S. (2020). Evaluation of serum amino acids and non-enzymatic antioxidants in drug-naïve first-episode major depressive disorder. BMC psychiatry, 20, 1-8.

Islam, M. R., Ali, S., Karmoker, J. R., Kadir, M. F., Ahmed, M. U., Nahar, Z., ... & Islam, M. S. (2020). Evaluation of serum amino acids and non-enzymatic antioxidants in drug-naïve first-episode major depressive disorder. BMC psychiatry, 20, 1-8.

Israel, J., 2006. Remission in depression: definition and initial treatment approaches. Journal of Psychopharmacology, 20(3_suppl), pp.5-10.

Jia, Q.-fang et al. (2020) “The role of lipoprotein profile in depression and Cognitive Performance: A network analysis,” Scientific Reports, 10(1). Available at: https://doi.org/10.1038/s41598-020-77782-9.

Jia, Y., Hui, L., Sun, L., Guo, D., Shi, M., Zhang, K., ... & Zhu, Z. (2023). Association between human blood metabolome and the risk of psychiatric disorders. Schizophrenia Bulletin, 49(2), 428-443.

Jimenez-Fernandez, S., Gurpegui, M., Diaz-Atienza, F., Pérez-Costillas, L., Gerstenberg, M., & Correll, C. U. (2015). Oxidative stress and antioxidant parameters in patients with major depressive disorder compared to healthy controls before and after antidepressant treatment: results from a meta-analysis. The Journal of clinical psychiatry, 76(12), 13705.

Juruena, M. F., Cleare, A. J., & Young, A. H. (2017). The role of early life stress in HPA axis and anxiety. In Neurobiology of Stress (Vol. 2, pp. 25-33). Academic Press .

Kageyama, Y., Kasahara, T., Morishita, H., Mataga, N., Deguchi, Y., Tani, M., Kuroda, K., Hattori, K., Yoshida, S., Inoue, K. and Kato, T., 2016. Search for plasma biomarkers in drug-free patients with bipolar disorder and schizophrenia using metabolome analysis. Psychiatry and Clinical Neurosciences, 71(2), pp.115-123.

Kageyama, Y., Kasahara, T., Nakamura, T., Hattori, K., Deguchi, Y., Tani, M., Kuroda, K., Yoshida, S., Goto, Y., Inoue, K. and Kato, T., 2017. Plasma Nervonic Acid Is a Potential Biomarker for Major Depressive Disorder: A Pilot Study. International Journal of Neuropsychopharmacology, 21(3), pp.207-215.

Kandola, A. et al. (2019) ‘Physical activity and depression: Towards understanding the antidepressant mechanisms of physical activity’, Neuroscience & Biobehavioral Reviews, 107, pp. 525–539. doi:10.1016/j.neubiorev.2019.09.040.

Kandola, A. et al. (2019) ‘Physical activity and depression: Towards understanding the antidepressant mechanisms of physical activity’, Neuroscience & Biobehavioral Reviews, 107, pp. 525–539. doi:10.1016/j.neubiorev.2019.09.040.

Kawamura, N., Shinoda, K., Sato, H., Sasaki, K., Suzuki, M., Yamaki, K., et al. (2018). Plasma metabolome analysis of patients with major depressive disorder. Psychiatry Clin. Neurosci. 72, 349–361. doi: 10.1111/pcn.12638

Keaton, S. A., Arnetz, J., Jamil, H., Dhalimi, A., Stemmer, P. M., Ruden, D. M., Yamin, J., Achtyes, E., Smart, L., Brundin, L., & Arnetz, B. B. (2021). IL-10: A possible

immunobiological component of positive mental health in refugees. Comprehensive psychoneuroendocrinology, 8, 100097. https://doi.org/10.1016/j.cpnec.2021.100097

Kendall, K., Van Assche, E., Andlauer, T., Choi, K., Luykx, J., Schulte, E., & Lu, Y. (2021). The genetic basis of major depression. Psychological Medicine, 51(13), 2217-2230. doi:10.1017/S0033291721000441

Kenis, G., & Maes, M. (2002). Effects of antidepressants on the production of cytokines. International Journal of Neuropsychopharmacology, 5(4), 401-412.

Kim, E. Y., Lee, J. W., Lee, M. Y., Kim, S. H., Mok, H. J., Ha, K., ... & Kim, K. P. (2018). Serum lipidomic analysis for the discovery of biomarkers for major depressive disorder in drug-free patients. Psychiatry research, 265, 174-182.

Kim, S. Y., Shin, S. Y., Park, S. J., Im, J. P., Kim, H. J., Lee, K. M., ... & Choi, H. K. (2023). Changes in fecal metabolic and lipidomic features by anti-TNF treatment and prediction of clinical remission in patients with ulcerative colitis. Therapeutic Advances in Gastroenterology, 16, 17562848231168199.

Kirkpatrick, B., Garcia-Rizo, C., Tang, K., Fernandez-Egea, E. and Bernardo, M., 2010. Cholesterol and triglycerides in antipsychotic-naive patients with nonaffective psychosis. Psychiatry Research, 178(3), pp.559-561.

Kirmayer, L. J., & Minas, H. (2023). The future of cultural psychiatry: an international perspective. Medical Anthropology, 135-143.

Kirwan, J. A. (2023). Translating metabolomics into clinical practice. Nature Reviews Bioengineering, 1-2.

Kofler, M., Schiefecker, A. J., Gaasch, M., Sperner-Unterweger, B., Fuchs, D., Beer, R., ... & Helbok, R. (2019). A reduced concentration of brain interstitial amino acids is associated with depression in subarachnoid hemorrhage patients. Scientific Reports, 9(1), 2811.

Koike, S., Bundo, M., Iwamoto, K., Suga, M., Kuwabara, H., Ohashi, Y., Shinoda, K., Takano, Y., Iwashiro, N., Satomura, Y., Nagai, T., Natsubori, T., Tada, M., Yamasue, H. and Kasai, K., 2014. A snapshot of plasma metabolites in first-episode schizophrenia: a capillary electrophoresis time-of-flight mass spectrometry study. Translational Psychiatry, 4(4), pp.e379-e379.

Kong, Y., Chen, Z., Feng, X., Zuo, Y., & Zhang, J. (2023). Gut microbiota and metabolome in sporadic Creutzfeldt–Jakob disease. Journal of Neurology, 1-12.

Korherr, C., Hofmeister, R., Wesche, H., & Falk, W. (1997). A critical role for interleukin-1 receptor accessory protein in interleukin-1 signaling. European journal of immunology, 27(1), 262-267.

KR;, F. (2021) Introduction to lipids and Lipoproteins, National Center for Biotechnology Information. U.S. National Library of Medicine. Available at: https://pubmed.ncbi.nlm.nih.gov/26247089/ (Accessed: March 16, 2023).

Kujawski, S., Wong, M., & Clements, J. (2018). The impact of blood glucose and diet on mood and cognition. Nutritional Neuroscience, 21(9), 637-645. https://doi.org/10.1080/1028415X.2017.1340383

Kurano, M., Iso-O, N., Hara, M., Noiri, E., Koike, K., Kadowaki, T., & Tsukamoto, K. (2011). Plant sterols increased IL-6 and TNF-α secretion from macrophages, but to a lesser extent than cholesterol. Journal of atherosclerosis and thrombosis, 18(5), 373-383.

Kuwano, N., Kato, T. A., Setoyama, D., Sato-Kasai, M., Shimokawa, N., Hayakawa, K., ... & Kanba, S. (2018). Tryptophan-kynurenine and lipid related metabolites as blood biomarkers for first-episode drug-naïve patients with major depressive disorder: an exploratory pilot case-control study. Journal of affective disorders, 231, 74-82.

Laboute, T., Zucca, S., Holcomb, M., Patil, D. N., Garza, C., Wheatley, B. A., ... & Martemyanov, K. A. (2023). Orphan receptor GPR158 serves as a metabotropic glycine receptor: mGlyR. Science, 379(6639), 1352-1358.

Lakhan, S. E., & Vieira, K. F. (2008). Nutritional therapies for mental disorders. Nutrition journal, 7(1), 1-8.

Lakshmi Reddy, P., Khanna, S., Subhash, M., Channabasavanna, S. and Sridhara Rama Rao, B., 1992. CSF amine metabolites in depression. Biological Psychiatry, 31(2), pp.112-118.

Lee, M. J., McLean, K. E., Kuo, M., Richardson, G. R., & Henderson, S. B. (2023). Chronic diseases associated with mortality in British Columbia, Canada during the 2021 western North America extreme heat event. GeoHealth, 7(3), e2022GH000729.

Lehto, S. M., Niskanen, L., Tolmunen, T., Hintikka, J., Viinamäki, H., Heiskanen, T., ... & Koivumaa-Honkanen, H. (2010). Low serum HDL-cholesterol levels are associated with long symptom duration in patients with major depressive disorder. Psychiatry and clinical neurosciences, 64(3), 279-283.

Lehto, S. M., Niskanen, L., Tolmunen, T., Hintikka, J., Viinamäki, H., Heiskanen, T., ... & Koivumaa-Honkanen, H. (2010). Low serum HDL-cholesterol levels are associated with long symptom duration in patients with major depressive disorder. Psychiatry and clinical neurosciences, 64(3), 279-283.

Lehto, S. M., Niskanen, L., Tolmunen, T., Hintikka, J., Viinamäki, H., Heiskanen, T., ... & Koivumaa-Honkanen, H. (2010). Low serum HDL-cholesterol levels are associated with long symptom duration in patients with major depressive disorder. Psychiatry and clinical neurosciences, 64(3), 279-283.

Lemaitre, R. N., Yu, C., Hoofnagle, A., Hari, N., Jensen, P. N., Fretts, A. M., Umans, J. G., Howard, B. V., Sitlani, C. M., Siscovick, D. S., King, I. B., Sotoodehnia, N., & McKnight, B. (2018). Circulating Sphingolipids, Insulin, HOMA-IR, and HOMA-B: The Strong Heart Family Study. Diabetes, 67(8), 1663–1672. https://doi.org/10.2337/db17-1449

Lenze, E. J., Mulsant, B. H., Shear, M. K., Schulberg, H. C., Dew, M. A., Begley, A. E., ... & Reynolds III, C. F. (2000). Comorbid anxiety disorders in depressed elderly patients. American Journal of Psychiatry, 157(5), 722-728.

Li, C., Wang, A., Wang, C., Ramamurthy, J., Zhang, E., Guadagno, E. and Trakadis, Y., 2018. Metabolomics in patients with psychosis: A systematic review. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 177(6), pp.580-588.

Li, Y., Wu, L., Chen, C., Wang, L., Guo, C., Zhao, X., et al. (2020). Serum metabolic profiling reveals the antidepressive effects of the total iridoids of valeriana jatamansi jones on chronic unpredictable mild stress mice. Front. Pharmacol. 11, 338. doi: 10.3389/fphar.2020.00338

Li, Z., Ruan, M., Chen, J., & Fang, Y. (2021). Major depressive disorder: advances in neuroscience research and translational applications. Neuroscience bulletin, 37, 863-880.

Liao, Y. et al. (2019) “Efficacy of omega-3 pufas in depression: A meta-analysis,” Translational Psychiatry, 9(1). Available at: https://doi.org/10.1038/s41398-019-0515-5.

Libby, P., Ridker, P. M., & Maseri, A. (2002). Inflammation and atherosclerosis. Circulation, 105(9), 1135-1143.

Liemburg, E., Sibeijn-Kuiper, A., Bais, L., Pijnenborg, G., Knegtering, H., van der Velde, J., Opmeer, E., de Vos, A., Dlabac-De Lange, J., Wunderink, L. and Aleman, A., 2016. Prefrontal NAA and Glx Levels in Different Stages of Psychotic Disorders: a 3T 1H-MRS Study. Scientific Reports, 6(1).

Lima Giacobbo, B., Doorduin, J., Klein, H. C., Dierckx, R. A., Bromberg, E., & de Vries, E. F. (2019). Brain-derived neurotrophic factor in brain disorders: focus on neuroinflammation. Molecular neurobiology, 56, 3295-3312.

Lin, P. Y., Huang, S. Y., & Su, K. P. (2010). A meta-analytic review of polyunsaturated fatty acid compositions in patients with depression. Biological psychiatry, 68(2), 140-147.

Lin, S. J., Yen, H. T., Chen, Y. H., Ku, H. H., Lin, F. Y., & Chen, Y. L. (2003). Expression of interleukin-1β and interleukin-1 receptor antagonist in oxLDL-treated human aortic smooth muscle cells and in the neointima of cholesterol-fed endothelia-denuded rabbits. Journal of cellular biochemistry, 88(4), 836-847.

Liu, C., Wu, Y., Feng, G., Gao, X., Zhou, Y., Hou, W., Qin, X., Du, G. and Tian, J., 2015. Plasma-metabolite-biomarkers for the therapeutic response in depressed patients by the traditional Chinese medicine formula Xiaoyaosan: A 1H NMR-based metabolomics approach. Journal of Affective Disorders, 185, pp.156-163.

Liu, D., Ray, B., Neavin, D., Zhang, J., Athreya, A., Biernacka, J., Bobo, W., Hall-Flavin, D., Skime, M., Zhu, H., Jenkins, G., Batzler, A., Kalari, K., Boakye-Agyeman, F., Matson, W., Bhasin, S., Mushiroda, T., Nakamura, Y., Kubo, M., Iyer, R., Wang, L., Frye, M., Kaddurah-Daouk, R. and Weinshilboum, R., 2018. Beta-defensin 1, aryl hydrocarbon receptor and plasma kynurenine in major depressive disorder: metabolomics-informed genomics. Translational Psychiatry, 8(1).

