The interplay between the gut metabolome and immune system and their role on the efficacy of neoadjuvant chemotherapy in early breast cancer

Georgia Lambie, BSc Hons Genetics

School of Cancer Sciences

Submitted in fulfilment of the requirements for the

Degree of Master of Science by Research

Word Count: 19,134

School of Cancer Sciences,

College of Medical, Veterinary & Life Sciences

University of Glasgow

Abstract

Promising results have identified the gut microbiome and immune system as important modulators of response to cancer therapies. This provided a rationale for the current study to explore their combined roles on influencing response to neoadjuvant chemotherapy (NACT) in early-breast cancer patients enrolled in the NEO-MICROBE BREAST clinical study. This study aimed to characterise the composition of metabolites within the gut by conducting LCMS analysis on stool samples from cancer patients. A second aim was to start the process of immunophenotyping peripheral blood mononuclear cells through optimisation of a novel 10-colour flow cytometry panel.

Untargeted and targeted metabolomics data analysis of the stool samples profiled the gut metabolic environment of cancer patients and healthy volunteers highlighting several metabolites indicative of NACT response: elucidating alpha-ketoglutarate (aKG) and citrate as potential biomarkers for achievement of pathological complete response, associating guanine and uridine with non-response to NACT and uridine-5'-monophosphate (UMP) potentially linked to cancer status. In combination with metabolomic analysis of blood samples from the same cohort; glutamate demonstrated a correlation between higher stool- and blood- derived levels in TNBC patients, demonstrating potential direct involvement of glutamate in the response to NACT.

A 10-colour panel was generated and optimised via flow cytometry analysis of PBMCs from healthy blood donors and its ability to successfully identify T cell subsets associated with breast cancer was validated. Upon further validation to allow for optimal marker detection, this panel provides the foundation for future PBMC immunophenotyping of this cohort. Overall, this study strengthens the evidence that NACT efficacy may be influenced by the involvement of gut-derived metabolites and systemic modulation; and highlights the importance of the immune-profiling of breast cancer to allow for a comprehensive approach to improve NACT response via personalised treatment.

Table of Contents

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my primary supervisor, Prof. Iain MacPherson, for their support, guidance, and encouragement throughout this research project. Your mentorship has been crucial in forming this thesis and allowing me to develop and grow as a researcher.

I would also like to extend my thanks to Dr Johan Vande Voorde and his entire team at the Wolfson Wohl Cancer Research Centre, for their assistance in the stool metabolomics experiments. This lab provided an environment where I was encouraged to develop my knowledge and hone my practical skills in the field of cancer metabolomics.

A special thank you goes to Prof. Simon Milling and the members of the Milling Lab, especially Dr. Tom Fenton for their help and guidance with flow cytometry training for the optimisation of the flow cytometry panel.

I am also thankful to the entire team at the Human Nutrition Laboratory at Glasgow Royal Infirmary, especially Prof. Konstantinos Gerasimidis for their continued support throughout my research project and for always being willing to answer my many questions.

I would also like to express my gratitude to Dr. Kirsty Ross for allowing me to work on the NEO-Microbe study – without your support to embark on this project, I would not have been able to develop my knowledge and skills in clinical research. Furthermore, I am deeply grateful to the patients and healthy volunteers who generously donated their samples through a very difficult time in their life.

Finally, I wish to thank my family and friends, especially my Mum and Dad (Alison and Steve) for their constant unwavering support, patience, and encouragement throughout not only my postgraduate studies but my full academic journey. Your belief in me from the beginning to the end has sustained me this far. A special mention to my Aunt Lisa who has been undergoing breast cancer treatment for the majority of my project; you have inspired

me everyday to carry on in this field of work and reminded me how worthwhile it can truly be.

List of Abbreviations

ACN - Acetonitrile

aKG – Alpha-ketoglutarate

APC – Allophycocyanin

BMI – Body Mass Index

BC – Breast Cancer

BC T1 – Breast Cancer patients at Timepoint 1

CD – Cluster of Differentiation

CCR7 – C-C Chemokine Receptor Type 7

CTLA-4 – Cytotoxic T-Lymphocyte-Associated Protein 4

DNA – Deoxyribonucleic Acid

DMSO – Dimethyl Sulfoxide

DAMPs – Damage-Associated Molecular Patterns

dTMP – Deoxythymidine Monophosphate

ER – Estrogen Receptor

ER+ – Estrogen Receptor Positive

FACS – Fluorescence-Activated Cell Sorting

FBS – Fetal Bovine Serum

FMO – Fluorescence Minus One

FFPE – Formalin-Fixed Paraffin-Embedded

GLN – Glutamine

GLU – Glutamate

G6P – Glucose-6-Phosphate

HER2 – Human Epidermal Growth Factor Receptor 2

HER2+ – HER2 Positive

HMGB1 – High Mobility Group Box 1

HV – Healthy Volunteers

ICD – Immunogenic Cell Death

ICIs – Immune Checkpoint Inhibitors

IHC – Immunohistochemistry

IL-10 – Interleukin 10

LCMS – Liquid Chromatography Mass Spectrometry

MeOH – Methanol

MAIT – Mucosal-Associated Invariant T cells

NACT – Neoadjuvant Chemotherapy

NEO-MICROBE – NEO-MICROBE BREAST Study

n/a – Not Available / Not Applicable

NLR – Neutrophil-to-Lymphocyte Ratio

pCR – Pathological Complete Response

PBS – Phosphate Buffered Saline

PBMC – Peripheral Blood Mononuclear Cells

PE – Phycoerythrin

PERCP-Cy5.5 – Peridinin-Chlorophyll-Protein-Cyanine 5.5

PD-1 – Programmed Cell Death Protein 1

PD-L1 – Programmed Death-Ligand 1

PR – Progesterone Receptor

PR+ – Progesterone Receptor Positive

RNA – Ribonucleic Acid

RNA-seq – RNA Sequencing

SCFA – Short-Chain Fatty Acid

SI – Separation Index

TCA – Tricarboxylic Acid Cycle

TEMRA – T Effector Memory Cells Re-expressing CD45RA

TILs – Tumour-Infiltrating Lymphocytes

TNBC – Triple Negative Breast Cancer

Tregs – Regulatory T Cells

TYMS – Thymidylate Synthase

UMP – Uridine-5'-monophosphate

WHO – World Health Organization

1. Introduction

An introduction to Breast Cancer

Breast cancer has the highest incidence rate among all other malignancies worldwide, responsible for 12.5% of all new cancer diagnoses recorded globally in 2020, with an estimated 2.3 million new cases (1,2). 15% of the new cases of cancer in the UK were breast cancer, making it the most common cancer in the UK between 2017-2019 (3). Breast cancer is a heterogeneous disease and can be clinically classified into subtypes based on expression of various receptors; specifically, estrogen receptor positive (ER+) and/or progesterone receptor positive (PR+), human epidermal growth factor receptor 2 positive (HER2+) and triple-negative (TNBC). By determining the biological classification of breast cancer this can guide clinicians on treatment decisions and prognostic assessments (4). A plethora of risk factors have been associated with breast cancer over the years such as age, family history and inherited gene mutations, however 70% of women who are diagnosed with breast cancer are not shown to exhibit any of these risk factors (5). This is in accordance with the World Health Organisation who stated in March 2024 "roughly half of all breast cancers occur in women with no specific risk factors" (6).
The advent of immunohistochemistry (IHC) has allowed for more precise diagnosis of specific subtypes of breast cancer therefore benefiting patients, as the most appropriate therapy regimes can be selected, potentially leading to better outcomes (7). IHC is a powerful tool that can characterize breast cancer tissue, using monoclonal and polyclonal antibodies to detect specific antigens within the tissue (8).

1.1.1 ER/PR positive breast cancer

Estrogen receptor (ER) and progesterone receptor (PR) have been identified as important diagnostic and prognostic biomarkers of breast cancer. This is due to reports of high ER expression being characteristic of approximately 70-75% of invasive breast malignancies

(9). Furthermore, PR expression is regulated by ER therefore both have exhibited significantly high levels in breast cancer cells (10,11). ER+ tumours can be treated using targeted hormone therapies, such as tamoxifen and aromatase inhibitors. In a study looking at trends in breast cancer incidence in Scotland based on data of molecular subtypes between 1997 and 2016, it was found that cases of ER+ tumours increased for all ages of women but in particular those 50-69 years old showed the largest increases from 1997 to 2011(12). This elevated trend in ER+ cases in postmenopausal women is believed to be due to a variety of factors including changes in reproductive and lifestyle habits that have been associated with contributing to higher estrogen levels such as childbirth later in life, reduced breastfeeding rates and increased rates of obesity. The study also revealed that ER+ tumours showed a better prognosis than ER- tumours, consequential of earlier diagnosis being typical of ER+ tumours. ER+ breast cancer is also shown to progress at a slower rate than ER- and when treated with appropriate endocrine therapy, has better long-term survival rates. Similar to ER+ breast cancer, PR+ breast cancer exhibits favourable prognostic features such as lower tumour grade and better response to endocrine therapy resulting in improved disease-free survival. A study conducting a meta-analysis of over 13,000 patients linked PR expression in ER+ tumours with improved disease-free survival compared to ER+PR- tumours (13). Further research is ongoing to gain a greater understanding of the molecular characteristics of ER+ breast cancer, contributing to targeted therapies such as CDK4/6 inhibitors and selective estrogen receptor degraders (SERDs) being developed (12).

1.1.2 HER2 positive breast cancer

$HER2^+$ breast cancer is a subset of the breast cancer classified by overexpression of the HER2 protein which subsequently promotes proliferation of cancer cells. HER2+ breast cancer comprises 15-20% of all breast cancer cases and is connected to aggressive disease

progression and poor patient prognosis (12). However, due to the evolution of the breast cancer treatment landscape using targeted biological therapies that inhibit the HER2 protein; the overall survival rate of patients has been significantly improved within the last two decades (14). Trastuzumab is the most commonly prescribed HER2-targeted therapy in combination with chemotherapy. This treatment has proven to significantly improve disease free survival and allow management of a once fatal diagnosis, however, recently resistance to HER2-targeted therapies has a presented a clinical challenge. This resistance is resultant of both HER2-dependent and independent resistance mechanisms such as truncated HER2 receptor forms and alternative receptor tyrosine kinases such as IGF-1R being activated, deeming trastuzumab ineffective and continuation of cell survival and proliferation (15). Various next-generation therapies such as Pertuzumab and Ado-trastuzumab emtansine (T-DM1) have been developed to combat these resistance mechanisms. Combination strategies of HER2 inhibitors as well as inhibitors of downstream signaling pathways (e.g. PI3K) is key focus of current HER2+ cancer research, all in an effort to overcome resistance (12,15). Accurate identification of HER2 status of each patient is crucial in effective use of targeted interventions for optimal patient outcome and improvement of disease recurrence. About 2/3 of HER2+ cancers are ER-positive and 1/3 are ER-negative.

1.1.3 Triple negative breast cancer

Triple negative breast cancer (TNBC) can be defined as breast cancer lacking expression of ER, PR and HER2. It accounts for around 15-20% of breast cancer cases overall (16) and is most commonly seen in women harbouring a BRCA1 mutation. BRCA mutation carriers represent approximately 12% of triple negative breast cancers and 80% of BRCA1 associated breast tumours are triple negative breast cancers. TNBC has been found to be

more aggressive than the other subsets, due to the nature of its high proliferation rate and likelihood to have metastasised at initial diagnosis; therefore, resulting in poorer patient prognosis and higher chance of cancer recurrence (17). Chemotherapy is the standard primary treatment for TNBC, however it common for this subset to show resistance to conventional chemotherapeutic drugs, resulting in disease reoccurrence and further progression (18). There has been recent development in alternative treatments for TNBC due to advances in research identifying potential therapeutic targets; including poly (ADP-ribose) polymerase (PARP), androgen receptor as well as immune checkpoint inhibition. PARP is associated with DNA repair leading to recovery of damaged cells; therefore, it was first hypothesised that targeted inhibition of these proteins would block the repair of DNA damage in cancer cells consequential of chemotherapy, driving cell death and reducing chemotherapeutic resistance (19). PARP inhibitors have also shown efficacy in treatment against breast cancer with a BRCA mutation as well as TNBC cancers (20). Further research is still required to gain a deeper understanding of the heterogeneity of TNBC, with the aim of developing individualised treatment strategies to improve disease free survival for patients.

1.2 Neoadjuvant Systemic Therapy

Neoadjuvant chemotherapy (NACT) is widely used approach to treatment in various solid tumour malignancies such as breast, rectal, upper G.I and ovarian cancer, and is defined as drug treatment given prior to surgery to resect the primary tumour.

1.2.1 Advantages and Disadvantages of Neoadjuvant Chemotherapy

This approach is favourable as it can be beneficial for patients as tumour shrinkage as a result of the NACT may downstage the tumour allowing for a less aggressive surgery (21), making this particularly beneficial for patients with large tumours who could be

reconsidered for breast-conserving procedures after NACT instead of a full mastectomy (22). It is also beneficial for the clinician to assess how the individual responds to certain drugs, which may influence adjuvant treatment plans and allow to predict patient outcomes (23). Clinical studies have revealed that a pathological complete response (pCR) can be achieved in HER2+ and TNBC subtypes. Pathological complete response (pCR) can be defined as "the absence of any residual invasive cancer in the breast and axillary lymph nodes" (24). Achieving pCR is strongly associated with improved disease-free survival rates (25). Combination therapy of NACT with targeted therapies such as trastuzumab in HER2+ breast cancer has demonstrated improved long-term outcomes for patients and significant pCR rate increases (26). On the other hand, the disadvantage of NACT is there is currently no way of predicting patient response to the chemotherapy, which in turn allows time for further tumour proliferation and potential risk of cancerous cells metastasising if poor response is seen. Cases of poor response rate to NACT have been identified in various bladder cancers which have been correlated with advancement of stage, tumour growth, metastases, and reduced survival rate (27).

In recent years, several studies have demonstrated NACT's ability to achieve a tumour pathologic complete response (pCR) more frequently in various cancers, specifically breast cancer patients with specific subtypes of breast cancer (i.e. TNBC and HER2+) (28). Findings such as these allow for progression towards today's NACT regimes becoming more personalized to the individual, to achieve the best possible outcome (29)

1.2.2 Neoadjuvant Chemotherapy (NACT) in Breast Cancer

Neoadjuvant chemotherapy (NACT) is recognized as a primary treatment for breast cancer patients with operable tumours and has been shown to improve patients' treatment options for large, or inoperable tumours, as the therapy reduces tumour size (30). Quantifying pCR within patients allows for rapid monitoring of response to chemotherapy; therefore, this

model provides the basis for other trials evaluating the efficacy of novel therapies using pCR as a surrogate marker for long term outcomes (disease free survival and overall survival) (28,31). Meta-analysis carried out over various trials has shown that in patients with triple- negative and $HER2^+$ breast cancer receiving trastuzumab, there was a significant association between pCR and long-term outcomes, thus these subtypes have the highest sensitivity to NACT (31). More recently, immune cells, specifically peripheral blood monocytes, and the cytokine IL-10 have been shown to be potentially useful predictive markers for pCR in breast cancer patients, in the NACT setting (32).

1.2.3 Tumour Immune Environment and NACT

Tumour infiltrating lymphocytes (TILs) are pro-inflammatory lymphocytes that have invaded the tumour tissue residing in both the tumour and the surrounding microenvironment. Typically, they are made up of CD8+ and CD4+ T cells, as well as B-cells, natural killer, and dendritic cells (33). Extensive clinical studies have found that a higher density of TILs is associated with increased pathological complete response (pCR). These studies have highlighted TILs predictive value as a potential biomarker for response to NACT (34).

1.2.4 Pathological Complete Response Rate to NACT in Breast Cancer

Various clinical trials have demonstrated the clinical benefits of NACT, including tumour size reduction resulting in increased cases of breast-conserving surgery (BCS) instead of a mastectomy, which can have longer recovery time and pose more risk of incurring problems (35). Despite these advantages, complete or partial response is only achieved in 30% of patients after NACT (36). For example, a study with a cohort of 13,939 women looked at evaluating pCR rates in women with various subsets of breast cancer after undergoing NACT. Results showed overall only 19% of the patients achieved pCR, with

patients with HER2+ showing the highest pCR rate (38.7%), and TNBC subtype was shown to be nearly 100 times more likely to achieve pCR than luminal A (ER+ and/or PR+, HER2-) (37).

The factors determining why a patient does or does not achieve pCR are not yet fully understood therefore; this highlights the need to investigate other potential factors involved such as the immune system and gut microbiome and their association with pCR to NACT. By characterizing NACT responders/non-responders, it may allow for treatment plans to be tailored to the individual more effectively, thereby allowing for increased efficacy of neoadjuvant chemotherapy and patient outcomes.

1.3 The Immune System and Cancer

The immune system's involvement in cancer was first recognized by William B. Coley. He is widely known as the father of immunotherapy for his research which involved stimulating strong immune responses through injecting patients diagnosed with inoperable sarcomas with a cocktail of various bacteria referred to as "Coley's toxins;" subsequently inducing sepsis within patients and initiating immune response (38). In the last decade, the idea that the immune system plays a vital role in almost all types of cancers has shown to be true with its promising potential to be utilized in a therapeutic manner. This development has transformed the basis of primary cancer treatments, distinguishing immunotherapies as a key therapeutic strategy for cancer patients (39). One key concept recognised is immune surveillance, which is pivotal in recognizing and destroying cancer cells before tumour formation occurs. This important interaction between the immune system and cancer cells, is also seen between tumour cells that have evaded this so-called immune surveillance through a process called cell-mediated immune escape (40), subsequently allowing for tumour formation. Cell's ability to escape this process and

therefore resulting in failure of immunosurveillance is now a recognised hallmark of cancer (41).

1.4 Immune System and treatment response

For years, research into the immune systems role in the efficacy of treatments for cancer patients has been neglected. More recently, this field of the immune systems contribution to therapy response has been more widely acknowledged, allowing for discoveries into the immunological aspects of current treatments as well as the application of novel immunotherapies in combination with standard treatments.

1.4.1 Immune Checkpoint Modulation

In more recent years, a deeper understanding of T-cell immune checkpoints and the effects of targeting inhibitory immune receptors such as cytotoxic T-lymphocyte antigen-4 (CTLA-4), programmed cell death-1 (PD-1) and programmed cell death ligand-1 (PD-L1), has further revolutionized the field of cancer-immunotherapy (42). Upregulation of immune checkpoints serve as negative regulators within the immune system, whereby they inhibit the cytotoxic response of immune T cells to prevent destruction of healthy cells within the body.

Cancer cells utilize this immunosuppressive network so that it becomes advantageous, thus escaping detection and subsequent destruction. Tumour cells corrupt the immune system by creating an immunosuppressive microenvironment through upregulation of cell surface immune checkpoint receptors (CTLA-4, PD-L1). Upregulation of these receptors has shown to cause a lack of T cell function, leading to an impaired immune response resulting in an immunosuppressive state being established (43). Targeted inhibition of these immune checkpoints using antibodies to block binding with their partner proteins (figure 1.1), has proven to have a therapeutic effect for various cancers, including breast cancer. Immunotherapy drugs known as immune checkpoint inhibitors (ICIs) have been approved

by the FDA for clinical use. Various checkpoints have been researched as targets, however those with the most promising clinical outcomes are CTLA-4, PD-1 and PD-L1 (44).

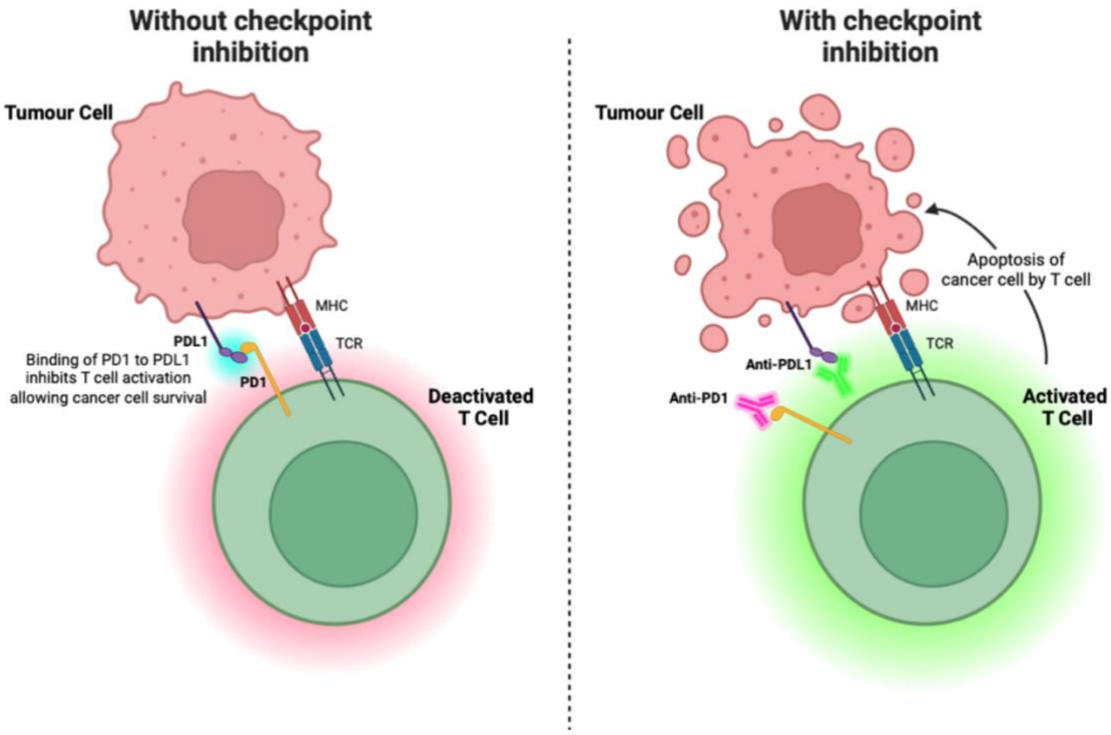


Figure 1.1 Immune Checkpoint Modulation of T cells. Upon binding of PD1 to PDL1 T cells activity is inhibited and therefore the cancer cells escape death and continue to proliferate. Using checkpoint inhibitors such as anti-PD1 and anti-PDL1 to block their subsequent binding, allowing the immune cells cytotoxic response to be activated and cell death of the tumour cell.

1.4.2 Cancer Chemotherapy

Chemotherapy has shown to not only have a direct cytotoxic effect on cancer cells but also has a significant involvement in both immunosuppressive and immunostimulatory pathways. For example, the chemotherapeutic agent anthracycline induces an adaptive immune response that is mediated by dendritic cell maturation and cytotoxic T cell activation; resultant of immunogenic death of tumour cells whereby danger-associated molecular pattern (DAMPs), such as calreticulin and HMGB1 are released in response to the chemotherapies (45). On the other hand, high-dose glucocorticoids have demonstrated

immunosuppressive effects by impairing cell function of dendritic cells and cytokine production.

There have been recent developments that receptiveness to immune-mediated attack is improved by the modulation of tumour microenvironment in response to chemotherapy. A study showed that efficacy of immune checkpoint inhibitors can be increased as antigen-presenting cells are recruited and activated upon chemotherapy-induced immunogenic cell death (ICD) increasing anti-tumour T cell activity, creating an optimal tumour microenvironment for ICIs (46). Overall, this provides substantial evidence that chemotherapy plays a dual role in killing cancer cells and modifying the tumour microenvironment, therefore therapeutic efficacy could be improved with the synergy of chemotherapy alongside immunotherapy as a more standardised method of care.

1.4.3 HER2-targeted treatment

HER2+ tumours are often associated with better response to therapy due to them being more immunogenic than other breast cancer subtypes as a result of them having more tumour infiltrating lymphocytes (TILs). The targeted anti-HER2 antibody trastuzumab has a multipurpose role in inhibiting HER2 as well as upregulating immune effector cell activity (i.e. natural killer cells) thus signalling the death of tumour cells. Research has suggested that the immune microenvironment is critical in determining therapeutic response; for example, anti-HER2 therapy resistance has been linked to immunosuppressive mechanisms within the tumour microenvironment (47).

1.5 Immunophenotyping of Peripheral Blood Mononuclear Cells (PBMCs)

Research on various subtypes of cancer shows differences in sensitivity to chemotherapy, therefore suggesting that neoadjuvant chemotherapy doesn't have the same effect in all breast cancer patients. Further personalization of chemotherapy, using IHC to immunophenotype each patient could provide a more patient-specific treatment and better clinical outcomes. Peripheral blood mononuclear cells (PBMCs) have been proven to be an effective biomarker material for identifying and characterizing immune cells in cancer patients. Flow cytometry analysis is a very common technique for immunophenotyping PBMCs. Flow cytometry can detect levels of immune cell populations of interest through the binding of fluorophore-labelled antibodies to their antigen-targets on the surface of the immune cell in PBMC sub-populations (48).

In a previous study observing PBMCs expressing the CD45 receptor, it was suggested that TNBC cells exploit CD45 function to suppress immune response. It was shown that CD45 may be a potential therapeutic target as treatment with C24D reversed this inhibitory immune effect, by binding to CD45 and inducing its subsequent signal transduction pathway, killing TNBC cells (49).

In a study analysing CD56 as a predictive biomarker for immunotherapy sensitivity in breast cancer, it was highlighted that CD56 is involved in regulating NK-mediated cytotoxicity. Interestingly reduction of CD56 expression in breast cancer tissues is thought to be a mechanism to escape NK-immune response (50).

With various studies suggesting that the immune system and varying PBMC subsets play a significant role in breast cancer, this highlights the need to determine if changes in leukocyte subsets are correlated to NACT sensitivity and pCR rates in breast cancer patients.

1.6 Metabolomics and Cancer

Metabolomics can be defined as the study of the metabolites within a biological system. Advancements in analytical technologies has allowed for a deeper understanding of the various effects of metabolites, at both local and systemic level, which has sparked an interest within the field of cancer research. This interest stems from the desire to provide insight into how certain metabolites may be associated with disease status, progression, and therapy response (51). Cancer cells rely on dysregulation of cellular metabolism to survive and proliferate rapidly. Metabolomics allows for the identification of biomarkers of cancer as well as disease related metabolomic pathways which can present novel therapeutic vulnerabilities (52). The reprogramming of energy metabolisms a recognised hallmark of cancer which involves cancer cells utilising glycolysis for energy, even when oxygen is not restricted (53). Metabolomics can also identify potential biomarkers of cancer that could be utilised as diagnostic tools as well as providing insight on the efficacy of treatments (54). This provides an example of how tumour cells' adaptive behaviour to exploit metabolic pathways, presents a need for these mechanisms to be understood.

In an effort to gain insight into tumour biology, disease progression and response to treatment the analytical approach of metabolomics can be an extremely useful tool, and various studies have employed either untargeted or targeted data analysis to reveal metabolic alterations in cancer. Untargeted data analysis aims to measure all detectable metabolites within a given sample (55). Conversely, targeted metabolomics uses a bank of pre-defined metabolites and measures specific metabolites within a sample (56).

1.7 The Gut Microbiome and Cancer

1.7.1 Gut Dysbiosis in Cancer

The human microbiota is composed of a vast consortium of bacteria and other microorganisms that all play a role in overall human health. The gastrointestinal (GI) tract is home to the largest number of bacteria compared to the rest of the human organs, with 70% of the human body's microbes residing in the colon alone (57). With the growing interest in the gut microbiota in humans and how it influences systemic homeostasis, it has been conceptualized that gut microbiota may have an important role in cancer development and studies to improve knowledge of this theory may have therapeutic benefits (5). Bacterial metabolites (such as short chain fatty acids (SCFA), lactate, pyruvate, secondary bile acids and metabolites of amino acids) have been analysed to establish microbial biomarkers linked to increased risk of cancer (58). Further assessment of the fecal metabolome via techniques such as liquid chromatography mass spectrometry (LCMS), can produce a functional readout of the gut microbial metabolism (59), allowing identification of specific metabolites involved in microbial-cancer interactions.

This subject is still not fully understood, however more evidence is emerging to support the hypothesis that changes in the gut microbiome are strongly linked to disease development and progression of breast cancer. Studies in murine models have demonstrated a link between microbial metabolites and breast cancer (57). Miko et al reported the bacterial metabolite, lithocholic acid, has anti-proliferative effects on breast cancer cells through modulating TGR5 receptor activity via the human circulatory system (60).

1.7.2 Microbial Metabolites in Cancer Treatment

Recently the focal point of research has shifted to the interplay between the gut microbiota and the efficacy of cancer therapy (61). Researchers have found that bacterial communities residing in the GI tract are related to the activity of the hosts immune system. Along with

the advances in immunotherapy for breast cancer, this has created the shift to investigate the regulation of innate or adaptive immunity via the gut microbiota and how this may influence therapy response. Studies focusing mainly on checkpoint inhibitors anti-PD-1 and anti CTLA-4, have identified specific bacterial species that may have potential in predicting patient response to respective immunotherapies in melanoma patients (62–65). These findings have provoked studies observing the modulation of cancer immunotherapies through targeted interventions in the gut microbiota. Changes to dietary fiber intake and probiotic supplementation were assessed in melanoma patients to investigate the impact of these dietary modifications to immunotherapy response. Clinical evidence demonstrated that significantly improved progression-free survival was observed in 128 patients reporting a high fiber diet (66). Fecal microbiota transplantation (FMT) has also become a popular method to alter a person's gut microbiome, e.g. with the objective to improve immunotherapy efficacy (67,68). Potential in vivo mechanisms that have been described have been; changes in T-cell activity pathways, increased memory CD8+ T cells and CD4+ T cells in the tumour microenvironment and increased activity of mucosal-associated invariant T (MAIT) cells (66,68).

Beyond immunotherapy, studies have demonstrated that the gut metabolome is strongly associated with response to chemotherapy. Recent studies have indicated that gut microbes play a role in drug metabolism which therefore influences the rate and extent of systemic availability of chemotherapeutic drugs as well as potential drug toxicity. For example, a recent study by Alexander et al., 2021 recognised that the integral probiotic members of the gut microbiota, *Lactobacillus* and *Bifidobacterium* have the ability to reduce the gastrointestinal toxicity of fluoropyrimidine chemotherapy which subsequently improves the anti-tumour efficacy of this drug (69). Furthermore, another study indicated the composition of the gut microbiome influences response to chemotherapy in pancreatic cancer via modulation of immune cell infiltration in the tumour microenvironment (70).

1.8 The Gut Microbiome and Breast Cancer

Despite the intriguing results observed on the influence of the gut microbiome on other diseases and cancers, there is limited data on the functional relationship between the gut microbiome and breast cancer. It has been established that there is significant variability of bacterial species comprising the gut microbiome among different subtypes of cancers, as well as clinical characteristics such as breast cancer stage and grade (71). Dysbiosis may also be responsible for tumorigenesis in the breast, as changes to the bacterial composition of the gut can lead to alterations of immune responses that promote tumour progression (72).

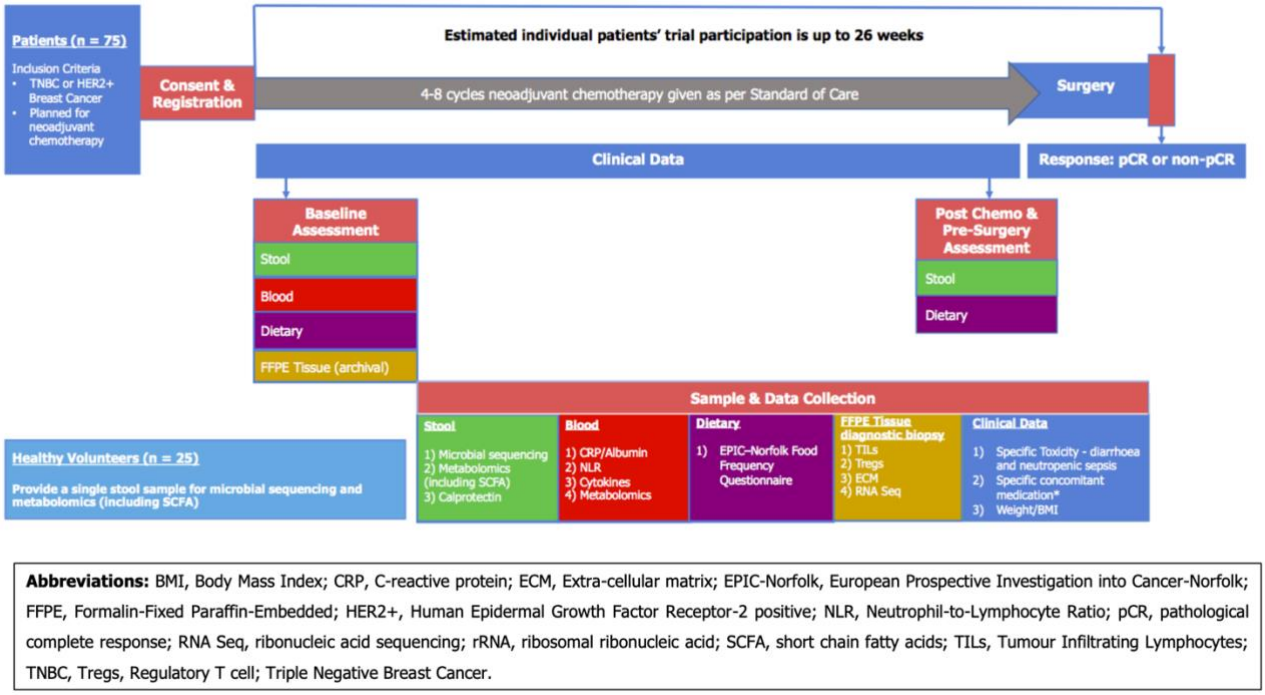
1.9 The Gut Microbiome as a predictive marker for NACT response in Breast Cancer

A previous pilot study recruited 21 patients diagnosed with early breast cancer and due to commence neoadjuvant chemotherapy for early breast cancer at the Beatson West of Scotland Cancer as well as 21 healthy volunteers from August 2017 to June 2019. Baseline gut microbial richness was significantly lower in patients without subsequent pCR compared to those patients with pCR or healthy volunteers. Results revealed that pCR outcome was associated with the diversity and composition of the gut microbiome. It was highlighted that patients with a higher abundance of beneficial gut microbial taxa such as Akkermansia and Ruminococcaceae, were more likely to achieve pCR. Increased abundance of SCFA producing genera were present in patients without pCR and mean concentrations of butyrate and propionate were higher in patients without subsequent pCR vs. patients with pCR (73). It is thought this could be due to these microbes influencing systemic immune activation by creating a favourable immunometabolic environment which subsequently improves the patient's ability to respond to therapy. In contrast, patients who exhibited an enrichment of pro-inflammatory bacteria were often non-

responders to NACT, thus not achieving pCR(74). These findings support the idea that gut microbiota strongly influences NACT efficacy, which drove the inception of the NEO-MICROBE BREAST clinical study to gain a deeper insight into this topic.

