Doctoral theses at NTNU, 2021:70

Julia Debik

Metabolic characterization of breast cancer for improved precision medicine

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Medicine and Health Sciences
Department of Circulation and Medical Imaging

NTNU
Norwegian University of
Science and Technology

Julia Debik

Metabolic characterization of breast cancer for improved precision medicine

Thesis for the Degree of Philosophiae Doctor

Trondheim, March 2021

Norwegian University of Science and Technology
Faculty of Medicine and Health Sciences
Department of Circulation and Medical Imaging

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Medicine and Health Sciences
Department of Circulation and Medical Imaging

ISBN 978-82-326-6712-3 (printed ver.)
ISBN 978-82-326-6662-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:70

Printed by NTNU Grafisk senter

Sammendrag

Brystkreft er kreftformen som rammer flest kvinner, 1 av 12 kvinner vil bli diagnostisert med brystkreft før fylte 75 år i Norge, og forekomsten fortsetter å øke. Det er en heterogen og kompleks sykdom, pasienter med lik diagnose kan respondere ulikt på samme behandling, og dermed ha forskjellig utfall. Det trengs mer kunnskap om sykdommen for å kunne utvikle mer persontilpasset behandling og bedre metoder for å overvåke behandlingsrespons. Prognosen for de fleste brystkreftpasienter er god, med best prognose ved tidlig oppdagelse. Derfor trengs det også mer kunnskap om de tidlige biologiske mekanismene bak dannelsen av brystkreft, slik at kvinner med høy risiko kan identifiseres tidlig og tilbys tettere oppfølgning, noe som potensielt kan bidra til redusert forekomst av avansert sykdom.

Kreftceller har et endret energiomsetning (metabolisme) i forhold til vanlige, friske celler. Raskt voksende kreftceller omdanner næringsstoffer til biomasse samtidig som de må opprettholde en høy energiproduksjon. Denne prosessen kan observeres ved å måle konsentrasjonen av små molekyler, kalt metabolitter, som er aktive komponenter av cellenes energiomsetning. Den metodiske tilnærmingen som benyttes for å måle metabolittene kalles metabolomikk, og kan gjøres blant annet ved magnetisk resonans spektroskopi (MRS), hvor et bredt panel av metabolitter observeres samtidig. Denne metoden har for eksempel vist at metabolske profiler av vevsprøver fra brystkreftpasienter kan si noe om prognosen til pasientene.

Hovedmålet i denne avhandlingen har vært å identifisere prognostiske og prediktive biomarkører for brystkreft gjennom en metabolsk tilnærming. For å kunne identifisere robuste biomarkører er det avgjørende å vite hvordan pre-analytiske prosesser kan påvirke metabolittene vi måler. I en biobank blir biologisk materiale oppbevart i fryst form, ofte over mange år. Hvor mange ganger disse prøvene er blitt tint og fryst igjen før analyse kan variere. Det er derfor viktig å vite effekten av slike sykluser på metabolittene, for å kunne tolke resultatene riktig. Artikkel II i denne avhandlingen er en metodeartikkel, hvor det er blitt undersøkt hvordan metabolitter målt i serum og urin, og lipoprotein partikler målt i serum, blir påvirket av gjentatte fryse- og tinesykluser. Denne studien viste at det ikke observeres særlige systematiske effekter av opptil 5 fryse og tine sykluser, noe som betyr at MRS er en god metode for analyser av biobank-prøver.

I Artikkel I ble de metabolske effektene av behandling med neoadjuvant kjemoterapi i brystkreftpasienter undersøkt, både i vevsbiopsier og i serum. Tilgangen til to typer biologisk materiale i denne studien gjorde det mulig å undersøke korrelasjonsmønstre mellom metabolitter målt i vev og i serum, i tillegg til innad i hver type biologisk materiale. Svake korrelasjoner ble observert mellom konsentrasjonene av samme metabolitter målt både i vev og i serum. Studien viste også at de metabolske profilene av vevsprøvene, men ikke profilene fra serum, kunne predikere overlevelse.

Dette skyldes mest sannsynlig at den metabolske profilen av serum gir et mer helhetlig bilde av pasientens tilstand fordi blodet sirkulerer gjennom alle vev og organer i kroppen, mens den metabolske profilen til en vevsprøve beskriver mer direkte hva som foregår i selve svulsten.