Liu, M., Zhang, X., Du, X., Fang, Z., Liu, Z., Xu, Y., Zheng, P., Xu, X., Cheng, P., Huang, T., Bai, S., Zhao, L., Qi, Z., Shao, W. and Xie, P., 2015. Severe disturbance of glucose metabolism in peripheral blood mononuclear cells of schizophrenia

patients: a targeted metabolomic study. Journal of Translational Medicine, 13(1).

Liu, M., Zheng, P., Liu, Z., Xu, Y., Mu, J., Guo, J., Huang, T., Meng, H. and Xie, P., 2014. GC-MS based metabolomics identification of possible novel biomarkers for schizophrenia in peripheral blood mononuclear cells. Mol. BioSyst., 10(9), pp.2398-2406.

Liu, X., Li, J., Zheng, P., Zhao, X., Zhou, C., Hu, C., ... & Xu, G. (2016). Plasma lipidomics reveals potential lipid markers of major depressive disorder. Analytical and bioanalytical chemistry, 408, 6497-6507.

Liu, X., Liu, C., Tian, J., Gao, X., Li, K., Du, G. and Qin, X., 2020. Plasma metabolomics of depressed patients and treatment with Xiaoyaosan based on mass spectrometry technique. Journal of Ethnopharmacology, 246, p.112219.

Liu, X., Zheng, P., Zhao, X., Zhang, Y., Hu, C., Li, J., … & Xie, P. (2015). Discovery and validation of plasma biomarkers for major depressive disorder classification based on liquid chromatography–mass spectrometry. Journal of proteome research, 14(5), 2322-2330 .

Liu, Y., Ho, R. C. M., & Mak, A. (2012). Interleukin (IL)-6, tumour necrosis factor alpha (TNF-α) and soluble interleukin-2 receptors (sIL-2R) are elevated in patients with major depressive disorder: a meta-analysis and meta-regression. Journal of affective disorders, 139(3), 230-239.

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., ... & Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. The Lancet, 396(10248), 413-446.

Lobo-Silva, D., Carriche, G. M., Castro, A. G., Roque, S., & Saraiva, M. (2016). Balancing the immune response in the brain: IL-10 and its regulation. Journal of neuroinflammation, 13(1), 1-10.

Lombardi, A. L., Manfredi, L., & Conversi, D. (2022). How does IL-6 change after combined treatment in MDD patients? A systematic review. Brain, Behavior, & Immunity-Health, 100579.

MacAndrew, C. (1982) ‘An examination of the relevance of the Individual Differences (a-trait) formulation of the tension-reduction theory to the etiology of alcohol abuse in young males’, Addictive Behaviors, 7(1), pp. 39–45. doi:10.1016/0306-4603(82)90023-5.

Maes, M. (2008). The cytokine hypothesis of depression: inflammation, oxidative & nitrosative stress (IO&NS) and leaky gut as new targets for adjunctive treatments in depression. Neuroendocrinology letters, 29(3), 287-291.

Maes, M., Smith, R., Christophe, A., Vandoolaeghe, E., Gastel, A. V., Neels, H., ... & Meltzer, H. Y. (1997). Lower serum high-density lipoprotein cholesterol (HDL-C) in major depression and in depressed men with serious suicidal attempts: relationship with immune-inflammatory markers. Acta Psychiatrica Scandinavica, 95(3), 212-221.

Mamalakis, G. et al. (2006) 'Depression and adipose and serum cholesteryl ester polyunsaturated fatty acids in the survivors of the seven countries study population of Crete', European Journal of Clinical Nutrition, 60(8), pp. 1016–1023. doi:10.1038/sj.ejcn.1602413.

Mantantzis, K., Schlaghecken, F., Sünram-Lea, S. I., & Maylor, E. A. (2019). Sugar rush or sugar crash? A meta-analysis of carbohydrate effects on mood. Neuroscience & Biobehavioral Reviews, 101, 45-67.

Mantovani, A., Dinarello, C. A., Molgora, M., & Garlanda, C. (2019). Interleukin-1 and related cytokines in the regulation of inflammation and immunity. Immunity, 50(4), 778-795.

Martins-de-Souza, D., 2014. Proteomics, metabolomics, and protein interactomics in the characterization of the molecular features of major depressive disorder. Dialogues in Clinical Neuroscience, 16(1), pp.63-73.

Marx, W., Manger, S. H., Blencowe, M., Murray, G., Ho, F. Y. Y., Lawn, S., ... & O'Neil, A. (2023). Clinical guidelines for the use of lifestyle-based mental health care in major depressive disorder: World Federation of Societies for Biological Psychiatry (WFSBP) and Australasian Society of Lifestyle Medicine (ASLM) taskforce. The World Journal of Biological Psychiatry, 1-54.

Marx, W., Penninx, B. W., Solmi, M., Furukawa, T. A., Firth, J., Carvalho, A. F., & Berk, M. (2023). Major depressive disorder. Nature Reviews Disease Primers, 9(1), 44.

Mayengbam, S. S., Singh, A., Pillai, A. D., & Bhat, M. K. (2021). Influence of cholesterol on cancer progression and therapy. Translational oncology, 14(6), 101043.

McEvoy, J., Baillie, R., Zhu, H., Buckley, P., Keshavan, M., Nasrallah, H., Dougherty, G., Yao, J. and Kaddurah-Daouk, R., 2013. Lipidomics Reveals Early Metabolic

Changes in Subjects with Schizophrenia: Effects of Atypical Antipsychotics. PLoS ONE, 8(7), p.e68717.

McGinty, J., Sayeed Haque, M. and Upthegrove, R., 2018. Depression during first episode psychosis and subsequent suicide risk: A systematic review and meta-analysis of longitudinal studies. Schizophrenia Research, 195, pp.58-66.

McGuffin, P., Katz, R. and Rutherford, J. (1991) “Nature, nurture and depression: A twin study,” Psychological Medicine, 21(2), pp. 329–335. Available at: https://doi.org/10.1017/s0033291700020432.

McNamara, R.K. et al. (2018) “Role of polyunsaturated fatty acids in human brain structure and function across the lifespan: An update on neuroimaging findings,” Prostaglandins, Leukotrienes and Essential Fatty Acids, 136, pp. 23–34. Available at: https://doi.org/10.1016/j.plefa.2017.05.001.

Meltzer, H., 2002. The Role of Serotonin in Schizophrenia. In: K. Davis, D. Charney, J. Coyle and C. Nemeroff, ed., Neuropsychopharmacology: The Fifth edition of Progress, 5th ed. [online] Philadelphia, Pennsylvania: Lippincott, Williams and Wilkins, pp.823-826. Available at: <https://acnp.org/wp-content/uploads/2017/11/CH58_819-832.pdf> [Accessed 10 July 2020].

Merritt, K., McCutcheon, R. A., Aleman, A., Ashley, S., Beck, K., Block, W., ... & 1H-MRS in Schizophrenia Investigators de la Fuente-Sandoval Camilo 11 12 van Amelsvoort Thérèse 7 McGuire Philip K. 5. (2023). Variability and magnitude of brain glutamate levels in schizophrenia: a meta and mega-analysis. Molecular psychiatry, 1-10.

Mesquita, A. R., Correia-Neves, M., Roque, S., Castro, A. G., Vieira, P., Pedrosa, J., & Sousa, N. (2008). IL-10 modulates depressive-like behavior. Journal of psychiatric research, 43(2), 89-97.

Micó, J., Rojas-Corrales, M., Gibert-Rahola, J., Parellada, M., Moreno, D., Fraguas, D., Graell, M., Gil, J., Irazusta, J., Castro-Fornieles, J., Soutullo, C., Arango, C., Otero, S., Navarro, A., Baeza, I., Martínez-Cengotitabengoa, M. and González-Pinto, A., 2011. Reduced antioxidant defense in early onset first-episode psychosis: a case-control study. BMC Psychiatry, 11(1).

Mitani, H., Shirayama, Y., Yamada, T., Maeda, K., Ashby, C. and Kawahara, R., 2006. Correlation between plasma levels of glutamate, alanine and serine with

severity of depression. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 30(6), pp.1155-1158.

Mittli, D., Tukacs, V., Ravasz, L., Csősz, É., Kozma, T., Kardos, J., ... & Kékesi, K. A. (2023). LPS-induced acute neuroinflammation, involving interleukin-1 beta signaling, leads to proteomic, cellular, and network-level changes in the prefrontal cortex of mice. Brain, Behavior, & Immunity-Health, 28, 100594.

Moffett, J.R. et al. (2020) 'Acetate revisited: A key biomolecule at the nexus of metabolism, epigenetics and oncogenesis—part 1: Acetyl-COA, acetogenesis and acyl-COA short-chain synthetases', Frontiers in Physiology, 11. doi:10.3389/fphys.2020.580167.

Molina, V., Sánchez, J., Sanz, J., Reig, S., Benito, C., Leal, I., Sarramea, F., Rebolledo, R., Palomo, T. and Desco, M., 2007. Dorsolateral prefrontal N-acetyl-aspartate concentration in male patients with chronic schizophrenia and with chronic bipolar disorder. European Psychiatry, 22(8), pp.505-512.

Morales-Medina, J. C., Dumont, Y., & Quirion, R. (2010). A possible role of neuropeptide Y in depression and stress. Brain research, 1314, 194-205.

Moriguchi, S. et al. (2018) 'Glutamatergic neurometabolite levels in major depressive disorder: A systematic review and meta-analysis of Proton Magnetic Resonance Spectroscopy Studies', Molecular Psychiatry, 24(7), pp. 952–964. doi:10.1038/s41380-018-0252-9.

Moriguchi, S., Takamiya, A., Noda, Y., Horita, N., Wada, M., Tsugawa, S., Plitman, E., Sano, Y., Tarumi, R., ElSalhy, M., Katayama, N., Ogyu, K., Miyazaki, T., Kishimoto, T., Graff-Guerrero, A., Meyer, J., Blumberger, D., Daskalakis, Z., Mimura, M. and Nakajima, S., 2018. Glutamatergic neurometabolite levels in major depressive disorder: a systematic review and meta-analysis of proton magnetic resonance spectroscopy studies. Molecular Psychiatry, 24(7), pp.952-964.

Müller, C. P., Reichel, M., Mühle, C., Rhein, C., Gulbins, E., & Kornhuber, J. (2015). Brain membrane lipids in major depression and anxiety disorders. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 1851(8), 1052-1065.

Murphy, J. A., Sarris, J., & Byrne, G. J. (2017). A review of the conceptualisation and risk factors associated with treatment-resistant depression. Depression research and treatment, 2017.

Nagasawa, M., Murakami, T., Tomonaga, S., & Furuse, M. (2012). The impact of chronic imipramine treatment on amino acid concentrations in the hippocampus of mice. Nutritional Neuroscience, 15(5), 26-33.

Najjar, S., Pearlman, D., Alper, K., Najjar, A. and Devinsky, O., 2013. Neuroinflammation and psychiatric illness. Journal of Neuroinflammation, 10(1).

Nedic Erjavec, G., Konjevod, M., Nikolac Perkovic, M., Svob Strac, D., Tudor, L., Barbas, C., Grune, T., Zarkovic, N. and Pivac, N., 2018. Short overview on metabolomic approach and redox changes in psychiatric disorders. Redox Biology, 14, pp.178-186.

Nestel, P., Shige, H., Pomeroy, S., Cehun, M., Abbey, M., & Raederstorff, D. (2002). The n−3 fatty acids eicosapentaenoic acid and docosahexaenoic acid increase systemic arterial compliance in humans. The American journal of clinical nutrition, 76(2), 326-330.

Newgard, C., 2017. Metabolomics and Metabolic Diseases: Where Do We Stand?. Cell Metabolism, 25(1), pp.43-56.

Ng, A. et al. (2018) 'IL-1β, IL-6, TNF- α and CRP in elderly patients with depression or alzheimer's disease: Systematic review and meta-analysis', Scientific Reports, 8(1). doi:10.1038/s41598-018-30487-6.

Ng, S. M., & Lee, T. M. (2020). The mediating role of hardiness in the relationship between perceived loneliness and depressive symptoms among older. Aging & mental health, 24(5), 805-810.

Nouwen, A., Nefs, G., Caramlau, I., Connock, M., Winkley, K., Lloyd, C. E., ... & European Depression in Diabetes (EDID) Research Consortium. (2011). Prevalence of depression in individuals with impaired glucose metabolism or undiagnosed diabetes: a systematic review and meta-analysis of the European Depression in Diabetes (EDID) Research Consortium. Diabetes care, 34(3), 752-762.

Ogawa, S., Hattori, K., Sasayama, D., Yokota, Y., Matsumura, R., Matsuo, J., Ota, M., Hori, H., Teraishi, T., Yoshida, S., Noda, T., Ohashi, Y., Sato, H., Higuchi, T., Motohashi, N. and Kunugi, H., 2015. Reduced cerebrospinal fluid ethanolamine concentration in major depressive disorder. Scientific Reports, 5(1).

Ogłodek, E. (2022). Changes in the Serum Levels of Cytokines: IL-1β, IL-4, IL-8 and IL-10 in Depression with and without Posttraumatic Stress Disorder. Brain Sciences, 12(3), 387.

Olusi, S. O., & Fido, A. A. (1996). Serum lipid concentrations in patients with major depressive disorder. Biological psychiatry, 40(11), 1128-1131.

Onaolopo, A (2021). “Glutamate and depression: Reflecting a deepening knowledge of the gut and brain effects of a ubiquitous molecule”, World Journal Psychiatry. 11(7), pp.297-315.