1.10 NEO-MICROBE BREAST Study

The NEO-MICROBE BREAST study (ISRCTN13877559) was then initiated to prospectively test the hypothesis that the composition and function of the gut microbiome contributes to achieving pCR and to explore putative mechanisms including modulation of the immune microenvironment and alterations in circulating metabolites and cytokines. Patients were recruited prior to commencing neoadjuvant chemotherapy and stool and blood samples obtained as shown in figure 1.2. In addition, a cohort of healthy volunteers was recruited.



Abbreviations: BMI, Body Mass Index; CRP, C-reactive protein; ECM, Extra-cellular matrix; EPIC-Norfolk, European Prospective Investigation into Cancer-Norfolk; FFPE, Formalin-Fixed Paraffin-Embedded; HER2+, Human Epidermal Growth Factor Receptor-2 positive; NLR, Neutrophil-to-Lymphocyte Ratio; pCR, pathological complete response; RNA Seq, ribonucleic acid sequencing; rRNA, ribosomal ribonucleic acid; SCFA, short chain fatty acids; TILs, Tumour Infiltrating Lymphocytes; TNBC, Tregs, Regulatory T cell; Triple Negative Breast Cancer.

Figure 1.2. Overview of study design and sample collection timeline for the NEO-MICROBE BREAST cancer trial. The study involves 75 breast cancer patients (with TNBC or HER2$^+$ subtypes) undergoing 4–8 cycles of neoadjuvant chemotherapy followed by surgery, with individual participation lasting up to 26 weeks. Clinical data and biospecimens are collected at baseline and post-chemotherapy (pre-surgery).

The co-primary objectives of the NEO-MICROBE BREAST study were (1) to test the association of gut microbial taxonomic richness with efficacy of neoadjuvant chemotherapy in early breast cancer, (2) to investigate the association of Short Chain Fatty Acid (SCFA) levels with efficacy of neoadjuvant chemotherapy in early breast cancer, and (3) to investigate the association of bacterial species abundance with the efficacy of neoadjuvant chemotherapy in early breast cancer. At the time of writing, analysis is ongoing, and these results have not yet been reported.

The samples collected within the NEO-MICROBE BREAST clinical study provide a rich resource for further study. The main aims of the current work were to:

1) To optimise a protocol for performing metabolomic analysis of stool samples.
2) To investigate differences in stool metabolomics in patients with cancer versus healthy volunteers and in patients with pathological complete response (pCR) versus no pathological complete response (non-pCR).
3) To validate a panel for flow cytometry analysis of NEO-MICROBE BREAST PBMCs with the aim of identifying metabolites correlated with specific peripheral blood immune phenotypes.

2. Materials and Methods

2.1 Sample Collection

This work utilised samples and data previously collected within the NEO-MICROBE BREAST study. Patients eligible for NEO-MICROBE BREAST were female, age 16 or older and planned to commence neoadjuvant chemotherapy for HER2-positive or triple negative breast cancer. Potential patients eligible for participation in the NEO-MICROBE BREAST study were identified by their own clinical care team within the NHS. Once notified of the study the patients were provided with an information sheet and consent form. They were given at least 24 hours to make a decision before committing to take part in the study and were notified of their ability to leave the study at any time. Exclusion criteria included unequivocal evidence of metastatic disease or any other malignancy in the preceding 24 months, which was treated with systemic therapy. In addition, patients were not eligible of they had a history of gastrointestinal (GI) disorders including inflammatory bowel disease, irritable bowel syndrome (if severe or on regular medication) or persistent infectious gastroenteritis, colitis, or gastritis, or persistent or chronic diarrhoea of unknown aetiology or clostridium difficile infection. Major gastrointestinal surgery except for appendicectomy and cholecystectomy was also an exclusion. Treatment with systemic corticosteroids (intravenous or oral) or other immunosuppressive therapy within 28 days prior to Cycle 1 of neoadjuvant chemotherapy was also not allowed. Twenty-five healthy volunteers with an approximate age and body mass index (BMI) matching the patient cohort were recruited. This recruitment process was carried out using methods such as word of mouth and REC-approved advertising material placed on posters (placed in NHS Greater Glasgow and Clyde locations and Glasgow City Council libraries) and social media (e.g Beatson Cancer Charity website). Healthy volunteers were female, aged 16 or over, and were taking no medications requiring regular medical consultations. Eligibility

also required no history of breast cancer at any time or of any other malignancy requiring systemic therapy in the previous 24 months. Exclusion criteria related to GI conditions and use of immunosuppression were identical to the patient population.

The NEO-MICROBE BREAST Study, including the exploratory analyses reported here for breast cancer patients and healthy volunteers, was approved by West of Scotland Research Ethics Committee 4 (ref: 21/WS/0078)

2.1.1 Blood Collection

Upon patients attending their pre-assessment oncology clinic as part of the NEO-MICROBE study they provided a baseline blood sample (total volume 50ml) prior to chemotherapy, for later analysis including metabolomics and immune marker analysis. Bloods were taken from the antecubital fossa and if possible, via the venflon already sited to reduce further venepuncture to the patients. Various tubes of blood were collected (2 x 6ml EDTA, 2 x 6ml LiHep) by the research team at oncology outpatient's clinics at four sites within the NHS GGC, NHS Lanarkshire, and NHS Forth Valley, and labelled accordingly to the patients assigned study ID number. Samples were then transported by driver to the lab at the Glasgow Royal Infirmary (GRI) for processing.

2.1.2 Stool Collection

Stool samples were collected at two timepoints; Timepoint 1- Baseline before starting NACT and Timepoint 2 – after completion of NACT and before surgery. Samples were collected at home in a pre-weighed plastic tub and stored in a cool-bag or mini-freezer provided. Samples were then transported to the Human Nutrition Lab at GRI for

processing. The patients provided a baseline sample (T1) no earlier than 28 days pre-chemotherapy, but ideally this was collected roughly 14 days prior. Patients provided a second stool sample after completion of chemotherapy and prior to surgery (T2), following the same protocol as the collection of T1. Stool samples for healthy volunteers were collected as above at a single timepoint.

2.2 Stool sample preparation

Stool samples were processed at the human nutrition lab at the GRI. The frozen sample was left out at room temperature to thaw and the tub containing the sample was weighed to calculate the sample weight. The sample was homogenised using an electric blender before transferring specific weights of sample into various labelled tubes and stored in -80°C or -20°C for later analysis. The sample was then freeze-dried before weighing out and transferring the specific weight the necessary for later LCMS analysis (25mg).

2.3 Liquid Chromatography-Mass Spectrometry

Stool samples were prepared using a methanol-based extraction protocol adapted and optimised based on the methodology decribed by Clarke et al. (2022) (75), whereby 25mg of freeze-dried stool sample from NEO-MICROBE patients was weighed out into 2ml microtubes prefilled with beads and extracted using a 100% cold MeOH solvent. The cells were then disrupted using a Precellys Homogenizer (Bertin Instruments) to ensure complete homogenisation. After centrifugation, the supernatant was recovered and ACN:H20 solution was added to obtain a final solvent ratio of MeOH:ACN:H20 of 50:30:20 before transferring to glass LCMS vials for analysis. Metabolomic analysis was completed using a Thermo-Ultimate 3000 high-performance liquid-chromatography (HPLC) system, equipped with a ZIC-pHILIC column (SeQuant; 150 mm by 2.1 mm, 5 µm; Merck KGaA), with a ZIC-pHILIC guard column (SeQuant; 20 mm by 2.1 mm) for

metabolite seperation.Quality control for background interference and potential sources of contamination solvent blanks were prepared and incorporated into the LCMS workflow at the beginning of every run. As this was the first time stool metabolomics had been performed at the Beatson Institute for Cancer Research instrument blanks were run intermittently between samples, serving as a control to assess for carryover and ensuring robustness of results. Furthermore, the reliability of metabolites detected was enhanced by running Plasmax medium mimicing the metabolic composition of human plasma alongside the sample cohort. Each NEO-Microbe stool sample was analysed in a single technical replicate.

2.4 Data analysis

Raw LCMS data was processed using Thermo Tracefinder 4.1 for targeted data analysis, and data were visualised using Metabolite Autoplotter v2.6, which generated PCA plots and bar plots of individual metabolites of interest; allowing for visualisation of patterns in metabolite profiles and identification of potential biomarkers of breast cancer. If required, further statistical analysis was performed using Microsoft Excel or Graphpad Prism, which was also used to generate volcano plots. Thermo Compound Discoverer was used for untargeted data analysis.

2.6 PBMC Extraction

The whole blood sample was diluted with sterile phosphate buffered saline (PBS) and layered over 4ml histopaque. Peripheral blood mononuclear cells (PBMC) were isolated by histopaque density centrifugation, creating 4 distinct layers. The buffy coat layer contained the PBMCs was carefully collected and washed in PBS. After spinning, the pellet was the resuspended in PBS and the total cell yield was calculated. The PBMCs were stained

within the same day, or they were cryopreserved in a medium of 20% fetal bovine serum (FBS) and 10% dimethyl sulfoxide (DMSO).

2.7 PBMC Flow Cytometry Staining

2.7.1 Viability Staining

Approximately 500,000 cells per sample were incubated with a live/dead master mix (according to the lab protocol at Milling Lab, Sir Graeme Davies Building; PBS, eFluor780 fixable viability dye (ThermoFisher) in PBS at 1000:1 dilution) and blocked with 5µl/ml Human TruStrain FcX (Biolegend) and incubated on ice for 15 mins in the dark. The viability dye eFluor780 was chosen due to its compatibility with the fixation/permeabilisation protocols and minimal spectral overlap with other fluorophores used in the panel. Following incubation cells were washed with PBS and resuspended in 500µl PBS.

2.7.2 Extracellular Staining

For surface marker staining, 500,000 cells were stained with the correlating master mix of fluorochrome-conjugated monoclonal antibodies in 500µl PBS and incubated at room temperature for 30mins in the dark.

2.7.3 Intracellular Staining

The cells were first fixed and permeabilized using the Foxp3/Transcription Factor Staining Buffer kit (eBioscience, ThermoFisher Scientific) and incubated for 45mins in the dark at room temperature. The cells were then resuspended in 500µl 1x permeabelisation buffer and centrifuged; 5mins, 400 x g, at room temperature. The pellet was the resuspended in 100µl 1 x permeabelisation buffer and 1µl of each fluorochrome-conjugated monoclonal antibody was added and incubated for 30mins at room temperature. The cells were washed twice with 1 x permeabelisation buffer and resuspended in 200µl FACS buffer for analysis.

2.8 Flow Cytometry analysis

PBMCs were prepared for flow cytometry analysis to assess surface and intracellular marker expression. The trypan blue exclusion test was used to assess cell viability within the population. Data acquisition was carried out on a BD LSRFortessa Cell Analyser, with 50,000 events being recorded per sample. Compensation was carried out using single stained UltraComp eBeads Plus Compensation Beads (Invitrogen). Gating strategies were implemented using fluorescence minus one (FMO) controls.

2.9 Quantitative Analysis of Flow Cytometry Data – Flowjo

The data obtained from the flow cytometry analysis of PBMCs was analyses using FlowJo software (version 10.10.0) whereby FCS files were imported. The compensation data from the single-stained comp beads was applied for each fluorochrome and gating strategies were used to exclude debris and doublets. Further gating strategies were applied to determine cell populations of interest allowing these populations to be quantified.

2.10 Antibodies Used

	Marker	Fluorophore	Company	Clone	Laser	Channel
1	Viability	eFluor780	ThermoFisher Scientific		640nm	780/60 755LP
2	CD45	BB515	BD Biosciences	HI30	488nm	530/30
3	CD45	Spark Blue 574	BioLegend	2D1	488nm	530/30 505LP
4	CD45RA	BV510	BioLegend	HI100	405nm	510/50
5	CD45RA	BUV395	BD Biosciences	5H9	355nm	379/28
6	CD3	BV785	BioLegend	OKT3	405nm	780/60
7	CD3	BV510	BioLegend	SK7	405nm	525/50 502LP
8	CD4	PE-Cy7	BioLegend	OKT4	561nm	780/60
9	CD4	BUV805	BD Biosciences	SK3	355nm	820/60 770LP
10	CD8	BV421	BioLegend	RPA-T8	405nm	450/50
11	CD8	Pacific Blue	BioLegend	SK1	405nm	450/50
12	CD25	APC	BioLegend	BC96	640nm	670/14
13	CCR7	PerCP-Cy5.5	BioLegend	G043H7	488nm	695/40
14	CCR7	PE	BioLegend	G043H7	561nm	585/15 570LP
15	FoxP3	PE	BioLegend	206D	488NM	585/42
16	FoxP3	BB700	BD Biosciences	236A/E7	488nm	695/40 655LP
17	CTLA-4	AF700	NovusBiologicals	BNI3	640nm	730/45 685LP
18	PD1	BUV395	BD Biosciences	MIH4	355nm	379/28
19	HLA-DR	BV605	BioLegend	L243	405nm	605/20 595LP

Table 1. All antibodies used for flow cytometry antibody optimisation to characterise immune cell populations. Each row details the target marker, conjugated fluorophore, supplier, antibody clone, excitation laser, and corresponding detection filter(s). Multiple conjugates were included for key markers (e.g., CD45, CD3, CD4, CD8, FoxP3, CCR7) to facilitate comparison and panel optimisation.

3. Results : Optimisation of novel stool metabolomics protocol

Recently the comprehensive approach of using the gut microbiota as a diagnostic and predictive tool to response of therapies, specifically in NACT in breast cancer, has proven to show promising results. Similarly, the fecal metabolome which is dependent on the composition of the gut microbiome may also offer predicting information as well as possible mechanistic insights. Studies have shown specific metabolite signatures to be valuable biomarkers for stratifying responders and non-responders to NACT (76).

3.1 Optimisation of stool metabolomics protocol

To obtain an efficient protocol for extraction of metabolites from stool samples for metabolomic analysis, we first followed the protocol used in the study by Kelly PE et al. (77) and carried out optimisation of the protocol as stool metabolomics had not previously been carried out in our facility. Optimisation was completed through work to determine the appropriate conditions for optimal extraction.

3.1.1 Evaluating the drying stage in stool metabolomics protocol

Freeze-dried faecal samples from 3 healthy individuals were pulverized and 50mg was weighed out before adding 100% MeOH at a 50mg/ml. The samples were vortexed, and the supernatant was recovered and centrifuged for 10mins at high speed. After spinning, the supernatant was recovered and aliquoted for treatment in five separate conditions to determine the most suitable condition for metabolomic analysis. All conditions ended up with a final proportion of solvents in coherence with the CRUK Scotland Institute's extraction solution previously established for pHILIC chromatography; MeOH:ACN:H20 50:30:20 (Table 2).

Sample conditions A-D were placed under N2 to evaporate the MeOH to dryness using a MULTIVAP nitrogen evaporator block heater before the dry pellet was reconstituted in a MeOH: ACN:H20 solution at various volumes to explore the effects of concentrating the sample. As condition E did not contain a drying stage, a solution containing only ACN:H20 (60:40) was added to the solute. All samples were vortexed and transferred to glass vials for LCMS analysis. As samples B-D began to precipitate prior to LCMS analysis, these samples were excluded from the experiment.

As conditions B-D were excluded this allowed a direct comparison between A and E, comparing the effects of drying down(A) the sample prior to LCMS compared to not (E). After inspecting the full scan extracted ion chromatogram obtained from this experiment it can be seen from figure 3.2 that there was a large peak in the first few minutes suggesting a high output at the start of the chromatography which may suggest non-polar substances that have not been optimally removed during extraction.

Untargeted analysis of samples A and E also showed that there was no benefit to drying down samples before LCMS analysis. This can be visualised on a PCA plot (figure 3.3) showing little separation between the conditions for all samples analysed.

When targeted analysis was carried out on specific metabolites of interest, it was again found that for the majority of metabolites there was little apparent variation between the two conditions (figure 3.4a), with only 6 metabolites among the 61 analysed appearing discordant according to drying conditions(figure 3.4b). From this it was concluded that the drying down stage is not necessary during extraction for this specific protocol.

Condition	Solvent Used	Drying Condition	Solvent Volume added to end up with Beatson Solution
A	MeOH:ACN:H20	under N2	400ul
B	MeOH:ACN:H20	under N2	200ul
C	MeOH:ACN:H20	under N2	100ul
D	MeOH:ACN:H20	under N2	50ul
E	ACN:H20	N/A	200ul

Table 2. The various conditions used in the metabolite extraction. Table 2 shows the five conditions selected to determine the best for optimal extraction of metabolites from freeze-drier stool samples. Conditions A-D were dried down, whereas condition E was not dried down.

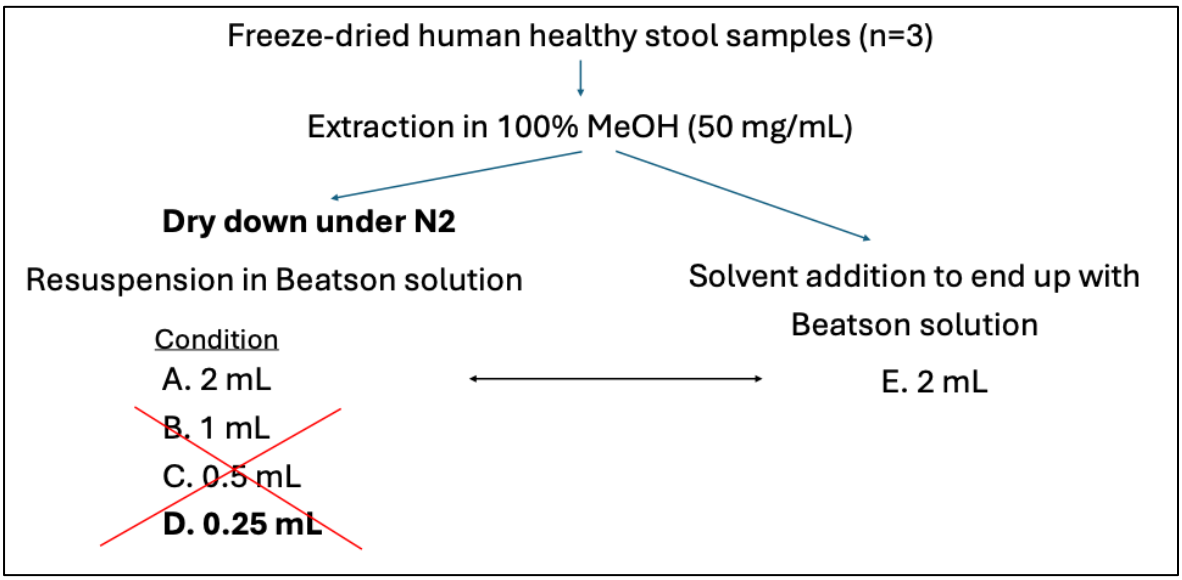


Figure 3.1 Schematic of experimental protocol of metabolomic analysis of stool samples evaluating optimal extraction methods and comparing drying methods. Figure 3.1 is a flowchart of the protocols followed for each sample extraction method to determine the optimal sample weight for metabolite extraction as well as to compare if a drying down stage using nitrogen was necessary. Conditions B-D have been crossed out due to them being excluded from the experiment as they began to precipitate prior to LCMS analysis.

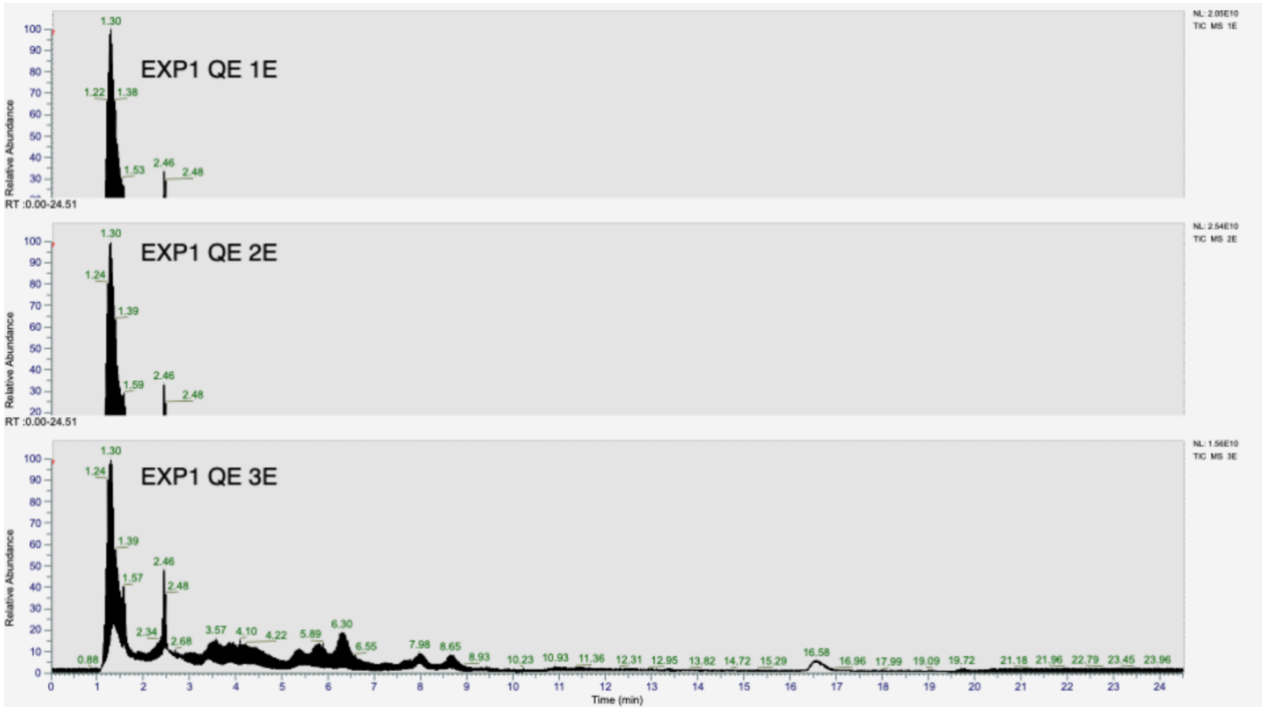


Figure 3.2 Chromatogram of untargeted metabolomic analysis of stool samples from three different individuals that have been dried down under N2. . Visual output of the signal detected from the stool samples obtained from 3 seperate healthy individuals (1E/2E/3E); all of which had not undergone the drying down stage during processing prior to LCMS analysis. Each peak represents different components and metabolites within the sample. It can be suggested the large peak emitted between 1-2 minutes is due to unwanted substances that have failed to be removed during extraction.

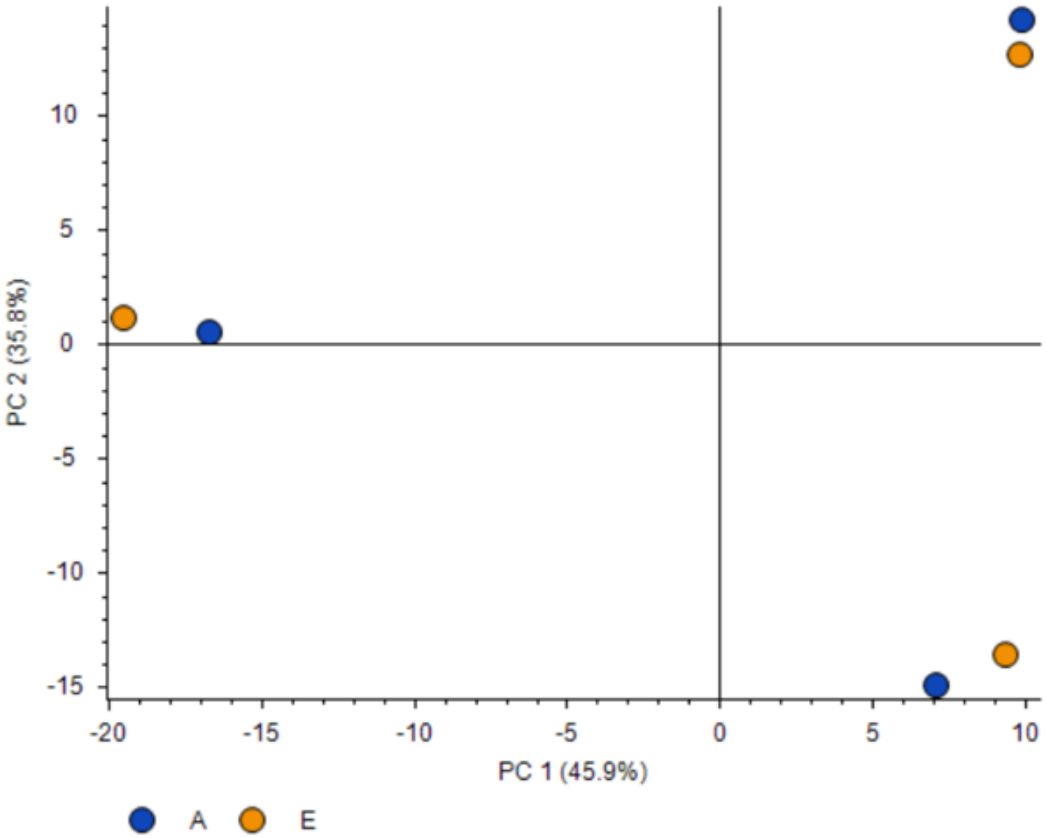


Figure 3.3 PCA plot of targeted metabolomic analysis comparing stool sample drying methods during metabolite extraction protocol. . The PCA plot represents the variation between the two sample conditions for each sample group. It can be seen in figure 3.2 that for each sample group the dots are close by therefore it can be suggested there is little variation introduced between the two extraction conditions

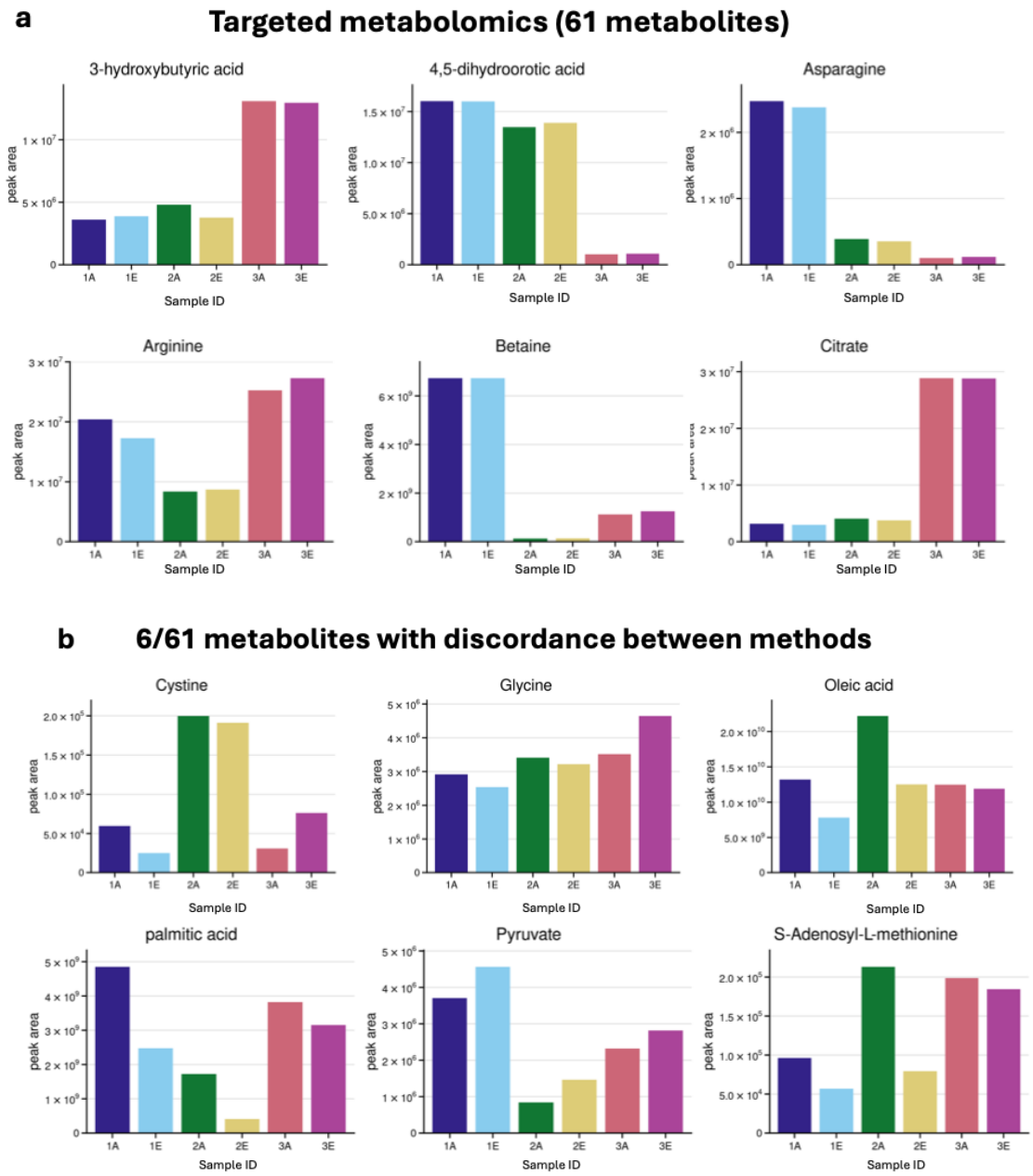


Figure 3.4 Bar plots representing peak area of metabolites from targeted metabolomic analysis comparing drying down methods in triplicate. The barplot describes the levels of metabolites detected in stool samples which have been dried down (A) and not dried down (E) from three different individuals. **a)** Selected representative metabolites with little variation in the peak area detectedbetween the two extraction conditions.. **b)** metabolites (n=6) with discordance between extraction conditions.

3.1.2 Determining optimal conditions for metabolite extraction

A metabolite extraction was carried out by adding 100% MeOH to a freeze-dried faecal sample from one healthy individual that had been weighed out in 2ml tubes containing ceramic beads at three weights in triplicates. Upon extraction in MeOH various desired concentrations were obtained (50mg/ml, 25mg/ml, and 10mg/ml). An ACN:H20 solution was then added to these samples to obtain three final dilutions per sample (½, ¼, 1/8), at a ratio to obtain a final concentration in coherence with that of the CRUK Scotland extraction solution. The samples were then centrifuged for 10mins at high speed and 100ul of supernatant was recovered from each sample to undergo a hexane cleanup attempting to maximally remove the apolar fraction. The rest of the supernatant was recovered and transferred to glass vials for LCMS analysis as such. To each extract undergoing the hexane prep, 300ul of hexane was added, samples were vortexed and spun at 6°C, 3mins, high speed. A glass pasteur pipette was used to recover the bottom layer of the extract and transferred to glass vials for LCMS analysis (figure 3.4).

The results of the untargeted analysis on the LCMS experiment 2 showed that there was more variation between samples at higher concentrations (i.e. 25mg/ml (50mg, 1_2, no_hex) and 12.5mg/ml (25mg, 1_2, no_hex))) compared to lower concentrations such as 6.25mg/ml (50mg, 1_8, no_hex) and 3.125mg/ml (25mg, 1_8, no_hex). This trend of lower concentrations correlating with reduced variability between sample replicates can be visualized in figure 3.5. The samples outlined in red were selected for further analysis (targeted) as they were clusters of samples of the same final concentration however had various initial concentrations prior to solvent addition (i.e. final concentration of 6.25mg/ml = 50mg 1/8, 25mg ¼, 10mg ½). The other samples were chosen to have a lower concentration comparison (3.125mg/ml) as well as being able to see the effect of the

hexane prep. All of the samples chosen for targeted analysis were also chosen for their reliability of results between replicates due to their reduced variability.