I Artikkel III ble cirka 2400 serumprøver av friske kvinner fra HUNT2 studien analysert, hvorav halvparten senere utviklet brystkreft. I denne studien fant vi assosiasjoner mellom fremtidig brystkreft og en rekke variabler knyttet til ulike egenskaper av lipoproteiner. Variablene var assosiert med en signifikant økning i risiko, men var ikke sterke nok til å utvikle en robust modell for prediksjon av fremtidig brystkreft.

Samlet sett viser avhandlingen at metabolomikk har stor nytteverdi innen brystkreftforskning og kan være et verktøy for utvikling av kliniske biomarkører for forbedret persontilpasset diagnostikk og behandling.

Kandidat: Julia Debik

Institutt: Institutt for sirkulasjon og bildediagnostikk

Hovedveileder: Guro F. Giskeødegård

Biveiledere:

Tone F. Bathen, Institutt for sirkulasjon og bildediagnostikk, Fakultet for medisin og helsevitenskap, NTNU Trondheim

Hao Wang, Institutt for datateknologi og informatikk, Fakultet for informasjonsteknologi og elektronikk, NTNU Gjøvik

Finansieringskilde: NTNU

Acknowledgement

The work presented in this thesis has been carried out at the MR Cancer group, Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), between January 2018 and November 2020.

I would like to thank all women who have participated in the NeoAva and HUNT2 studies, and donors from the blood bank, making this work possible.

I would like to express my gratitude and appreciation to my main supervisor, Guro F. Giskeødegård. Guro, I feel fortunate to have had the opportunity to work under your supervision, and I admire you as a researcher and a person. You have always been accessible and supportive for discussing my scientific work, but also as a friend with shared interests. You have inspired me to search new challenges in order to continuously improve as a person. Without your support and encouragement, I would not have succeeded. I would like to thank my co-supervisor, Tone F. Bathen, which has contributed with her expertise and high experience in the field. Tone, thank you for being engaged in my research and for providing guidance throughout all of my projects. I appreciate the inclusive way in which you lead the group, and for always celebrating our successes and for giving encouragement when things get tough. I would also like to thank my co-supervisor, Hao Wang, for taking time to discuss deep learning approaches for data analysis.

Besides my supervisors, a special acknowledgement goes to all my current colleagues: Alexandros, Alicja, Bendik, Christine, Christopher, Daniel, Debbie, Ellen Marie, Elise, Feng, Gabriel, Ingerid, Kaia, Line, Maria Karoline, Marco, Maren, Matteo, Mattijs, May-Britt, Mohammed, Sebastian, Shanti, Sissel, Siver, Therese, Torfinn Torill, Tina, Trygve, and also former colleagues Liv, Hanna Maja, Maria TG, Christina, Kirsten, Neil, Leslie and Tonje. It has been inspiring to work in such a resourceful and intra-disciplinary group, with a diverse set of skills. I am grateful for useful feedback received when presenting at our group meetings and I look forward to continuing working with you!

I would also like to thank all of my co-authors and collaborators, for providing insightful comments and suggestions for improving my work. In specific, Olav Engebråten and his research team in Oslo, for the collaboration on Paper I, and Bruker Biospin, for collaborations on Paper II and Paper III.

My sincere thanks goes to my family. I would like to thank my children Isabella, Aleksander and Sebastian. You never stop surprising me, you make me laugh, and I am very proud of each one of you. You have been my biggest motivation for doing my best, for working efficiently, and have helped to balance out the stress. Last, but not least, I would like to thank my husband, Karol, for being my best

friend and a steady rock during this emotional roller coaster journey. Thank you for your patience, but also for reacting when my work-life balance was getting too skewed towards work.

Julia Debik

Trondheim, November 2020

Summary

Breast cancer is the most common cancer type among women, 1 of 12 women will be diagnosed with breast cancer before turning 75 years in Norway, and the incidence rate is continuously increasing. It is a heterogeneous and complex disease, and patients with the same diagnosis respond differently to the same treatment, and may thus have a different outcome. There is a need for more knowledge about the disease in order to develop a more personalized treatment regime, and a minimally-invasive tool for monitoring treatment response. The prognosis is good for the majority of breast cancer patients, but is highly dependent on the stage of the disease at the time of diagnosis. Therefore, there is also a need for more knowledge about the early biological mechanisms driving the cancer formation, to identify women at a high risk for developing the disease, which may be given a closer follow-up, thus potentially reducing the incidence rate of severe cases.