Orešič, M., Tang, J., Seppänen-Laakso, T., Mattila, I., Saarni, S., Saarni, S., Lönnqvist, J., Sysi-Aho, M., Hyötyläinen, T., Perälä, J. and Suvisaari, J., 2011. Metabolome in schizophrenia and other psychotic disorders: a general population-based study. Genome Medicine, 3(3), p.19.

Ovaskainen, Y., Koponen, H., Jokelainen, J., Keinänen-Kiukaanniemi, S., Kumpusalo, E., & Vanhala, M. (2009). Depressive symptomatology is associated with decreased interleukin-1 beta and increased interleukin-1 receptor antagonist levels in males. Psychiatry Research, 167(1-2), 73-79.

Ovaskainen, Y., Koponen, H., Jokelainen, J., Keinänen-Kiukaanniemi, S., Kumpusalo, E., & Vanhala, M. (2009). Depressive symptomatology is associated with decreased interleukin-1 beta and increased interleukin-1 receptor antagonist levels in males. Psychiatry Research, 167(1-2), 73-79.

Pai, Jennifer K., Tobias Pischon, Jing Ma, JoAnn E. Manson, Susan E. Hankinson, Kaumudi Joshipura, Gary C. Curhan et al. “Inflammatory markers and the risk of coronary heart disease in men and women.” New England Journal of Medicine 351, no. 25 (2004): 2599-2610.

Pal, M.M. (2021) ‘Glutamate: The master neurotransmitter and its implications in chronic stress and mood disorders’, Frontiers in Human Neuroscience, 15. doi:10.3389/ fnhum.2021.722323.

Pal, P., Pramanik, S., & Ray, S. (2021). Disorders of gastrointestinal motility in diabetes mellitus: an unattended borderline between diabetologists and gastroenterologists. Diabetes.

Palsson, E., Sellgren, C., Ryden, E., Kizza, R., Pelanis, A., Zetterberg, H., ... Landen, M. (2017). Cerebrospinal fluid monoamine metabolite profiles in bipolar disorder, ADHD, and controls. Journal of Neural Transmission (Vienna), 124(9), 1135–1143. https://doi.org/10.1007/s00702-017- 1746-3.

Pan, J., Xia, J., Deng, F., Liang, W., Wu, J., Yin, B., Dong, M., Chen, J., Ye, F., Wang, H., Zheng, P. and Xie, P., 2018. Diagnosis of major depressive disorder based on

changes in multiple plasma neurotransmitters: a targeted metabolomics study. Translational Psychiatry, 8(1).

Parekh, A. et al. (2017) “The role of lipid biomarkers in major depression,” Healthcare, 5(1), p. 5. Available at: https://doi.org/10.3390/healthcare5010005.

Partonen, T., Haukka, J., Virtamo, J., Taylor, P. R., & Lönnqvist, J. (1999). Association of low serum total cholesterol with major depression and suicide. The British Journal of Psychiatry, 175(3), 259-262.

Pasco, J.A. et al. (2008) ‘Tobacco smoking as a risk factor for major depressive disorder: Population-based study’, British Journal of Psychiatry, 193(4), pp. 322–326. doi:10.1192/bjp.bp.107.046706.

Paul, E.R. et al. (2023) ‘Towards a multilevel model of major depression: Genes, immuno-metabolic function, and cortico-striatal signaling’, Translational Psychiatry, 13(1). doi:10.1038/s41398-023-02466-7.

Pedrini, M., Cao, B., Nani, J. V. S., Cerqueira, R. O., Mansur, R. B., Tasic, L., ... & Brietzke, E. (2019). Advances and challenges in development of precision psychiatry through clinical metabolomics on mood and psychotic disorders. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 93, 182-188.

Pedrini, M., Massuda, R., Fries, G., de Bittencourt Pasquali, M., Schnorr, C., Moreira, J., Teixeira, A., Lobato, M., Walz, J., Belmonte-de-Abreu, P., Kauer-Sant’Anna, M., Kapczinski, F. and Gama, C., 2012. Similarities in serum oxidative stress markers and inflammatory cytokines in patients with overt schizophrenia at early and late stages of chronicity. Journal of Psychiatric Research, 46(6), pp.819-824.

Peng, Y. F., Xiang, Y., & Wei, Y. S. (2016). The significance of routine biochemical markers in patients with major depressive disorder. Scientific reports, 6(1), 34402.

Peng, Y., Liu, J., Shi, L., Tang, Y., Gao, D., Long, J., & Liu, J. (2016). Mitochondrial dysfunction precedes depression of AMPK/AKT signaling in insulin resistance induced by high glucose in primary cortical neurons. Journal of neurochemistry, 137(5), 701-713.

Peng, Y.-F., Xiang, Y. and Wei, Y.-S. (2016b) ‘The significance of routine biochemical markers in patients with major depressive disorder’, Scientific Reports, 6(1). doi:10.1038/srep34402.

Penninx, B. W., Beekman, A. T., Honig, A., Deeg, D. J., Schoevers, R. A., Van Eijk, J. T., & Van Tilburg, W. (2001). Depression and cardiac mortality: results from a community-based longitudinal study. Archives of general psychiatry, 58(3), 221-227.

Penttinen, J. (1995). Hypothesis: low serum cholesterol, suicide, and Interleukin-2. American journal of epidemiology, 141(8), 716-718.

Persons, J. E., & Fiedorowicz, J. G. (2016). Depression and serum low-density lipoprotein: a systematic review and meta-analysis. Journal of affective disorders, 206, 55-67.

Pitsillou, E., Bresnehan, S. M., Kagarakis, E. A., Wijoyo, S. J., Liang, J., Hung, A., & Karagiannis, T. C. (2020). The cellular and molecular basis of major depressive disorder: towards a unified model for understanding clinical depression. Molecular Biology Reports, 47(1), 753-770

Polderman, T. J., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nature genetics, 47(7), 702-709. :

Postila, P.A. and Róg, T. (2019) “A perspective: Active role of lipids in neurotransmitter dynamics,” Molecular Neurobiology, 57(2), pp. 910–925. Available at: https://doi.org/10.1007/s12035-019-01775-7.

Pradhan, A. D., Manson, J. E., Rossouw, J. E., Siscovick, D. S., Mouton, C. P., Rifai, N., ... & Ridker, P. M. (2002). Inflammatory biomarkers, hormone replacement therapy, and incident coronary heart disease: prospective analysis from the Women's Health Initiative observational study. Jama, 288(8), 980-987.

Prince, N., Stav, M., Cote, M., Chu, S. H., Vyas, C. M., Okereke, O. I., ... & Kelly, R. S. (2023). Metabolomics and Self-Reported Depression, Anxiety, and Phobic Symptoms in the VA Normative Aging Study. Metabolites, 13(7), 851.

Qiu, W. et al. (2021) “Update on the relationship between depression and neuroendocrine metabolism,” Frontiers in Neuroscience, 15. Available at: https://doi.org/10.3389/fnins.2021.728810.

Rabe-Jabłońska, J., & Poprawska, I. (2000). Levels of serum total cholesterol and LDL-cholesterol in patients with major depression in acute period and remission. Medical Science Monitor, 6(3), CR539-CR547.

Rai, D. et al. (2013) ‘Country- and individual-level socioeconomic determinants of depression: Multilevel cross-national comparison’, British Journal of Psychiatry, 202(3), pp. 195–203. doi:10.1192/bjp.bp.112.112482.

Rai, D., Heuvelman, H., Dalman, C., Culpin, I., Lundberg, M., Carpenter, P., & Magnusson, C. (2018). Association between autism spectrum disorders with or without intellectual disability and depression in young adulthood. JAMA network open, 1(4), e181465-e181465.

Ramirez-Bermudez, J., Ruiz-Chow, A., Perez-Neri, I., Soto-Hernandez, J., Flores-Hernandez, R., Nente, F., Montes, S. and Rios, C., 2008. Cerebrospinal fluid homovanillic acid is correlated to psychotic features in neurological patients with delirium. General Hospital Psychiatry, 30(4), pp.337-343.

Recchia, F. et al. (2022) ‘Comparative effectiveness of exercise, antidepressants and their combination in treating non-severe depression: A systematic review and network meta-analysis of Randomised Controlled Trials’, British Journal of Sports Medicine, 56(23), pp. 1375–1380. doi:10.1136/bjsports-2022-105964.

Rhee, E. P., Cheng, S., Larson, M. G., Walford, G. A., Lewis, G. D., McCabe, E., Yang, E., Farrell, L., Fox, C. S., O'Donnell, C. J., Carr, S. A., Vasan, R. S., Florez, J. C., Clish, C. B., Wang, T. J., & Gerszten, R. E. (2011). Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. The Journal of clinical investigation, 121(4), 1402–1411. https://doi.org/10.1172/JCI44442

Roohafza, H., Sadeghi, M., Afshar, H., Mousavi, G., & Shirani, S. (2010). Evaluation of lipid profile in patient with major depressive disorder and generalized anxiety disorder. ARYA Atherosclerosis Journal, 1(1).

Ross, B. M., Seguin, J., & Sieswerda, L. E. (2007). Omega-3 fatty acids as treatments for mental illness: which disorder and which fatty acid?. Lipids in health and disease, 6(1), 1-19.

Ross, R. E., VanDerwerker, C. J., Saladin, M. E., & Gregory, C. M. (2023). The role of exercise in the treatment of depression: biological underpinnings and clinical outcomes. Molecular Psychiatry, 28(1), 298-328.

Rozali, N. L., Azizan, K. A., Singh, R., Jaafar, S. N. S., Othman, A., Weckwerth, W., & Ramli, U. S. (2023). Fourier transform infrared (FTIR) spectroscopy approach combined with discriminant analysis and prediction model for crude palm oil

authentication of different geographical and temporal origins. Food Control, 146, 109509.

Ryan, K. et al. (2020) “Tryptophan metabolite concentrations in depressed patients before and after electroconvulsive therapy” Brain Behaviour Immunology, 83, pp.153-162. Available at https://doi.org/10.1016/j.bbi.2019.10.005.

Rybakowski, J., 2014. Suicidal Behavior in Schizophrenia may be Related to Low Lipid Levels. Medical Science Monitor, 20, pp.1486-1490.

Sadeghi, M., Roohafza, H., Afshar, H., Rajabi, F., Ramzani, M., Shemirani, H., & Sarafzadeghan, N. (2011). Relationship between depression and apolipoproteins A and B: a case-control study. Clinics, 66, 113-117.

Sadeghi, O., Keshteli, A. H., Afshar, H., Esmaillzadeh, A., & Adibi, P. (2021). Adherence to Mediterranean dietary pattern is inversely associated with depression, anxiety and psychological distress. Nutritional neuroscience, 24(4), 248-259.

Sallis, J. F., Floyd, M. F., Rodríguez, D. A., & Saelens, B. E. (2012). Role of built environments in physical activity, obesity, and cardiovascular disease. Circulation, 125(5), 729-737.

Saul H;Gursul D;Cassidy S;Duffy L;Lewis G;Lewis G; (2022) Almost half of people on long term antidepressants can stop without relapsing, BMJ (Clinical research ed.). Available at: https://pubmed.ncbi.nlm.nih.gov/35764321/ (Accessed: 16 May 2023).

Scanlon, S. M., Williams, D. C., & Schloss, P. (2001). Membrane cholesterol modulates serotonin transporter activity. Biochemistry, 40(35), 10507-10513.

Schuch, F. B., Vancampfort, D., Richards, J., Rosenbaum, S., Ward, P. B., & Stubbs, B. (2018). Exercise as a treatment for depression: A meta-analysis adjusting for publication bias. Journal of Psychiatric Research, 100, 115-121.

Serafini, G., Adavastro, G., Canepa, G., Capobianco, L., Conigliaro, C., Pittaluga, F., Murri, M., Valchera, A., De Berardis, D., Pompili, M., Lindqvist, D., Brundin, L. and Amore, M., 2017. Abnormalities in Kynurenine Pathway Metabolism in Treatment-Resistant Depression and Suicidality: A Systematic Review. CNS & Neurological Disorders - Drug Targets, 16(4).

Setoyama, D., Kato, T., Hashimoto, R., Kunugi, H., Hattori, K., Hayakawa, K., Sato-Kasai, M., Shimokawa, N., Kaneko, S., Yoshida, S., Goto, Y., Yasuda, Y., Yamamori,

H., Ohgidani, M., Sagata, N., Miura, D., Kang, D. and Kanba, S., 2016. Plasma Metabolites Predict Severity of Depression and Suicidal Ideation in Psychiatric Patients-A Multicenter Pilot Analysis. PLOS ONE, 11(12), p.e0165267.

Sevincok, L., Buyukozturk, A., & Dereboy, F. (2001). Serum lipid concentrations in patients with comorbid generalized anxiety disorder and major depressive disorder. The Canadian Journal of Psychiatry, 46(1), 68-71.

Shibata, H., Kumagai, S., Watanabe, S., & Suzuki, T. (1999). Relationship of serum cholesterols and vitamin E to depressive status in the elderly. Journal of epidemiology, 9(4), 261-267.

Shin, J. Y., Suls, J., & Martin, R. (2008). Are cholesterol and depression inversely related? A meta-analysis of the association between two cardiac risk factors. Annals of Behavioral Medicine, 36(1), 33-43.

Simić, K., Miladinović, Z., Todorović, N., Trifunović, S., Avramović, N., Gavrilović, A., ... & Mandić, B. (2023). Metabolomic Profiling of Bipolar Disorder by 1H-NMR in Serbian Patients. Metabolites, 13(5), 607.

Simon, G. E., Von Korff, M., Saunders, K., Miglioretti, D. L., Crane, P. K., Van Belle, G., & Kessler, R. C. (2006). Association between obesity and psychiatric disorders in the US adult population. Archives of general psychiatry, 63(7), 824-830.