Targeted analysis was carried out on the samples selected from the untargeted metabolomic results. Figure 3.6a shows the targeted metabolomic results which demonstrate the levels of individual metabolites detected by LCMS for each dilution selected. At the higher concentration (6.25mg/ml) there is more variability in results between the various dilutions compared to the lower concentration of 3.125mg/ml. Samples that received the hexane cleanup showed positive results in figure 3.6c whereby the levels of various fatty acids (non-polar) were reduced, and the signal intensity for lactate (polar) was higher. However, as exemplified in Figure 3.6a hexane prep did not improve detection for the majority of the polar metabolites (apart from lactate). In addition, the large signal of obtained early during chromatography was not markedly reduced post hexane prep (Figure 3.7). Hence this extra step which can introduce errors during sampling handling was not carried forward for future analyses. Figure 3.6b indicates that typical hydrophilic metabolites of interest remain well detected at the lower concentration (>1 x 10^6 AUC).

In conclusion, based on the previous experiments we determined that in order to achieve optimal results the most appropriate condition for metabolite extraction was 25mg/ml,1/8 (final conc = 3.125mg/ml) without the drying stage before extraction and the hexane prep. This condition provided a stool weight that allowed for easy sample handling while still demonstrating robust results. It was determined the drying stage before metabolite extraction was not required as results were consistent with samples that did not undergo drying. Furthermore, the hexane prep did not remove the high output seen at the start of the column during LCMS analysis, therefore it was deemed an unnecessary step (Figure 3.7).

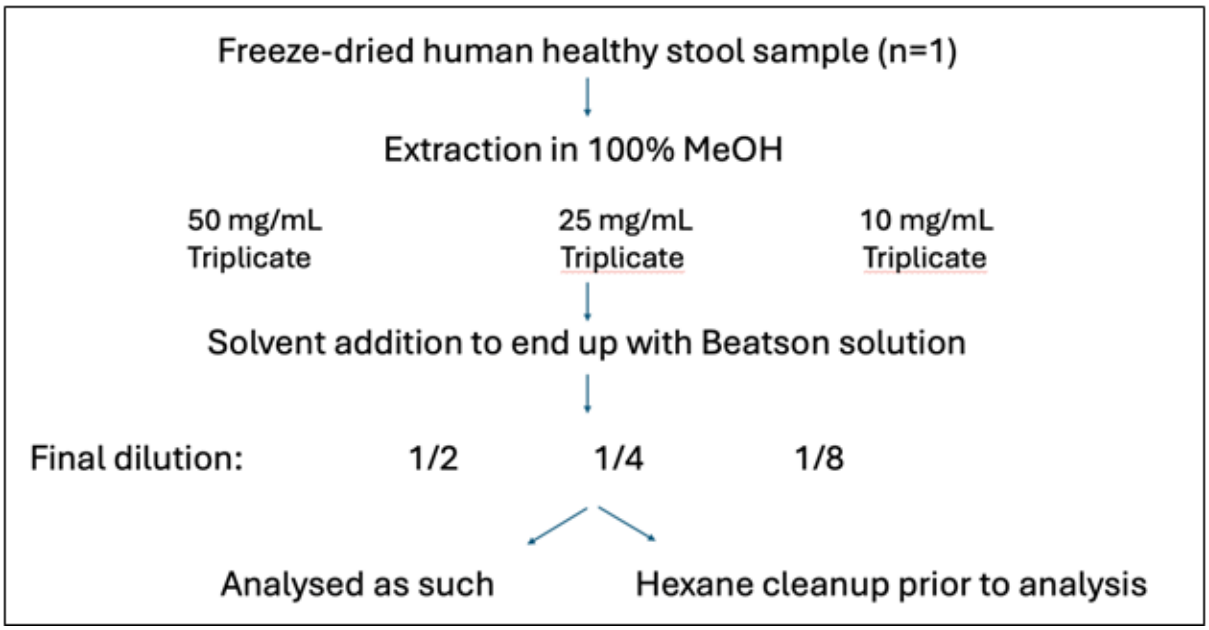


Figure 3.5 Schematic oftargeted metabolomic analysis protocol to evaluate optimal metabolite extraction condition and the benefit of the hexane clean up step prior to LCMS analysis. . Figure 3.4 is a flowchart of the protocols followed for each sample extraction method. Three different weights of stool sample from the same individual was prepared in triplicate to obtain various final concentrations, to assess the optimal metabolite detection output as well as, the easiest method for sample handling.

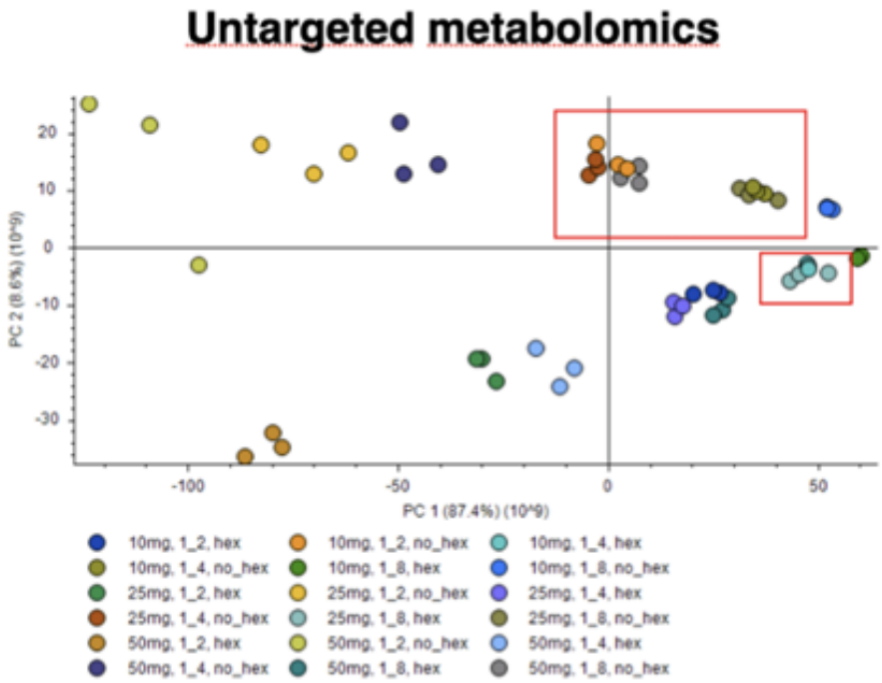


Figure 3.6 PCA plot from untargeted metabolomic analysis of each stool sample condition comparing various concentrations and hexane clean up methods for optimal metabolite extraction. The various colours of dots represent each sample dilution and whether they have received the hexane clean-up. There is a trend at lower concentrations the sample replicates become closer together, ultimately reducing variability.

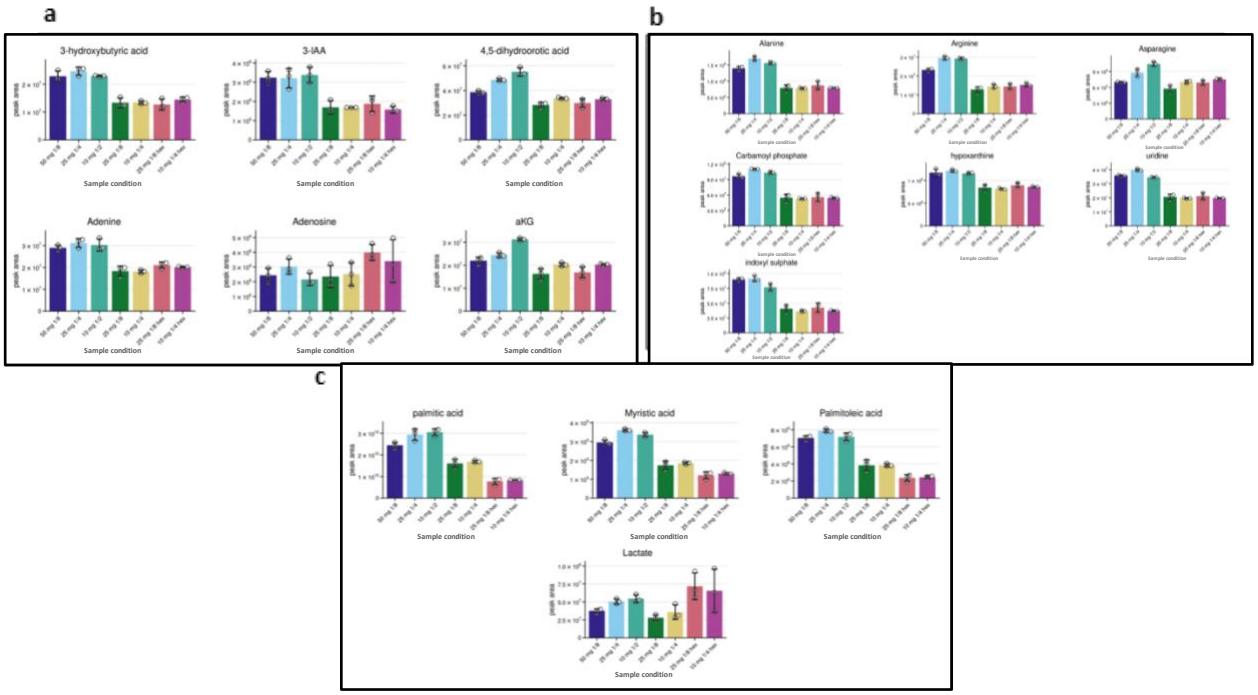


Figure 3.7Bar plots representing peak area of metabolites from targered metabolomic analysis comparing two separate sample concentrations and hexane clean up methods. A) The bar charts describe the variability of each metabolite for the two concentrations 6.25mg/ml and 3.125mg/ml with hexane prep, and 3.125mg/ml without hexane. B) Indicates that at lower concentrations there is still a good quality detection of metabolites of interest. C) Shows the variability between no hexane and hexane prep in the 3.125mg/ml samples.

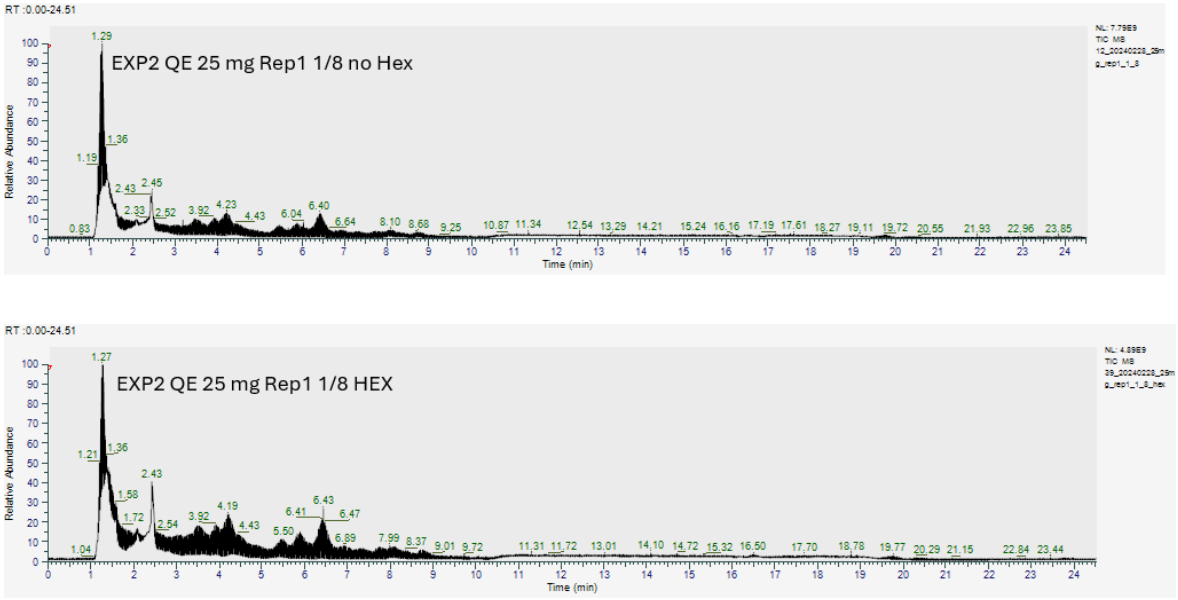


Figure 3.8 Chromatogram of untargeted metabolomic analysis assessing the effect of hexane clean up on metabolite detection using LCMS. Visual output of the signal detected from the stool samples with and without hexane prep obtained from LCMS analysis. Each peak represents different components and metabolites within the sample. It can be suggested the large peak emitted between 1-2 minutes is due to unwanted substances that have failed to be removed during extraction. The figure suggests that the hexane prep step is not necessary as the large peak is still visualised.

4. Results : LCMS Analysis of NEO-MICROBE Stool Samples

Targeted and untargeted analysis of stool metabolomics was carried out to elucidate any potential metabolites that may be involved in modulating patient response to NACT as well as their role as biomarkers to predict therapeutic outcomes.

From November 2021 to August 2023, 65 patients consented to enter NEO-MICROBE BREAST study, one patient subsequently withdrew consent. Patient characteristics are shown in Table 2; patients (n=64) had a median BMI of 27kg/m2 and were largely of white ethnicity (92.2%). Half of the cohort were pre-menopausal with most having no co-morbidities and TNBC was the most common subtype seen within the group, as well as nearly all having grade 3 invasive ductal carcinoma. The median baseline tumour size was 34.5mm. Fifty-three of the 64 patients had negative baseline clinical nodal status with stage II being the most common clinical TNM stage within the population (68.75%). Twenty-five healthy volunteers with a median age of 51.5 years old (range of 26y-68y) and a median BMI of 23.8 kg/m2 (range of 19.3kg/m2-28.2kg/m2) were also recruited.

4.1 NEO-MICROBE Patient Clinical Data

	All Patients (n=64)	**Pathological Complete Response (n=33)**	**Non-Pathological Complete Response (n=31)**
Age – median (range)	51 (29-74)	54 (32-73)	50 (29-74)
BMI (kg/m2) – median (range)	27 (18.1- 48.9)	26.7 (19.1-38.7)	28.2 (18.1-48.9)
• Underweight <18.5	1 (1.6%)	0	1 (3.2%)
• Healthy 18.5-25	23 (35.9%)	13 (39.4%)	10 (32.3%)
• Overweight 25-30	17 (26.6%)	9 (27.2%)	8 (25.8%)
• Obese >30	21 (32.8%)	11 (33.3%)	10 (32.3%)
• Morbidly Obese >40	2 (3.1%)	0	2 (6.5%)
Ethnicity			
• White	59 (92.2%)	30 (90.9%)	29 (93.5%)
• Black African	2 (3.1%)	1 (3%)	1 (3.2%)
• Asian Pakistani	1 (1.6%)	1 (3%)	0
• Asian Chinese	1 (1.6%)	1 (3%)	0
• Arab	1 (1.6%)	0	1 (3.2%)
Menopausal status			
• Pre- or peri-menopausal	32 (50%)	18 (54.5%)	14 (45.2%)
• Post-menopausal	32 (50%)	15 (45.5%)	17 (54.8%)
Co-morbidities			
• gastrointestinal	5 (7.8%)	2 (6.1%)	3 (9.7%)
• diabetes	4 (6.25%)	1 (3%)	3 (9.7%)
• autoimmune	3 (4.68%)	1 (3%)	2 (6.5%)
Breast cancer subtype			
• HER2+/ER+	24 (37.5%)	9 (27.3%)	15 (48.3%)
• HER2+/ER-	14 (21.9%)	10 (30.3%)	4 (12.9%)
• TNBC	26 (40.6%)	14 (42.4%)	12 (38.7%)
Histological Subtype			
• Invasive Ductal carcinoma	62 (96.9%)	31 (93.9%)	31 (100%)
• Invasive Lobular carcinoma	2 (3.1%)	2 (6.1%)	0
Histological grade			
• Grade 1	1 (1.6%)	0	1 (3.2%)
• Grade 2	23 (35.9%)	13 (39.3%)	10 (32.3%)
• Grade 3	40 (62.5%)	20 (60.6%)	20 (64.5%)
Baseline tumour size, (mm) • Median (range)	34.5 (10-112)	36 (10-112)	31 (15-110)
Baseline nodal status			
• Negative	53 (82.8%)	26 (78.8%)	27 (87.1%)
• Positive	11 (17.2%)	7 (21.2%)	7 (12.9%)
Clinical TNM Stage			
• I	5 (7.8%)	0	5 (16.1%)

• II • III	44 (68.75%) 15 (23.4%)	24 (72.7%) 9 (27.3%)	20 (64.5%) 6 (19.4%)

Table 2: Baseline Demographics of All Evaluable Patients and Segregated by Pathological Complete Response Outcome

4.2 Untargeted Analysis

To obtain a comprehensive view of the gut microbiome and its impact on response we first started with untargeted analysis of the stool metabolomics.

4.2.1 Healthy Volunteers versus cancer patients

Principal component analysis (PCA) was conducted on untargeted metabolomic profiles to gain an understanding on the variation in gut metabolites between breast cancer patients (BC) and healthy volunteers (HV). The plot shows the distribution between the two group along the first two principal components with PC1 accounting for 10.9% total variance and 7.8% accounted for by PC2. The PCA plot (Figure 4.1) exhibits significant overlap in the metabolomic profiles of the two groups, revealing no clear separation between the data points. This finding suggests that individual variation exceeds any variation between HV and BC patients.

Further statistical analysis (T-test) was carried out to generate the volcano plot (Figure 4.2) to present the differences observed in metabolic composition between HV and BC patients. Metabolites that are modulated exhibited a fold change (FC) >1 (red/green dots). Any metabolites with FC > 1 and statistically significant are coloured in the red square (increased levels) and those that are downregulated (<-1) are highlighted by the green square. A total of 97 metabolites (55 upregulated, 42 downregulated) were significantly altered based on cancer status. These specific metabolites may have predictive value in cancer status. Overall, these findings identify considerable differences in metabolomic profiles between healthy individuals and those with breast cancer, signifying that necessity for further analysis to identify these altered metabolites.

4.2.2 PCR vs non-PCR– whole population

Next, we conducted untargeted metabolomic profile of cancer patients at timepoint one who achieved pathological complete response (pCR) and those who did not (non-pCR). Principal component analysis was completed to understand the metabolic variation between the two groups and to explore if such variation has an influence on chemosensitivity. The PCA plot (Figure 4.3) demonstrates the separation along the first two principal components (PC1, 11.9% total variance and PC2, 7.6% total variance). The plot exhibits no distinct separation between the two groups which suggests that, in general, metabolic profiles of responders and non-responders are similar. The data points are widely dispersed across the plot which indicates a high degree of metabolic heterogeneity.

Subsequent analysis was carried out to generate the volcano plot (Figure 4.4) in order to understand the abundance of metabolites that showed a FC > 1 (red) or lower (FC<-1) (green) levels in non-pCR patients compared to pCR patients. A total of three metabolites showed statistically significantly higher levels in non-responders, as well as meeting the cutoff of FC > 1 (red gate) and 6 metabolites demonstrated statistically significantly lower abundance and a FC < -1 (green gate) in non-responders compared with responders. Overall, this analysis identifies potential metabolic biomarkers of NACT response ($p < 0.05$) as well as other metabolites shown by red and green dots that may be contributing to the variance driving response to treatment.

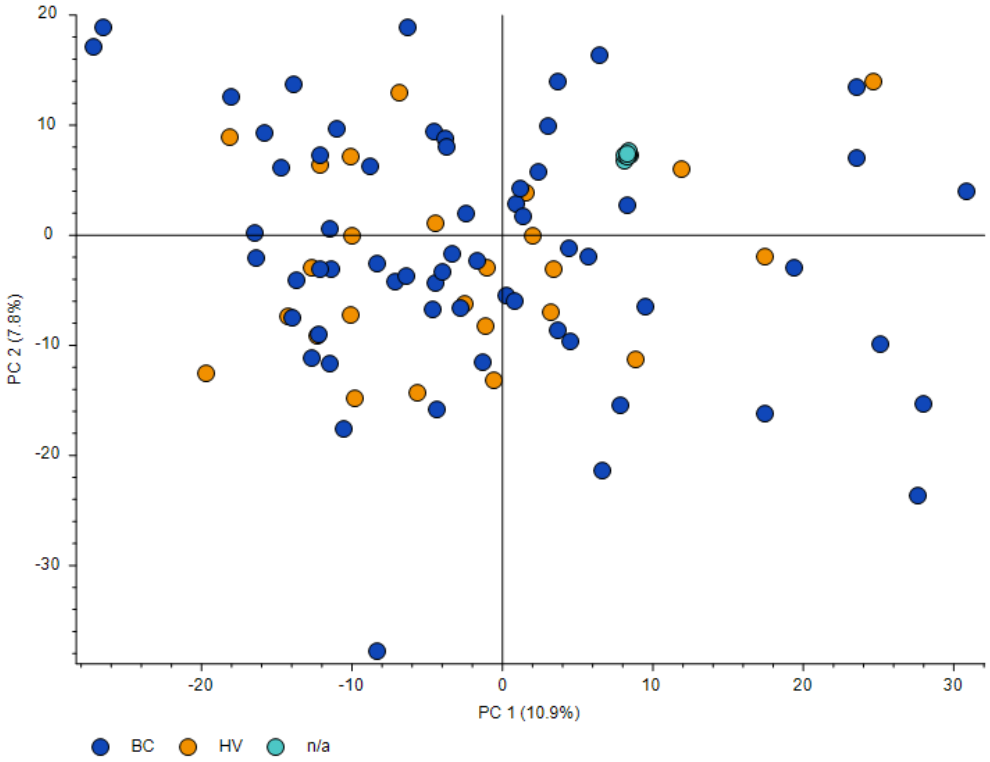


Figure 4.1 Principal component analysis (PCA) of untargeted stool metabolomic profiles from healthy volunteers (HV)(n=25) and breast cancer patients at timepoint 1 (BC T1) (n=64). Each point represents an individual sample. The plot demonstrates clear overlap between the sample groups, indicating that variance is not primarily driven by disease status.

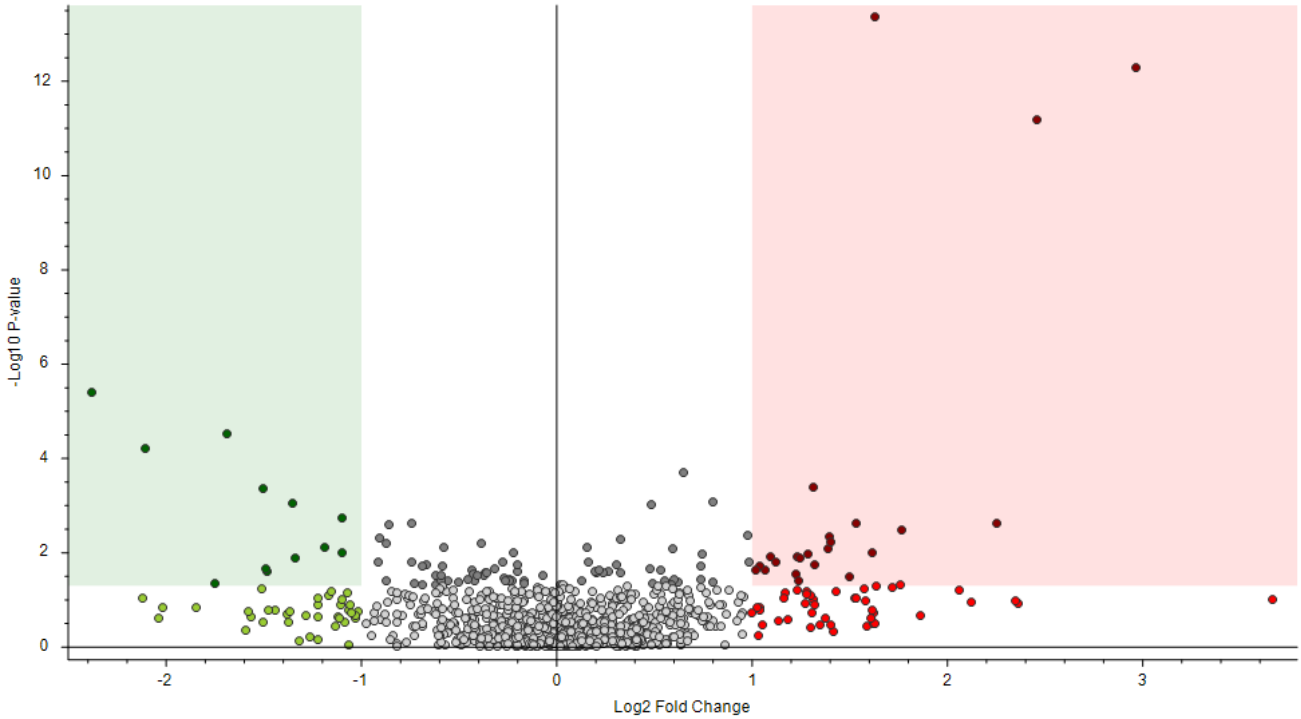


Figure 4.2. Volcano plot showing the relationship between untargeted metabolomic profiles p-values (T-test) and fold changes among patients with breast cancer at timepoint one (BC T1) and healthy individuals (HV). The effect size of the two groups is plotted on the x-axis on a logarithmic scale to base 2 ($Log_2(FC)$). The -log10 p-values are plotted on the y-axis. Red gates demonstrate the metabolites that demonstrated statistically significantly higher level in BC T1 patients and a FC > 1 and statistically significantly lower levels in BC T1 patients and a FC < -1 are described by the green gate. Each dot describes a specific metabolite.

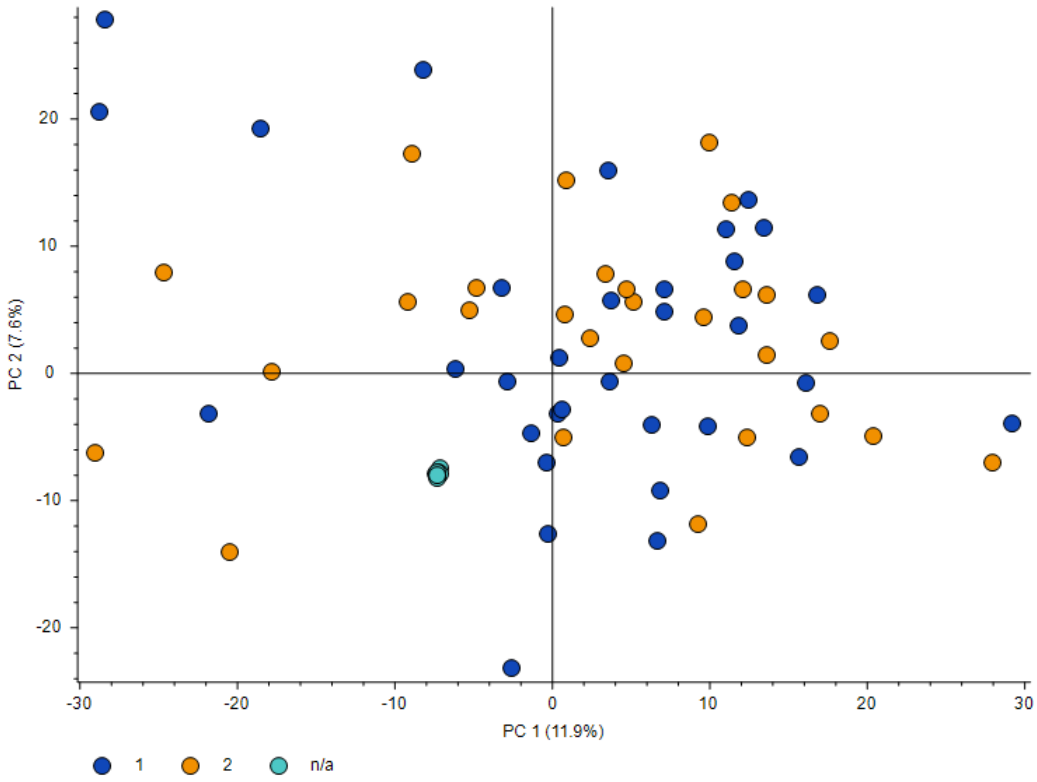


Figure 4.3. Principal component analysis (PCA) of untargeted stool metabolomic profiles from responders (1) (pCR) and non-responders to NACT (2) (non-PCR). Each point represents an individual sample. The plot demonstrates clear overlap between the sample groups, indicating that variance is not primarily driven by response to NACT.

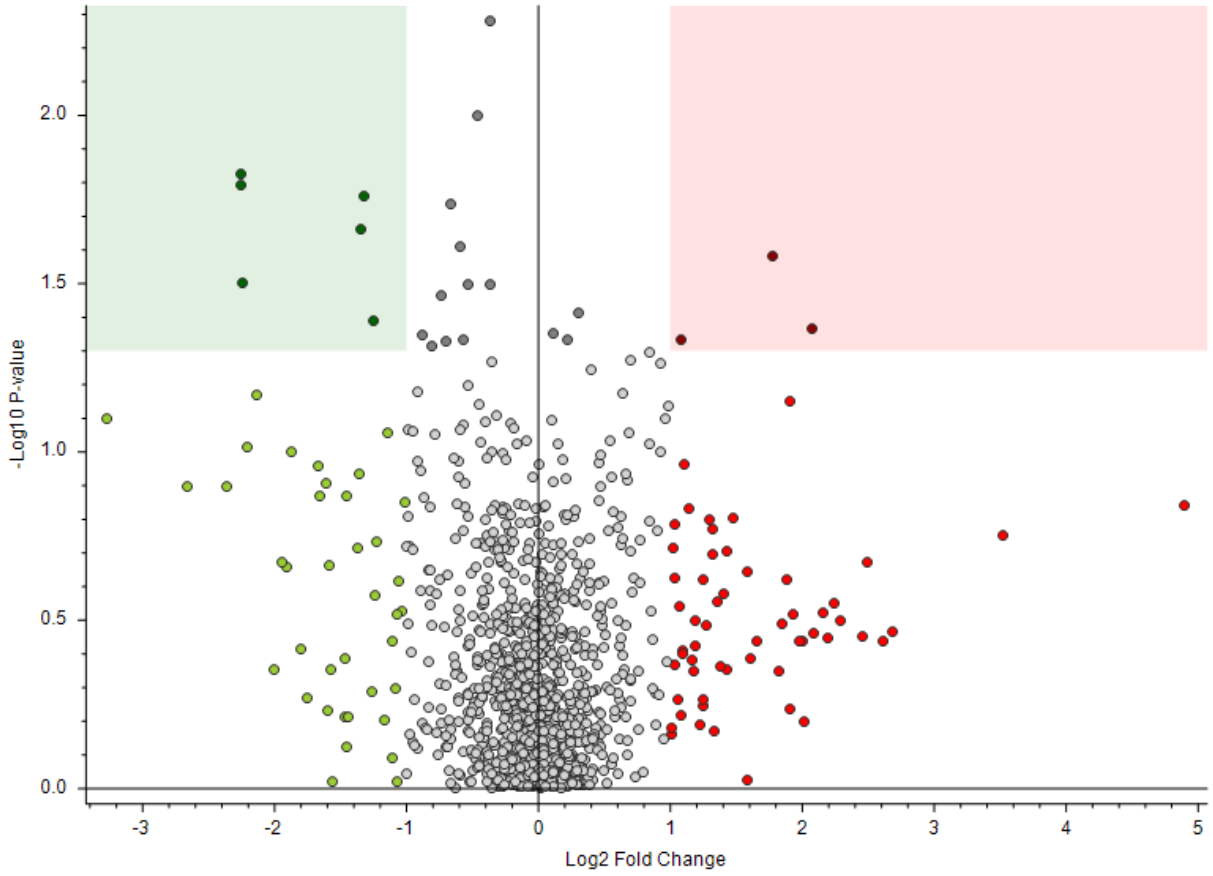


Figure 4.4. Volcano plot showing the relationship between untargeted metabolomic profiles p-values and fold changes among patients who did not achieve pathological complete response (non-PCR) versus those who did achieve pCR (PCR). The effect size of the two groups is plotted on the x-axis on a logarithmic scale to base 2 ($Log_2(FC)$). The -log10 p-values (T-test) are plotted on the y-axis. Red gates demonstrate the metabolites that demonstrated statistically significantly higher level in non-responders also with a FC > 1 and statistically significantly lower levels in non-pCR patients and with a FC < -1 are described by the green gate. Each dot describes a specific metabolite.

4.3 Targeted Analysis

Targeted analysis of the stool metabolomics data was performed to understand how a predefined panel of cancer-relevant metabolites influenced pCR in patients from the NEO-MICROBE study. Defining metabolites within the gut that may be linked to the efficacy of NACT provides an insight into the mechanistic role of these metabolites impact on drug metabolism and potential immune modulation via systemic circulation. Global analyses were first performed on the entire study populations before stratification by subtype to determine subtype-specific metabolic trends.

4.3.1 Healthy volunteer's vs patients with cancer

Targeted analyses were performed on selected metabolites of interest (n=76) known to be linked to cancer to investigate potential differences between HV and BC T1 at the compound level. Principal component analysis (PCA) was performed to understand the variations between healthy volunteers (HV) and cancer patients stool metabolomic profiles at T1 (pre-chemotherapy) (BC T1). The PCA plot (Figure 4.5) shows no clear separation between the two sample groups, thus highlighting that there is no strong variance between healthy volunteers and cancer patient's pre-treatment. Considerable overlap can be visualised along the first two principal components (PC1 and PC2), which accounted for 33.3% explained variance overall. This suggests that there are similarities in metabolic profiles between individuals with cancer and healthy individuals, thus variation is not disease driven. As PCA only considers dominant variance, additional targeted metabolite specific analysis was conducted.