Cancer cells have a metabolism radically different than normal, healthy cells. Quickly growing cancer cells need to convert nutrients to biomass while maintaining a high energy production. This process may be observed by measuring the concentrations of small molecules, called metabolites, which are active components of the cell cycle. Metabolic profiling can be performed by magnetic resonance spectroscopy (MRS), where a range of metabolites can be observed simultaneously. This techniques has for example shown that metabolic profiles in tissue biopsies from breast cancer patients can provide information about the prognosis of the patients.

The main aim of this thesis has been to search for prognostic and predictive biomarkers, using a metabolomics approach. In order to identify robust biomarkers it is crucial to know how the pre-analytical processes may influence the metabolites we measure. In a biobank, biological material is stored frozen, often for many years. How many times these samples have been thawed and frozen prior to analysis can be variable, and it is therefore important to known the effect of such cycles on the metabolites to correctly evaluate findings based on samples from biobanks. Paper II in this thesis is a methodological paper, in which the effect of repeated freeze and thaw cycles on metabolites measured in serum and urine, and lipoprotein particles measured in serum has been assessed. This study showed that there was a small accumulated effect of up to five freeze and thaw cycles, which means that MRS is a good method for analyzing samples from biobanks.

In Paper I the metabolic effects of neoadjuvant treatment in breast cancer patients were investigated, both in tissue biopsies and in serum samples. The availability of two types of biological samples in this study made it possible to investigate correlations between metabolites measured in tissue and in serum, and in each type of medium alone. Weak correlations were observed between the same metabolites measured in tissue and serum samples. The study also showed that tissue metabolic

profiles, but not serum metabolic profiles, could predict survival. This is probably because the serum metabolic profile gives a more whole picture of the current state of a patient, as blood circulates through all tissues and organs in the body, while the tissue metabolic profile describes ongoing metabolic processes in the tumor directly.

In Paper III approximately 2400 serum samples of healthy women from the HUNT2 study were analyzed, of which half later developed breast cancer. In this study we found associations between future breast cancer and multiple lipoprotein parameters. Variables significantly associated with an increase in the risk of developing breast cancer, were not strong enough to develop a robust model for prediction of future breast cancer.

In total, this thesis has shown that metabolomics is a useful tool in breast cancer research, and may have a future role in the development of clinical biomarkers for improved personalized diagnostics and treatment.

Abbreviations

Apo	Apolipoprotein
BC	Breast cancer
CH	Cholesterol
CNN	Convolutional neural network
CPMG	Carr-Purcell-Meiboom-Gill
CV	Coefficient of variation or Cross validation
CVD	Cardiovascular disease
DL	Deep learning
DNA	Deoxyribonucleic acid
DRFS	Distant relapse-free survival
ER	Estrogen receptor
FC	Free cholesterol
FID	Free induction decay
FTC	Freeze-thaw cycle
GBM	Gradient boosting machine
HDL	High-density lipoprotein
HER-2	Human epidermal growth factor receptor 2
ICC	Intraclass correlation coefficient
IDL	Intermediate-density lipoprotein
JRES	J-resolved Spectroscopy
LABC	Locally advanced breast cancer
LDL	Low-density lipoprotein
LOD	Limit of detection
LR	Logistic regression
LV	Latent variable
ML	Machine learning
MR	Magnetic resonance
MS	Mass spectroscopy
NAC	Neoadjuvant chemotherapy
NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser effect spectroscopy
PC	Principal component
PCA	Principal component analysis

pCR	Pathologic complete response
PgR	Progesterone receptor
PL	Phospholipids
PLS	Partial least squares
PLS-DA	Partial least squares discriminant analysis
ppm	Parts per million
QC	Quality-control
RCB	Residual cancer burden
RCT	Randomized controlled trial
RF	Random forest
RNA	Ribonucleic acid
SGD	Stochastic gradient descent
TG	Triglycerides
TNBC	Triple negative breast cancer
TP	Time point
VIP	Variable importance in projection
VLDL	Very low-density lipoprotein

List of papers

Paper I

Assessing treatment response and prognosis by serum and tissue metabolomics in breast cancer patients.