Skorobogatov, K., De Picker, L., Verkerk, R., Coppens, V., Leboyer, M., Mueller, N., & Morrens, M. (2021). Brain versus blood: A systematic review on the concordance between peripheral and central kynurenine pathway measures in psychiatric disorders. Frontiers in Immunology, 12, 716980.

Skovlund, C.W. et al. (2016) ‘Association of hormonal contraception with depression’, JAMA Psychiatry, 73(11), p. 1154. doi:10.1001/jamapsychiatry.2016.2387.

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., & Siuzdak, G. (2020). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical Chemistry, 78(3), 779-787.

Smith, J. K., & Johnson, A. R. (2020). The relationship between blood glucose levels and anxiety and depression symptoms: A quantitative analysis. Journal of Mental Health, 45(3), 321-335.

Spector, T. (2021) TwinsUK Data and Samples, TwinsUK. Available at: https://twinsuk.ac.uk/resources-for-researchers/data-samples/ (Accessed: 15 May 2023).

Speed, M.S. et al. (2019) Investigating the association between body fat and depression via Mendelian randomization [Preprint]. doi:10.1101/539601.

Steegmans, P. H., Hoes, A. W., Bak, A. A., van der Does, E., & Grobbee, D. E. (2000). Higher prevalence of depressive symptoms in middle-aged men with low serum cholesterol levels. Psychosomatic medicine, 62(2), 205-211.

Stephan, K., Friston, K. and Frith, C., 2009. Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-monitoring. Schizophrenia Bulletin, 35(3), pp.509-527.

Stewart, J. C., Rand, K. L., Muldoon, M. F., & Kamarck, T. W. (2009). A prospective evaluation of the directionality of the depression–inflammation relationship. Brain, behavior, and immunity, 23(7), 936-944.

Stuart, M. J., & Baune, B. T. (2012). Depression and type 2 diabetes: inflammatory mechanisms of a psychoneuroendocrine co-morbidity. Neuroscience & Biobehavioral Reviews, 36(1), 658-676.

Stubbs, B., Vancampfort, D., Rosenbaum, S., Firth, J., Cosco, T., Veronese, N., ... & Schuch, F. B. (2017). An examination of the anxiolytic effects of exercise for people with anxiety and stress-related disorders: A meta-analysis. Psychiatry Research, 249, 102-108.

Sun, S., Yang, S., Mao, Y., Jia, X., & Zhang, Z. (2015). Reduced cholesterol is associated with the depressive-like behavior in rats through modulation of the brain 5-HT1A receptor. Lipids in Health and Disease, 14(1), 1-11.

Suvitaival, T., Mantere, O., Kieseppä, T., Mattila, I., Pöhö, P., Hyötyläinen, T., Suvisaari, J. and Orešič, M., 2016. Serum metabolite profile associates with the development of metabolic co-morbidities in first-episode psychosis. Translational Psychiatry, 6(11), pp.e951-e951.

Svensson, J., Svanborg, C., Plavén-Sigray, P., Kaldo, V., Halldin, C., Schain, M. and Lundberg, J., 2021. Serotonin transporter availability increases in patients recovering from a depressive episode. Translational Psychiatry, 11(1).

Swendsen, J. D., Merikangas, K. R., Canino, G. J., Kessler, R. C., Rubio-Stipec, M., Angst, J., ... & Wittchen, H. U. (2009). The comorbidity of alcoholism with anxiety and depressive disorders in four geographic communities. Comprehensive Psychiatry, 40(2), 106-113.

Tan, H., Chen, W., Liu, Q., Yang, G. and Li, K., 2018. Pectin Oligosaccharides Ameliorate Colon Cancer by Regulating Oxidative Stress- and Inflammation-Activated Signaling Pathways. Frontiers in Immunology, 9.

Tasic, L., Pontes, J., Carvalho, M., Cruz, G., Dal Mas, C., Sethi, S., Pedrini, M., Rizzo, L., Zeni-Graiff, M., Asevedo, E., Lacerda, A., Bressan, R., Poppi, R., Brietzke, E. and Hayashi, M., 2017. Metabolomics and lipidomics analyses by 1 H nuclear magnetic resonance of schizophrenia patient serum reveal potential peripheral biomarkers for diagnosis. Schizophrenia Research, 185, pp.182-189.

Terao, T., Iwata, N., Kanazawa, K., Takano, T., Takahashi, N., Hayashi, T., & Sugawara, Y. (2000). Low serum cholesterol levels and depressive state in human dock visitors. Acta Psychiatrica Scandinavica, 101(3), 231-234.

Thomas, A.J. et al. (2005) ‘Increase in interleukin-1β in late-Life Depression’, American Journal of Psychiatry, 162(1), pp. 175–177. doi:10.1176/appi.ajp.162.1.175.

Tian, J., Peng, G., Gao, X., Zhou, Y., Xing, J., Qin, X. and Du, G., 2014. Dynamic analysis of the endogenous metabolites in depressed patients treated with TCM formula Xiaoyaosan using urinary 1H NMR-based metabolomics. Journal of Ethnopharmacology, 158, pp.1-10.

Tian, J., Peng, G., Wu, Y., Zhou, J., Xiang, H., Gao, X., Zhou, Y., Qin, X. and Du, G., 2016. A GC–MS urinary quantitative metabolomics analysis in depressed patients treated with TCM formula of Xiaoyaosan. Journal of Chromatography B, 1026, pp.227-235.

Ting, E.Y.-C., Yang, A.C. and Tsai, S.-J. (2020) ‘Role of interleukin-6 in depressive disorder’, International Journal of Molecular Sciences, 21(6), p. 2194. doi:10.3390/ijms21062194.

Torres, R. J. D. A., Luchini, A., Barberini, L. Y., Precoma, L., Torres, C. L. D. A., Torres, R. A. D. A., ... & Precoma, D. B. (2014). Expression of TNF-α and IL-6 cytokines in the choroid and sclera of hypercholesterolemic rabbits. Arquivos Brasileiros de Oftalmologia, 77, 168-172.

Trpkovic, A., Resanovic, I., Stanimirovic, J., Radak, D., Mousa, S. A., Cenic-Milosevic, D., ... & Isenovic, E. R. (2015). Oxidized low-density lipoprotein as a biomarker of cardiovascular diseases. Critical reviews in clinical laboratory sciences, 52(2), 70-85.

Trygg, J., & Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). Journal of Chemometrics: A Journal of the Chemometrics Society, 16(3), 119-128.

Tsapakis, E. and Travis, M., 2002. Glutamate and psychiatric disorders. Advances in Psychiatric Treatment, 8(3), pp.189-197.

Tsui-Pierchala, B. A., Encinas, M., Milbrandt, J., & Johnson, E. M. (2002). Lipid rafts in neuronal signaling and function. Trends in neurosciences, 25(8), 412-417.

Uguz, S., Bozdemir, N., Güzel, R., Burgut, R., Saatçi, E., & Akpinar, E. (2004). The relationship between cholesterol levels and depression in the elderly.

Vajargah, M. F., Mansouri, K., & Rezaei, J. (2014). The combination of chemometrics methods with UV-Visible spectroscopy as an efficient tool for quantitative analysis of furosemide in pharmaceutical products. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 123, 120-127.

van Himbergen, T. M., Beiser, A. S., Ai, M., Seshadri, S., Otokozawa, S., Au, R., Thongtang, N., Wolf, P. A., & Schaefer, E. J. (2012). Biomarkers for insulin resistance and inflammation and the risk for all-cause dementia and alzheimer disease: results from the Framingham Heart Study. Archives of neurology, 69(5), 594–600. https://doi.org/10.1001/archneurol.2011.670

van Reedt Dortland, A.K. et al. (2009) “Associations between serum lipids and major depressive disorder,” The Journal of Clinical Psychiatry, 71(06), pp. 729–736. Available at: https://doi.org/10.4088/jcp.08m04865blu.

Vinagre, J. C., Vinagre, C. G., Pozzi, F. S., Slywitch, E., & Maranhão, R. C. (2013). Metabolism of triglyceride-rich lipoproteins and transfer of lipids to high-density lipoproteins (HDL) in vegan and omnivore subjects. Nutrition, Metabolism and Cardiovascular Diseases, 23(1), 61-67.

Wagner, C.J. et al. (2019) “LDL cholesterol relates to depression, its severity, and the prospective course,” Progress in Neuro-Psychopharmacology and Biological Psychiatry, 92, pp. 405–411. Available at: https://doi.org/10.1016/j.pnpbp.2019.01.010.

Wallace, C. J., & Milev, R. (2017). The effects of probiotics on depressive symptoms in humans: a systematic review. Annals of general psychiatry, 16(1), 1-10.

Walther, A., Cannistraci, C. V., Simons, K., Durán, C., Gerl, M. J., Wehrli, S., & Kirschbaum, C. (2018). Lipidomics in Major Depressive Disorder. Frontiers in psychiatry, 9, 459. https://doi.org/10.3389/fpsyt.2018.00459

Wang, C. et al. (2021) 'Dietary epa-enriched phospholipids alleviate chronic stress and lps-induced depression- and anxiety-like behavior by regulating immunity and neuroinflammation', Molecular Nutrition & Food Research, 65(17), p. 2100009. doi:10.1002/mnfr.202100009.

Wang, D., Zhang, S., Li, L., Liu, X., Mei, K., & Wang, X. (2010). Structural insights into the assembly and activation of IL-1β with its receptors. Nature immunology, 11(10), 905-911.

Wang, F., Guo, L., Zhang, T., Cui, Z., Wang, J., Zhang, C., ... & Peng, Z. (2022). Alterations in plasma lipidomic profiles in adult patients with schizophrenia and major depressive disorder. Medicina, 58(11), 1509.

Wang, H., He, Y., Sun, Z., Ren, S., Liu, M., Wang, G., & Yang, J. (2022). Microglia in depression: An overview of microglia in the pathogenesis and treatment of depression. Journal of Neuroinflammation, 19(1), 132.

Wang, H., Xu, J., Lazarovici, P., Quirion, R. and Zheng, W., 2018. cAMP Response Element-Binding Protein (CREB): A Possible Signaling Molecule Link in the Pathophysiology of Schizophrenia. Frontiers in Molecular Neuroscience, 11.

Westerhuis, J. A., Hoefsloot, H. C., Smit, S., Vis, D. J., Smilde, A. K., van Velzen, E. J., ... & van Duijnhoven, J. P. (2008). Assessment of PLSDA cross-validation. Metabolomics, 4(1), 81-89.

Whipp, A. M., Heinonen-Guzejev, M., Pietiläinen, K. H., van Kamp, I., & Kaprio, J. (2022). Branched-chain amino acids linked to depression in young adults. Frontiers in Neuroscience, 16, 935858.

Whitney, E. J., Krasuski, R. A., Personius, B. E., Michalek, J. E., Maranian, A. M., Kolasa, M. W., ... & Gotto Jr, A. M. (2005). A randomized trial of a strategy for increasing high-density lipoprotein cholesterol levels: effects on progression of coronary heart disease and clinical events. Annals of internal medicine, 142(2), 95-104.

Wiklund, S., Johansson, E., Sjöström, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., ... & Trygg, J. (2008). Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. Analytical chemistry, 80(1), 115-122.

Winter, G., Hart, R. A., Charlesworth, R. P., & Sharpley, C. F. (2018). Gut microbiome and depression: what we know and what we need to know. Reviews in the Neurosciences, 29(6), 629-643.

Wu, H., Tang, D., Zhao, X., Yuan, G., & Su, X. (2019). Molecular imaging. Nuclear Medicine in Oncology: Molecular Imaging and Target Therapy, 153-176.

Wu, Z., Yu, H., Tian, Y., Wang, Y., He, Y., Lan, T., ... & Xie, P. (2022). Non-targeted metabolomics profiling of plasma samples from patients with major depressive disorder. Frontiers in Psychiatry, 12, 810302.

Wu, Z., Zhao, P., Long, Z., Li, J., Yang, G., Zhang, Q., Duan, G. and Li, H., 2019. Biomarker screening for antenatal depression in women who underwent caesarean section: a matched observational study with plasma Lipidomics. BMC Psychiatry, 19(1).

Wysokiński, A., Strzelecki, D. and Kłoszewska, I., 2015. Levels of triglycerides, cholesterol, LDL, HDL and glucose in patients with schizophrenia, unipolar depression and bipolar disorder. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 9(3), pp.168-176.

Xu, K., Hu, D., Liu, Y., Han, Y., Guo, X., Teng, F., & Zhou, Y. (2019). Relationship of suicidal ideation with demoralization, depression, and anxiety: a study of cancer patients in mainland China. The Journal of nervous and mental disease, 207(5), 326-332.

Xuan, J., Pan, G., Qiu, Y., Yang, L., Su, M., Liu, Y., Chen, J., Feng, G., Fang, Y., Jia, W., Xing, Q. and He, L., 2011. Metabolomic Profiling to Identify Potential Serum Biomarkers for Schizophrenia and Risperidone Action. Journal of Proteome Research, 10(12), pp.5433-5443.

Yang, J. L., Liu, D. X., Jiang, H., Pan, F., Ho, C. S., & Ho, R. C. (2016). The effects of high-fat-diet combined with chronic unpredictable mild stress on depression-like behavior and leptin/leprb in male rats. Scientific reports, 6(1), 35239.