Overall, both cohorts showed similar distribution in metabolic levels with no statistically significant difference in peak area between HV and BC 1 for most metabolic compounds analysed. Figure 4.6 highlights three metabolites that are representative of most metabolites analysed; 3-indoleacetic acid (3-IAA), demonstrated no statistically significant difference in levels between the two sample groups (p=0.74). The same trend was evident

for Aspartate and Adenine, with no statistical significance observed; p=0.39 and p=0.55, respectively. Some inter-individual variability was evident across all metabolites however overall; relatively consistent levels were seen across individuals. This is demonstrated with the tight clustering of data points in 3-IAA (Figure 4.2), excluding a few elevated outliers which is consistent across most metabolites. These findings suggest that for most metabolites analysed they were not predictive of cancer status, as the peak area levels were not differentially abundant at baseline (BC T1) versus HV.

Metabolites showing statistically significant differences between the two cohorts are shown in Figure 4.7.

Linoleic acid displays significant difference with a p-value of 0.024, with a peak area for HV being significantly higher than BC T1. Marked separation of the individual points distribution between the two groups can be visualised, suggesting that cancer status is influencing different metabolic profiles here.

Oleic acid exhibits a similar pattern whereby a significant difference (p=0.02) in peak areas between HV and BC T1 is highlighted, with healthy individuals showing a higher level of the compound. Similar distribution of data points is also shown, with distinct clustering of points being consistently lower in BC T1 versus HV.

UMP displays substantial separation, with UMP peak area being considerably higher in BC T1 patients than healthy individuals; this is demonstrated through a highly significant p-value of 0.00094. Distribution of individual points suggests that HV had mostly extremely low or absent peak areas, with only one outlier presenting a peak area of $1x10^6$. This finding suggests that UMP could potentially be acting as a driver in the complex mechanisms involved in cancer development and progression.

Similar to UMP, the metabolic compound Cystine showed a noticeably higher peak area in BC T1 patients than HV. A significant difference in cystine levels was demonstrated with a p-value of 0.024. Cystines higher abundance in cancer patients is emphasized by the

skewed distribution of individual data points that can be visualised; suggesting that cystine in general is higher in BC T1 versus HV however outlying data points are pulling distribution towards higher peak areas, thus contributing to the significant difference between the two groups.

As UMP showed highly significant elevated levels in BC T1 vs HV it was hypothesised UMP may be associated to breast cancer status in a subset of individuals, therefore further targeted analyses were performed to gain a better understanding of clinical factors influence on UMP levels (Figure 4.8).

Figure 4.8a presents the correlation between UMP and menopausal status in breast cancer patients at T1. No significant difference ($p=0.5009$) was displayed between patients that are pre/peri-menopausal and menopausal cancer patients, suggesting that menopausal status does not modulate UMP levels in breast cancer patients.

The correlation between UMP levels and age of clinical patients of the NEO-MICROBE study at T1 is illustrated by XY plot (figure 4.8b). A weak positive correlation is demonstrated by the correlation coefficient ($r = 0.3374$) and p-value ($p = 0.0730$); however, this was not statistically significant.

Similar to the other clinical factors, no association between BMI of patients at T1 and UMP levels is shown. Figure 4.8c displays boxplots representing the comparison of UMP levels between patients with low UMP and elevated UMP, however no statistically significant difference was indicated ($p = 0.5033$).

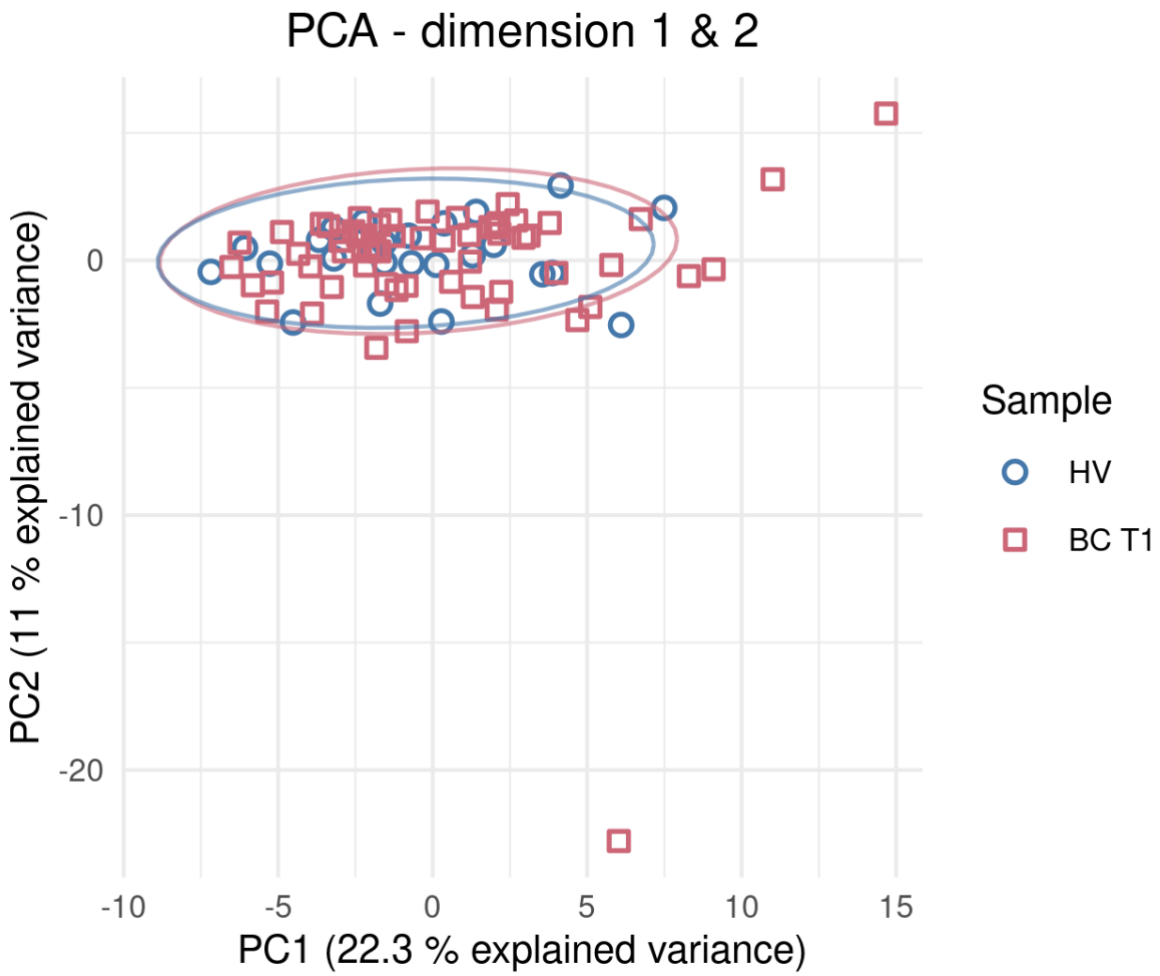


Figure 4.5. Principal component analysis (PCA) of stool metabolomic data from healthy volunteers (HV)(n=25) and breast cancer patients at timepoint 1 (BC T1) (n=64). Each point represents an individual sample. The plot demonstrates clear clustering between the sample groups, indicating that variance is not primarily driven by disease status.

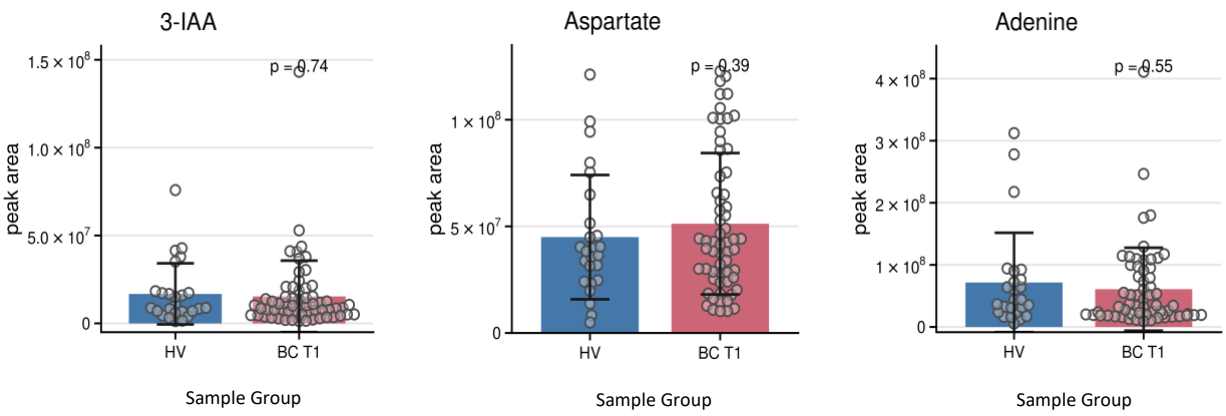


Figure 4.6. Boxplots demonstrating the peak areas of metabolites (3-IAA, Aspartate and Adenine) representative of the overall metabolites of interest that were analysed across healthy individuals (HV)(blue) and breast cancer patients from the NEO-MICROBE study at timepoint 1 (BC T1) (red). The graph also displays statistical comparisons (p-value) (T-test) highlighted the statistically significant differences between the two groups.

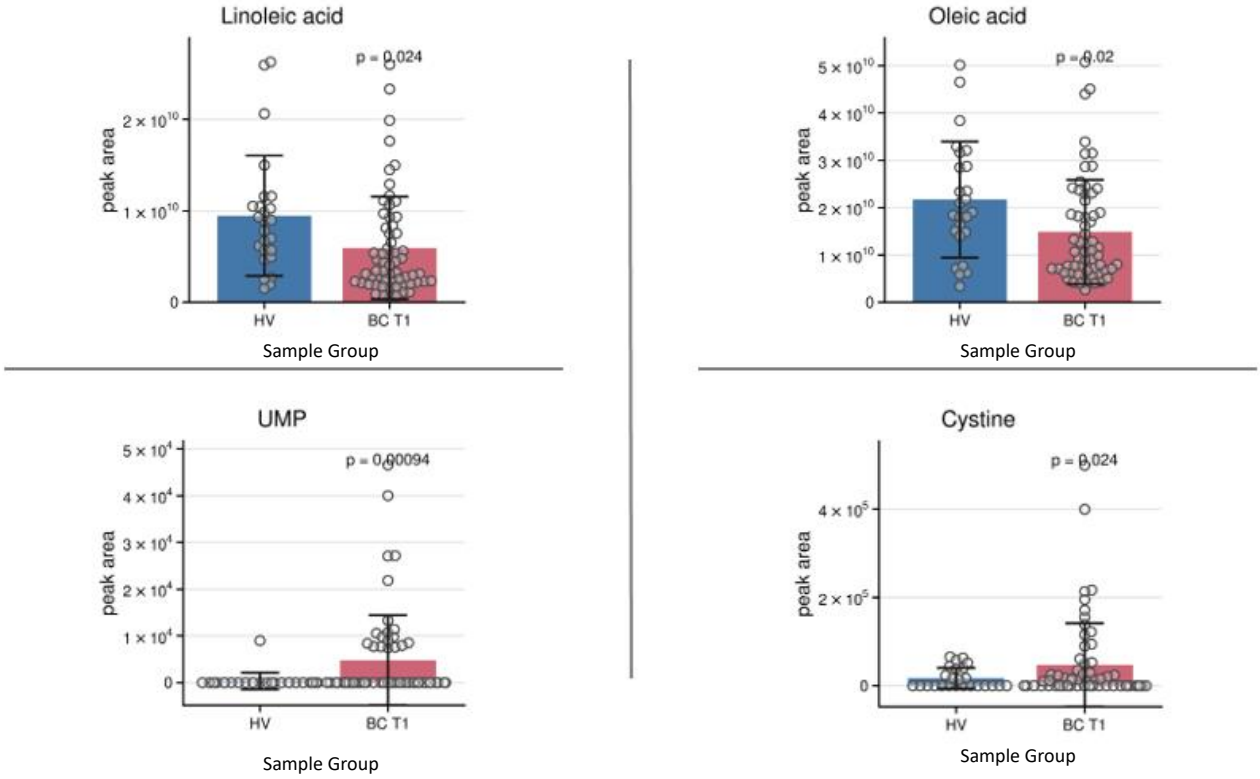


Figure 4.7. Boxplots demonstrating the peak areas of metabolites (Linoleic acid, Oleic acid, UMP and Cystine) that showed statistical significance across healthy individuals (HV)(blue) and breast cancer patients from the NEO-MICROBE study at timepoint 1 (BC T1) (red). Significant statistical differences are displayed by the p-value (T-test).

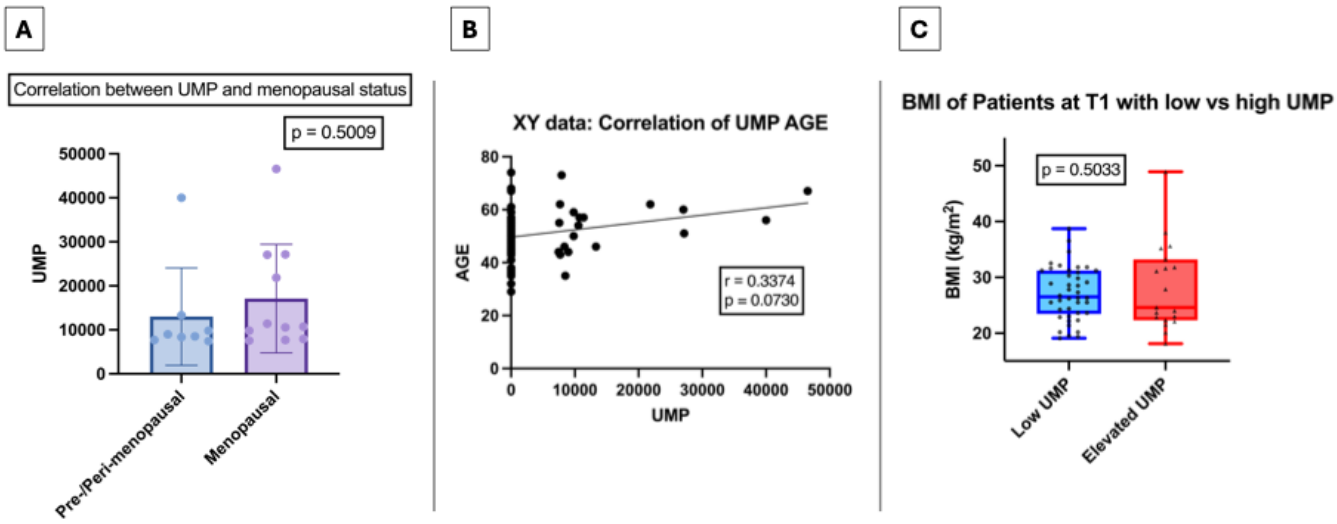


Figure 4.8. Correlation plots between UMP in breast cancer patients at T1 and clinical factors. A two-tailed T-test was carried out to calculate the p-value (A) Comparison of UMP levels between pre-/peri- menopausal and menopausal individuals showed no significant difference (p=0.5009). (B) Correlation analysis between UMP levels and patient age revealed a weak, non-significant positive correlation (r = 0.3374, p = 0.0730). (C) Body mass index (BMI) at timepoint 1 (T1) was compared between patients with low versus elevated UMP levels, showing no significant difference (p=0.5033).

4.3.2 PCR vs Non-PCR– whole population

We first examined the metabolic profiles of cancer patients at T1 who achieved a pathological complete response (pCR)(n=33) versus those who did not achieve pCR (n=28) after NACT. By assessing the gut metabolome of each individual with reference to pCR we aimed to identify potential biomarkers of neoadjuvant treatment response. Initial targeted analysis was conducted on the whole population (BC T1) to gain a broader understanding of treatment efficacy overall and how the gut microbiome may influence this response.

A Principal Component Analysis (PCA) was carried out to assess the variability of metabolic profiles between pCR and non-pCR patients. Figure 4.9 demonstrates the separation along the first two principal components between pCR (22.4% total variance explained) and non-pCR (12.8% total variance explained) groups. The pCR data points (circles) are more tightly clustered suggesting less variability within this group compared to non-pCR patients show points (squares) indicating a broader distribution of data points. Overall, there is some overlap of the groups suggesting similarities in metabolic profiles however some points are distinguishable between the groups, demonstrating a degree of metabolic difference potentially influencing response to NACT.

Statistical analysis was performed on metabolomic data generated from LCMS. The volcano plot in Figure 4.10 presents the variance of metabolites between pCR vs non-pCR

patients. A two-tailed T test was conducted on non-pCR vs pCR patients and log10 transformation was applied to calculate the -Log10(p-value) (y-axis), representing the level of significance of each metabolite. The magnitude of differences between the two groups for each metabolite is represented on a logarithmic scale to base 2 ($Log_2(FC)$)(x-axis). The metabolites that showed significant results and a FC > 1 are highlighted in red, with the metabolites in the upper left/right corners showing statistical significance (aKG and dTMP). The volcano plot highlights that aKG has higher levels in patients who achieved pCR and dTMP levels are more abundant in non-pCR patients. Visualisation of the data in Figure 4.10 allowed for identification of potential metabolites that are involved in driving response to NACT.

Following this, the data was analysed using the software tool Metabolite Autoplotter to measure metabolite abundance in cancer patients at T1 who achieved pCR versus non-pCR. Results indicated that most metabolites did not exhibit statistically significant differences (p-value ≤ 0.05) between the two groups (Figure 4.11, representative of all non-significant metabolites). The distribution of peak areas for individual points showed similar trends across the majority of metabolites, with a few outliers which may be consequent of subtype-specific trends.

The metabolites that did show statistically significant differences in peak area between the two groups included aKG, guanine and dTMP (Figure 4.12). The significant difference in dTMP levels (A) is represented by the p-value of 0.043, with non-PCR patients displaying higher levels. The relatively tight clustering of the data points for the pCR group shows that peak area was consistently low across the sample group. A broader spread of points is visible in the non-pCR group, suggesting more variability in dTMP levels within patients who did not achieve pCR. This variability may be explained by subtype driven metabolic

patterns. Guanine (B) also demonstrated a statistically significant difference (p=0.042), with reduced levels of guanine in pCR patients compared to non-pCR; thus, highlighting depleted guanine levels as a potential biomarker for treatment response. Distribution of individual points shows similar patterns to dTMP, with more variance seen in non-pCR patients. In the pCR group, clustering of points can be visualised however there is a small subset of patients with higher peak area of guanine. Lastly, aKG revealed the most striking results, demonstrating significantly higher levels in patients who achieved a complete pathological response, with a p-value of 0.0022.

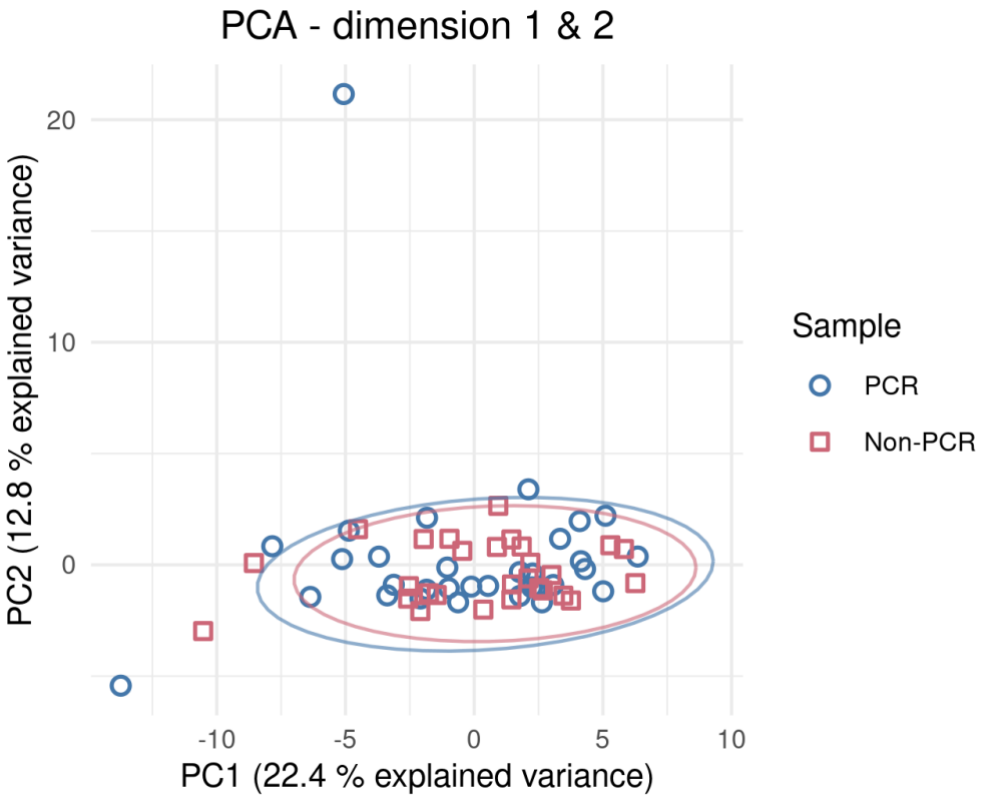


Figure 4.9. Principal component analysis (PCA) of stool metabolomic data from cancer patients at T1 who achieved pathological complete response after NACT (PCR)(n=33) and breast cancer patients who did not achieve pCR (Non-PCR)(n=28). Each point represents an individual sample.

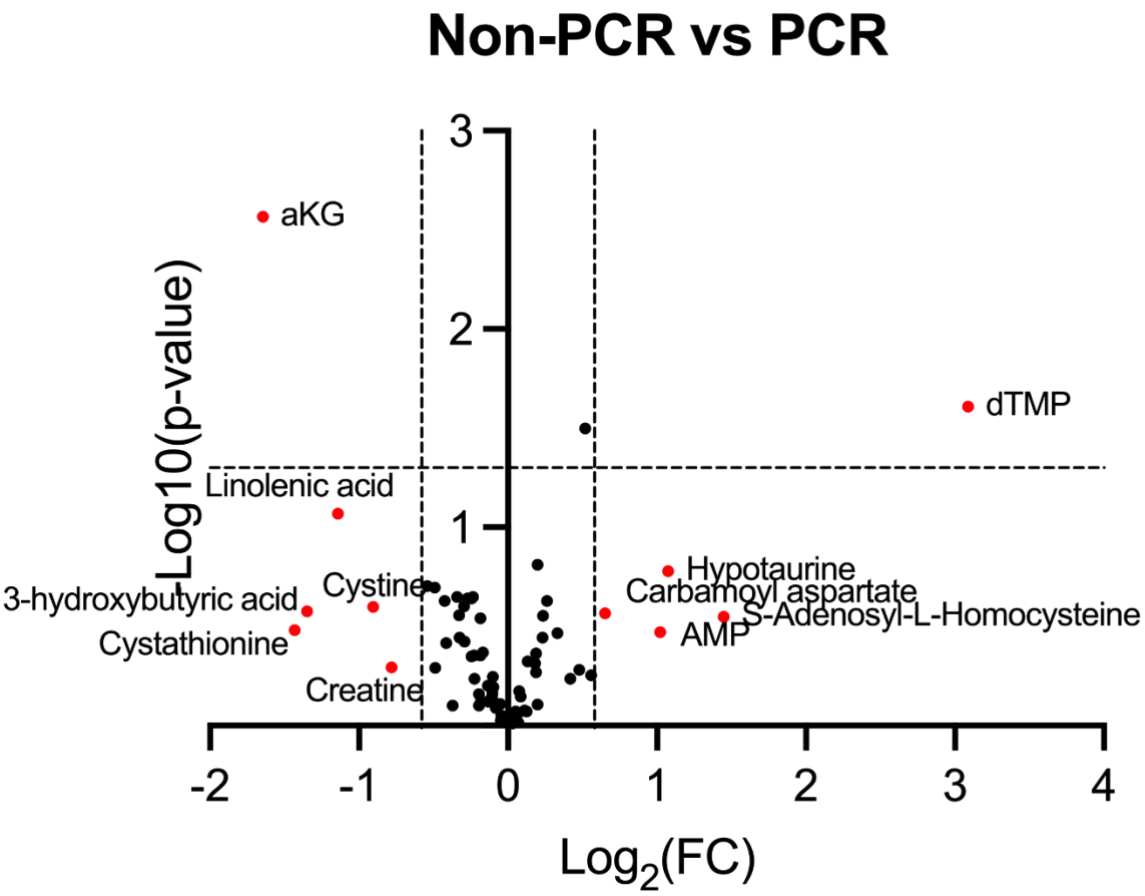


Figure 4.10. Volcano plot showing the relationship between metabolites p-values and fold changes among patients who did not achieve pathological complete response (Non-PCR) versus those who did achieve pCR (PCR). The effect size of the two groups is plotted on the x-axis on a logarithmic scale to base 2 (Log_2(FC)). The -log10 p-values are plotted on the y-axis (T-test). Gridlines act as cut off points for results of significance; metabolites of significance are highlighted by red dots.

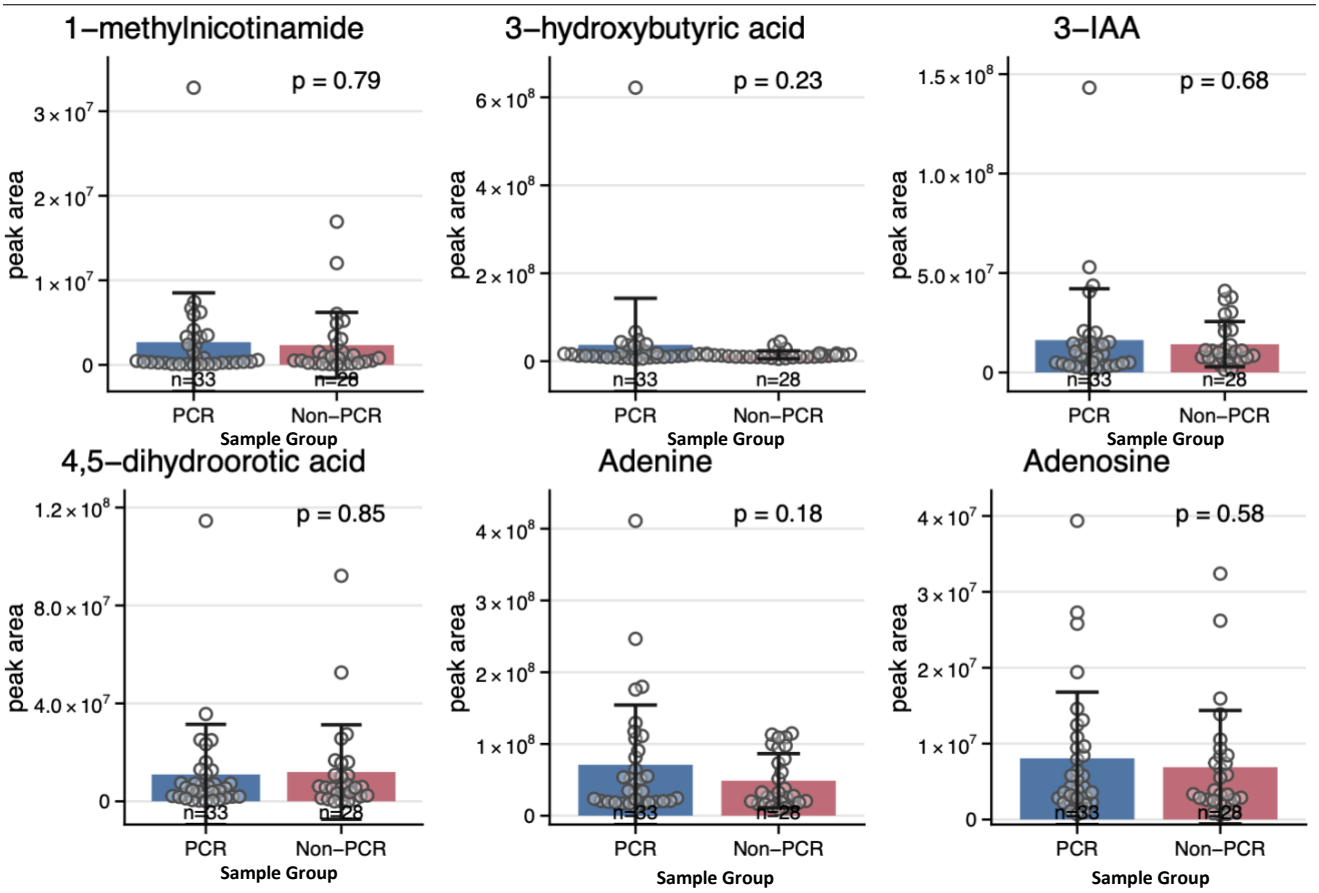


Figure 4.11. Boxplots representative of all metabolites analysed demonstrates the peak areas of metabolites (1-methylnicotinamide, 3-hydroxybutyric acid, 3-IAA, 4,5-dihydroorotic acid, Adenine, Adenosine) that did not show statistical significance across breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=33) and those who did not (Non-PCR)(red, n=28). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

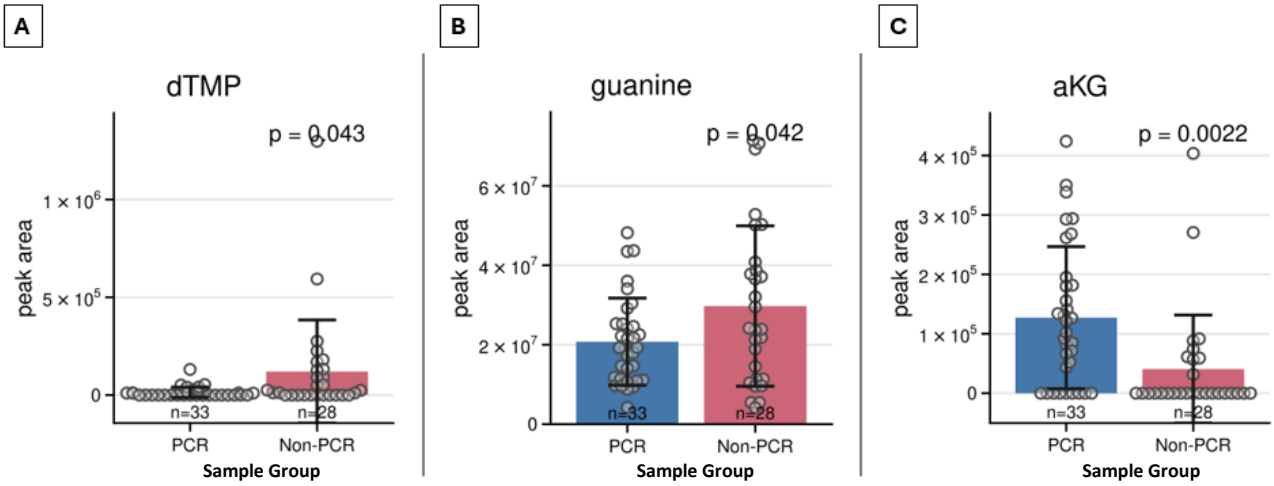


Figure 4.12. Boxplots demonstrating the peak areas of metabolites (dTMP (A), Guanine (B), aKG (C)) that exhibited statistical significance across breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=33) and those who did not (Non-PCR)(red, n=28). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

Targeted Analysis in Breast Cancer Subtypes

After carrying out a comprehensive analysis of the whole population to obtain an understanding of how metabolite profile has an influence on response rate to NACT in breast cancer patients, we next examined subtype-specific responses. By exploring metabolic profiles of patients based on their subtype and how they respond to NACT, this enables detection of potential metabolic drivers of NACT response unique to breast cancer subtype; subsequently allowing for potential therapeutic targets to be recognised and development of personalised treatments.

Table 3 shows the demographics of patients stratified into subtypes of breast cancer. Patients demonstrate fairly similar median ages and bmi across HER2+/ER+, HER2+/ER- and TNBC; median age = 54y, 56y, 45.5y, and median BMI = 26.8kg/m2, 30.4kg/m2, 24.8kg/m2, respectively. The majority of patients were white with no co-morbidities across all subtypes. Patients were more commonly post-menopausal in HER2+ ER+ and ER- groups, compared to pre-menopausal in TNBC. Histology of the tumour highlighted nearly all patients had invasive ductal carcinoma in all subtypes with a fairly even split between grade 2 and 3 histological grade however in TNBC the majority of patients (88.5%) were grade 3. Patients had similar median tumour size (mm) at baseline (34.5mm, 46mm, 30.3mm), with the majority of patients in each subtype presenting negative baseline nodal status and a clinical TNM stage of II. Anthracycline and taxane was the primary treatment in HER2+ER+ and ER-, however contrastingly the addition of platinum alongside anthracycline and taxane was more commonly used in TNBC. Patients with TNBC were largely recruited prior to routine use of the immune checkpoint inhibitor, pembrolizumab with only one patient having received this immunotherapeutic drug.