Julia Debik, Leslie R. Euceda, Steinar Lundgren, Olav Engebraaten, Øystein Garred, Elin Borgen, Hedda von der Lippe Gythfeldt, Tone F. Bathen and Guro F. Giskeødegård

Journal of Proteome Research: 2019 Oct 4; 18(10):3649-3660. doi: 10.1021/acs.jproteome.9b00316

Paper II

Effect of repeated freeze-thaw cycles on NMR measured lipoproteins and metabolites in biofluids.

Feng Wang[*], Julia Debik[*], Trygve Andreassen, Leslie R. Euceda, Tonje H. Haukaas, Claire Cannet, Hartmut Schäfer, Tone F. Bathen[#], Guro F. Giskeødegård[#]; [*]shared first authorship; [#]shared last authorship

Journal of Proteome Research: 2019 Oct 4; 18(10): 3681-3688. doi: 10.1021/acs.jproteome.9b00343

Paper III

Serum metabolic profiling for assessment of breast cancer risk in women participating in the HUNT2 study.

Julia Debik, Hartmut Schaefer, Trygve Andreassen, Feng Wang, Fang Fang, Claire Cannet, Manfred Spraul, Tone F. Bathen, Guro F. Giskeødegård

Manuscript

Table of contents

1 Introduction

The human body is made of cells, in which complex biological processes take place continuously. Cells can grow and divide, allowing for replacement of worn out cells. Cell division is governed by a series of tightly regulated events, called the cell cycle [2]. The cell cycle involves the replication of deoxyribonucleic acid (DNA). The DNA then separates into two sets and the cell divides its cytoplasm, forming two new cells. While normal cells are strictly controlled by regulatory signals, cancer cells are capable of avoiding these mechanisms, thus cancer cells exhibit uncontrolled growth and proliferation.

Cancer refers to a high collection of diseases, which can occur in different organs of the human body, with a high complexity and variety in characteristics [3, 4]. Nevertheless, six essential alternations necessary for malignant growth which are common traits or capabilities of the disease, have been described. These are referred to as the Hallmarks of Cancer, first described by Hanahan and Weinberg in 2000 [5]. During tumor development, cancer cells become capable of (1) sustaining proliferative signaling, (2) evading growth suppressors, (3) resisting cell death, (4) enabling replicative immortality, (5) inducing angiogenesis, and (6) activating invasion and metastasis. The hallmarks of cancer have later been extended by two characteristics: (7) deregulation of cellular energetics and (8) avoiding immune destruction [1], as illustrated in Figure 1.1.

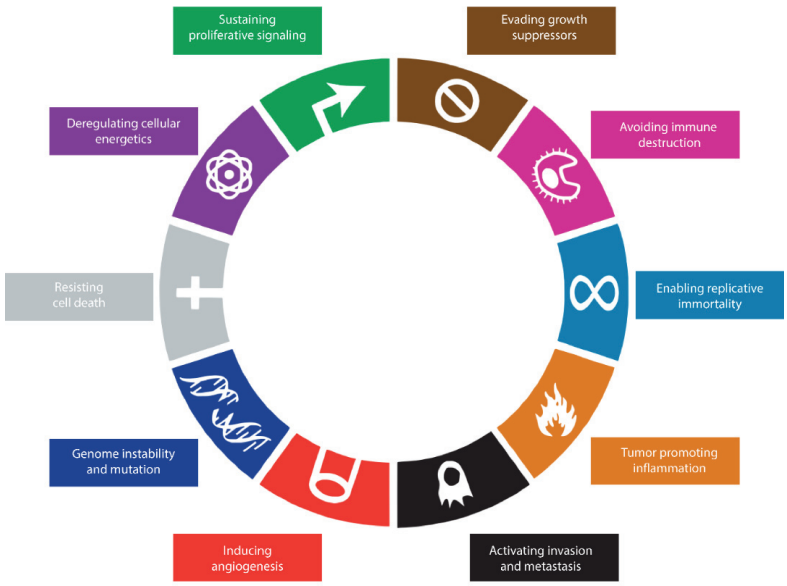


Figure 1.1 The Hallmarks of cancer. Figure reproduced and modified with permission, Hanahan D, Weinberg RA [1].