Yang, J., Chen, T., Sun, L., Zhao, Z., Qi, X., Zhou, K., Cao, Y., Wang, X., Qiu, Y., Su, M., Zhao, A., Wang, P., Yang, P., Wu, J., Feng, G., He, L., Jia, W. and Wan, C.,

2011. Potential metabolite markers of schizophrenia. Molecular Psychiatry, 18(1), pp.67-78.

Yang, J., Guo, H., Sun, D., Duan, J., Rao, X., Xu, F., Manyande, A., Tang, Y., Wang, J. and Wang, F., 2020. Elevated glutamate, glutamine and GABA levels and reduced taurine level in a schizophrenia model using an in vitro proton nuclear magnetic resonance method. American Journal of Translational Research. 11(9). Pp.5919-5931.

Yang, X., Sun, L., Zhao, A., Hu, X., Qing, Y., Jiang, J., Yang, C., Xu, T., Wang, P., Liu, J., Zhang, J., He, L., Jia, W. and Wan, C., 2017. Serum fatty acid patterns in patients with schizophrenia: a targeted metabonomics study. Translational Psychiatry, 7(7), pp.e1176-e1176.

Yoshikawa, A., Nishimura, F., Inai, A., Eriguchi, Y., Nishioka, M., Takaya, A., Tochigi, M., Kawamura, Y., Umekage, T., Kato, K., Sasaki, T., Ohashi, Y., Iwamoto, K., Kasai, K. and Kakiuchi, C., 2018. Mutations of the glycine cleavage system genes possibly affect the negative symptoms of schizophrenia through metabolomic profile changes. Psychiatry and Clinical Neurosciences, 72(3), pp.168-179.

Yoshimi, N., Futamura, T., Kakumoto, K., Salehi, A., Sellgren, C., Holmén-Larsson, J., Jakobsson, J., Pålsson, E., Landén, M. and Hashimoto, K., 2016. Blood metabolomics analysis identifies abnormalities in the citric acid cycle, urea cycle, and amino acid metabolism in bipolar disorder. BBA Clinical, 5, pp.151-158.

Yudkoff, M., Daikhin, Y., Melø, T. M., Nissim, I., Sonnewald, U., & Nissim, I. (2007). The ketogenic diet and brain metabolism of amino acids: relationship to the anticonvulsant effect. Annu. Rev. Nutr., 27, 415-430.

Zagorski, N. (2018) ‘Low acetyl-l-carnitine levels may be linked to depression’, Psychiatric News, 53(19). doi:10.1176/appi.pn.2018.9b5.

Zhang, A. H., Ma, Z. M., Sun, H., Zhang, Y., Liu, J. J., Wu, F. F., ... & Wang, X. J. (2019). Metabolomics reveals the effect of Xue-Fu-Zhu-Yu decoction on blood stasis syndrome via UPLC-Q-TOF/MS analysis. Biomedical Chromatography, 33(12), e4687.

Zhang, A., Sun, H. and Wang, X., 2012. Serum metabolomics as a novel diagnostic approach for disease: a systematic review. Analytical and Bioanalytical Chemistry, 404(4), pp.1239-1245.

Zhang, L., Bijker, M. S., & Herzog, H. (2011). The neuropeptide Y system: pathophysiological and therapeutic implications in obesity and cancer. Pharmacology & therapeutics, 131(1), 91-113.

Zhang, Q. et al. (2022) 'Low cholesterol is not associated with depression: Data from the 2005-2018 National Health and Nutrition Examination Survey', Lipids in Health and Disease, 21(1). doi:10.1186/s12944-022-01645-7.

Zhang, X. H., Liu, R. J., Zheng, J. J., Qing, X. D., Yang, K. L., Zhang, Y. Q., ... & Nie, J. F. (2023). Authentication of the production season of Xinyang Maojian green tea using two-dimensional fingerprints coupled with chemometric multivariate calibration and pattern recognition analysis. LWT, 176, 114556.

Zheng, H., Tumin, D., Qian, Z. (2015). Obesity and mortality risk: New findings from body mass index trajectories. American Journal of Epidemiology, 181(5), 328-335.

Zheng, H., Zheng, P., Zhao, L., Jia, J., Tang, S., Xu, P., et al. (2017). Predictive diagnosis of major depression using NMR-based metabolomics and least-squares support vector machine. Clinica Chimica Acta 464, 223–227. doi: 10.1016/j.cca.2016.11.039

Zheng, P., Chen, J., Zhou, C., Zeng, L., Li, K., Sun, L., Liu, M., Zhu, D., Liang, Z. and Xie, P., 2016. Identification of sex-specific urinary biomarkers for major depressive disorder by combined application of NMR- and GC–MS-based metabonomics. Translational Psychiatry, 6(11), pp.e955-e955.

Zheng, P., Wang, Y., Chen, L., Yang, D., Meng, H., Zhou, D., Zhong, J., Lei, Y., Melgiri, N. and Xie, P., 2013. Identification and Validation of Urinary Metabolite Biomarkers for Major Depressive Disorder. Molecular & Cellular Proteomics, 12(1), pp.207-214.

Zheng, W., Kollmeyer, J., Symolon, H., Momin, A., Munter, E., Wang, E., ... & Merrill Jr, A. H. (2006). Ceramides and other bioactive sphingolipid backbones in health and disease: lipidomic analysis, metabolism and roles in membrane structure, dynamics, signaling and autophagy. Biochimica et Biophysica Acta (BBA)-Biomembranes, 1758(12), 1864-1884.

Zhou, X., Liu, L., Lan, X., Cohen, D., Zhang, Y., Ravindran, A., Yuan, S., Zheng, P., Coghill, D., Yang, L., Hetrick, S., Jiang, X., Benoliel, J., Cipriani, A. and Xie, P., 2018. Polyunsaturated fatty acids metabolism, purine metabolism and inosine as potential independent diagnostic biomarkers for major depressive disorder in children and adolescents. Molecular Psychiatry, 24(10), pp.1478-1488.

Zhu, Y., Jha, S. C., Shutta, K. H., Huang, T., Balasubramanian, R., Clish, C. B., & Kubzansky, L. D. (2022). Psychological Distress and Metabolomic Markers: A Systematic Review of Posttraumatic Stress Disorder, Anxiety, and Subclinical Distress. Neuroscience & Biobehavioral Reviews, 104954.

Zięba, A., Matosiuk, D., & Kaczor, A. A. (2023). The Role of Genetics in the Development and Pharmacotherapy of Depression and Its Impact on Drug Discovery. International Journal of Molecular Sciences, 24(3), 2946.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. Clinical Chemistry, 39(4), 561-577.

.

Appendices

Appendix-1

Table 1 Peripheral metabolites and organic acids found at abnormal levels in depression compared to a control condition (healthy volunteers or medicated/treated depressed people in remission). MDD = major depressive disorder

Metabolite	Reference	Analytical Platform	Biofluid	Subjects	Depression vs control
α-ketoglutarate	(Pan et al., 2018)	GC-MS/ LC-MS	Plasma, CSF	49 MDD subjects, 40 healthy controls	Increased
α-ketoglutarate	(Zheng et al., 2013)	H-NMR	Urine	82 MDD patients, 82 healthy controls	Increased
α-ketoglutarate	(Tian et al., 2014)	H-NMR	Urine	21 depressed	
α-ketoglutarate	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Decreased (for men)
3-hydroxybutyrate	(Setoyama et al., 2016)	LC-MS	Plasma	26 unmedicated with depressive mood, 23 medicated with MDD	Decreased
3-hydroxybutyrate	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Decreased (for women)
3-hydroxyphenylacetate	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Decreased (for women)
3-hydroxyphenylacetate	(Zheng et al., 2013)	H-NMR	Urine	82 MDD patients, 82 healthy	Decreased

				controls	
4-HPA/4-HB	(Bhattacharyya et al., 2019)	LCECA	Plasma	803 MDD patients	Decreased
Alanine	(Tian et al., 2014)	NMR	Urine	21 depressed, 21 depressed after treatment	Increased
Alanine	(Tian et al., 2016)	GC-MS	Urine	25 depressed, 33 healthy	Increased
Alanine	(Liu et al., 2015)	H-NMR	Plasma	16 depressed, 16 healthy	Increased
Alanine	(Chen et al., 2018)	NMR/GC-MS	Urine	48 depressed, 48 healthy controls	Increased
Alanine	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Increased
Alanine	(Zheng et al., 2013)	H-NMR	Urine	82 MDD patients, 82 healthy controls	Decreased
Alanine	(Ding et al., 2014)	GC-MS	Plasma	46 MDD patients, 25 healthy controls	Increased
Ascorbic acid	(Hashimoto et al., 2017)	GC-MS	CSF	28 elderly depressed, 18 healthy controls	Increased
Azelaic acid	(Chen et al., 2018)	NMR, GC-MS	Urine	48 depressed, 48 healthy controls	Increased
Azelaic acid	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Increased (for women)
Cholesterol sulphate	(Wu et al., 2019)	UPLC-MS	Plasma	66: 33 with antenatal depression and 33 without	Decreased
Citrate	(Tian et al., 2014)	H-NMR	Urine	21 depressed, 21 depressed after	Increased

				treatment	
Citrate	(Liu et al., 2015)	H-NMR	Plasma	16 depressed, 16 healthy	Increased
Citrate	(Tian et al., 2016)	GC-MS	Urine	25 depressed, 33 healthy	Increased
Citrate	(Setoyama et al., 2016)	LC-MS	Plasma	26 unmedicated with depressive mood, 23 medicated with MDD	Increased
Citrate	(Zheng et al., 2013)	H-NMR	Urine	82 MDD patients, 82 healthy controls	Increased
Creatine	(Tian et al., 2014)	H-NMR	Urine	21 depressed, 21 depressed after treatment	Decreased
Creatinine	(Setoyama et al., 2016)	LC-MS	Plasma	26 unmedicated with depressive mood, 23 medicated with MDD	Decreased
Creatinine	(Zheng et al., 2013)	H-NMR	Urine	82 MDD patients, 82 healthy controls	Decreased
Dimethylamine	(Tian et al., 2014)	NMR	Urine	21 depressed	Increased
Ethanolamine (EA)	(Ogawa et al., 2015)	CE-TOF-MS	CSF	52 MDD patients, 54 healthy controls	Increased
GABA	(Pan et al., 2018)	GC-MS/ LC-MS	Plasma, CSF	49 MDD subjects, 40 healthy controls	Increased
GABA	(Setoyama et al., 2016)	LC-MS	Plasma	26 unmedicated with depressive mood, 23 medicated with MDD	Decreased
Glucose	(Liu et al., 2015)	H-NMR	Plasma	16 depressed, 16 healthy	Increased

Glutamate	(Martins-de-Souza, 2014)	GC-MS	Plasma	9 elderly MDD, 11 remitted, 10 healthy	Increased
Glutamine	(Liu et al., 2015)	H-NMR	Plasma	16 depressed, 16 healthy	Decreased
Glycerolphosphocholine	(Zheng et al., 2012)	H-NMR	Urine	82 MDD patients, 82 healthy controls	Decreased
Glycine	(Liu et al., 2015)	H-NMR	Plasma	16 depressed, 16 healthy	Increased
Glycine	(Chen et al., 2018)	NMR/GC-MS	Urine	48 depressed, 48 healthy controls	Increased
Glycine	(Ding et al., 2014)	GC-MS	Plasma	46 MDD patients, 25 healthy controls	Increased
Hippurate	(Tian et al., 2016)	GC-MS	Urine	25 depressed, 33 healthy	Increased
Hippurate	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Decreased (for women)
Hippurate	(Zheng et al., 2013)	H-NMR	Urine	82 MDD patients, 82 healthy controls	Decreased
HVA	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Decreased
Indoxyl sulphate	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Increased
Inosine	(Zhou et al., 2018)	UPLC-MS	Plasma	84 depressed, 50 healthy controls	Decreased
Kynurenine	(Pan et al., 2018)	GC-MS/ LC-MS	Plasma, CSF	49 MDD subjects, 40	Decreased

				healthy controls	
Kynurenine	(Ryan et al., 2020)	LC-MS	Plasma	94 MDD patients, 57 healthy controls	Decreased
Kynurenine	(Liu et al., 2018)	HPLC	Plasma	290 MD patients	Decreased
Lactate	(Tian et al., 2014)	H-NMR	Urine	21 depressed, 21 depressed after treatment	Increased
Lactate	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Increased (for women)
Lactate	(Ding et al., 2014)	GC-MS	Plasma	46 MDD patients, 25 healthy controls	Increased
Lactate	(Liu et al., 2015)	H-NMR	Plasma	16 depressed, 16 healthy	Decreased
Lipids	(Liu et al., 2015)	H-NMR	Plasma	16 depressed, 16 healthy	Increased
Oxalic and stearic acid	(Liu et al., 2020)	GC-MS	Plasma	17 depressed, 17 healthy	Increased
Palmitic, oleic, capric, dodecanoic acids	(Zhou et al., 2018)	UPLC-MS	Plasma	84 depressed, 50 healthy controls	Decreased
Phenylalanine	(Liu et al., 2015)	H-NMR	Plasma	16 depressed, 16 healthy	Increased
Phenylalanine	(Tian et al., 2016)	GC-MS	Urine	25 depressed, 33 healthy	Decreased
Phenylalanine	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Decreased (for men)
Phosphatidylcholine	(Wu et al., 2019)	UPLC-MS	Plasma	66: 33 with antenatal depression and 33 healthy	Decreased
Phosphatidylcholine	(Liu et al.,	H-NMR	Plasma	16 depressed, 16	Increased

	2015)			healthy controls	
Serotonin (5-HT)	(Bhattacharyya et al., 2019)	LCECA	Plasma	290 MDD patients	Decreased
Sphingomyelin	(Czysz et al., 2019)	FIA-MS	Plasma	159 depressed (at baseline, and 83 available for follow-up)	Increased
Taurine	(Tian et al., 2014)	H-NMR	Urine	21 depressed	Decreased
Taurine	(Zheng et al., 2013)	H-NMR	Urine	82 MDD patients, 82 healthy controls	Increased
TMAO	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Decreased (for women)
TMAO	(Zheng et al., 2012)	H-NMR	Urine	82 MDD patients, 82 healthy controls	Decreased
Tryptophan	(Pan et al., 2018)	GC-MS/ LC-MS	Plasma, CSF	49 MDD subjects, 40 healthy controls	Decreased
Tryptophan	(Ryan et al., 2020)	LC-MS	Plasma	94 MDD patients, 57 healthy controls	Decreased
Tyrosine	(Tian et al., 2016)	GC-MS	Urine	25 depressed, 33 healthy	Decreased
Tyrosine	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75 healthy controls)	Decreased
Valine	(Liu et al., 2020)	GC-MS	Plasma	17 depressed, 17 healthy	Decreased
Valine	(Zheng et al., 2016)	NMR/GC-MS	Urine	93 MDD patients (43 women vs 48 healthy controls, 50 men vs 75	Decreased (for men)

				healthy controls)	
Valine,	(Liu et al., 2015)	H-NMR	Plasma	16 MDD patients, 16 healthy	Increased
Xanthurenic acid	(Tian et al., 2014)	H-NMR	Urine	21 MDD patients, 21 depressed after treatment	Decreased
EPA, DHA	(Borsini et al., 2020)	UPLC-MS	Plasma	22 MDD patients, 23 healthy controls	Decreased

Appendix-2

Table 2. Peripheral metabolites and organic acids found at abnormal levels in schizophrenia/ psychosis compared to a control condition (healthy volunteers or medicated/treated depressed people in remission).