	HER2+/ER+ (n=24)	HER2+/ER- (n=14)	TNBC (n=26)
Age – median (range)	54 (43-68)	56 (32-74)	45.5 (29-67)
BMI (kg/m2) – median (range)	26.8 (21.5-42.1)	30.4 (19.4-48.9)	24.8 (18.1-38.7)
• Underweight <18.5	0	0	1 (3.8%)
• Healthy 18.5-25	7 (29.2%)	4 (28.6%)	12 (46.1%)
• Overweight 25-30	10 (41.7%)	3 (21.4%)	4 (15.4%)
• Obese >30	6 (25%)	6 (42.9%)	9 (34.6%)
• Morbidly Obese >40	1 (4.2%)	1 (7.1%)	0
Ethnicity			
• White	23 (95.8%)	13 (92.9%)	23 (88.5%)
• Black African	0	1 (7.1%)	1 (3.8%)
• Asian Pakistani	0	0	1 (3.8%)
• Asian Chinese	1 (4.2%)	0	0
• Arab	0	0	1 (3.8%)
Menopausal status			
• Pre- or peri-menopausal	9 (37.5%)	4 (28.6%)	19 (73.1%)
• Post-menopausal	15 (62.5%)	10 (71.4%)	7 (26.9%)
Co-morbidities			
• gastrointestinal	3 (12.5%)	0	2 (7.7%)
• diabetes	1 (4.2%)	1 (7.1%)	2 (7.7%)
• autoimmune	1 (4.2%)	1 (7.1%)	1 (3.8%)
Histology			
• Invasive Ductal carcinoma	23 (95.8%)	13 (92.9%)	26 (100%)
• Invasive Lobular carcinoma	1 (4.2%)	1 (7.1%)	0
Histological grade			
• Grade 1	1 (4.2%)	0	0
• Grade 2	13 (54.2%)	7 (50%)	3 (11.5%)
• Grade 3	10 (41.7%)	7 (50%)	23 (88.5%)
Baseline tumour size, (mm)			
• Median (range)	34.5 (18-112)	46 (25-110)	30.3 (10-73)
Baseline nodal status			
• Negative	18 (75%)	11 (78.6%)	24 (92.3%)
• Positive	6 (25%)	3 (21.4%)	2 (7.7%)
Clinical TNM Stage			
• I	3 (12.5%)	0	2 (7.7%)
• II	12 (50%)	12 (85.7%)	20 (76.9%)
• III	9 (37.5%)	2 (14.3%)	4 (15.4%)
Chemotherapy			
• Anthracycline / taxane	17 (70.8%)	9 (64.3%)	4 (15.4%)

• anthracycline / taxane with platinum	0	0	21 (80.8%)
• Docetaxel + carboplatin	6 (25%)	3 (21.4%)	0
• Taxane monotherapy	0	1 (7.1%)	0
• Docetaxel + cyclophosphamide	1 (4.2%)	1 (7.1%)	1 (3.8%)
Immunotherapy			
• pembrolizumab	0	0	1 (3.8%)
HER2-targeted therapy			
• Trastuzumab + pertuzumab	24 (100%)	13 (92.9%)	0
• Trastuzumab	0	1 (7.1%)	0

Table 3: Baseline Demographics and Neoadjuvant Treatment by Breast Cancer Subtype

4.3.3 PCR vs Non-PCR – HER2+

Principal component analysis was conducted to examine the variation in metabolic profiles of HER2+ breast cancer patients who achieved pCR (PCR)(n=19) and those who did not (Non-PCR) (n=17) (Figure 4.13). The separation between the two groups is along the first two principal components (PC1 and PC2, 39.8 explained total variance overall). Little to no separation is seen between the two groups, suggesting there is little variation in metabolic profiles between pCR and non-pCR patients.

A two-tailed T-test was conducted to determine the difference between the average peak area of the metabolites analysed in HER2+ patients who achieved pCR and those who did not. The effect size was also calculated, and log transformation was carried out on both of these statistics; the results of this statistical analysis can be visualised in Figure 4.14. The volcano plot indicates metabolites of significance by highlighting them with red dots (x = <0.58 or >0.58); furthermore, metabolites with a -log10 p-value >1.3 showed statistical significance. Some noteworthy statistically significant metabolites included aKG, guanosine, uridine, and inosine, suggesting potential association with response to NACT in HER2+ breast cancer.

As the next phase we carried on from the previous metabolomic analysis in Metabolite Autoplotter to visualise the difference in peak areas of metabolites between pCR and non-PCR HER2+ patients. The majority of metabolites showed no association to response of NACT in HER2+ breast cancer patients (Figure 4.15).

Five metabolites showed statistically significant differences between HER2+ patients who achieved pCR and those who did not; aKG, Citrate, Guanosine, Inosine and Uridine (Figure 4.12). Inosine (A), uridine (B) and guanosine (D) all showed elevated levels in non-pCR patients (p = 0.042, 0.028 and 0.021, respectively). All of these metabolites

showed similar patterns in distribution of individual points as well with clustering towards low/absent levels in the pCR group and wider spread of point in non-pCR patients. Citrate (E) showed a slightly statistically significant difference with a p-value of 0.053. Patients who achieved a pCR were shown to have higher abundance of citrate in their gut, with some individuals showing an unusually high peak area; this may be contributing to the significantly higher levels of citrate observed relative to non-pCR patients that demonstrated little variance within the population. The metabolite with the highest statistical significance was aKG (C), demonstrating higher levels in the pCR patients compared to non-pCR patients ($p=0.0073$). Clustering of individual data points is presented in non-pCR patients, suggesting that aKG is consistently low in most individuals within the group. In contrast, patients within the pCR group show a large distribution of points signifying high levels of variance within the group, with the majority of them being higher than non-pCR patients. Overall, it can be suggested that these metabolites which are all involved in the same nucleotide metabolic pathway may play a role in response to NACT in HER2+ malignancies.

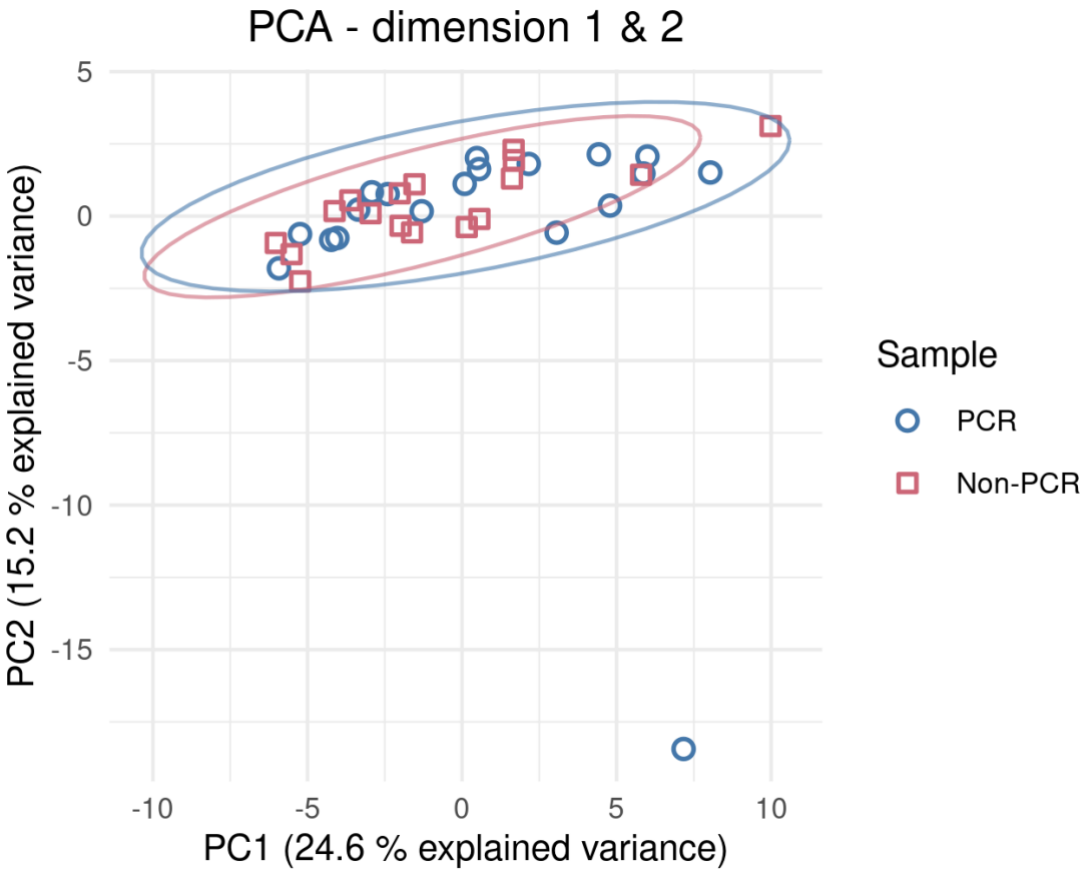


Figure 4.13. Principal component analysis (PCA) of stool metabolomic data from HER2+ cancer patients at T1 who achieved pathological complete response after NACT (PCR)(n=19) and breast cancer patients who did not achieve pCR (Non-PCR) (n=17). Each point represents an individual sample.

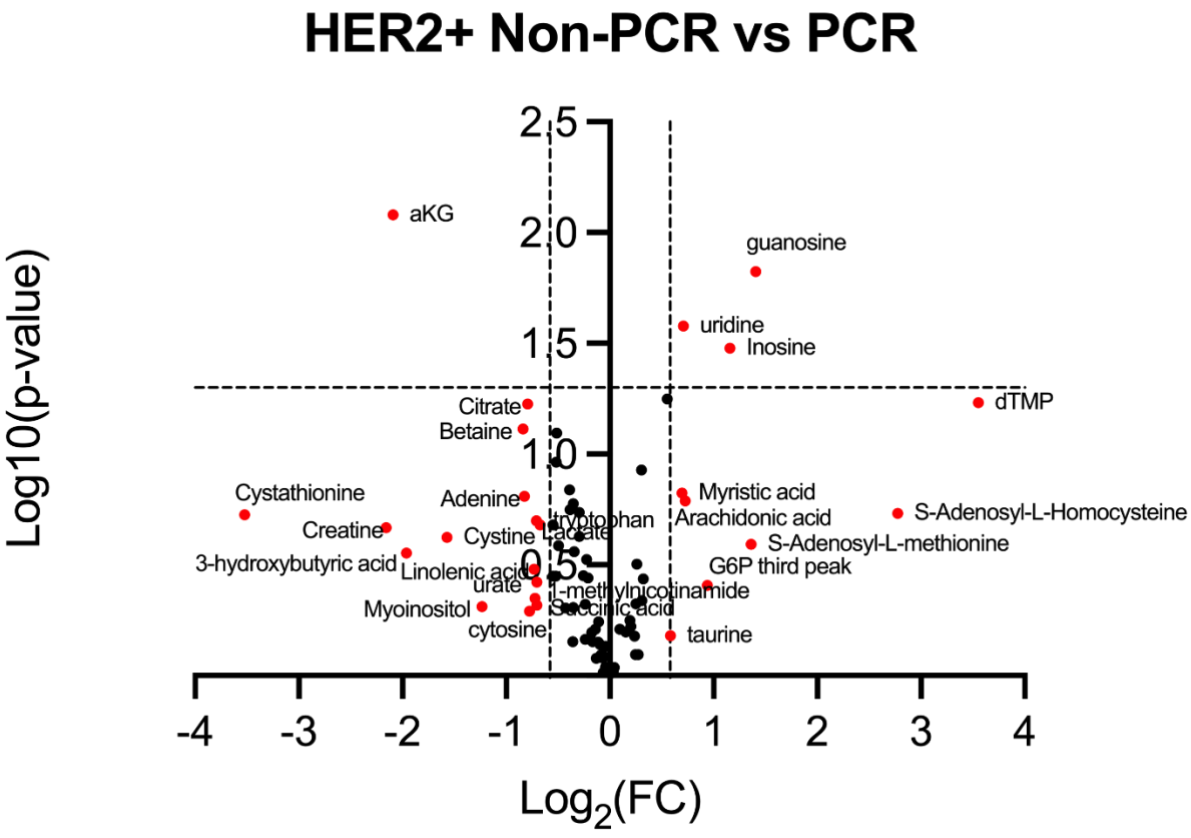


Figure 4.14. Volcano plot showing the relationship between metabolites p-values and fold changes among HER2+ patients who did not achieve pathological complete response (Non-PCR) versus those who did achieve pCR (PCR). The effect size of the two groups is plotted on the x-axis on a logarithmic scale to base 2 ($Log_2(FC)$). The -log10 p-values are plotted on the y-axis (T-test).

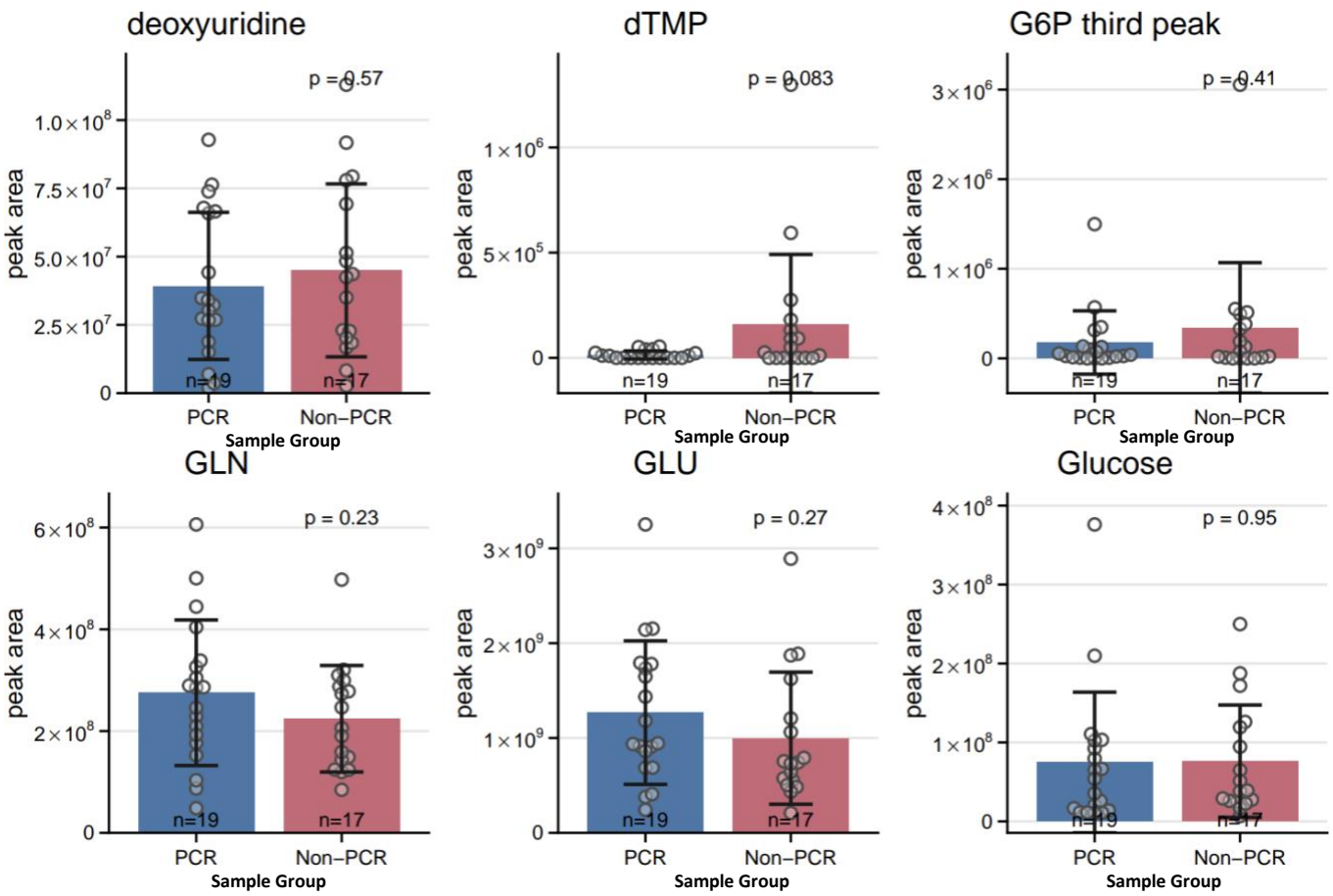


Figure 4.15. Boxplots demonstrating the peak areas of metabolites (deoxyuridine, dTMP, Glucose-6-Phosphate (G6P) third peak, GLN, GLU and Glucose) that did not show statistical significance across HER2+ breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=19) and those who did not (Non-PCR)(red, n=17). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

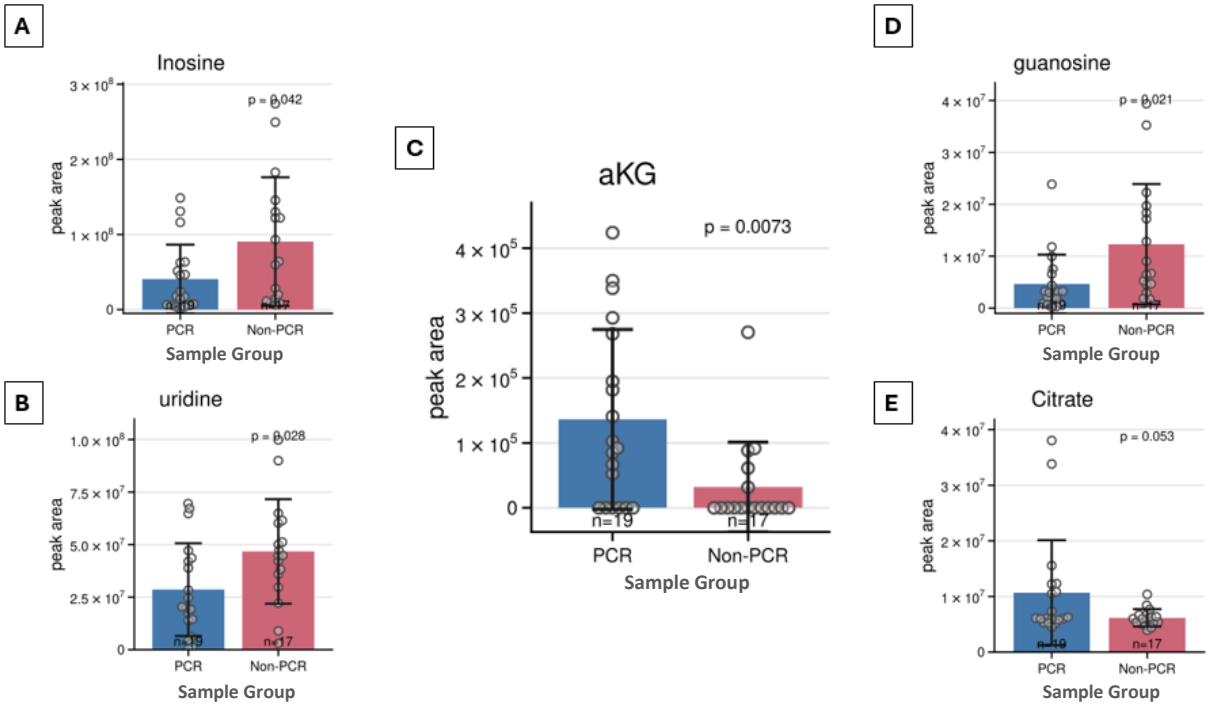


Figure 4.16. Boxplots demonstrating the peak areas of the five metabolites (Inosine (A), Uridine (B), aKG (C), Guanosine (D), Citrate (E)) that exhibited statistical significance across HER2+ breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=19) and those who did not (Non-PCR)(red, n=17). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

4.3.4 PCR vs Non-PCR – HER2+ER+

Principal component analysis (PCA) was carried out to understand the differences between the two populations; HER2+ER+ patients at T1 who achieved pCR compared to those who did not (non-pCR) after treatment. Figure 4.17 shows that there was no separation between the two groups suggesting that the gut microbiome of the HER+ER+ patients who achieved pCR and those who did not (non-pCR) were very similar.

After various statistical analysis (as per other targeted analysis) to understand any differences in metabolic profiles between the two groups (pCR and non-pCR) in HER2+ER+ breast cancer patients. Figure 4.18 demonstrates the metabolites that showed significant difference (red dots) between the two sample groups; six metabolites exhibited statistically significant differences including Guanosine, Inosine, Myristic acid, Citrate, Cystine and Uridine. This is emphasized in Figure 4.20, which demonstrates analysis of the difference in peak level of each metabolite specific to HER2+ER+ breast cancer subtype. The bar plots displayed in Figure 4.19 highlight that for most of the metabolites analysed (n=76) there was no significant change in metabolic profiles between individuals who responded completely to NACT (pCR) and those who did not (non-pCR). This lack of variation is consistent with the patterns of distribution of individual points seen between the two groups, with the exception of a few outliers (Figure 4.19).

The bar plot in Figure 4.20 indicates that three metabolites; Myristic acid (A), guanosine (B) and inosine (C) all exhibit a statistically significant difference in peak area, with elevated levels in HER+ER+ patients who did not achieve pathological response to NACT. In contrast, HER+ER+ patients who achieved pCR exhibited statistically significant levels of Citrate (D) and L-pyroglutamic acid (E). Cystine appeared to be statistically significant in Figure 4.20, with elevated levels in PCR patients; however, this did not reach the

threshold of $p<0.05$ when statistical analysis was performed using Metabolite Autoplotter rather than Microsoft Excel (Figure 4.20 (E). HER2+ER+ patients who achieved pCR still exhibit higher levels of cystine however the majority of the population show low or absent levels, this is highlighted by the clustering of individual data points towards the bottom of the y-axis in the group.

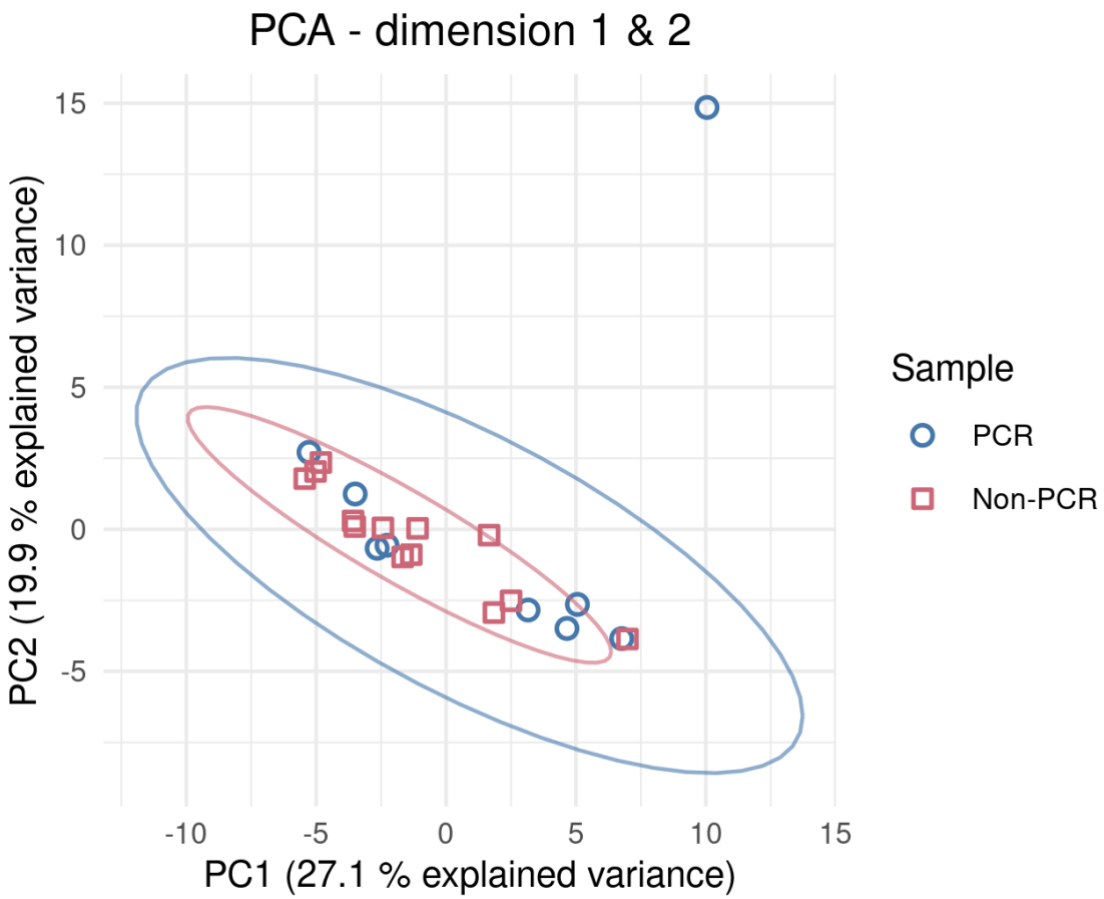


Figure 4.17. Principal component analysis (PCA) of stool metabolomic data from HER2+ER+ cancer patients at T1 who achieved pathological complete response after NACT (PCR)(n=9) and breast cancer patients who did not achieve pCR (Non-PCR) (n=13). Each point represents an individual sample.

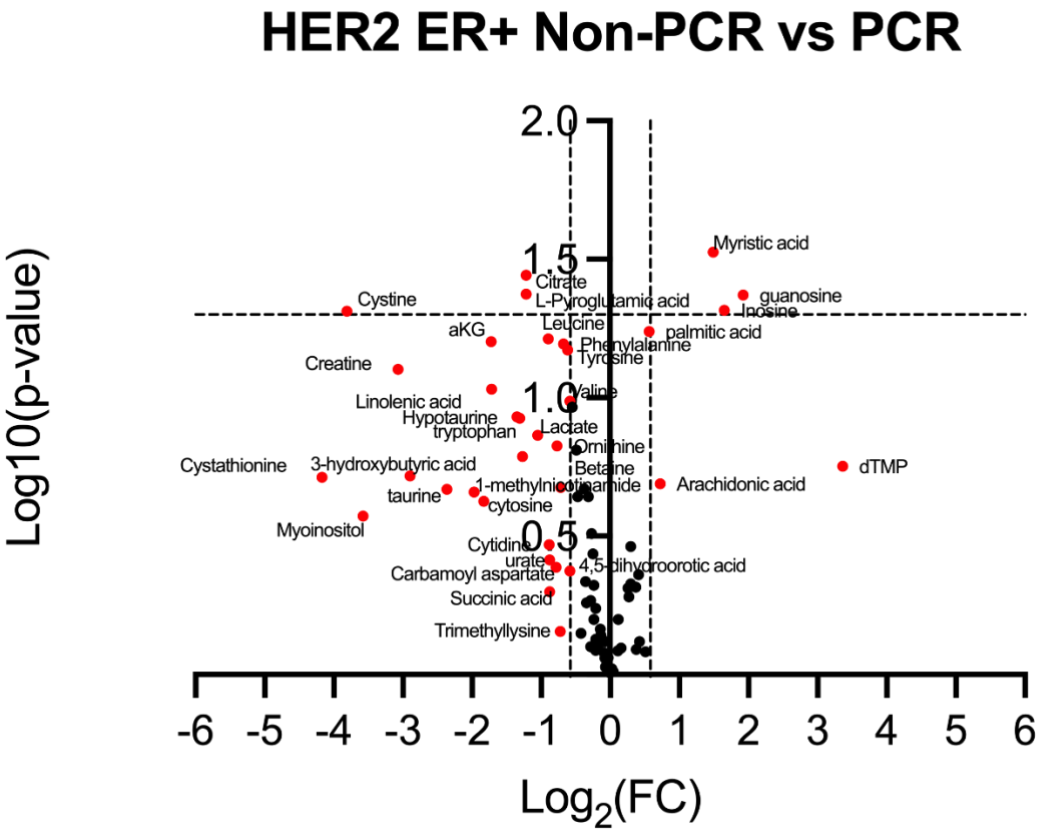


Figure 4.18. Volcano plogts showing targeted analysis of stool metabolomics. Volcano plot showing the relationship between metabolites p-values (T-test) and fold changes among HER2+ER+ patients who did not achieve pathological complete response (Non-PCR) versus those who did achieve pCR (PCR). The effect size of the two groups is plotted on the x-axis on a logarithmic scale to base 2 ($Log_2(FC)$). The -log10 p-values are plotted on the y-axis.

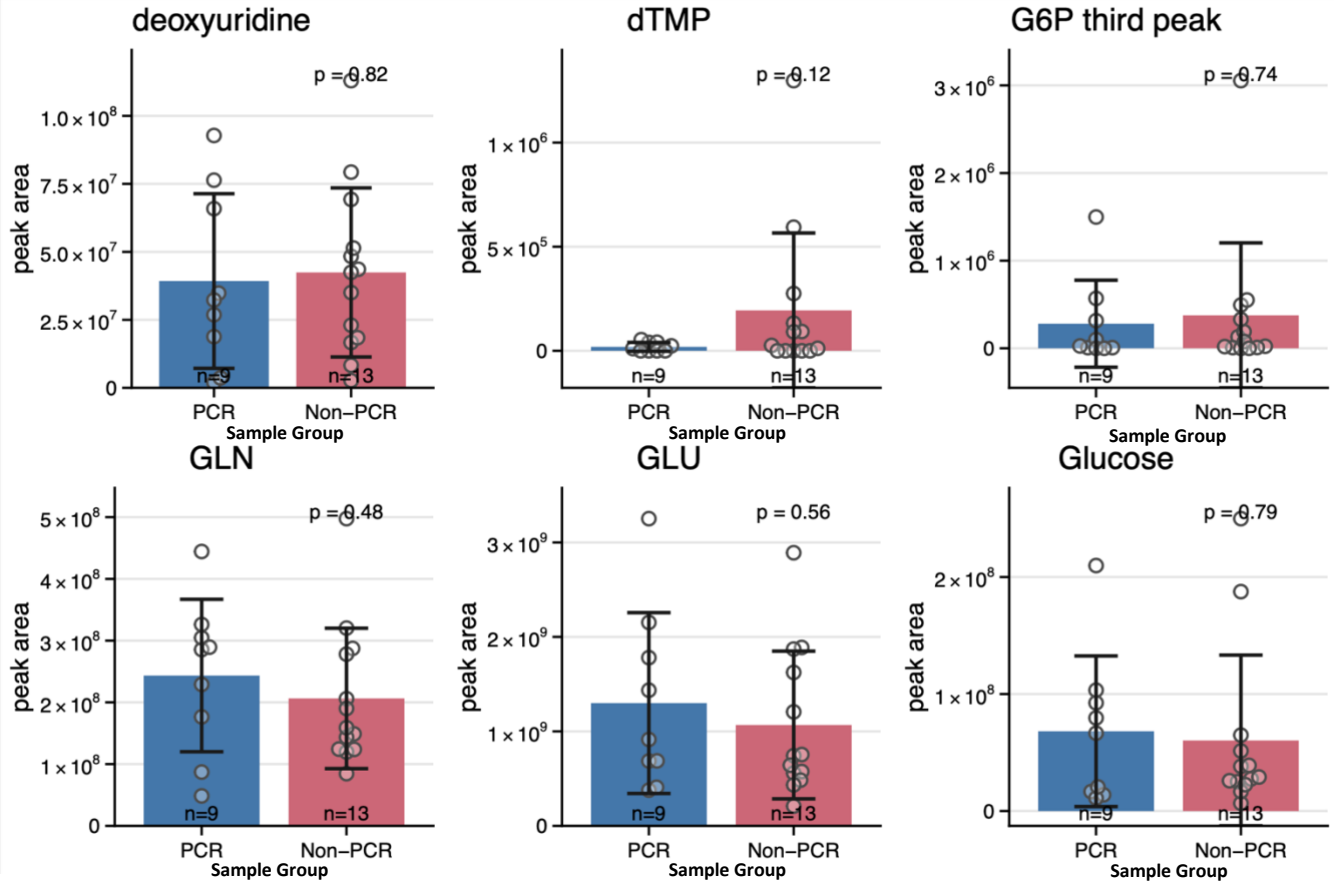


Figure 4.19. Comparison of peak areas of metabolites according to response. Boxplots representing the peak areas of metabolites (deoxyuridine, dTMP, G6P third peak, GLN, GLU and Glucose) that did not show statistical significance across HER2+ER+ breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=9) and those who did not (Non-PCR)(red, n=13). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

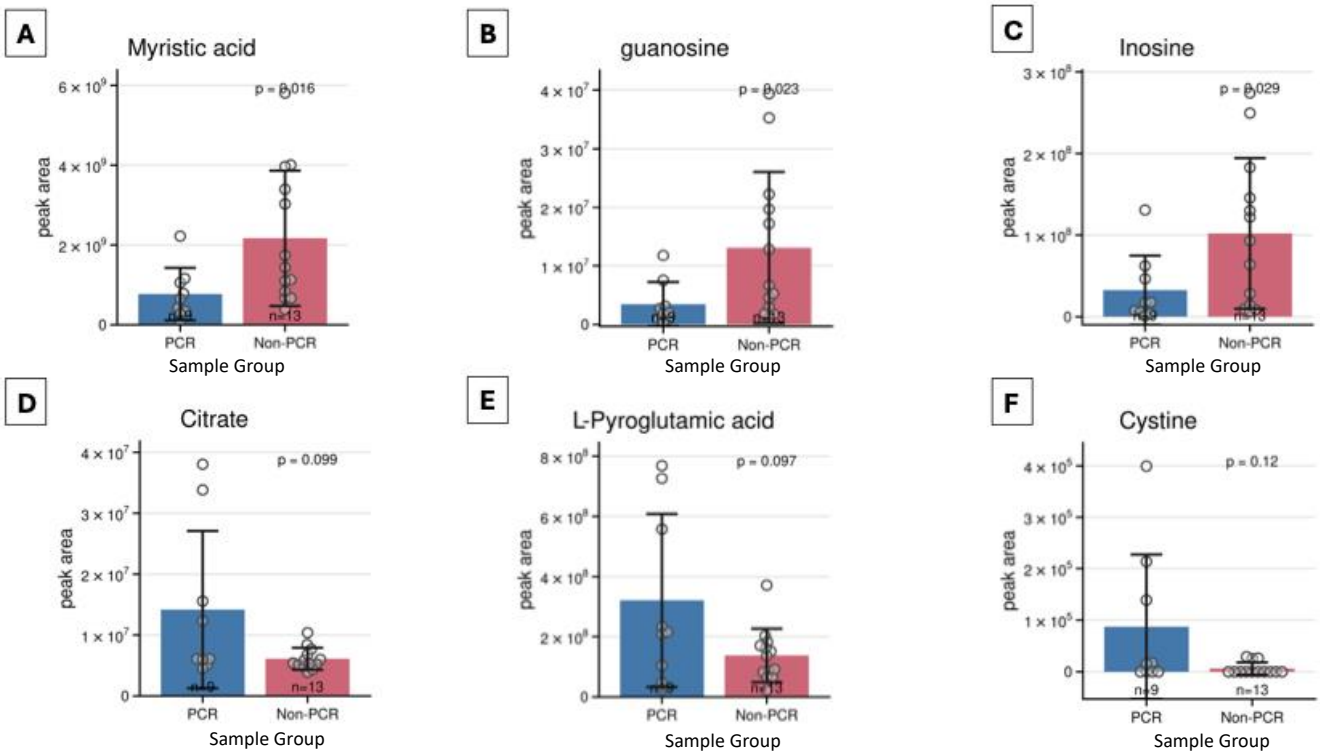


Figure 4.20. Boxplots demonstrating the peak areas of the five metabolites (Myristic acid (A), Guanosine (B), Inosine (C), Citrate (D), L-Pyroglutamic acid (E), Cystine (F)) that exhibited statistical significance across HER2+ER+ breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=9) and those who did not (Non-PCR)(red, n=13). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

4.3.5 PCR vs Non-PCR – HER2+ER-

A Principal Component Analysis was carried out on stool samples using metabolite autoplotter software from HER2+ER- breast cancer patients at T1 to understand the variance in metabolomic profiles between patients who achieve pathological complete response (PCR) and those who did not achieve pCR (Non-PCR). Figure 4.21 describes the differences in metabolites present in the gut microbiome of patients in both sample groups; the separation is shown along the first two principal components (PC1 explaining 29.9% total variance, PC2 explaining 14% total variance). The PCA plot suggests that there is separation between the populations suggesting the metabolite profiles of pCR compared to non-pCR are distinct. There is a slight overlap seen suggesting that there may be some sort of similarities in metabolites found in the patients across the two groups.