1.1 Breast cancer

Breast cancer is the most frequently diagnosed cancer among women in Norway and worldwide [4]. There were 3623 new cases among Norwegian women in 2018 and breast cancer comprises more than 20% of female cancers [7].

The mortality rates of breast cancer have decreased during the last years, however, the incidence rate remains increasing [6, 7]. There are many known potential risk factors associated with the development of breast cancer, however, there is no method available to assess an individuals' overall risk [8-10]. A minimally-invasive method for personalized risk stratification and early detection would be valuable to decrease the incidence rate and for evaluation of treatment applied at an early stage of the cancer formation.

The five-year survival of breast cancer patients is estimated to be 90.7% in Norway. It is however difficult to predict each cancer patient's outcome. Patients with the same diagnosis may have different response to treatment [11, 12]. It is therefore crucial to characterize breast cancer heterogeneity as well as response to treatment.

1.1.1 Anatomy of the breast

The breast contains a complex network of lobules, lobes (groups of lobules) and ducts, surrounded by adipose tissue [13], as illustrated in Figure 1.2. In nursing women, milk is produced in the lobules, which are connected to ducts that transport the milk to the nipple. Surrounding the breast are lymph nodes and vessels, containing immune cells, which fight harmful substances and germs that enter the body. The breast tissue undergoes changes throughout a life cycle, signaled by growth factors, cytokines and hormones [14, 15]. The majority of breast cancers originate from the lobules or ducts. Ductal carcinoma in situ (DCIS) is a precancerous condition characterized by the presence of abnormal cells in the ducts. Similarly, lobular carcinoma in situ (LCIS) is a precancerous condition which originates in the lobe [16]. LCIS is much rarer than DCIS, however it are associated with a greater risk of developing an invasive cancer, which spreads from its origin and infiltrates the surrounding tissue [17]. The basement membrane is a thin, dense sheet of extracellular matrix, between epithelial tissues and the underlying connective tissue [18]. It provides structural support to cells, divides tissues into compartments and acts as a platform for cell signaling [19]. If the basement membrane has not been broken, the cancer is classified as carcinoma in situ, invasive otherwise. Invasive carcinoma of no special type (NST), previously called ductal carcinomas, make up about 70-80% of breast cancer cases

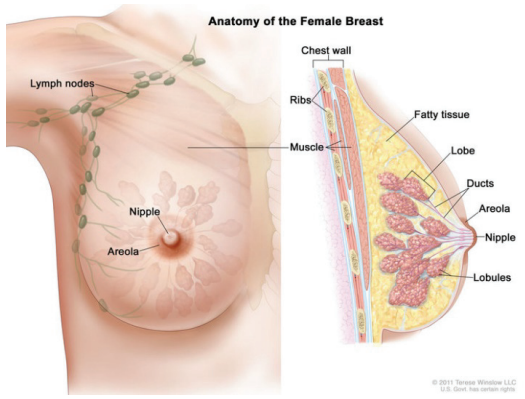


Figure 1.2 The anatomy of the female breast. The female breast consists mainly of complex networks of lobules and ducts, surrounded by adipose tissue. Surrounding the female breast are lymph nodes and vessels. Figure reproduced with permission from Terese Winslow LLC.

in Norway, while 10-20 % are invasive lobular carcinomas [20]. In addition there are several rarer types of malignant tumors: sarcomatoid carcinoma, phyllodes tumors and sarcomas, for which specific treatment regimens exist.

Cancer originating in the breast can metastasize to the bone, lungs or liver through hematogenous dissemination, or to local lymph nodes through lymphogenous spread. The term locally advanced breast cancer (LABC) is used to describe breast cancer that has progressed locally in the absence of distant metastasis [21]. One or more of the following criteria must be met for a breast cancer to be classified as LABC, given that there is no distant metastasis: 1) the tumor is more than 5 cm in size, 2) the tumor has a direct extension to the chest wall or the skin, and 3) the tumor has spread to lymph nodes in areas near the breast [20].