Metabolite	Reference	Analytical Platform	Biofluid	Subjects	SCZ/Psychosis vs control
α-ketoglutarate	(Cai et al., 2012)	UPLC-MS and NMR	Plasma, urine	32 schizophrenics, 31 healthy controls	Increased
α-ketoglutarate	(Xuan et al., 2011)	GC-MS	Serum	18 schizophrenics, 18 controls	Decreased
α-ketoglutarate	(Tasic et al., 2017)	H-NMR	Plasma	26 schizophrenics, 27 healthy controls	Increased
α-ketoglutarate	(Liu et al., 2014)	GC-MS	PMNC	69 schizophrenics, 85 controls	Decreased
γ-glutarate	(Xuan et al., 2011)	GC-MS	Serum	18 schizophrenics, 18 controls	Decreased
2-hydroxybutyrate	(Kageyama et al., 2016)	CE-TOFMS	Plasma	17 schizophrenics, 17 healthy controls	Increased
3-hydroxybutyrate	(Fukushima et al.,	HPLC	Plasma	25 schizophren	Decreased

	2014)			ics, 27 healthy controls	
3-hydroxybutyrate	(Yang et al., 2011)	GC-MS	Serum	112 schizophrenics, 110 controls	Increased
3-hydroxykynurenine (3-OHKY)	(Condray et al., 2011)	LCEC	Plasma	25 FEP patients,	Increased
Acetate	(Dietrich Mouszalska et al., 2012)	NMR	CSF, serum	82 schizophrenics, 70 controls	Decreased
Arachidonic acid	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Decreased
Arachidonic acid	(Yang et al., 2017)	UPLC-MS	Plasma	110 schizophrenics, 109 healthy controls	Increased
Arginine	(Fournier et al., 2014)	UPLC-MS	Plasma	30 FEP patients, 20 healthy controls	Decreased
Aspartate	(Yoshikawa et al., 2018)	CE-TOFMS	Plasma	474 schizophrenics, 475 healthy controls	Increased
Benzoic, nonanoic, perillic, betaine acids	(Koike et al., 2014)	CE-TOFMS	Plasma	30 FEP patients, 38 healthy	Decreased

				controls	
Citrate	(Cai et al., 2012)	UPLC-MS and NMR	Plasma, urine	32 schizophrenics, 31 healthy controls	Increased
Citrate	(Xuan et al., 2011)	GC-MS	Serum	18 schizophrenics, 18 controls	Decreased
Citrate	(Liu et al., 2015)	GC-MS	Plasma	55 schizophrenics, 55 controls	Decreased
Creatine	(Kageyama et al., 2016)	CE-TOFMS	Plasma	17 schizophrenics, 17 healthy controls	Decreased
Creatine	(Koike et al., 2014)	CE-TOFMS	Plasma	30 FEP patients, 38 healthy controls	Increased
Creatine	(Fournier et al., 2014)	UPLC-MS	Plasma	30 FEP patients, 20 healthy controls	Increased
D-serine (derivative of glycine)	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Decreased
GABA	(Tasic et al., 2017)	H-NMR	Plasma	26 schizophrenics, 27 healthy	Increased

				controls	
Glucose	(Xuan et al., 2011)	GC-MS	Serum	18 schizophrenics, 18 controls	Decreased (FEP)
Glucose	(Holmes et al., 2006)	HMRS	CSF	82 schizophrenics, 70 controls	Decreased (FEP)
Glucose	(Huang et al., 2007)	H-NMR	CSF	54 FEP patients, 70 healthy controls	Increased
Glucose, fructose, ribose-5-phosphate, succinic acid	(Liu et al., 2015)	GC-MS	Plasma	55 schizophrenics, 55 controls	Increased
Glutamate	(Buckley, 2012)	UPLC-MS	Plasma	45 schizophrenics, 57, non-affective psychosis, 37 affective psychosis and controls	Increased
Glutamate	(Yoshikawa et al., 2018)	CE-TOFMS	Plasma	474 schizophrenics, 475 healthy controls	Increased
Glutamate	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Increased
Glutamate	(Oresic et	UPLC-MS	Plasma	45	Increased

	al., 2011)	and GC-TOFMS		schizophrenics, 37 healthy controls	
Glutamate	(Garip et al., 2019)	LC-MS	Plasma	41 schizophrenics, 35 healthy controls	Increased
Glutamate	(Nagai et al., 2017)	LC-MS	Plasma	19 schizophrenics, 19 healthy controls	Increased
Glutamine	(He at al., 2012)	FIA-MS	Plasma	265 schizophrenics, 216 healthy controls	Decreased
Glutathione	(Mico et al., 2011)	Spectrophotometric assays	Plasma	49 schizophrenics, 102 controls	Decreased
Glutathione and its precursor g-L-Glutamyl-cysteine	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Decreased
Glycerate	(Liu et al., 2015)	GC-MS	Plasma	55 schizophrenics, 55 controls	Increased
Glycerate, eicosenoic acid	(Yang et al., 2011)	GC-MS	Serum	112 schizophrenics, 110 controls	Increased

Glycine	(Tasic et al., 2017)	H-NMR	Plasma	26 schizophrenics, 27 healthy controls	Increased
Guanine	(Tasic et al., 2017)	H-NMR	Plasma	26 schizophrenics, 27 healthy controls	Increased
Histidine	(He at al., 2012)	FIA-MS	Plasma	265 schizophrenics, 216 healthy controls	Decreased
Homovanillic acid (HVA)	(Palsson et al., 2017)	HPLC	CSF	113 controls, 175 schizophrenics	Increased
HVA	(Ramirez-Bermudez et al., 2008)	HPLC	CSF	51 psychotic cases	Increased
IL-6, IL-10	(Pedrini et al., 2012)	ELISA	Serum	112 schizophrenics, 110 controls	Increased
Kynurenine	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Increased

Lactate	(Dietrich Mouszalska et al., 2012)	NMR	CSF, serum	82 schizophrenics, 70 controls	Decreased
Lactate	(Holmes et al., 2006)	HMRS	CSF	82 schizophrenics, 70 controls	Decreased
Lactate	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Increased
Lactate	(Huang et al., 2007)	H-NMR	CSF	54 FEP patients, 70 healthy controls	Decreased
Linoleic acid	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Decreased
linoleic acid, eicosenoic acid, pentadecanoic acid	(Suvitaival et al., 2016)	GC-MS	Plasma	36 FEP patients, 19 healthy controls	Decreased
Lipoproteins: LDL, VLDL, HDL	(Cai et al., 2012)	NMR	Plasma	11 psychotic, 11 control	Decreased
Mannitol	(Tasic et al., 2017)	H-NMR	Plasma	26 schizophrenics, 27 healthy controls	Increased
NAA	(Molina et al., 2007)	H-MRS	CSF	24 psychotic,	Decreased

				10 controls	
NAA	(Xuan et al., 2011)	GC-MS	Serum	18 schizophrenics, 18 controls	Decreased
Nervonic acid (similar to oleic acid)	(Kageyama et al., 2017)	LC-MS	Plasma	17 schizophrenics, 19 healthy controls	Increased
Palmitic, linoleic, oleic, stearic acids	(Xuan et al., 2011)	GC-MS	Serum	18 schizophrenics, 18 controls	Decreased
Palmitic, linoleic, stearic oleic acids	(Yang et al., 2017)	UPLC-MS	Plasma	110 schizophrenics, 109 healthy controls	Increased
p-aminobenzoic acid (PABA)	(Tasic et al., 2017)	H-NMR	Plasma	26 schizophrenics, 27 healthy controls	Increased
Pantothenate	(Tasic et al., 2017)	H-NMR	Plasma	26 schizophrenics, 27 healthy controls	Increased
Phosphatidylcholine	(Suvitaival et al., 2016)	GC-MS	Plasma	36 FEP patients, 19 healthy controls	Decreased
Phosphatidylcholine	(He at al., 2012)	FIA-MS	Plasma	265 schizophrenics, 216	Decreased

				healthy controls	
Pregnanediol	(Cai et al., 2012)	UPLC-MS, NMR	Plasma, urine	32 schizophrenics, 31 healthy controls	Increased
Proline	(Fournier et al., 2014)	UPLC-MS	Plasma	30 FEP patients, 20 healthy controls	Increased
Proline, phenylalanine	(Buckley, 2012)	UPLC-MS	Plasma	45 schizophrenics, 57 non-affective psychosis, 37 affective psychosis and controls	Decreased
Pyroglutamic acid	(Yoshikawa et al., 2018)	CE-TOFMS	Plasma	474 schizophrenics, 475 healthy controls	Increased
Serotonin	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Decreased
Taurine	(Cai et al., 2012)	NMR	Urine	11 schizophrenics, 11 controls	Decreased
Threonine, tyrosine (derivative of phenylalanine)	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy	Decreased

				controls	
Triglycerides	(Suvitaival et al., 2016)	GC-MS	Plasma	36 FEP patients, 19 healthy controls	Increased
Triglycerides	(Buckley, 2012)	UPLC-MS	Plasma	45 schizophrenics, 57 non-affective psychosis, 37 affective psychosis and controls	Increased
Tryptophan	(Xuan et al., 2011)	GC-MS	Serum	18 schizophrenics, 18 controls	Decreased
Tryptophan	(Fukushima et al., 2014)	HPLC	Plasma	25 schizophrenics, 27 healthy controls	Increased
Tryptophan	(Fournier et al., 2014)	UPLC-MS	Plasma	30 FEP patients, 20 healthy controls	Decreased
Tyrosine	(Fournier et al., 2014)	UPLC-MS	Plasma	30 FEP patients, 20 healthy controls	Decreased

Appendix-3

Table 3: Comparison of results stratified according to class of metabolite

Metabolite Class	Schizophrenia	Depression
Fatty acids (triglycerides, phospholipids and cholesterol esters), fatty acid metabolism and lipoproteins**:** eicosenoic acid, 3-hydroxybutyrate, linoleic acid, oleic acid, palmitic stearic acid**:** high-density lipoprotein (HDL), low-density lipoprotein (LDL), very low-density lipoprotein (VLDL), lipid, lipoprotein, phosphatidylcholine, pantothenate	- Triglycerides upregulated in 2 studies - 3-hydroxybutyrate upregulated in 1 study, downregulated in another study - Linoleic acid downregulated in 3 studies, upregulated in 1 study - Oleic acid upregulated in 1 study, downregulated in another study - Arachidonic acid downregulated in 1 study, upregulated in another study - Phosphatidylcholine downregulated in 2 studies - VLDL, HDL, LDL downregulated in 1 study - Pantothenate upregulated in 1 study	- Oxalic and stearic acid downregulated in 1 study - Phosphatidylcholine downregulated in 1 study, upregulated in another study - Inosine downregulated in 1 study - Palmitic, oleic, dodecanoic and capric acid downregulate din 1 study - Sphingomyelin upregulated in 1 study - Lipids upregulated in 1 study
Carbohydrate metabolism: acetate, lactate, pyruvate, glucose	- Glucose upregulated in 4 studies, with 2 studies specifically looking at FEP patients - Acetate downregulated in 1 study - Lactate downregulated in 2 studies but	- Glucose upregulated in 1 study - Lactate upregulated in 2 studies, downregulated in 1 study

	upregulated in 1	
Amino acid metabolism: glutamate, cystine, ornithine, arginine, glutamine, histidine, valine, trimethylamine-N-oxide (TMAO), inositol phosphate, myoinositol, N-acetyl aspartate (NAA), aspartate, glycine, ethanolamine, azelaic acid, tyrosine	- Glutamate upregulated in 4 studies - NAA downregulated in 2 studies - Proline and phenylalanine downregulated in 2 studies - Pyroglutamic acid upregulated in 1 study - Aspartate upregulated in 1 study - Glutamine downregulated in 1 study - Histidine downregulated in 1 study - Glycine upregulated in 1 study but its derivative D-serine downregulated in another study - α-ketoisovaleric acid (valine metabolite) downregulated in 1 study - Guanine downregulated in 1 study - Tyrosine downregulated in 2 studies - Mannitol upregulated in 1 study - PABA upregulated in 1 study - Arginine downregulated in 1	- Glutamine downregulated in 1 study - Glutamate upregulated in 1 study - Phenylalanine upregulated in 1 study, downregulated in another study - Glycine and leucine upregulated in 1 study - Alanine upregulated in 6 studies - Dimethylamine upregulated in 1 study - Tyrosine downregulated in 1 study - Valine upregulated in 1 study and downregulated in another study - 4-hydroxyphenylacetic acid and 4-hydroxybenzoic acid downregulated in 1 study - Ethanolamine downregulated in 1 study - TMAO downregulated in 2 studies, upregulated in 1 study - Azelaic acid upregulated in 2 studies - Tyrosine downregulated in 2 studies