Statistical analysis (p-value and fold change calculated) was conducted on the stool metabolomics data in excel from LCMS analysis, which was then inputted into software (GraphPad Prism) for visualisation. The volcano plot (Figure 4.22) demonstrates the results of the statistical analysis, highlighting the metabolites that exhibited significant differences between the two groups in red. The metabolites that showed statistically significant differences in gut metabolite composition are seen in the upper right section of the plot (Cytosine, Guanine, Uridine and Cytidine).

Further analysis using Prism software was carried out to understand the exact statistical difference in specific metabolites (n=76) between patients who achieved pCR and those who did not across the HER2+ER- population. For the majority of metabolites (Figure 4.23) no statistically significant differences were demonstrated in peak area of metabolites between the sample groups. This suggests that for the metabolites similar metabolic

profiles were seen in patients across the two therefore they do not appear to have an influence on response to NACT.

The exact statistical value of difference was calculated for the metabolites that showed elevated levels in non-PCR patients in the volcano plot. Figure 4.24 demonstrates that Uridine (A) ($p = 0.018$), Guanine (C) ($p = 0.0068$) and Cytosine (D) ($p = 0.016$) presented with statistical significantly higher levels in non-pCR HER2+ER- patients. Glutamate (GLU)(B) exhibited a statistically significant difference, with elevated levels seen in HER2+ER- patients who achieved pCR. This metabolic result was not shown in figure 4.24 which may be a consequence of the p value being 0.05, making it only just significant.

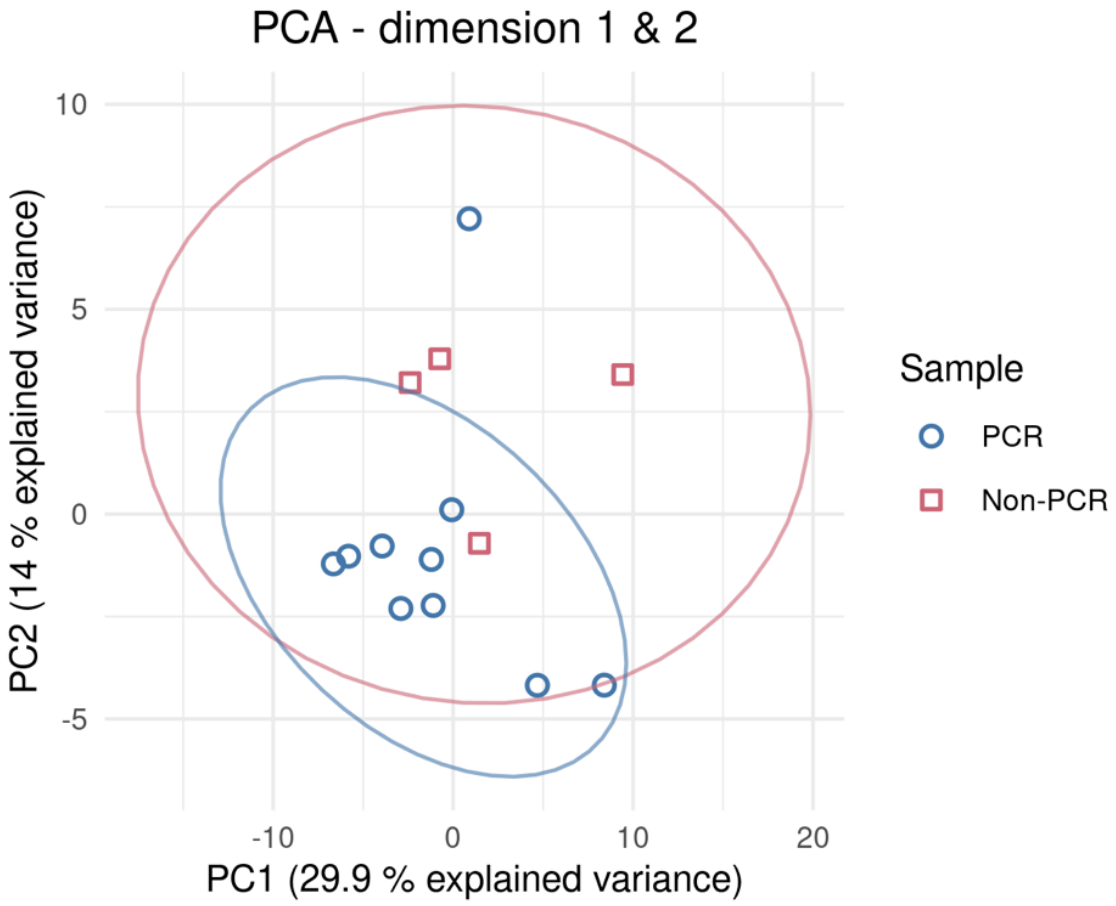


Figure 4.21. Principal component analysis (PCA) of stool metabolomic data from HER2+ER- cancer patients at T1 who achieved pathological complete response after NACT (PCR)(n=10) and breast cancer patients who did not achieve pCR (Non-PCR) (n=4). Each point represents an individual sample.

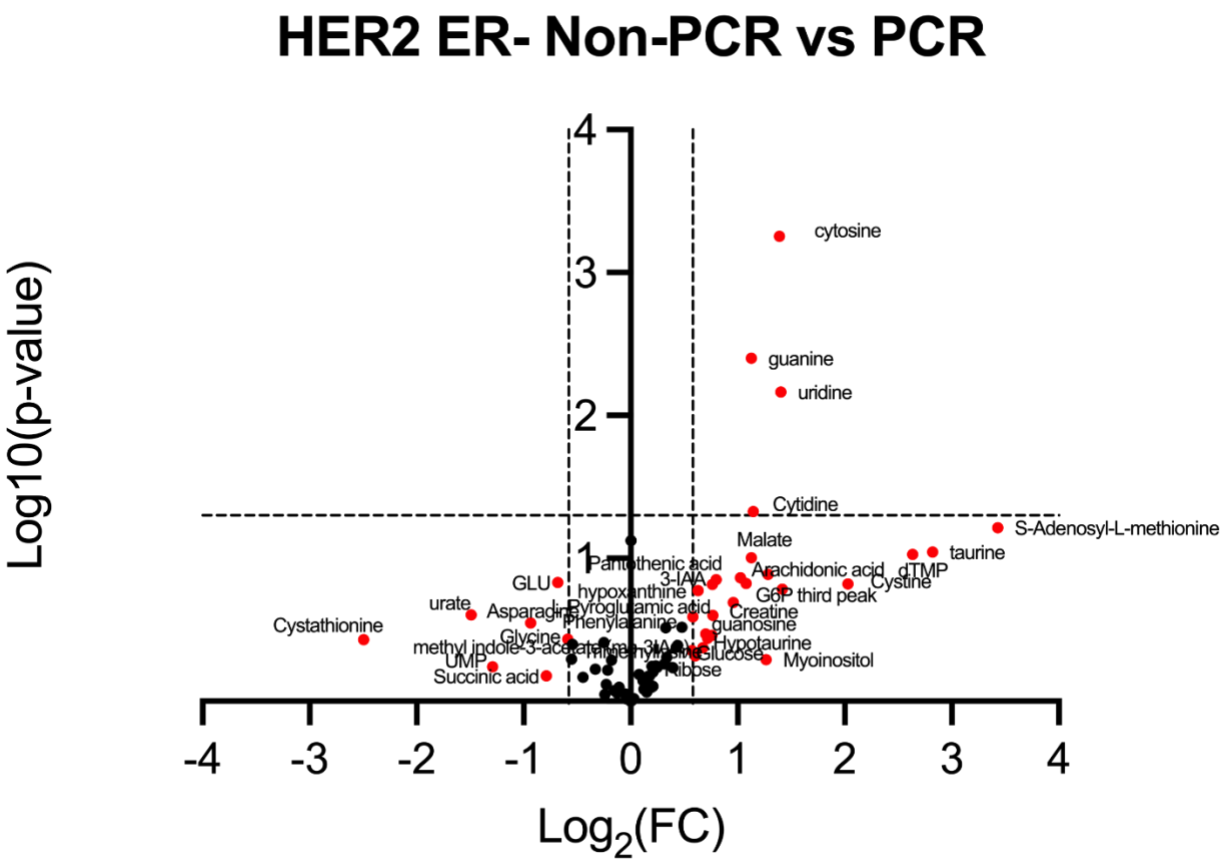


Figure 4.22. Volcano plot showing the relationship between metabolites p-values and fold changes among HER2+ER- patients who did not achieve pathological complete response (Non-PCR) versus those who did achieve pCR (PCR). The effect size of the two groups is plotted on the x-axis on a logarithmic scale to base 2 ($Log_2(FC)$). The -log10 p-values are plotted on the y-axis (T-test).

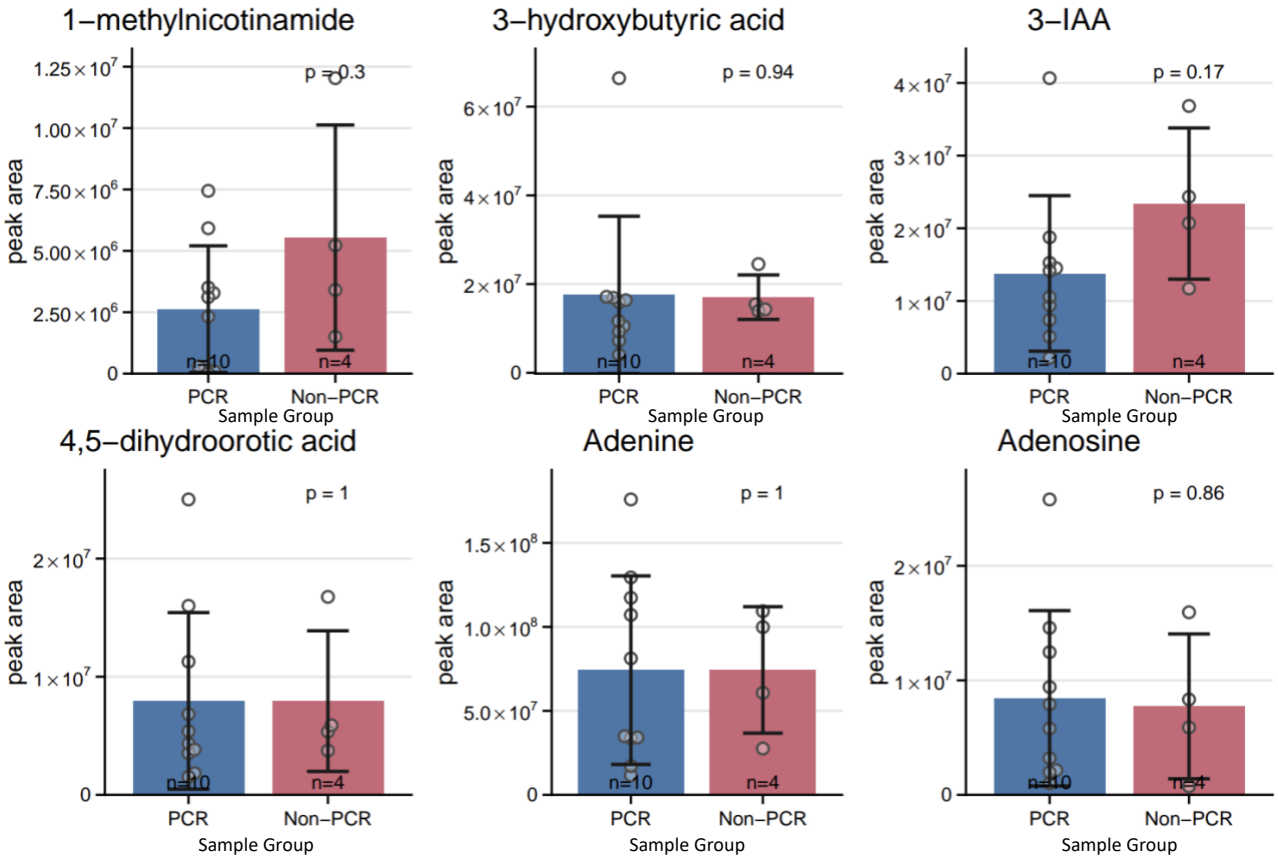


Figure 4.23. Boxplots demonstrating the peak areas of metabolites (1-methylnicotinamide, 3-hydroxybutyric acid, 3-IAA, 4,5-dihydroorotic acid, Adenine, Adenosine) that did not show statistical significance across HER2+ER+ breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=10) and those who did not (Non-PCR)(red, n=4). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value.

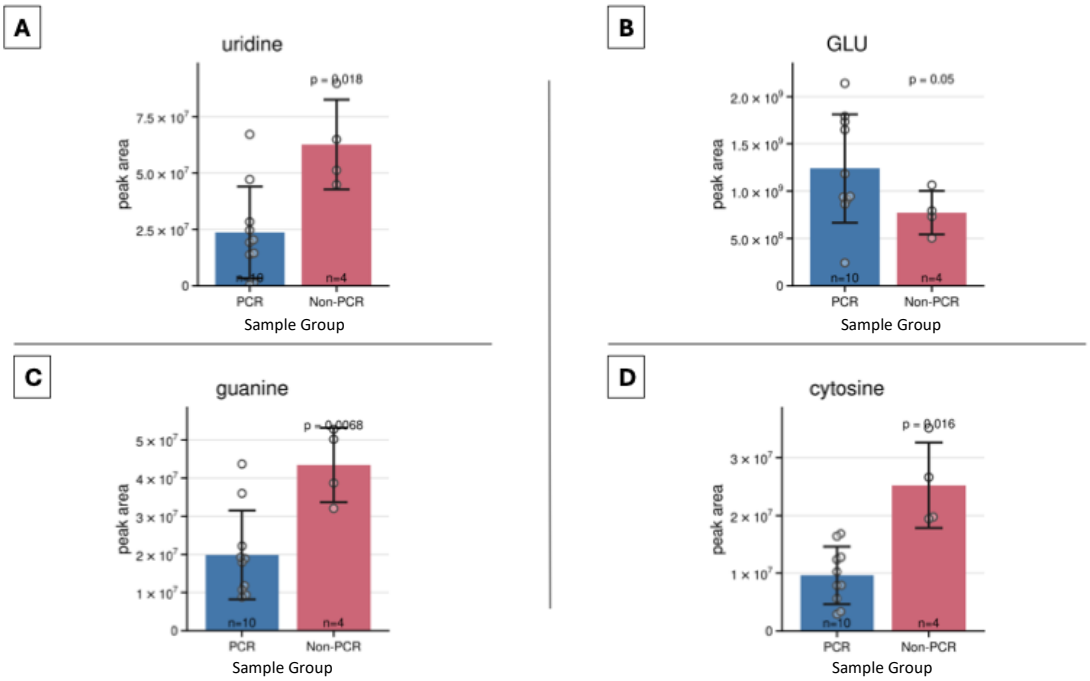


Figure 4.24. Boxplots demonstrating the peak areas of the five metabolites (Uridine (A), GLU (B), Guanine (C), Cytosine (D)) that exhibited statistical significance across HER2+ER+ breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=10) and those who did not (Non-PCR)(red, n=4). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

4.3.6 PCR vs Non-PCR – TNBC

A principal component analysis carried out using Metabolite Autoplotter of the stool metabolomics data from stool samples of TNBC breast cancer patients at T1 demonstrated little to no separation between the two sample groups (PCR and non-PCR). Figure 4.25 presents the separation between the groups along the first two principal components, with 26.4 total variance explained by PC1 and 13.2% total variance explained by PC2. The plot presents a high degree of overlap in metabolites with the exception of a few outliers, suggesting there is little variation in metabolomic profiles in TNBC patients who achieved pCR and those who did not.

Statistical analysis using Graphpad Prism on the two sample groups further reiterated that the majority of metabolites analysed demonstrated no significant difference between the two groups and subsequently highlight no influence on response to NACT. The volcano plot (Figure 4.26) presents the metabolites that were significantly different between TNBC pCR and non-pCR patients in red, with the compounds of statistical significance including Aspartate, Creatine, Glutamate, Inosine, Linoleic acid, Oleic acid, Stearic acid, and Succinic acid.

The boxplots shown in Figure 4.27 highlight the metabolites that did not show statistically significant differences between the two groups after further statistical analysis. Figure 4.28 demonstrates the eight metabolites that did show statistically significant differences in levels between TNBC patients who achieved pCR and those who did not; including Aspartate (A)($p = 0.016$), Creatine (B)($p = 0.052$), Glutamate (C)($p = 0.041$), Inosine (D)($p = 0.018$), Linoleic acid (E)($p = 0.031$), Oleic acid (F)($p = 0.027$), Stearic acid (G)($p = 0.035$) and Succinic acid (H)($p = 0.066$). Although succinic acid appeared to be of statistical significance on the volcano plot (Figure 4.28), when the p-value was calculated via Metabolite Autoplotter, it did not reach the threshold for significance.

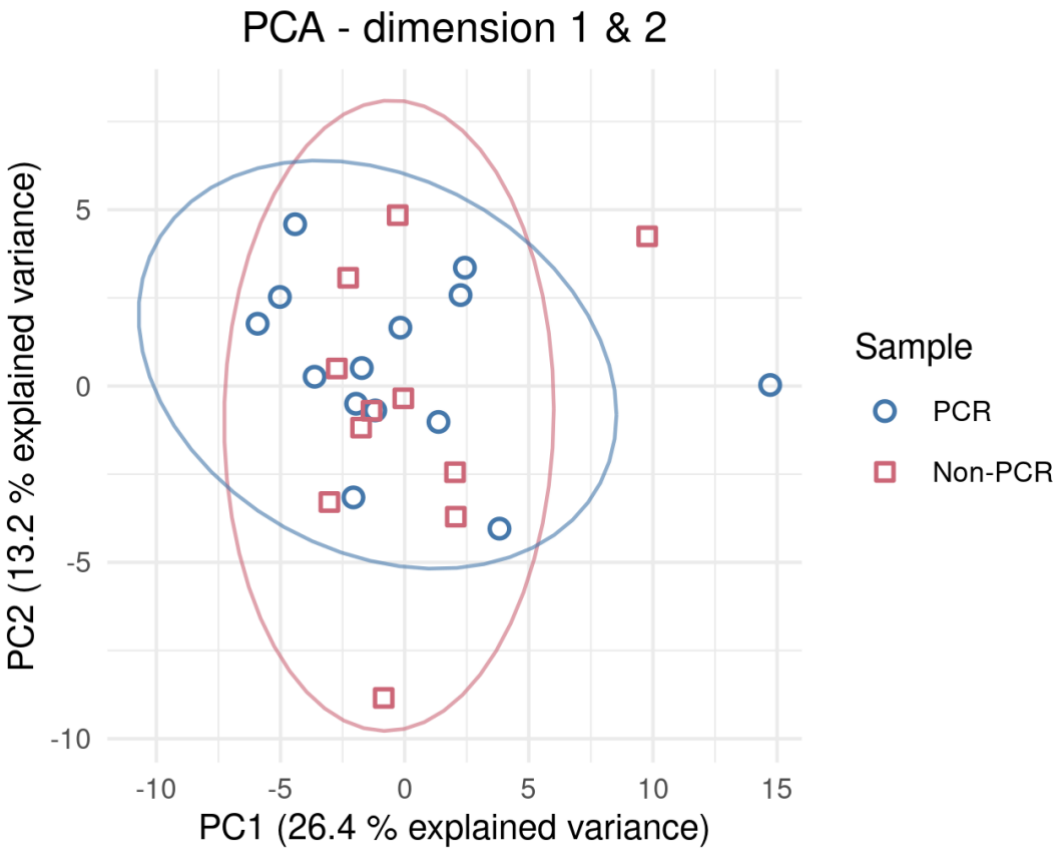


Figure 4.25. Principal component analysis (PCA) of stool metabolomic data from TNBC cancer patients at T1 who achieved pathological complete response after NACT (PCR)(n=14) and breast cancer patients who did not achieve pCR (Non-PCR) (n=11). Each point represents an individual sample.

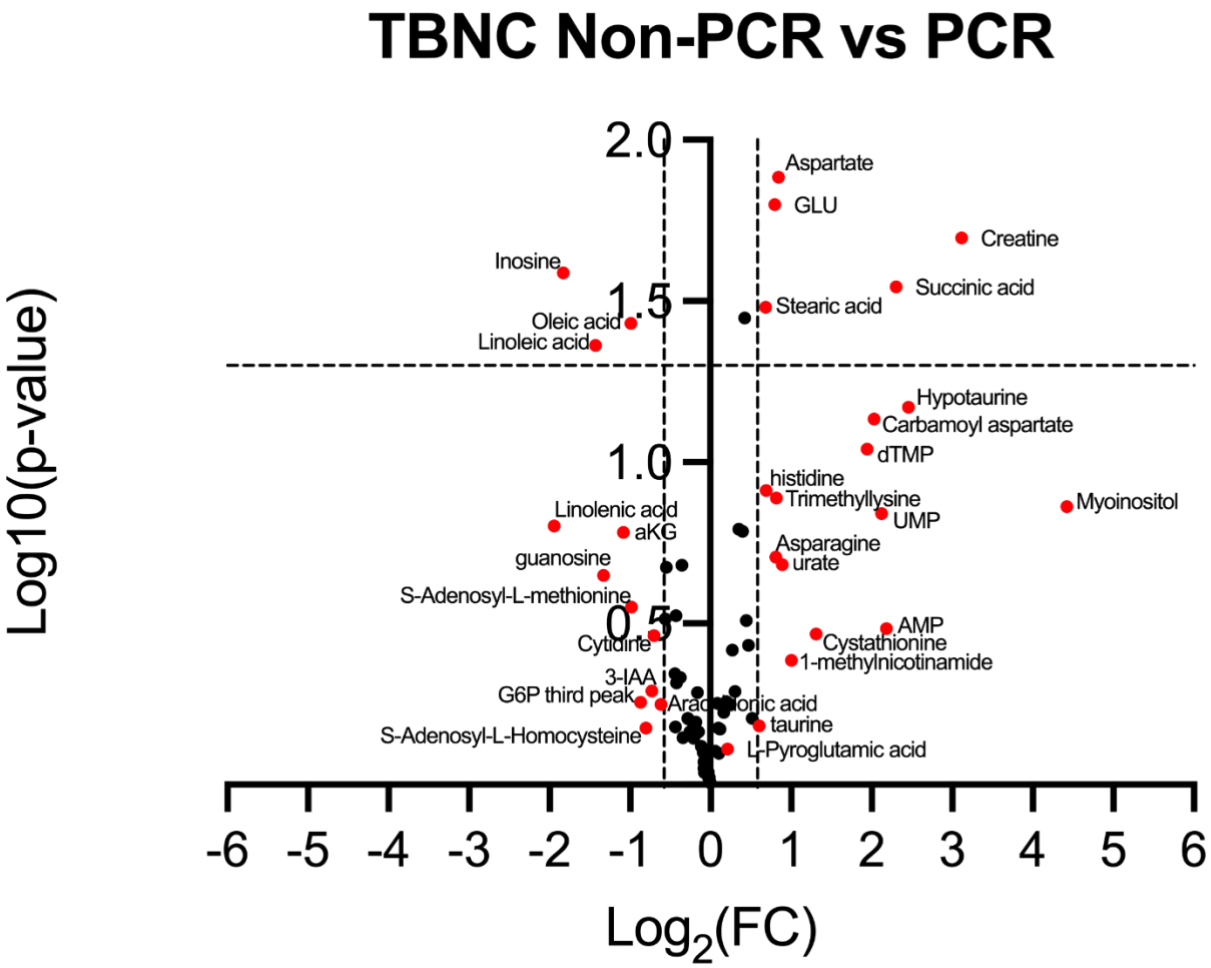


Figure 4.26. Volcano plot showing the relationship between metabolites p-values and fold changes among TNBC patients who did not achieve pathological complete response (Non-PCR) versus those who did achieve pCR (PCR). The effect size of the two groups is plotted on the x-axis on a logarithmic scale to base 2 ($Log_2(FC)$). The -log10 p-values are plotted on the y-axis (T-test).

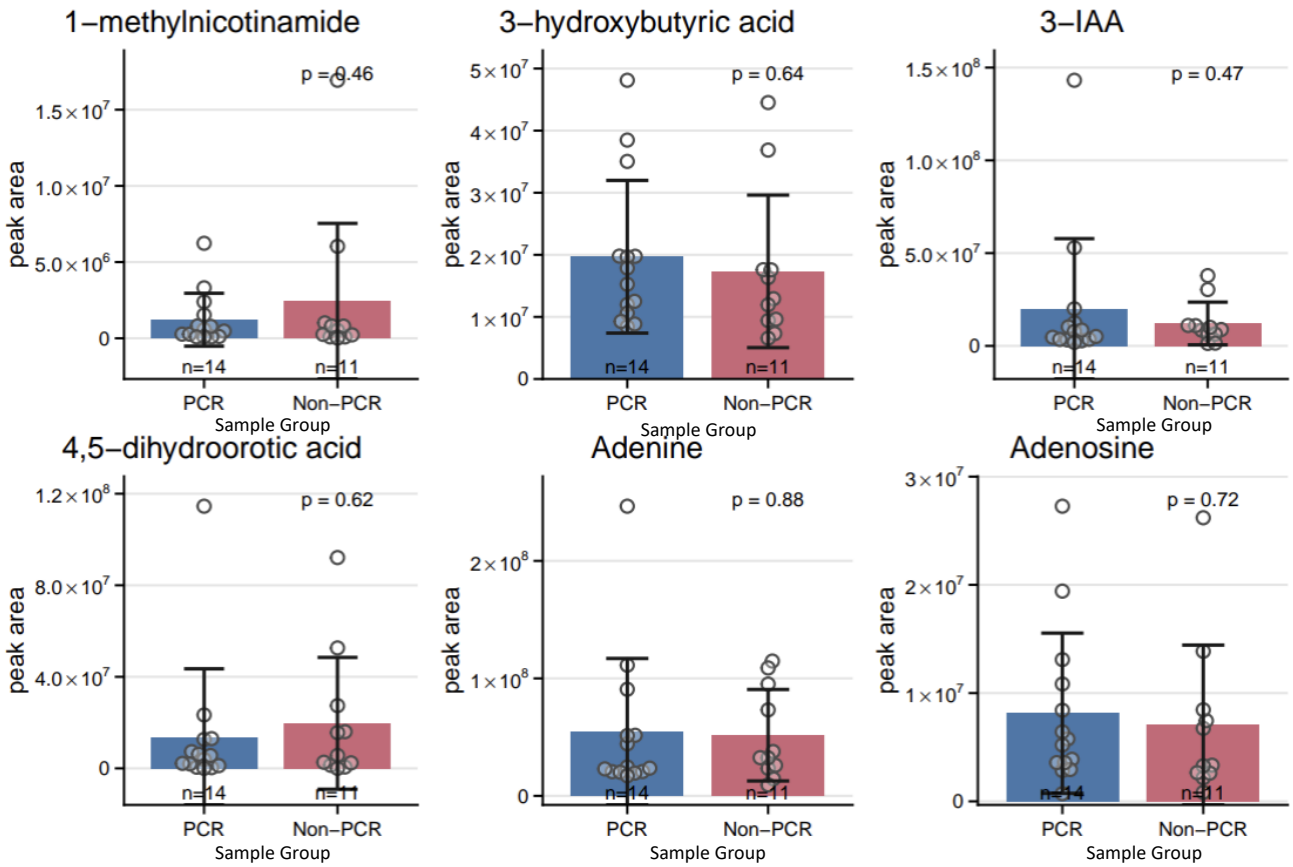


Figure 4.27. Boxplots demonstrating the peak areas of metabolites (1-methylnicotinamide, 3-hydroxybutyric acid, 3-IAA, 4,5-dihydroorotic acid, Adenine, Adenosine) that did not show statistical significance across TNBC breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=14) and those who did not (Non-PCR)(red, n=11). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

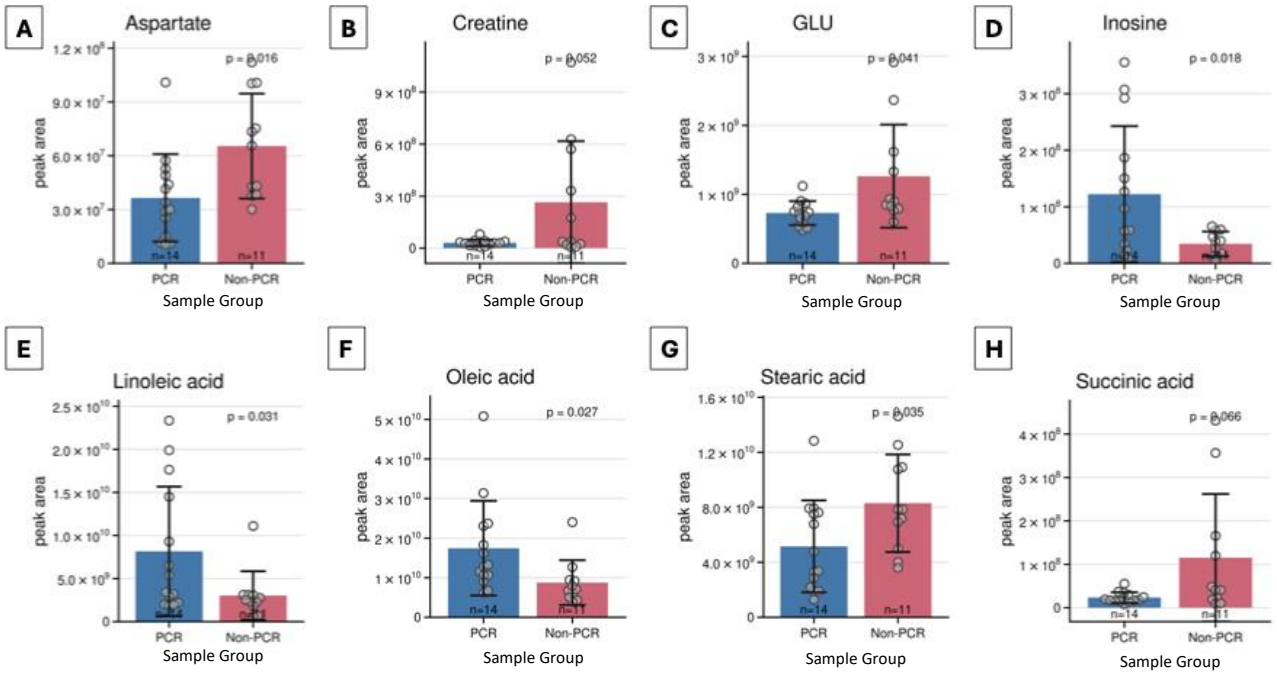


Figure 4.28. Boxplots demonstrating the peak areas of the five metabolites (Aspartate (A), Creatine (B), GLU (C), Inosine(D), Linoleic acid (E), Oleic acid (F), Stearic acid (G), Succinic acid (H)) that exhibited statistical significance across HER2+ER+ breast cancer patients from the NEO-MICROBE study at timepoint 1 that achieved pathological complete response(PCR)(blue, n=14) and those who did not (Non-PCR)(red, n=11). Individual data points represent peak area levels per individual. Significant statistical differences are displayed by the p-value (T-test).