1.1.2 Known risk factors in breast cancer

The most significant risk factor for developing breast cancer is gender, as less than one percent of all breast cancer cases develop in men [8, 22, 23]. Breast cancer is an age-related disease, thus the second biggest risk factor is age, while about 5-10% of breast cancers are thought to be hereditary, caused by abnormal genes passed from parent to child [24]. Some of the risk factors, such as age, family history and medical history, are beyond the control of an individual, however other risk factors may be controlled, and preventive actions are possible [9, 25]. Common for some of the risk factors is that they are associated with the hormone estrogen, and higher estrogen levels increase the risk of developing breast cancer. Overweight is associated with the risk of developing breast cancer, especially for postmenopausal women [9]. Most estrogens are produced in the ovaries until menopause, while in postmenopausal women the ovaries cease to produce estrogen and estrogens mainly come from fat tissues, which produce and store estrogen [26, 27]. Taking combined hormone

replacement therapy, or estrogen alone, for several years can also increase the risk. Women who have had a full-term pregnancy or have their first child before 30 have a lower risk of breast cancer compared to women who gave birth after the age of 30 [28]. The risk decreases further with multiple full-term pregnancies and breast feeding [8, 29]. Smoking can increase the risk of developing breast cancer, especially among women who started smoking at adolescent or peri-menarcheal ages and women with a family history of breast cancer [30]. Alcohol consumption can increase the risk of developing breast cancer due to alcohol-induced hormonal dysregulations [30-33]. A healthy lifestyle, incorporating regular physical exercise and a diet rich in vegetables, omega 3, and low amounts of trans-fats, is associated with a lower risk of developing breast cancer [10, 34].

1.1.3 Breast cancer diagnosis and treatment

Triple diagnostics is the common approach for diagnosing breast cancer in Norway. It consists of a clinical examination, image diagnostics and a needle biopsy [20]. Mammography is an x-ray examination of the breast, and is offered as a screening tool to identify cancer in women aged 50-69 years in Norway. Ultrasound and magnetic resonance imaging (MRI) may in some cases aid as a supplement to mammography. The needle biopsy is used for a preoperative histological diagnosis. The stage of the breast cancer is classified, where the tumor size, degree of spread to lymph nodes and distant metastasis are considered, referred to as the TNM system [20]. The stages are T0 if no primary tumor, Tis for carcinoma in situ, otherwise T1-T4, with increasing tumor size. Number and location of lymph node metastasis may be classified into N0-N3, where increasing number is a higher degree of lymph node involvement. Distant metastasis is classified as M0 or M1, depending on its absence or present, respectively. These variables make up the TNM classification of the tumor, and define the stage (I-IV) of the tumor, where a higher stage means increased advancement of the tumor (size, spread to lymph nodes and distant metastasis). TNM classification is either clinical or pathological (pTNM), depending on whether it has been performed before or after surgery, respectively. The tumor is classified as primary operable (stage I or II) or inoperable (stage II if tumor > 5 cm in diameter, III or IV) [35].

The choice of treatment regime is based on the preoperative examination, comorbidity and dialog with the patient. Primary treatment includes surgery, with the removal of the tumor and sometimes lymph nodes. The sentinel node is the primary lymph node or lymph nodes into which drains a tumor [36]. If sentinel lymph node biopsy reveals cancer, these and the remaining lymph nodes, are almost always removed. In LABC neoadjuvant chemotherapy (NAC) is necessary prior to the surgical removal of the tumor, which has the purpose of shrinking and downstaging the tumor. Secondary treatment following surgery is recommended in approximately 90% of the cases, which includes adjuvant treatment with chemotherapy and/or endocrine treatment (hormone therapy). The purpose of this

treatment is to reduce the risk of relapse, however certain patients may have a low-risk breast cancer for which adjuvant treatment is not justified due to its associated toxicities [37]. Endocrine treatment and chemotherapy are systemic treatments, while local radiation therapy is in addition given to patients which had breast-conserving surgery or with spread to the lymph nodes.