	study	
Antioxidants: taurine, glutathione, ascorbic acid	- Glutathione downregulated in 2 studies - Taurine downregulated in 1 study	- Taurine downregulated in 1 study - Ascorbic acid upregulated in 1 study
Tryptophan metabolism: tryptophan, serotonin, xanthurenic acid, kynurenine, GABA	- Tryptophan downregulated in 2 studies but upregulated in 1 study - Kynurenine upregulated in 1 study - Serotonin downregulated in 1 study - GABA upregulated in 1 study	- Serotonin downregulated in 1 study - Tryptophan downregulated in 1 study - Kynurenine downregulated in 2 studies - GABA upregulated in 1 study, downregulated in another
Vitamin E metabolism: γ-tocopherol, α-tocopherol	- α-tocopherol downregulated in 1 study - γ-tocopherol downregulated in 1 study	- None found
Dopamine metabolism: homovanillic acid (HVA)	- HVA upregulated in 2 studies	- HVA downregulated in 1 study
Other pathways: Hippurate, creatine, creatinine, pregnanediol, 3-hydroxybutyrate, acetoacetate, 3-hydroxyphenylacetate	- Pregnanediol upregulated in 1 study - Creatine downregulated in 1 study, upregulated in 2 studies - 3-hydroxybutyrate downregulated in 1 study, upregulated in another study	- Creatine downregulated in 1 study - Creatinine downregulated in 2 studies - Hippurate downregulated in 2 studies, upregulated in 1 study - 3-hydroxybutyrate downregulated in 2 studies

		- 3-hydroxyphenylacetate downregulated in 2 studies
Tricarboxylic acid cycle: citrate, α-ketoglutarate, glycerate	- Citrate upregulated in 1 study and downregulated in 2 other studies - α-ketoglutarate upregulated in 1 study and downregulated in another study - Glycerate upregulated in 2 studies	- Citrate upregulated in 4 studies - α-ketoglutarate upregulated in 2 studies downregulated in 1 study

R code used to generate forest plots in R Studio, using citrate as an example

```
citratedata<-read.csv("/Users/alexandralim/Documents/R data/CITRATE.csv")
citrateresults<- metacont(N_MDD,
                          Mean_MDD,
                          SD_MDD,
                          N_Comp,
                          Mean_Comp,
                          SD_Comp,
                          data= citratedata,
                          studlab= Author_and_date,
                          comb.fixed=FALSE,
                          comb.random= TRUE,
                          sm="SMD")

forest(citrateresults)
```

Figure 5: Forest plot of standardised mean difference (SMD) in alanine levels found in the plasma or urine of patients with schizophrenia and healthy controls excluding Tian et al 2016 study

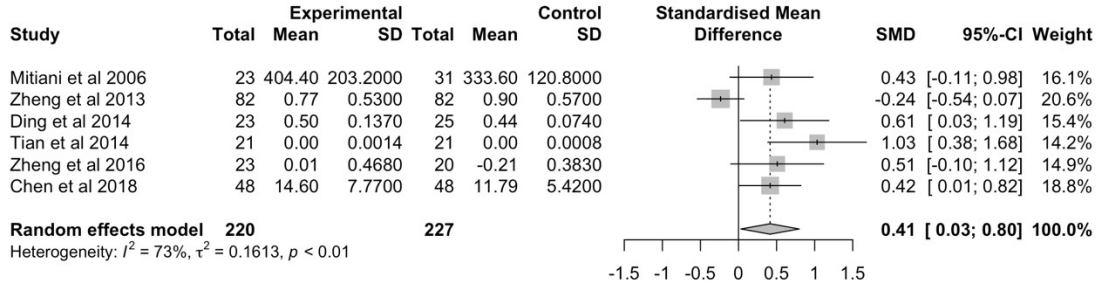
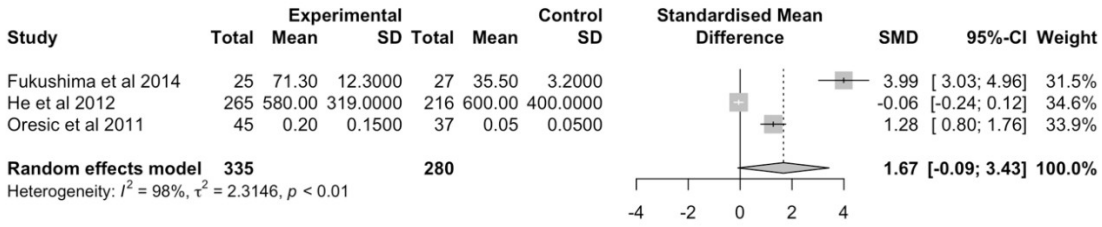


Figure 6: Forest plot of standardised mean difference (SMD) in glutamate levels found in the plasma or urine of patients with schizophrenia and healthy controls excluding Yoshikawa and Garip studies



Appendix 4

Hospital Anxiety and Depression Scale (HADS)

Hospital Anxiety and Depression Scale (HADS)

Tick the box beside the reply that is closest to how you have been feeling in the past week.
Don't take too long over you replies: your immediate is best.

D	A		D	A	
		I feel tense or 'wound up':			**I feel as if I am slowed down:**
	3	Most of the time	3		Nearly all the time
	2	A lot of the time	2		Very often
	1	From time to time, occasionally	1		Sometimes
	0	Not at all	0		Not at all
		I still enjoy the things I used to enjoy:			**I get a sort of frightened feeling like 'butterflies' in the stomach:**
0		Definitely as much		0	Not at all
1		Not quite so much		1	Occasionally
2		Only a little		2	Quite Often
3		Hardly at all		3	Very Often
		I get a sort of frightened feeling as if something awful is about to happen:			**I have lost interest in my appearance:**
	3	Very definitely and quite badly	3		Definitely
	2	Yes, but not too badly	2		I don't take as much care as I should
	1	A little, but it doesn't worry me	1		I may not take quite as much care
	0	Not at all	0		I take just as much care as ever
		I can laugh and see the funny side of things:			**I feel restless as I have to be on the move:**
0		As much as I always could		3	Very much indeed
1		Not quite so much now		2	Quite a lot
2		Definitely not so much now		1	Not very much
3		Not at all		0	Not at all
		Worrying thoughts go through my mind:			**I look forward with enjoyment to things:**
	3	A great deal of the time	0		As much as I ever did
	2	A lot of the time	1		Rather less than I used to
	1	From time to time, but not too often	2		Definitely less than I used to
	0	Only occasionally	3		Hardly at all
		I feel cheerful:			**I get sudden feelings of panic:**
3		Not at all		3	Very often indeed
2		Not often		2	Quite often
1		Sometimes		1	Not very often
0		Most of the time		0	Not at all
		I can sit at ease and feel relaxed:			**I can enjoy a good book or radio or TV program:**
	0	Definitely	0		Often
	1	Usually	1		Sometimes
	2	Not Often	2		Not often
	3	Not at all	3		Very seldom

Please check you have answered all the questions

Appendix-5

Correlations

		HADs	Age	BMI	Glucose	Insulin	Smoking	Alcohol	Interleukin10	Interleukin1B	Exercise	BP_Age	HC_Age	Income	Supplement	Depression
Age	Pearson Correlation	0.118														
	Sig. (2-tailed)	0.215														
	N	113														
BMI	Pearson Correlation	0.112	0.113													
	Sig. (2-tailed)	0.511	0.504													
	N	37	37													
Glucose	Pearson Correlation	-.221*	0.133	0.216												
	Sig. (2-tailed)	0.036	0.21	0.243												
	N	91	91	31												
Insulin	Pearson Correlation	0.092	.598**	-.639*	.634**											
	Sig. (2-tailed)	0.584	0	0.047	0											
	N	38	38	10	38											
Smoking	Pearson Correlation	.211*	.225*	0.008	-0.063	-0.127										
	Sig. (2-tailed)	0.027	0.019	0.962	0.562	0.455										
	N	109	109	36	88	37										
Alcohol	Pearson Correlation	.194*	0.154	-0.064	0.089	.391*	0.126									
	Sig. (2-tailed)	0.04	0.103	0.705	0.401	0.015	0.19									
	N	113	113	37	91	38	109									
Interleukin10	Pearson Correlation	-.191*	0.072	-0.181	0.047	-0.104	0.07	-.196*								
	Sig. (2-tailed)	0.043	0.45	0.285	0.658	0.534	0.468	0.038								
	N	113	113	37	91	38	109	113								
Interleukin1B	Pearson Correlation	-.204*	0.124	-0.096	0.088	-0.093	0.081	-0.173	.782**							
	Sig. (2-tailed)	0.03	0.189	0.571	0.407	0.579	0.405	0.067	0							
	N	113	113	37	91	38	109	113	113							
Exercise	Pearson Correlation	-.207*	-0.007	0.066	0.041	0.32	-0.135	.198*	-0.064	-0.056						
	Sig. (2-tailed)	0.039	0.942	0.712	0.718	0.061	0.189	0.048	0.528	0.582						
	N	100	100	34	81	35	96	100	100	100						
BP_Age	Pearson Correlation	0.144	0.152	-0.155	0.097	.543**	.259*	0.038	-0.04	-0.06	-.251*					
	Sig. (2-tailed)	0.174	0.149	0.421	0.404	0.002	0.015	0.718	0.703	0.571	0.024					
	N	91	91	29	76	30	87	91	91	91	81					
HC_Age	Pearson Correlation	-0.025	0.082	0.156	0.092	0.126	0.107	0.048	-0.155	-0.109	-0.09	.411**				
	Sig. (2-tailed)	0.815	0.441	0.419	0.428	0.509	0.322	0.652	0.142	0.303	0.426	0				
	N	91	91	29	76	30	87	91	91	91	81	91				
Income	Pearson Correlation	0.201	-.499**	-0.256	-0.021	-0.351	-0.061	0.032	-0.171	-0.187	-0.14	-0.001	-0.013			
	Sig. (2-tailed)	0.057	0	0.18	0.855	0.057	0.572	0.767	0.104	0.077	0.214	0.995	0.902			
	N	91	91	29	76	30	87	91	91	91	81	91	91			
Supplement	Pearson Correlation	.219*	.279**	0.06	0.147	.419**	0.147	-0.081	0.004	0.017	-0.137	0.144	0.146	-0.146		
	Sig. (2-tailed)	0.027	0.005	0.736	0.189	0.01	0.148	0.42	0.969	0.869	0.191	0.197	0.192	0.191		
	N	102	102	34	82	37	98	102	102	102	92	82	82	82		
Depression	Pearson Correlation	.754**	0.099	0.216	-0.108	0.122	0.122	0.119	-0.179	-.193*	-.231*	0.122	-0.007	0.107	.251*	
	Sig. (2-tailed)	0	0.299	0.2	0.308	0.467	0.207	0.21	0.057	0.04	0.021	0.249	0.948	0.312	0.011	
	N	113	113	37	91	38	109	113	113	113	100	91	91	91	102	
Anxiety	Pearson Correlation	.755**	0.107	0.168	-0.121	0.286	0.102	.213*	-0.086	-0.132	-0.077	0.153	-0.014	0.127	0.189	.471**
	Sig. (2-tailed)	0	0.258	0.32	0.252	0.082	0.29	0.024	0.363	0.165	0.447	0.149	0.893	0.23	0.058	0
	N	113	113	37	91	38	109	113	113	113	100	91	91	91	102	113

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix-6

Pearson correlations for HADs with variables in 2-year lipid cohort

Pairs	Pearson Correlation	P-Value
HADs - Age	0.118	0.215
HADs - BMI	0.112	0.510
HADs - Glucose	-0.221	0.036
HADs - Insulin	0.092	0.584
HADs - Depression	0.754	0.000
HADs - Anxiety	0.755	0.000
HADs - Smoking	0.216	0.024
HADs - Alcohol	0.198	0.037
HADs - Interleukin10	-0.191	0.043
HADs - Interleukin1B	-0.204	0.030
HADs - Exercise	-0.207	0.039
HADs - HC_Age	0.144	0.174
HADs - BP_Age	0.030	0.778
HADs - Income	0.201	0.057
HADs - Supplement	0.219	0.027
HADs - Contra1	-0.117	0.258
HADs - Contra2	0.009	0.954