4.3.7 Stool and Blood Metabolomics

Metabolomics analysis was conducted previous to this study on baseline blood samples collected as part of the NEO-MICROBE study. We first identified metabolites that stratified for response in stool and then carried out correlation analysis of these metabolites on the metabolomic data generated from the blood and stool of the same patients within the NEO-MICROBE study. This analysis was conducted to understand how response to neoadjuvant chemotherapy is influenced by systemic pathways that are modulated by translocation of gut-derived metabolites into the bloodstream. By identifying potential biomarkers of chemotherapy efficacy both in the blood and stool, this would allow for personalised therapy to reduce any systemic immunity driven by circulating microbial metabolites.

We analysed all of the metabolites of interest from the targeted analysis that demonstrated statistically significant variations in metabolomic profiles between responders and non-responders in the whole patient cohort and subtype specific response. We first looked at the correlation between the metabolomic results seen in blood versus the stool for the metabolites of interest (n=16) within the whole population at T1. As presented in Figure 4.29, there were no correlations demonstrated for any of the metabolites with most of the metabolites exhibiting r-values close to zero and unsignificant p-values. Two of the metabolites suggested a weak correlation between stool and blood metabolite levels; including succinic acid ($r = 0.0627$, $p = 0.0517$) and myristic acid ($r = 0.2564$, $p = 0.0461$) which exhibited a nearly statistically significant correlation (i.e. p-value < 0.05). These findings suggest that for these specific metabolites, stool metabolite levels aren't consistent with circulating metabolite levels across the whole population.

Glutamate was the only metabolite that demonstrated a correlation between the stool and blood metabolomics data. To understand whether gut-derived glutamate (GLU) levels are reflected systemically, we conducted correlation analysis between stool and blood glutamate levels in the whole breast cancer cohort and TNBC patients at T1. A statistically significant result was seen in TNBC patients, exhibiting a positive correlation in gut-derived GLU and circulating GLU levels (Figure 4.30). The p-value of 0.0134 demonstrates that the elevated levels of GLU seen in non-responders in TNBC patients (Figure 4.28) is moderately correlated ($r = 0.4879$) to higher levels of GLU in the blood of the same individuals. The scatter plot in Figure 4.30, highlights no statistically significant correlation ($r = -0.0591$, $p = 0.6508$) between gut and systemic GLU between responders and non-responders across the whole population. This suggests that the association is specific to TNBC subtype, which ultimately presents a potential biomarker for response to NACT in TNBC.

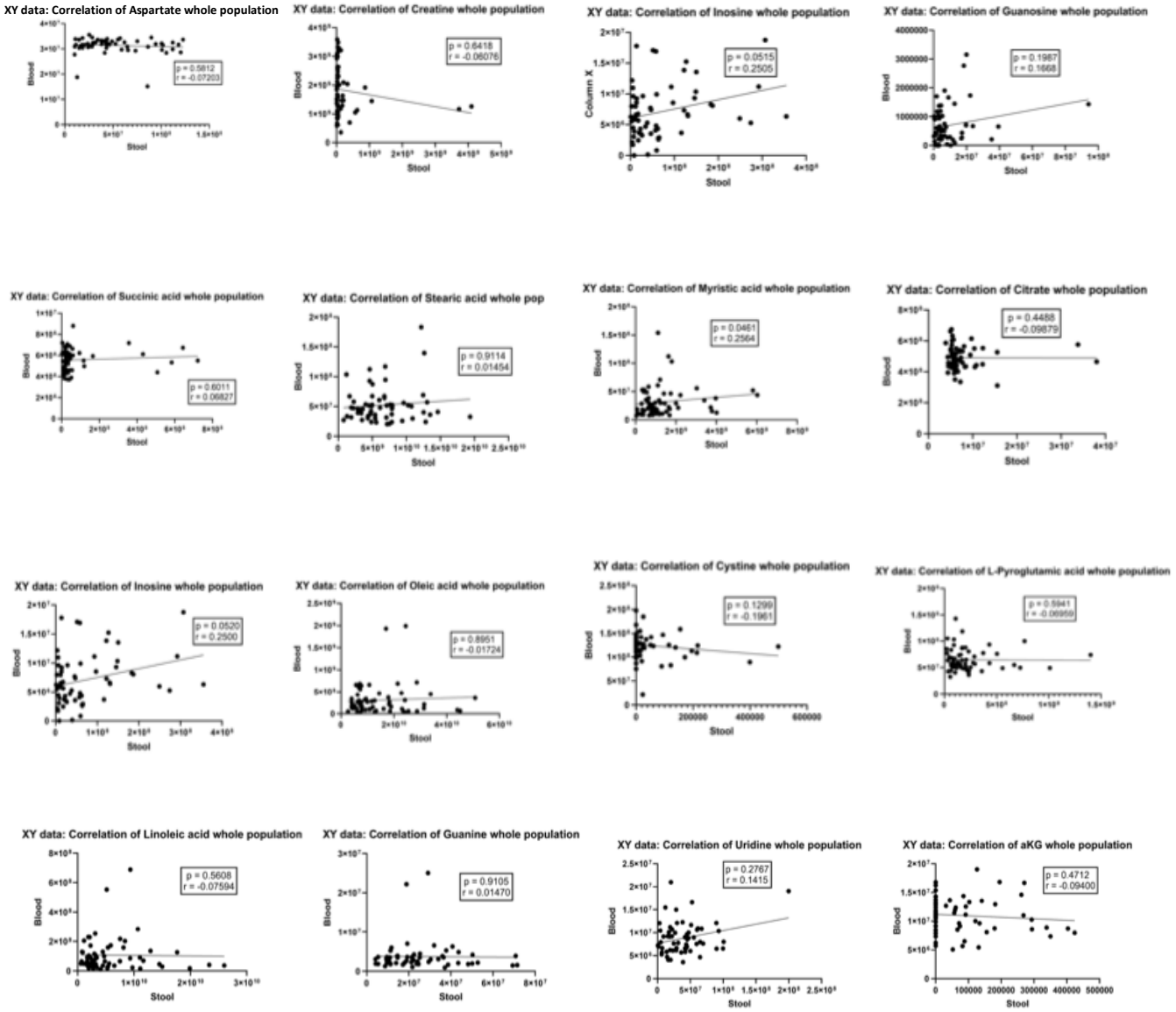


Figure 4.29. Correlation between stool and blood metabolite levels across the whole breast cancer population at T1. Scatter plots show spearman correlation analyses between stool and blood levels of metabolites of interest that were detected via targeted analysis and may be potentially implicated in treatment response. Each black point signifies a specific individual. The correlation coefficient (r) and p-values (p)(T-test) indicated the extent of correlation and the statistical significance between the two groups.

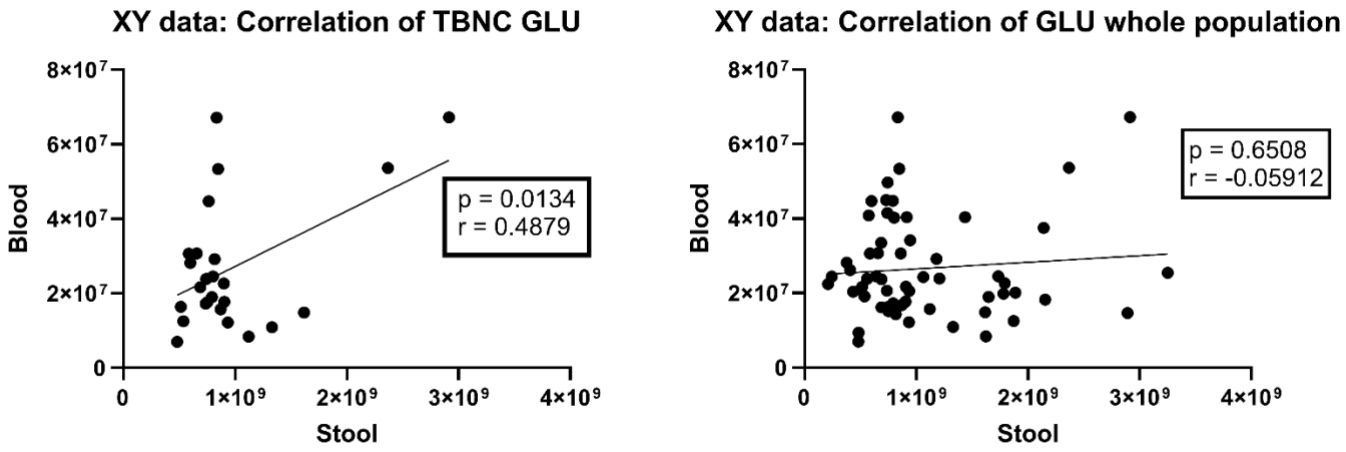


Figure 4.30. Correlation between stool and blood glutamate (GLU) levels in TNBC patients and across the whole breast cancer population. Scatter plots show spearman correlation analyses between stool-derived glutamate and circulatory glutamate in TNBC patient samples at timepoint 1 (Left) and in the whole breast cancer cohort (Right). The correlation co-efficient (r) and p-value (p)(T-test) is demonstrated on both graphs, with statistically correlation between stool and blood GLU seen in TNBC patients (p= 0.0134). Each black point represents an individual sample.

Discussion

Cancer cells adapt their metabolism to facilitate the uptake and utilization of various metabolic compounds such as amino acids to allow for survival and proliferation (78). To allow sustained proliferation, cancer cells rely on exogenous nutrient sources. Emerging evidence suggests that gut microbiome is a crucial source of metabolites for cancer cells to enable their altered metabolic processes (79). Furthermore, it has also been suggested that the gut microbiome not only be responsible for supplying cancer cells with nutrients but may contribute to modulating sensitivity and resistance to chemotherapy via metabolite signalling pathways (80).

Untargeted metabolomics revealed numerous metabolites that significantly associate with breast cancer, and with response to NACT. This highlights that gut-derived metabolites may serve as a biomarker for disease incidence and therapy selection and may be causally involved in these processes. We also investigated the targeted metabolomic profiles between HV and BC T1, with the aim of identifying metabolites that may be associated with cancer status. Unsurprisingly, we found no clear separation between the two sample groups upon principal component analysis (PCA), which suggests that in general there was little variation in metabolic profiles of cancer patients prior to treatment and healthy individuals. While tumorigenesis and metastasis are often linked to reprogramming of metabolic pathways (81), metabolic plasticity aims to maintain systemic metabolic homeostasis. Indeed, when we investigated the correlation between metabolites from the Beatson Cancer Institute database with known involvement in cancer and disease status, no statistically significant differences were observed for most metabolites. Of the 76 metabolites analysed four metabolites showed significant differences between healthy volunteers and breast cancer patients. The two fatty acids, linoleic acid and oleic acid highlighted metabolite-specific differences, with low abundance of both metabolites

recognised in the BC T1 group. Both of these fatty acids are involved in important cellular metabolic processes such as energy synthesis, activation of signalling pathways and cellular functions. Studies have shown that linoleic acid is involved in tumorigenesis due to it being metabolised into arachidonic acid, which undergoes enzymatic conversion resulting in eicosanoid formation. These arachidonic-acid-derived eicosanoids are linked to activation of signalling pathways promoting tumour growth, metastasis, and angiogenesis (82); additionally, it has been observed in several cases that arachidonic acid is involved in breast cancer proliferation and infiltration (83). As we found linoleic acid levels to be reduced in breast cancer patients, it would be of interest to study whether this goes together with an increase in arachidonic acid.

Oleic acid, which is abundantly found in animal and vegetable fats and oils (for example, olive oil), is a key metabolite in lipid homeostasis, with involvement in lipid synthesis and storage (84). Research into oleic acid and breast cancer has noted its anti-cancer properties such as improving response to oxidative stress, inhibition of tumour cell growth and migration. Further research has also shown oleic acids effectiveness in reducing cancer cell survival and progression, via inhibiting phosphorylation of Akt (85). Further research into these fatty acid's role in lipid metabolism within breast cancer is necessary, allowing us to gain an understanding of their potential involvement in lipid metabolism reprogramming, which is a known hallmark of cancer (86).

In contrast to the previous metabolites, Cystine exhibited higher levels in BC T1 patients. Cystine is a precursor to cysteine, an amino acid that is often upregulated in cancer cells due to its redox regulation abilities through glutathione. The latter serves as an important mechanism to protect tumour cells from oxidative damage allowing for survival and growth (87). This metabolic pathway has also been recognised as potential therapeutic target whereby inhibiting cystine uptake results in oxidative stress in tumour cells which can improve response to treatment (88). These studies are therefore in accordance with our

findings that cystine levels are more abundant in cancer patients (BC T1) suggest this is in support of metabolic reprogramming of cancer cells. Overall, further research is required to understand cystines behaviour at the subtype level as well as its potential role in disease formation.

Uridine monophosphate (UMP) demonstrated the most striking results with elevated levels in BC T1 patients compared to HV ($p = 0.00094$). Highlighted by the skewed distribution of the individual data points indicating that UMP levels were extremely low or non-existent in healthy volunteers whereas a specific subset of BC T1 patients exhibited higher peak areas of UMP across the cohort. This demonstrates a significant finding as it suggests UMP as a potential biomarker of breast cancer. Previous research has shown UMP as a key nucleotide in cancer cells metabolism. Tumour cells require increased RNA synthesis in order to maintain rapid cell proliferation; therefore, they undergo metabolic reprogramming to upregulate nucleotides including UMP to fulfil the cells metabolic requirements (89). Our findings align with results of previous studies, suggesting that UMP is a driver in cancer development via upregulated supporting altered nucleotide metabolism in cancer cells. We hypothesised that clinical factors may be contributing the highly significant differences in UMP observed in the cancer cohort however, no significant correlation was found between any of the clinical factors (menopausal status, age, and BMI). Further research is required to reveal the source and clinical relevance of UMP in this subset of BC patients.

This study hypothesised that gut-derived metabolites may influence response to neoadjuvant chemotherapy (NACT) in breast cancer patients, specifically achieving pathological complete response (pCR). We first explored stool metabolomic profiles across all breast cancer patients within the NEO-MICROBE clinical study in the context of pCR, before studying breast cancer subtype specific results. Our results both align with previous

literature on the gut metabolism's role in breast cancer, as well as reveal novel metabolites as potential biomarkers for treatment response.

Analysis of the whole cohort (n=61) revealed that there was some degree of variation in stool metabolomic profiles between patients who achieved pCR and non-pCR patients. This finding suggests that this metabolic divergence between the two groups could be involved in efficacy of NACT. A key intermediate metabolite of the Krebs cycle, α-Ketoglutarate (aKG) showed notably elevated levels in patients who achieved pCR. These higher levels of aKG associated with improved response to NACT demonstrated in this study is coherent with previous studies that have shown aKG to have correlations to greater chemosensitivity through enhancing anti-tumour immune responses or metabolic rewiring (90,91). The elevated levels of dTMP (deoxythymidine monophosphate) and guanine in non-responders suggests a potential mechanistic role in influencing anti-response to NACT. dTMP is a nucleotide that plays a central role in DNA synthesis and is produced by thymidylate synthase (TYMS). TYMS is often a chemotherapeutic target whereby its inhibition causes subsequent reduction in dTMP levels, resulting in a cytotoxic response to cancer cells (92). The increased dTMP levels in non-pCR patients demonstrated in this study, may be explained by the existing evidence that elevated levels of dTMP are linked to treatment resistance (93). Guanine, a purine nucleobase involved in DNA repair is strongly linked to cancer proliferation due to cancer cells high rates of DNA replication. Previous studies have shown that resistance to DNA-damaging chemotherapeutic agents may be influenced by shifts in purine metabolism in glioblastoma (94). These findings support the results seen in this study of greater guanine abundance in non-responders to NACT, which may be described by upregulation of DNA repair driving chemoresistance. Subtype-specific analysis was conducted to gain a more refined understanding of stool metabolites influence on response to NACT per subtype. Metabolomic profiling of HER2+

stool samples identified five metabolites with potential association to pCR status. TCA cycle intermediates such as aKG and citrate exhibited elevated levels in HER2+ patients who achieved pCR, suggesting a link between energy metabolism and improvement of treatment efficacy. This result is consistent with the findings observed in the whole population as well as HER2+ER+ patients who also demonstrated statistically significantly higher levels of citrate in patients who achieved pCR. It could also be suggested that purine metabolism is correlated to resistance to NACT as elevated levels of purine nucleosides such as guanosine were demonstrated in HER2+ patients, and a greater abundance of inosine, guanosine was observed in HER2+ER+ patients; all of whom did not achieve pCR. This finding was further emphasized in HER2+ER- patients showing statistically significant elevated guanine, uridine, cytosine, and cytidine levels in non-responders. Statistically significantly elevated levels of glutamate (GLU) was noted in HER2+ER- and TNBC patients who achieved pCR, this may subsequent of GLU ability to modulate oxidative stress which supports tumour survival and proliferation, therefore suggesting GLU as a potential biomarker for breast cancer and its involvement in antioxidant protection could be a target for improving NACT sensitivity (95). The reoccurrence of the same metabolites across multiple subtypes highlights these as potential metabolic biomarkers of response to NACT in breast cancer, however further studies would be required to determine these as definitive hallmarks of NACT sensitivity.

The importance of precise validation methods was highlighted by metabolites such as cystine in HER2+ER+ and succinic acid in TNBC that demonstrated initial statistical significance in the volcano plot analysis however upon deeper inspection were deemed insignificant.

Overall, our findings demonstrate stool-derived metabolites as potential indicators of NACT response in breast cancer patients; with aKG and citrate presenting promising

results as biomarkers for positive response to NACT, and resistance may be associated with purine nucleotides. In order to validate these findings, subsequent research using larger patients cohorts would be required to determine the predictive ability of these gut metabolites to tumour response.

Within the field of cancer metabolomics, the relationship between gut-derived metabolites and systemic circulation is an established method to study associations between gut microbiota composition and disease. In this study, we aimed to understand this relationship in the context of treatment response. We observed a statistically significant positive correlation ($r = 0.4879$, $p = 0.0134$) between gut-derived glutamate (GLU) and circulatory GLU in TNBC patients at T1. Contrastingly, when observing correlation of GLU levels across the whole breast cancer patient cohort no significant correlation was seen therefore highlighting a subtype-specific metabolic link between gut-microbial metabolites and systemic metabolism as a potential biomarker for treatment response to NACT in TNBC.

TNBC is known for its aggressive clinical behaviour and distinct metabolic dependencies in order to maintain rapid proliferation and redox regulation via enhanced dependence on pathways such as glutaminolysis and other anaplerotic pathways. These have significant roles within the tricarboxylic acid (TCA) cycle to support energy production, rapid cell division and survival. Glutamate (GLU) is a key intermediate of the TCA cycle, whereby GLU is converted to aKG prior to entry of the TCA cycle (96). GLU also plays a vital role in redox regulation in cancer cells via glutathione synthesis, which may influence positive response to NACT by supporting oxidative stress (97). The rewiring of TNBCs metabolic system is a recognised hallmark of cancer, allowing for absorption of compounds more efficiently to support this aggressive cancers enhanced metabolic demands. This could explain why higher gut-derived GLU is correlated with increases in circulatory GLU, as it

is more easily transported into circulation in TNBC due to the altered absorption to further facilitate TNBC upregulated metabolism compared to other subtypes.

No significant correlation in GLU levels was identified across the whole population, which may be due to different metabolic pathway reliance in other subtypes that don't utilise GLU the same as TNBC (98). Furthermore, dietary influences were not taken into consideration which can lead to individual variability in gut microbiota composition and influence GLU levels.

In conclusion, our findings present evidence that elevated glutamate levels in stool are moderately correlated to higher circulating levels, thus driving non-response in TNBC patients however this is not corroborated in the broader breast cancer population within the study. This suggests that glutamate may have predictive value specific to TNBC, highlighting the importance of molecular subtype stratification when evaluating stool metabolites as potential biomarkers of NACT response.

5. Flow Cytometry Panel Validation

5.1 Antibody Titrations

A novel 12-colour panel (table 1) for human T cell markers was designed for routine use in the NEO-MICROBE BREAST trial based on immune cell populations that have been previously shown in literature to have relevance in breast cancer development, as well as response to therapy . This 12-colour panel was then assessed to determine optimal concentration of each marker-specific antibody. Optimal concentration can be defined as the best separation between positive and negative populations and a metric known as the Separation Index (SI) was used to evaluate staining results. Determining the optimal concentration will also minimise oversaturating populations which will increase non-specific binding, creating background and reducing the resolution of results.

	Marker	Fluorophore
1	Viability	eFluor780
2	CD45	BB515
3	CD45RA	BV510
4	CD3	BV785
5	CD4	PE-Cy7
6	CD8	BV421
7	CD25	APC
8	CCR7	PerCP-Cy5.5
9	FoxP3	PE
10	CTLA-4	AF700
11	PD1	BUV395
12	HLA-DR	BV605

Table 4. 12 colour panel for Flow Cytometry analysis. This table contains the markers selected for flow cytometry analysis on the PBMC patient samples from the NEO-MICROBE study. The fluorophores for each marker are also included.

Each antibody was titrated individually on PBMCs from a healthy individual using the same protocol established for flow cytometry analysis of NEO-MICROBE BC patient PBMCs, to maintain reproducibility. Human PBMCs were stained with each antibody at increasing dilution and an unstained sample was also included in each titration (figure 5.1). Samples were analysed on a BD LSR Fortessa using FACSDiva software version. The acquired data was analysed using FlowJo software version 10.10.0. Within each group, doublets and dead cells were excluded, gating only for lymphocytes and expression of each marker was assessed on this population. The SI (Median positive – Median negative/ (84%Negative – MedianNegative)/0.995) was calculated for each dilution per sample.

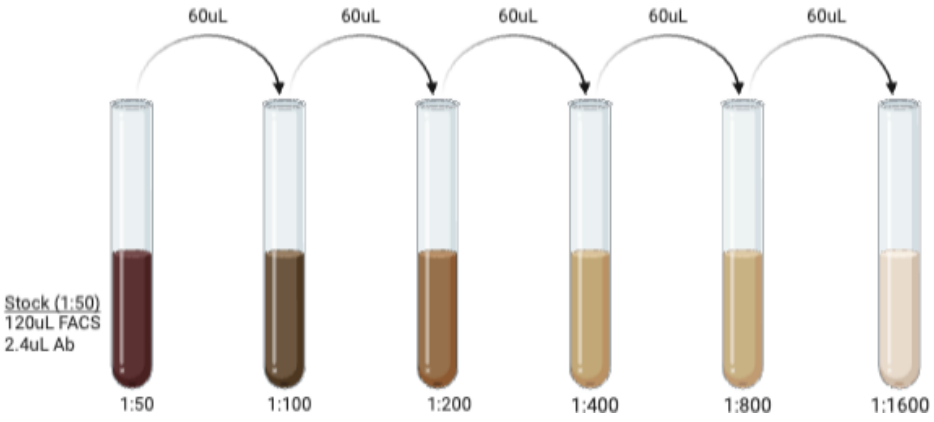


Figure 5.1. Serial dilution for antibody titrations. Figure 5.1 is a schematic of the dilution carried out for each antibody titration to determine optimal concentration. A stock dilution of 120ul FACS buffer and 2.4ul antibody (Ab) was used.

Figure 5.2 represents the concatenation of the datafiles from each antibody titration, with the optimal concentration highlighted in red. The dilution with the highest the separation index (SI) indicates optimal concentration as this shows the greatest separation between the negative and positive cell populations, thus giving the clearest result for positive staining signal and ultimately increasing robustness of results. Overall, these results will ensure optimal results are obtained when carrying out flow cytometry analysis on the 12-colour panel, while also ensuring cells are not being oversaturated with antibody which may produce ambiguous results. Figure 4 highlights that after multiple titration experiments of the antibodies for FOXP3 (PE), CD25 (APC) and CTLA-4 (AF700) there was little to no positive signal detected at all concentrations, suggesting these antibodies are not suitable.

5.2 Immunophenotyping PBMCs

After running the 12-colour panel (Table 4) using healthy blood several times, it has been identified that some antibodies are not being detected by the fortessa for various reasons. This is highlighted in figure 4 whereby several antibodies showed very low positive signals even at higher concentrations. From these results, it was concluded that changes to the selected antibodies must be made in order to detect the markers of interest (Table 5). The panel has also changed to a ten-colour panel, losing PD-1 and HLA-DR as it was not possible to optimise the panel with the inclusion of these markers within the available time frame

	Marker	Fluorophore
1	Viability	eFluor780
2	CD45	Spark Blue 574
3	CD45RA	BUV395
4	CD3	BV510
5	CD4	BUV805
6	CD8	Pacific Blue
7	CD25	APC
8	CCR7	PE
9	FoxP3	BB700
10	CTLA-4	AF700

Table 5. 10 colour panel for flow cytometry analysis. Table 3 shows the markers selected for flow cytomtery analysis and the new fluorophores selected.

5.3 Flow Cytometry analysis for T cell panel validation

Flow cytometry analysis was conducted to characterise T cell populations within peripheral blood mononuclear cells (PBMCs) from healthy volunteers. Blood samples collected from healthy volunteers was used for validation of the panel for future use on PBMC samples collected previously as part of the NEO-MICROBE study. Figure 5.3 demonstrates the gating strategy utilised to characterise the cells of interest. Lymphocytes were first gated based on forward and side scatter, and doublets were excluded to identify single cells. Cells were stained using eFluor780 to isolate live cells (76.1%) from dead cells (22%). Next, we gated on live cells for CD3+CD45+ T cells (21.6%), as CD3 is a pan-T cell marker and CD45 is expressed on all hematopoietic cells which we are interested in.

Within the total T cell population, we distinguished populations into CD4+ T cells (69.5%) and CD8+ T cells (28.1%). Of the CD4+ T cell population, 4.82% were Tregs (CD25+FoxP3+) and non-Tregs accounted for 93.8%. Among CD4+ non-Tregs, 10.3% showed a memory cell phenotype ($CD45RA^{-}CCR7^{+}$), and naïve cells accounted for 89.1% ($CD45RA^{+}CCR7^{+}$). CD4+ showed a distribution of naïve and memory cells of 58.8% and 41.2%, respectively.

Among the CD8+ T cell population, 11.2% were regulatory T cells (Tregs; CD25+FoxP3+) and non-Tregs accounted for 84.6% of the population. Further analysis divided the CD8+ non-Tregs into three subsequent populations; naïve cells (83.5%; $CD45RA^{+}CCR7^{+}$), memory cells (10.7%; $CD45RA^{-}CCR7^{+}$) and TEMRA cells (5.79%; $CD45RA^{+}CCR7^{-}$). TEMRA cells are a subset of T effector memory cells that re-express CD45RA+ and are more commonly found in CD8+ populations rather than CD4+. Of the CD8+ Tregs population, 62.5% were naïve and 31.2% were memory. The CTLA-4 marker exhibited very low expression so was not included in the analysis.

Overall, this panel and gating strategy will allow for a comprehensive understanding of T cell phenotypes among PBMCs of cancer patients within the study.

5.3.1 Fluorescence Minus One (FMO) Controls for T cell panel validation

Fluorescence Minus One (FMO) controls (Figure 5.4) were implemented to confirm the antibody is specific to its target marker and there is no overlap in the 10-colour panel, as well as to validate gating strategies. The unstained control highlighted there was no background fluorescence which would skew the results and confirmed no positive dead cells (97% live cells) as expected from the unstained sample. The gating for T cells (CD3+CD45+) was verified in the CD45 and CD3 FMOs, as there were no cells present within the T cells gate. The CD4 FMO validated the gating for T cell subset showing no CD4+ population (0%), however unexpectedly CD8+ cells were not excluded in the CD8 FMO which highlights an error. FMOs for CD25 and FOXP3 confirmed the accuracy of Treg and non-Treg identification, with 0% Tregs in both FMOs which confirms any Treg cells observed in the sample is correspondent of positive CD25 or FOXP3 status. Furthermore, naïve and memory gating strategies were validated using CD45RA and CCR7 FMOs, whereby little to no spillover was demonstrated.

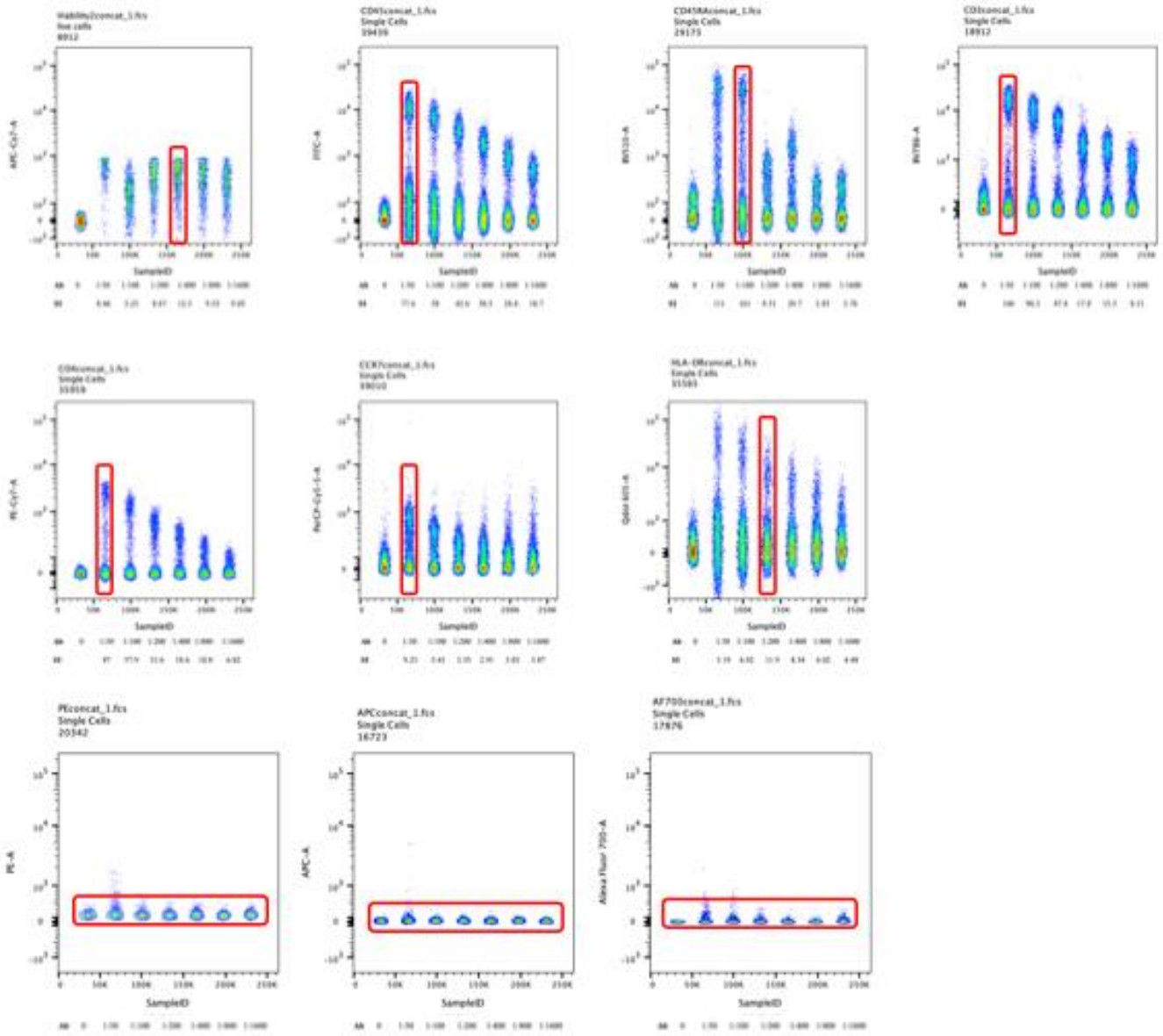


Figure 5.2. Flow cytometry dot plots of antibody titrations. Figure 5.2 shows the dot plots of each titration concentration for each antibody. The population above 10^3 is the positive population and those below 10^3 are negative. The concentration with the highest separation index is highlighted in a red box (the bottom row indicates the only populations detected were negative (below 10^3) demonstrated by the red box. Three antibodies(PE, APC, AF700) showed no seperation (can be seen on bottom row), as no positivation population were detected therefore demonstrating a handling error or that these fluorophores are not suitable for detecting these T cell populations.