Appendix-7

Sr. No.	Abbriviation	Name of lipid
1	XXL-VLDL-P	Concentration of chylomicrons and extremely large VLDL particles
2	XXL-VLDL-L	Total lipids in chylomicrons and extremely large VLDL
3	XXL-VLDL-PL	Phospholipids in chylomicrons and extremely large VLDL
4	XXL-VLDL-C	Total cholesterol in chylomicrons and extremely large VLDL
5	XXL-VLDL-CE	Cholesterol esters in chylomicrons and extremely large VLDL
6	XXL-VLDL-FC	Free cholesterol in chylomicrons and extremely large VLDL
7	XXL-VLDL-TG	Triglycerides in chylomicrons and extremely large VLDL
8	XL-VLDL-P	Concentration of very large VLDL particles
9	XL-VLDL-L	Total lipids in very large VLDL
10	XL-VLDL-PL	Phospholipids in very large VLDL
11	XL-VLDL-C	Total cholesterol in very large VLDL
12	XL-VLDL-CE	Cholesterol esters in very large VLDL
13	XL-VLDL-FC	Free cholesterol in very large VLDL
14	XL-VLDL-TG	Triglycerides in very large VLDL
15	L-VLDL-P	Concentration of large VLDL particles
16	L-VLDL-L	Total lipids in large VLDL
17	L-VLDL-PL	Phospholipids in large VLDL
18	L-VLDL-C	Total cholesterol in large VLDL
19	L-VLDL-CE	Cholesterol esters in large VLDL
20	L-VLDL-FC	Free cholesterol in large VLDL
21	L-VLDL-TG	Triglycerides in large VLDL
22	M-VLDL-P	Concentration of medium VLDL particles
23	M-VLDL-L	Total lipids in medium VLDL
24	M-VLDL-PL	Phospholipids in medium VLDL
25	M-VLDL-C	Total cholesterol in medium VLDL
26	M-VLDL-CE	Cholesterol esters in medium VLDL
27	M-VLDL-FC	Free cholesterol in medium VLDL
28	M-VLDL-TG	Triglycerides in medium VLDL
29	S-VLDL-P	Concentration of small VLDL particles
30	S-VLDL-L	Total lipids in small VLDL
31	S-VLDL-PL	Phospholipids in small VLDL
32	S-VLDL-C	Total cholesterol in small VLDL
33	S-VLDL-CE	Cholesterol esters in small VLDL
34	S-VLDL-FC	Free cholesterol in small VLDL
35	S-VLDL-TG	Triglycerides in small VLDL
36	XS-VLDL-P	Concentration of very small VLDL particles
37	XS-VLDL-L	Total lipids in very small VLDL
38	XS-VLDL-PL	Phospholipids in very small VLDL
39	XS-VLDL-C	Total cholesterol in very small VLDL
40	XS-VLDL-CE	Cholesterol esters in very small VLDL
41	XS-VLDL-FC	Free cholesterol in very small VLDL
42	XS-VLDL-TG	Triglycerides in very small VLDL
43	IDL-P	Concentration of IDL particles
44	IDL-L	Total lipids in IDL
45	IDL-PL	Phospholipids in IDL

46	IDL-C	Total cholesterol in IDL
47	IDL-CE	Cholesterol esters in IDL
48	IDL-FC	Free cholesterol in IDL
49	IDL-TG	Triglycerides in IDL
50	L-LDL-P	Concentration of large LDL particles
51	L-LDL-L	Total lipids in large LDL
52	L-LDL-PL	Phospholipids in large LDL
53	L-LDL-C	Total cholesterol in large LDL
54	L-LDL-CE	Cholesterol esters in large LDL
55	L-LDL-FC	Free cholesterol in large LDL
56	L-LDL-TG	Triglycerides in large LDL
57	M-LDL-P	Concentration of medium LDL particles
58	M-LDL-L	Total lipids in medium LDL
59	M-LDL-PL	Phospholipids in medium LDL
60	M-LDL-C	Total cholesterol in medium LDL
61	M-LDL-CE	Cholesterol esters in medium LDL
62	M-LDL-FC	Free cholesterol in medium LDL
63	M-LDL-TG	Triglycerides in medium LDL
64	S-LDL-P	Concentration of small LDL particles
65	S-LDL-L	Total lipids in small LDL
66	S-LDL-PL	Phospholipids in small LDL
67	S-LDL-C	Total cholesterol in small LDL
68	S-LDL-CE	Cholesterol esters in small LDL
69	S-LDL-FC	Free cholesterol in small LDL
70	S-LDL-TG	Triglycerides in small LDL
71	XL-HDL-P	Concentration of very large HDL particles
72	XL-HDL-L	Total lipids in very large HDL
73	XL-HDL-PL	Phospholipids in very large HDL
74	XL-HDL-C	Total cholesterol in very large HDL
75	XL-HDL-CE	Cholesterol esters in very large HDL
76	XL-HDL-FC	Free cholesterol in very large HDL
77	XL-HDL-TG	Triglycerides in very large HDL
78	L-HDL-P	Concentration of large HDL particles
79	L-HDL-L	Total lipids in large HDL
80	L-HDL-PL	Phospholipids in large HDL
81	L-HDL-C	Total cholesterol in large HDL
82	L-HDL-CE	Cholesterol esters in large HDL
83	L-HDL-FC	Free cholesterol in large HDL
84	L-HDL-TG	Triglycerides in large HDL
85	M-HDL-P	Concentration of medium HDL particles
86	M-HDL-L	Total lipids in medium HDL
87	M-HDL-PL	Phospholipids in medium HDL
88	M-HDL-C	Total cholesterol in medium HDL
89	M-HDL-CE	Cholesterol esters in medium HDL
90	M-HDL-FC	Free cholesterol in medium HDL
91	M-HDL-TG	Triglycerides in medium HDL
92	S-HDL-P	Concentration of small HDL particles
93	S-HDL-L	Total lipids in small HDL

94	S-HDL-PL	Phospholipids in small HDL
95	S-HDL-C	Total cholesterol in small HDL
96	S-HDL-CE	Cholesterol esters in small HDL
97	S-HDL-FC	Free cholesterol in small HDL
98	S-HDL-TG	Triglycerides in small HDL
99	XXL-VLDL-PL_%	Phospholipids to total lipds ratio in chylomicrons and extremely large VLDL
100	XXL-VLDL-C_%	Total cholesterol to total lipids ratio in chylomicrons and extremely large VLDL
101	XXL-VLDL-CE_%	Cholesterol esters to total lipids ratio in chylomicrons and extremely large VLDL
102	XXL-VLDL-FC_%	Free cholesterol to total lipids ratio in chylomicrons and extremely large VLDL
103	XXL-VLDL-TG_%	Triglycerides to total lipids ratio in chylomicrons and extremely large VLDL
104	XL-VLDL-PL_%	Phospholipids to total lipds ratio in very large VLDL
105	XL-VLDL-C_%	Total cholesterol to total lipids ratio in very large VLDL
106	XL-VLDL-CE_%	Cholesterol esters to total lipids ratio in very large VLDL
107	XL-VLDL-FC_%	Free cholesterol to total lipids ratio in very large VLDL
108	XL-VLDL-TG_%	Triglycerides to total lipids ratio in very large VLDL
109	L-VLDL-PL_%	Phospholipids to total lipds ratio in large VLDL
110	L-VLDL-C_%	Total cholesterol to total lipids ratio in large VLDL
111	L-VLDL-CE_%	Cholesterol esters to total lipids ratio in large VLDL
112	L-VLDL-FC_%	Free cholesterol to total lipids ratio in large VLDL
113	L-VLDL-TG_%	Triglycerides to total lipids ratio in large VLDL
114	M-VLDL-PL_%	Phospholipids to total lipds ratio in medium VLDL
115	M-VLDL-C_%	Total cholesterol to total lipids ratio in medium VLDL
116	M-VLDL-CE_%	Cholesterol esters to total lipids ratio in medium VLDL
117	M-VLDL-FC_%	Free cholesterol to total lipids ratio in medium VLDL
118	M-VLDL-TG_%	Triglycerides to total lipids ratio in medium VLDL
119	S-VLDL-PL_%	Phospholipids to total lipds ratio in small VLDL
120	S-VLDL-C_%	Total cholesterol to total lipids ratio in small VLDL
121	S-VLDL-CE_%	Cholesterol esters to total lipids ratio in small VLDL
122	S-VLDL-FC_%	Free cholesterol to total lipids ratio in small VLDL
123	S-VLDL-TG_%	Triglycerides to total lipids ratio in small VLDL
124	XS-VLDL-PL_%	Phospholipids to total lipds ratio in very small VLDL
125	XS-VLDL-C_%	Total cholesterol to total lipids ratio in very small VLDL
126	XS-VLDL-CE_%	Cholesterol esters to total lipids ratio in very small VLDL
127	XS-VLDL-FC_%	Free cholesterol to total lipids ratio in very small VLDL
128	XS-VLDL-TG_%	Triglycerides to total lipids ratio in very small VLDL
129	IDL-PL_%	Phospholipids to total lipds ratio in IDL
130	IDL-C_%	Total cholesterol to total lipids ratio in IDL
131	IDL-CE_%	Cholesterol esters to total lipids ratio in IDL
132	IDL-FC_%	Free cholesterol to total lipids ratio in IDL
133	IDL-TG_%	Triglycerides to total lipids ratio in IDL
134	L-LDL-PL_%	Phospholipids to total lipds ratio in large LDL
135	L-LDL-C_%	Total cholesterol to total lipids ratio in large LDL
136	L-LDL-CE_%	Cholesterol esters to total lipids ratio in large LDL

137	L-LDL-FC_%	Free cholesterol to total lipids ratio in large LDL
138	L-LDL-TG_%	Triglycerides to total lipids ratio in large LDL
139	M-LDL-PL_%	Phospholipids to total lipds ratio in medium LDL
140	M-LDL-C_%	Total cholesterol to total lipids ratio in medium LDL
141	M-LDL-CE_%	Cholesterol esters to total lipids ratio in medium LDL
142	M-LDL-FC_%	Free cholesterol to total lipids ratio in medium LDL
143	M-LDL-TG_%	Triglycerides to total lipids ratio in medium LDL
144	S-LDL-PL_%	Phospholipids to total lipds ratio in small LDL
145	S-LDL-C_%	Total cholesterol to total lipids ratio in small LDL
146	S-LDL-CE_%	Cholesterol esters to total lipids ratio in small LDL
147	S-LDL-FC_%	Free cholesterol to total lipids ratio in small LDL
148	S-LDL-TG_%	Triglycerides to total lipids ratio in small LDL
149	XL-HDL-PL_%	Phospholipids to total lipds ratio in very large HDL
150	XL-HDL-C_%	Total cholesterol to total lipids ratio in very large HDL
151	XL-HDL-CE_%	Cholesterol esters to total lipids ratio in very large HDL
152	XL-HDL-FC_%	Free cholesterol to total lipids ratio in very large HDL
153	XL-HDL-TG_%	Triglycerides to total lipids ratio in very large HDL
154	L-HDL-PL_%	Phospholipids to total lipds ratio in large HDL
155	L-HDL-C_%	Total cholesterol to total lipids ratio in large HDL
156	L-HDL-CE_%	Cholesterol esters to total lipids ratio in large HDL
157	L-HDL-FC_%	Free cholesterol to total lipids ratio in large HDL
158	L-HDL-TG_%	Triglycerides to total lipids ratio in large HDL
159	M-HDL-PL_%	Phospholipids to total lipds ratio in medium HDL
160	M-HDL-C_%	Total cholesterol to total lipids ratio in medium HDL
161	M-HDL-CE_%	Cholesterol esters to total lipids ratio in medium HDL
162	M-HDL-FC_%	Free cholesterol to total lipids ratio in medium HDL
163	M-HDL-TG_%	Triglycerides to total lipids ratio in medium HDL
164	S-HDL-PL_%	Phospholipids to total lipds ratio in small HDL
165	S-HDL-C_%	Total cholesterol to total lipids ratio in small HDL
166	S-HDL-CE_%	Cholesterol esters to total lipids ratio in small HDL
167	S-HDL-FC_%	Free cholesterol to total lipids ratio in small HDL
168	S-HDL-TG_%	Triglycerides to total lipids ratio in small HDL
169	VLDL-D	Mean diameter for VLDL particles
170	LDL-D	Mean diameter for LDL particles
171	HDL-D	Mean diameter for HDL particles
172	Serum-C	Serum total cholesterol
173	VLDL-C	Total cholesterol in VLDL
174	Remnant-C	Remnant cholesterol (non-HDL, non-LDL -cholesterol)
175	LDL-C	Total cholesterol in LDL
176	HDL-C	Total cholesterol in HDL
177	HDL2-C	Total cholesterol in HDL2
178	HDL3-C	Total cholesterol in HDL3
179	EstC	Esterified cholesterol
180	FreeC	Free cholesterol
181	Serum-TG	Serum total triglycerides
182	VLDL-TG	Triglycerides in VLDL
183	LDL-TG	Triglycerides in LDL
184	HDL-TG	Triglycerides in HDL

185	TotPG	Total phosphoglycerides
186	TG/PG	Ratio of triglycerides to phosphoglycerides
187	PC	Phosphatidylcholine and other cholines
188	SM	Sphingomyelins
189	TotCho	Total cholines
190	ApoA1	Apolipoprotein A-I
191	ApoB	Apolipoprotein B
192	ApoB/ApoA1	Ratio of apolipoprotein B to apolipoprotein A-I
193	TotFA	Total fatty acids
194	UnSat	Estimated degree of unsaturation
195	DHA	22:6, docosahexaenoic acid
196	LA	18:2, linoleic acid
197	FAw3	Omega-3 fatty acids
198	FAw6	Omega-6 fatty acids
199	PUFA	Polyunsaturated fatty acids
200	MUFA	Monounsaturated fatty acids; 16:1, 18:1
201	SFA	Saturated fatty acids
202	DHA/FA	Ratio of 22:6 docosahexaenoic acid to total fatty acids
203	LA/FA	Ratio of 18:2 linoleic acid to total fatty acids
204	FAw3/FA	Ratio of omega-3 fatty acids to total fatty acids
205	FAw6/FA	Ratio of omega-6 fatty acids to total fatty acids
206	PUFA/FA	Ratio of polyunsaturated fatty acids to total fatty acids
207	MUFA/FA	Ratio of monounsaturated fatty acids to total fatty acids
208	SFA/FA	Ratio of saturated fatty acids to total fatty acids

``````````````````The End``````````````````
``````````````````