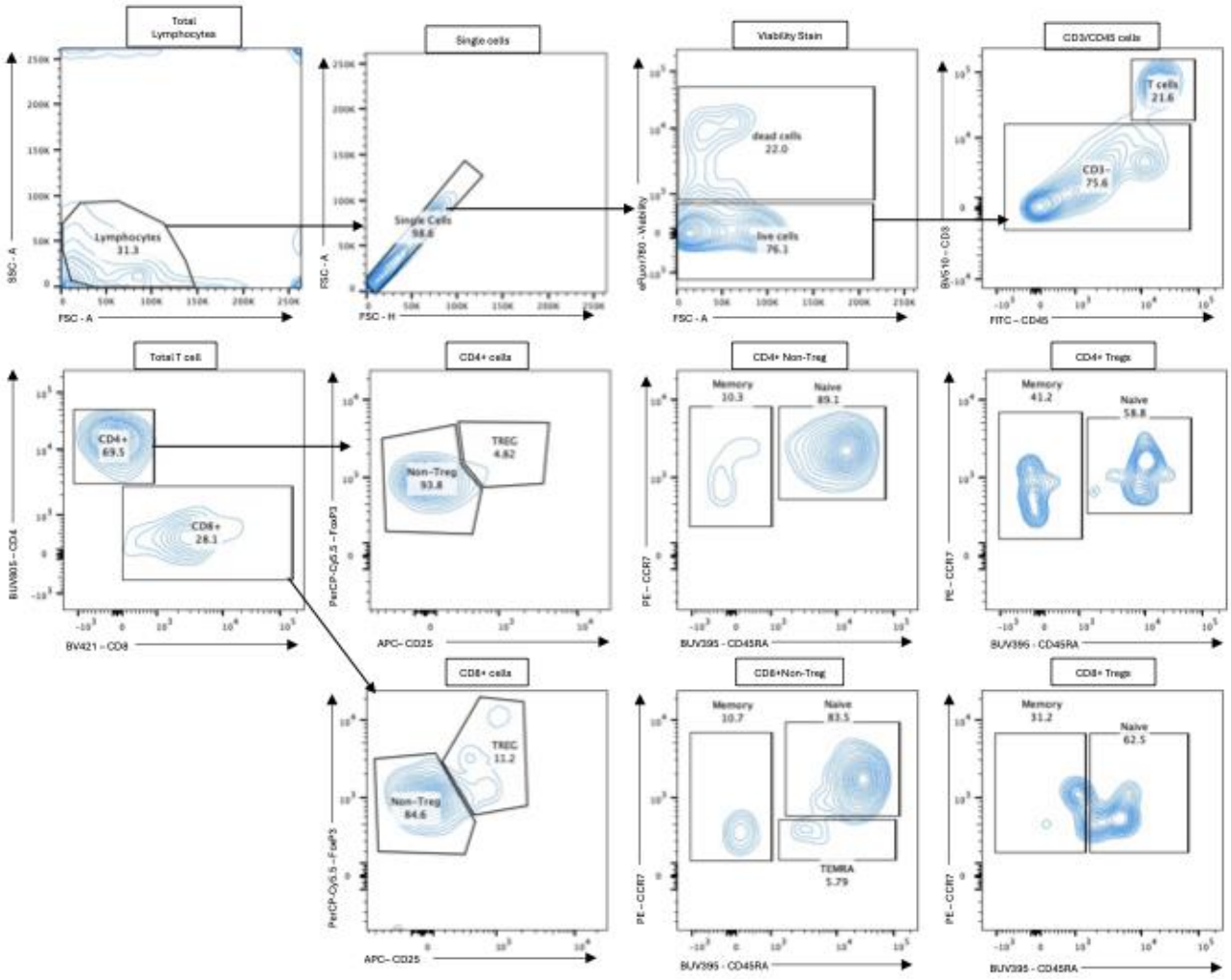


Figure 5.3 Flow Cytometry Analysis gating strategy to identify populations of T cell subsets. Flow cytometry analysis was conducted on PBMCs from healthy blood donors to validate the 10-colour panel for identification of T cell populations. Gating was carried out on full stained lymphocytes.

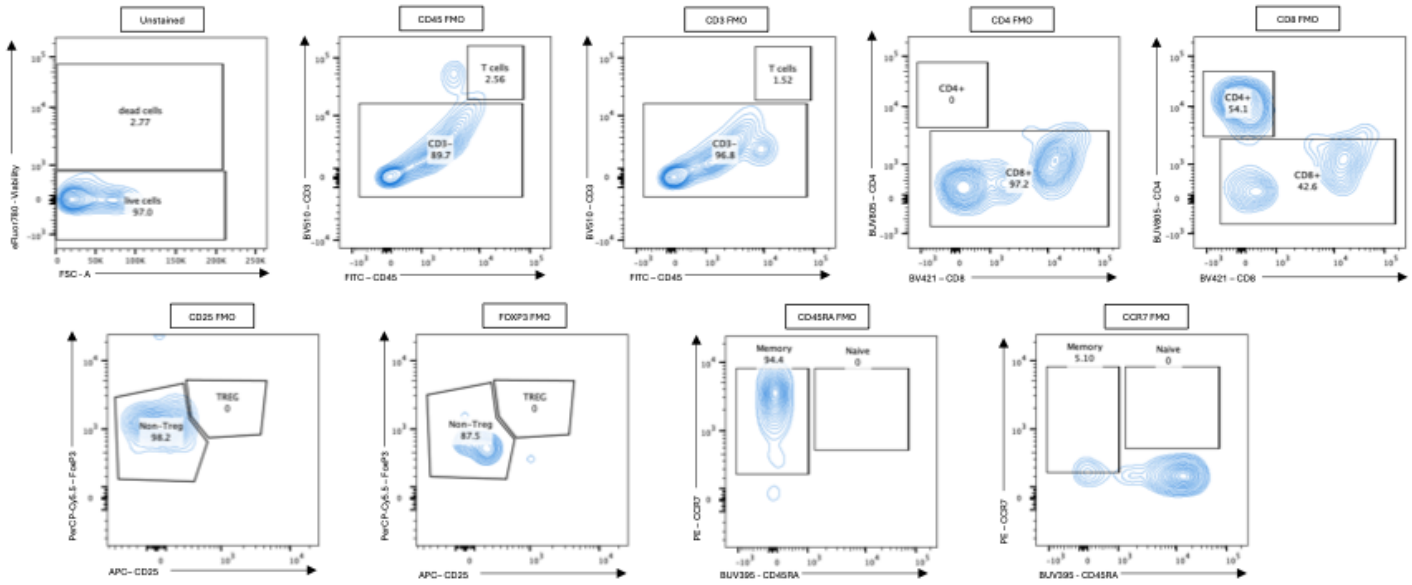


Figure 5.4. Fluorescence Minus One (FMO) controls for validation of T cell gating strategy and panel design. The contour plots show the FMO controls that allow for validation of the gating strategy applied to full stained samples and marker-specific fluorescence to identify the correct T cell population.

Discussion

In this study, a 12-colour panel was initially developed for flow cytometry analysis, antibodies were titrated for identification of optimal concentration before subsequently refining the panel to 10 colours while still allowing for optimal phenotyping of T cells within PBMCs from healthy blood donors. The objective was to generate and validate a robust panel capable of immunophenotyping PBMCs from breast cancer patients in the future to gain an understanding of the immune systems role in NACT response.

With the aim of determining optimal staining concentration for each specific marker, antibody titrations were conducted to ensure clear separation between each positively and negatively stained cell population. Figure 5.2 highlights that for the markers FOXP3 (PE), CD25 (APC) and CTLA-4 (AF700), there was no clear discrimination between positive and negative population; demonstrating no positive signal present at all. This finding suggests that these antibodies were not suitable to detect these markers therefore, it was necessary to revise the panel, and a decision was made to shift to a different 10-colour panel (Table 5) with the hopes of optimal detection of the markers of interest.

Healthy donor blood was used as a validation model of the revised 10-colour panel; whereby PBMCs were extracted and stained with their corresponding antibody. As demonstrated in Figure 5.3, the panel successfully identified T cells subsets including CD3+CD45+, CD4+ and CD8+ T cells, Tregs and non-Tregs, as well as subsequent classification of the Tregs populations into naïve, memory and TEMRA subsets. The validation of this panel in detection subsets such as TEMRA population is specifically beneficial in breast cancer immunology as CD8+TEMRA cells (CD45RA+CCR7-) have been associated with tumour infiltration and immune checkpoint therapy resistance due to altered effector functions (99). The use of FMO controls (Figure 5.40) examined the robustness of gating accuracy as well as ensuring correct identification of positive populations (100). Overall, the FMOs minimal signal detection of their respective

fluorophore, excluding the CD8 FMO which still demonstrated CD8+ cell population thus signifying a compensation error that determines the need for further validation. Additionally, there was very low signal observed for CTLA-4 which may be explained by the low expression levels typical of healthy cells therefore in vitro stimulation protocols may be required in the future to allow for a reliable signal to be detected (101). Furthermore, AF700 used to detect CTLA-4 is considered a fairly dim fluorophore therefore this in combination with a low expression of CTLA-4 could cause the low signals observed (102). An alternative fluorophore could be used to combat the lack of CTLA-4 signal.

In conclusion, the validation of the 10-colour panel facilitates future analysis of PBMCs from the NEO-MICROBE study to immunophenotype breast cancer patients and understand the influence on treatment efficacy. This study provides a solid foundation for future applications of the panel however some further optimisation and validation is required before utilisation in breast cancer immunology research.

6. Concluding remarks

This study aimed to investigate the interplay between the collective influence of the gut metabolome and the immune system on response to neoadjuvant chemotherapy (NACT) in early breast cancer, through analysis of samples from the NEO-MICROBE BREAST study. Combining insights from the metabolomic profiling of the stool and blood samples along with immunophenotyping of PBMCs in the future utilising the validated panel from this study, has potential to define metabolic and immune signatures involved in driving cancer status and influencing the efficacy of NACT.

We mainly focused on identifying stool-derived metabolites that showed potential as biomarkers of NACT response by conducting untargeted and targeted analysis on LCMS stool metabolomic data. Notably, several metabolites demonstrated significant association with achievement of pCR to NACT; including α-ketoglutarate (aKG), citrate, glutamate (GLU) and uridine monophosphate (UMP). The elevated levels of UMP observed in cancer patients compared to healthy individuals implies a potential role as a biomarker of cancer. It was suggested that energy metabolism may be linked to response to NACT in breast cancer patients due to elevated levels of the TCA cycle intermediates, aKG and citrate in patients who achieved pCR. On the other hand, enrichment of the metabolite's guanine and uridine in non-responders (non-pCR), and their involvement in purine metabolism suggests a potential link to purine metabolic pathways and resistance to NACT. Subtype-specific correlation analysis highlighted differential glutamate signatures in TNBC patients who achieved pCR, however this was not recognised across the whole cohort. This finding indicates that glutamate may be a valuable predictor of response to chemotherapy, specifically in TNBC patients.

Furthermore, with some future validation to ensure optimal antibody detection; the 10-colour panel flow cytometry generated in this study provides a robust foundation for future work to immunophenotype PBMCs of NEO-MICROBE patients, with the hope of

identifying immune profiles potentially modulated by gut metabolites recognised as significant to response in this study. Despite some limitations in detecting the antibodies for FOXP3 and CTLA-4, the panel successfully identified crucial T cell subsets related to breast cancer.

In order to generate more reliable and reproducible results a larger, distinct cohort is essential to validate the identified metabolic biomarkers and their influence on response. Additionally, a limiting factor in the flow cytometry analysis was reduced cell count which subsequently made marker detection more difficult therefore future validation work of the flow cytometry should use larger number of PBMCs when staining. It would also be beneficial to incorporate detailed dietary data for each patient in future studies to understand to what extent the metabolomic profiles of individuals are affected by dietary intake which could also reveal potential modifiable lifestyle factors that could enhance an individual's response to NACT. Despite the overall success of the project, it was not possible to complete immunophenotyping of the NEO-Microbe PBMC samples due to the time taken to design and validate a bespoke panel. In addition we were unable to include all of our planned markers due to technical challenges such as spectral overlap between fluorophores.

In conclusion, this study provides evidence in favour of the hypothesis that gut-derived metabolites and their systemic interplay are associated with modulation of response to NACT in breast cancer. Ultimately, understanding the collective influence of these identified metabolites alongside immune response provides insight into the mechanisms behind achieving pathological response, allowing for more effective, personalised treatment for early-stage breast cancer.

7. Bibliography

1. Worldwide cancer data | World Cancer Research Fund International [Internet]. WCRF International. [cited 2023 Nov 11]. Available from: https://www.wcrf.org/cancer-trends/worldwide-cancer-data/

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209–49.

3. Cancer Research UK [Internet]. 2015 [cited 2024 Jan 20]. Breast cancer statistics. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer

4. Orrantia-Borunda E, Anchondo-Nuñez P, Acuña-Aguilar LE, Gómez-Valles FO, Ramírez-Valdespino CA. Subtypes of Breast Cancer. In: Mayrovitz HN, editor. Breast Cancer [Internet]. Brisbane (AU): Exon Publications; 2022 [cited 2025 Feb 11]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK583808/

5. Nandi D, Parida S, Sharma D. The gut microbiota in breast cancer development and treatment: The good, the bad, and the useful! Gut Microbes. 2023 Dec 31;15(1):2221452.

6. Breast cancer [Internet]. [cited 2025 Feb 14]. Available from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

7. Duraiyan J, Govindarajan R, Kaliyappan K, Palanisamy M. Applications of immunohistochemistry. J Pharm Bioallied Sci. 2012;4(6):307.

8. Zaha DC. Significance of immunohistochemistry in breast cancer. World J Clin Oncol. 2014;5(3):382.

9. Zhang MH, Man HT, Zhao XD, Dong N, Ma SL. Estrogen receptor-positive breast cancer molecular signatures and therapeutic potentials (Review). Biomed Rep. 2014 Jan 1;2(1):41–52.

10. Nicolini A, Ferrari P, Duffy MJ. Prognostic and predictive biomarkers in breast cancer: Past, present and future. Semin Cancer Biol. 2018 Oct 1;52:56–73.

11. Hicks DG, Lester SC. Hormone Receptors (ER/PR). In: Hicks DG, Lester SC, editors. Diagnostic Pathology: Breast (Second Edition) [Internet]. Philadelphia: Elsevier; 2016 [cited 2025 Feb 11]. p. 430–9. (Diagnostic Pathology). Available from: https://www.sciencedirect.com/science/article/pii/B9780323377126500673

12. Distinct temporal trends in breast cancer incidence from 1997 to 2016 by molecular subtypes: a population-based study of Scottish cancer registry data | British Journal of Cancer [Internet]. [cited 2025 Apr 23]. Available from: https://www.nature.com/articles/s41416-020-0938-z

13. Boland MR, Ryan ÉJ, Dunne E, Aherne TM, Bhatt NR, Lowery AJ. Meta-analysis of the impact of progesterone receptor status on oncological outcomes in oestrogen receptor-positive breast cancer. Br J Surg. 2020 Jan;107(1):33–43.

14. Agostinetto E, Curigliano G, Piccart M. Emerging treatments in HER2-positive advanced breast cancer: Keep raising the bar. Cell Rep Med. 2024 June 18;5(6):101575.

15. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. Nat Rev Dis Primer. 2019 Sept 23;5(1):1–31.

16. A Review of Triple-Negative Breast Cancer - Roohi Ismail-Khan, Marilyn M. Bui, 2010 [Internet]. [cited 2025 Apr 30]. Available from: https://journals.sagepub.com/doi/10.1177/107327481001700305?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed

17. Triple-negative Breast Cancer | Details, Diagnosis, and Signs [Internet]. [cited 2025 Feb 14]. Available from: https://www.cancer.org/cancer/types/breast-cancer/about/types-of-breast-cancer/triple-negative.html

18. Medina MA, Oza G, Sharma A, Arriaga LG, Hernández Hernández JM, Rotello VM, et al. Triple-Negative Breast Cancer: A Review of Conventional and Advanced Therapeutic Strategies. Int J Environ Res Public Health. 2020 Mar;17(6):2078.

19. Medina MA, Oza G, Sharma A, Arriaga LG, Hernández Hernández JM, Rotello VM, et al. Triple-Negative Breast Cancer: A Review of Conventional and Advanced Therapeutic Strategies. Int J Environ Res Public Health. 2020 Mar;17(6):2078.

20. PARP inhibitors [Internet]. [cited 2025 Apr 30]. Available from: https://www.cancerresearchuk.org/about-cancer/treatment/targeted-cancer-drugs-immunotherapy/parp-inhibitors

21. Polanco TO, Moo TA, Nelson JA, Tokita HK. 23 - Perioperative Care of the Cancer Patient: Breast Procedures. In: Hagberg C, Gottumukkala V, Riedel B, Nates J, Buggy D, editors. Perioperative Care of the Cancer Patient [Internet]. New Delhi: Elsevier; 2023 [cited 2025 Feb 17]. p. 262–81. Available from: https://www.sciencedirect.com/science/article/pii/B9780323695848000232

22. Cain H, Macpherson IR, Beresford M, Pinder SE, Pong J, Dixon JM. Neoadjuvant Therapy in Early Breast Cancer: Treatment Considerations and Common Debates in Practice. Clin Oncol. 2017 Oct 1;29(10):642–52.

23. Janeway KA, Gorlick R, Bernstein ML. 22 - Osteosarcoma. In: Orkin SH, Fisher DE, Look AT, Lux SE, Ginsburg D, Nathan DG, editors. Oncology of Infancy and Childhood [Internet]. Philadelphia: W.B. Saunders; 2009 [cited 2025 Feb 17]. p. 871–910. Available from: https://www.sciencedirect.com/science/article/pii/B9781416034315000224

24. Laot L, Laas E, Girard N, Dumas E, Daoud E, Grandal B, et al. The Prognostic Value of Lymph Node Involvement after Neoadjuvant Chemotherapy Is Different among Breast Cancer Subtypes. Cancers. 2021 Jan;13(2):171.

25. Cain H, Macpherson IR, Beresford M, Pinder SE, Pong J, Dixon JM. Neoadjuvant Therapy in Early Breast Cancer: Treatment Considerations and Common Debates in Practice. Clin Oncol. 2017 Oct 1;29(10):642–52.

26. Spring LM, Bar Y, Isakoff SJ. The Evolving Role of Neoadjuvant Therapy for Operable Breast Cancer. J Natl Compr Canc Netw. 2022 June 1;20(6):723–34.

27. Tabayoyong W, Li R, Gao J, Kamat A. Optimal Timing of Chemotherapy and Surgery in Patients with Muscle-Invasive Bladder Cancer and Upper Urinary Tract Urothelial Carcinoma. Urol Clin North Am. 2018 May 1;45(2):155–67.

28. Spring LM, Fell G, Arfe A, Sharma C, Greenup R, Reynolds KL, et al. Pathologic Complete Response after Neoadjuvant Chemotherapy and Impact on Breast Cancer Recurrence and Survival: A Comprehensive Meta-analysis. Clin Cancer Res. 2020 June 15;26(12):2838–48.

29. IJMS | Free Full-Text | Triple-Negative Breast Cancer and Predictive Markers of Response to Neoadjuvant Chemotherapy: A Systematic Review [Internet]. [cited 2023 Dec 2]. Available from: https://www.mdpi.com/1422-0067/24/3/2969

30. Zhao Y, Schaafsma E, Cheng C. Gene signature-based prediction of triple-negative breast cancer patient response to Neoadjuvant chemotherapy. Cancer Med. 2020 Sept;9(17):6281–95.

31. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. The Lancet. 2014 July 12;384(9938):164–72.

32. Valdés-Ferrada J, Muñoz-Durango N, Pérez-Sepulveda A, Muñiz S, Coronado-Arrázola I, Acevedo F, et al. Peripheral Blood Classical Monocytes and Plasma Interleukin 10 Are Associated to Neoadjuvant Chemotherapy Response in Breast Cancer Patients. Front Immunol. 2020 July 9;11:1413.

33. Basu A, Ramamoorthi G, Jia Y, Faughn J, Wiener D, Awshah S, et al. Chapter Six - Immunotherapy in breast cancer: Current status and future directions. In: Wang XY, Fisher PB, editors. Advances in Cancer Research [Internet]. Academic Press; 2019 [cited 2024 Jan 19]. p. 295–349. (Immunotherapy of Cancer; vol. 143). Available from: https://www.sciencedirect.com/science/article/pii/S0065230X19300223

34. Asaoka M, Gandhi S, Ishikawa T, Takabe K. Neoadjuvant Chemotherapy for Breast Cancer: Past, Present, and Future. Breast Cancer Basic Clin Res. 2020 Jan 1;14:1178223420980377.

35. Choosing between breast conserving surgery (lumpectomy) or mastectomy [Internet]. [cited 2025 Feb 14]. Available from: https://www.cancerresearchuk.org/about-cancer/breast-cancer/treatment/surgery/lumpectomy-or-mastectomy

36. Kunnuru SKR, Thiyagarajan M, Martin Daniel J, Singh K B. A Study on Clinical and Pathological Responses to Neoadjuvant Chemotherapy in Breast Carcinoma. Breast Cancer Targets Ther. 2020 Nov;Volume 12:259–66.

37. Haque W, Verma V, Hatch S, Suzanne Klimberg V, Brian Butler E, Teh BS. Response rates and pathologic complete response by breast cancer molecular subtype following neoadjuvant chemotherapy. Breast Cancer Res Treat. 2018 Aug 1;170(3):559–67.

38. COLEY WB. THE TREATMENT OF INOPERABLE SARCOMA WITH THE ’MIXED TOXINS OF ERYSIPELAS AND BACILLUS PRODIGIOSUS.: IMMEDIATE AND FINAL RESULTS IN ONE HUNDRED AND FORTY CASES. J Am Med Assoc. 1898 Aug 27;XXXI(9):456–65.

39. Esfahani K, Roudaia L, Buhlaiga N, Del Rincon SV, Papneja N, Miller WH. A review of cancer immunotherapy: from the past, to the present, to the future. Curr Oncol. 2020 Apr;27(Suppl 2):S87–97.

40. Gaynor N, Crown J, Collins DM. Immune checkpoint inhibitors: Key trials and an emerging role in breast cancer. Semin Cancer Biol. 2022 Feb 1;79:44–57.

41. Jacqueline C, Finn OJ. Antibodies specific for disease-associated antigens (DAA) expressed in non-malignant diseases reveal potential new tumor-associated antigens (TAA) for immunotherapy or immunoprevention. Semin Immunol. 2020 Feb 1;47:101394.

42. Mellman I, Chen DS, Powles T, Turley SJ. The cancer-immunity cycle: Indication, genotype, and immunotype. Immunity. 2023 Oct;56(10):2188–205.

43. Alemohammad H, Najafzadeh B, Asadzadeh Z, Baghbanzadeh A, Ghorbaninezhad F, Najafzadeh A, et al. The importance of immune checkpoints in immune monitoring: A future paradigm shift in the treatment of cancer. Biomed Pharmacother. 2022 Feb 1;146:112516.

44. He X, Xu C. Immune checkpoint signaling and cancer immunotherapy. Cell Res. 2020 Aug;30(8):660–9.

45. Obeid M, Tesniere A, Ghiringhelli F, Fimia GM, Apetoh L, Perfettini JL, et al. Calreticulin exposure dictates the immunogenicity of cancer cell death. Nat Med. 2007 Jan;13(1):54–61.

46. Zitvogel L, Apetoh L, Ghiringhelli F, Kroemer G. Immunological aspects of cancer chemotherapy. Nat Rev Immunol. 2008 Jan;8(1):59–73.

47. Griguolo G, Pascual T, Dieci MV, Guarneri V, Prat A. Interaction of host immunity with HER2-targeted treatment and tumor heterogeneity in HER2-positive breast cancer. J Immunother Cancer. 2019 Mar 29;7(1):90.

48. Preglej T, Brinkmann M, Steiner G, Aletaha D, Göschl L, Bonelli M. Advanced immunophenotyping: A powerful tool for immune profiling, drug screening, and a personalized treatment approach. Front Immunol. 2023 Mar 24;14:1096096.

49. A novel role for an old target: CD45 for breast cancer immunotherapy - PMC [Internet]. [cited 2024 Jan 21]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8158046/

50. CD56 expression in breast cancer induces sensitivity to natural killer-mediated cytotoxicity by enhancing the formation of cytotoxic immunological synapse - PMC [Internet]. [cited 2024 Jan 21]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6584531/#

51. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol. 2016 July;17(7):451–9.

52. Schmidt DR, Patel R, Kirsch DG, Lewis CA, Vander Heiden MG, Locasale JW. Metabolomics in Cancer Research and Emerging Applications in Clinical Oncology. CA Cancer J Clin. 2021 July;71(4):333–58.

53. Hallmarks of Cancer: New Dimensions | Cancer Discovery | American Association for Cancer Research [Internet]. [cited 2025 May 22]. Available from: https://aacrjournals.org/cancerdiscovery/article/12/1/31/675608/Hallmarks-of-Cancer-New-DimensionsHallmarks-of

54. Beger RD. A Review of Applications of Metabolomics in Cancer. Metabolites. 2013 July 5;3(3):552–74.

55. Emerging applications of metabolomics in drug discovery and precision medicine | Nature Reviews Drug Discovery [Internet]. [cited 2025 May 2]. Available from: https://www.nature.com/articles/nrd.2016.32

56. Measuring the metabolome: current analytical technologies - PubMed [Internet]. [cited 2025 May 2]. Available from: https://pubmed.ncbi.nlm.nih.gov/15852128/

57. Zhang J, Xia Y, Sun J. Breast and gut microbiome in health and cancer. Genes Dis. 2021 Sept;8(5):581–9.

58. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. Nat Rev Microbiol. 2014 Oct;12(10):661–72.

59. Zierer J, Jackson MA, Kastenmüller G, Mangino M, Long T, Telenti A, et al. The fecal metabolome as a functional readout of the gut microbiome. Nat Genet. 2018 June;50(6):790–5.

60. Mikó E, Vida A, Kovács T, Ujlaki G, Trencsényi G, Márton J, et al. Lithocholic acid, a bacterial metabolite reduces breast cancer cell proliferation and aggressiveness. Biochim Biophys Acta BBA - Bioenerg. 2018 Sept;1859(9):958–74.

61. Fernandes MR, Aggarwal P, Costa RGF, Cole AM, Trinchieri G. Targeting the gut microbiota for cancer therapy. Nat Rev Cancer. 2022 Dec;22(12):703–22.

62. McCulloch JA, Davar D, Rodrigues RR, Badger JH, Fang JR, Cole AM, et al. Intestinal microbiota signatures of clinical response and immune-related adverse events in melanoma patients treated with anti-PD-1. Nat Med. 2022 Mar;28(3):545–56.

63. Lee KA, Thomas AM, Bolte LA, Björk JR, De Ruijter LK, Armanini F, et al. Cross-cohort gut microbiome associations with immune checkpoint inhibitor response in advanced melanoma. Nat Med. 2022 Mar;28(3):535–44.

64. Gharaibeh RZ, Jobin C. Microbiota and cancer immunotherapy: in search of microbial signals. Gut. 2019 Mar;68(3):385–8.

65. Shaikh FY, White JR, Gills JJ, Hakozaki T, Richard C, Routy B, et al. A Uniform Computational Approach Improved on Existing Pipelines to Reveal Microbiome Biomarkers of Nonresponse to Immune Checkpoint Inhibitors. Clin Cancer Res. 2021 May 1;27(9):2571–83.

66. Spencer CN, McQuade JL, Gopalakrishnan V, McCulloch JA, Vetizou M, Cogdill AP, et al. Dietary fiber and probiotics influence the gut microbiome and melanoma immunotherapy response. Science. 2021 Dec 24;374(6575):1632–40.

67. Davar D, Dzutsev AK, McCulloch JA, Rodrigues RR, Chauvin JM, Morrison RM, et al. Fecal microbiota transplant overcomes resistance to anti–PD-1 therapy in melanoma patients. Science. 2021 Feb 5;371(6529):595–602.

68. Routy B, Lenehan JG, Miller WH, Jamal R, Messaoudene M, Daisley BA, et al. Fecal microbiota transplantation plus anti-PD-1 immunotherapy in advanced melanoma: a phase I trial. Nat Med. 2023 Aug;29(8):2121–32.

69. Alexander JL, Wilson ID, Teare J, Marchesi JR, Nicholson JK, Kinross JM. Gut microbiota modulation of chemotherapy efficacy and toxicity. Nat Rev Gastroenterol Hepatol. 2017 June;14(6):356–65.

70. Riquelme E, Zhang Y, Zhang L, Montiel M, Zoltan M, Dong W, et al. Tumor Microbiome Diversity and Composition Influence Pancreatic Cancer Outcomes. Cell. 2019 Aug 8;178(4):795-806.e12.

71. Ingman WV. The Gut Microbiome: A New Player in Breast Cancer Metastasis. Cancer Res. 2019 July 15;79(14):3539–41.

72. Chen J, Douglass J, Prasath V, Neace M, Atrchian S, Manjili MH, et al. The microbiome and breast cancer: a review. Breast Cancer Res Treat. 2019 Dec;178(3):493–6.

73. Ross K, Nichols B, Papadopoulou R, Macleod M, Fraser J, Barrett S, et al. Gut Microbiota Predict Efficacy of Neoadjuvant Systemic Therapy in Patients with Early Breast Cancer. Clin Oncol. 2024 Apr 1;36(4):e107–8.

74. Ross K, Nichols B, Papadopoulou R, Macleod M, Fraser J, Barrett S, et al. Gut Microbiota Predict Efficacy of Neoadjuvant Systemic Therapy in Patients with Early Breast Cancer. Clin Oncol. 2024 Apr 1;36(4):e107–8.

75. Kelly PE, Ng HJ, Farrell G, McKirdy S, Russell RK, Hansen R, et al. An Optimised Monophasic Faecal Extraction Method for LC-MS Analysis and Its Application in Gastrointestinal Disease. Metabolites. 2022 Nov;12(11):1110.

76. Zhu J. New Metabolomic Insights into Cancer. Cancer J Sudbury Mass. 2024;30(5):301–6.

77. Kelly PE, Ng HJ, Farrell G, McKirdy S, Russell RK, Hansen R, et al. An Optimised Monophasic Faecal Extraction Method for LC-MS Analysis and Its Application in Gastrointestinal Disease. Metabolites. 2022 Nov 14;12(11):1110.

78. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science. 2009 May 22;324(5930):1029–33.

79. Yang S, Hao S, Ye H, Zhang X. Crosstalk between gut microbiota and cancer chemotherapy: current status and trends. Discov Oncol. 2024 Dec 24;15:833.

80. Zhu J. New Metabolomic Insights into Cancer. Cancer J Sudbury Mass. 2024;30(5):301–6.

81. DeBerardinis RJ, Chandel NS. Fundamentals of cancer metabolism. Sci Adv. 2016 May;2(5):e1600200.

82. Harizi H, Corcuff JB, Gualde N. Arachidonic-acid-derived eicosanoids: roles in biology and immunopathology. Trends Mol Med. 2008 Oct;14(10):461–9.

83. Koundouros N, Poulogiannis G. Reprogramming of fatty acid metabolism in cancer. Br J Cancer. 2020 Jan;122(1):4–22.

84. Lim JH, Gerhart-Hines Z, Dominy JE, Lee Y, Kim S, Tabata M, et al. Oleic Acid Stimulates Complete Oxidation of Fatty Acids through Protein Kinase A-dependent Activation of SIRT1-PGC1α Complex. J Biol Chem. 2013 Mar 8;288(10):7117–26.

85. Kim JS, Kim DK, Moon JY, Lee MY, Cho SK. Oleic acid inhibits the migration and invasion of breast cancer cells with stemness characteristics through oxidative stress-mediated attenuation of the FAK/AKT/NF-κB pathway. J Funct Foods. 2024 May 1;116:106224.

86. Fernández LP, Gómez de Cedrón M, Ramírez de Molina A. Alterations of Lipid Metabolism in Cancer: Implications in Prognosis and Treatment. Front Oncol. 2020 Oct 28;10:577420.

87. Bonifácio VDB, Pereira SA, Serpa J, Vicente JB. Cysteine metabolic circuitries: druggable targets in cancer. Br J Cancer. 2021 Mar 2;124(5):862–79.

88. Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. Cell. 2012 May 25;149(5):1060–72.

89. Ma J, Zhong M, Xiong Y, Gao Z, Wu Z, Liu Y, et al. Emerging roles of nucleotide metabolism in cancer development: progress and prospect. Aging. 2021 May 5;13(9):13349–58.

90. Carey BW, Finley LWS, Cross JR, Allis CD, Thompson CB. Intracellular α-ketoglutarate maintains the pluripotency of embryonic stem cells. Nature. 2015 Feb 19;518(7539):413–6.

91. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, et al. Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of α-Ketoglutarate-Dependent Dioxygenases. Cancer Cell. 2011 Jan 18;19(1):17–30.

92. Gonen N, Assaraf YG. Antifolates in cancer therapy: structure, activity and mechanisms of drug resistance. Drug Resist Updat Rev Comment Antimicrob Anticancer Chemother. 2012 Aug;15(4):183–210.

93. Sato Y, Tomita M, Soga T, Ochiai A, Makinoshima H. Upregulation of Thymidylate Synthase Induces Pemetrexed Resistance in Malignant Pleural Mesothelioma. Front Pharmacol. 2021;12:718675.

94. Zhou W, Yao Y, Scott AJ, Wilder-Romans K, Dresser JJ, Werner CK, et al. Purine metabolism regulates DNA repair and therapy resistance in glioblastoma. Nat Commun. 2020 July 30;11:3811.

95. DeBerardinis RJ, Cheng T. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. Oncogene. 2010 Jan 21;29(3):313–24.

96. Le A, Lane AN, Hamaker M, Bose S, Gouw A, Barbi J, et al. Glucose-independent glutamine metabolism via TCA cycling for proliferation and survival in B cells. Cell Metab. 2012 Jan 4;15(1):110–21.

97. DeBerardinis RJ, Cheng T. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. Oncogene. 2010 Jan 21;29(3):313–24.

98. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. Cell Metab. 2016 Jan 12;23(1):27–47.

99. Geginat J, Lanzavecchia A, Sallusto F. Proliferation and differentiation potential of human CD8+ memory T-cell subsets in response to antigen or homeostatic cytokines. Blood. 2003 June 1;101(11):4260–6.

100. Improved biocytin labeling and neuronal 3D reconstruction | Nature Protocols [Internet]. [cited 2025 May 14]. Available from: https://www.nature.com/articles/nprot.2011.449

101. CTLA-4 Control over Foxp3+ Regulatory T Cell Function | Science [Internet]. [cited 2025 May 14]. Available from: https://www.science.org/doi/10.1126/science.1160062

102. Alexa Fluor 700 Dye - UK [Internet]. [cited 2025 May 14]. Available from: https://www.thermofisher.com/uk/en/home/life-science/cell-analysis/fluorophores/alexa-fluor-700.html