Hormone receptors are proteins within and on the surface of certain cells that act like an on-off switch for a particular activity of the cell. When a signal molecule binds to its hormone receptor it induces a cascade of processes in the cell. Approximately 75% of breast cancers are estrogen receptor positive (ER+), meaning that at least 1% of tumor cells demonstrate positive nuclear staining by immunohistochemistry [38, 39], and the majority of these are also progesterone receptor positive (PgR positive) [40, 41]. Endocrine treatment is given receptor positive patients after a histopathological evaluation and assessment of the tumor's expression of ER and PgR [42]. Human epidermal growth factor receptor 2 (HER2) is a protein that promotes cellular growth and proliferation, and can be targeted by anti-HER2 treatment [43]. An amplification or overexpression of HER2 occurs in approximately 15-30% of breast cancers, and is associated with shorter disease-free and overall survival and a higher risk of recurrence compared to normal expressions of this protein [44].

Chemotherapy kills rapidly dividing cells, through disrupting the microtubule function, which are essential to cell division. As well as killing cancer cells, chemotherapy may impact healthy cells, especially rapidly-dividing ones, which include blood cells forming the bone marrow, hair cells, cells in the digestive tract and reproductive system [45]. There are different chemotherapy treatment regimes, of which anthracycline chemotherapy by fluorouracil, epirubicin and cyclophosphamide (FEC) or taxane is used in Norway.

Angiogenesis is the formation of new blood vessels from existing vasculature, and has an essential role for supplying nutrients and oxygen to rapidly growing tumors [46]. This can be therapeutically targeted by antiangiogenic treatment, such as Bevacizumab, which has the ability to inhibit the proangiogenic vascular endothelial growth factor (VEGF) [47]. When Bevacizumab binds to VEGF, the proteins function will be altered and the tumor will have reduced blood and thus nutrient supply. This drug is currently accepted for treatment of metastatic breast cancers only, because of the possibility of fatal adverse events including hemorrhage, pulmonary embolism and gastrointestinal tract perforation [48]. It is however possible that the benefits of Bevacizumab for some women with a locally advanced cancer are worth the risks of treatment, thus a biomarker would be valuable for identification of women who are most likely to benefit from this treatment. Bevacizumab is still used for other cancers, such as colorectal cancer [49]. The different breast cancer treatment strategies, based on current treatment guidelines in Norway are summarized in Table 1.1.

Table 1.1 Breast cancer treatment strategies based on current treatment guidelines in Norway [20].

		Operable tumor	
	Inoperable tumor		
	Before surgery	Surgery	After surgery
Treatment	Neoadjuvant treatment	Mastectomy or breast conserving surgery	Radiation therapy Endocrine therapy Chemotherapy Anti-Her2 treatment
Purpose	Tumor shrinkage and downstaging	Remove primary tumor and lymph nodes	Reduce recurrence

1.1.4 Treatment response criteria

Due to different treatment strategies, several response criteria of NAC treatment have emerged. Pathological complete response (pCR) has been considered the gold standard treatment outcome, and refers to complete disappearance of cancer cells at treatment completion. pCR has been associated with improved survival, however the association between pCR and long-term outcome varies between different breast cancer subtypes [50]. Residual cancer burden (RCB) is a continuous index, which combines pathologic measurements of the primary tumor (size and cellularity) and nodal metastases (number and size) [51]. RCB can be divided into four classes, where class 0 is equivalent to pathologic complete response (pCR). Another response criteria, based on anatomical measurements of the tumor, is the Responsive Criteria in Solid Tumors (RECIST), which has four response categories: complete response (CR), partial response (PR), stable disease (SB) and progressive disease (PG) [52]. CR and PR refer to complete disappearance of tumor and >30% tumor shrinkage, respectively, while PG describes an >20% increase in tumor and/or appearance of new lesions. SD refers to tumors whose size has not changed enough to quantify to PR or PD.

1.2 The omics of breast cancer

The omics cascade refers to the information flow, interactions and interrelations between the different omics levels: the genomics, transcriptomics, proteomics and metabolomics [53-55]. In all living cells, DNA is transcribed into RNA transcripts which are further translated into proteins. Proteins take part in molecular pathways, thus controlling metabolite levels. This flow of information from one omics level to another, illustrated in Figure 1.3, it is however also affected by additional factors, such as epigenetic alterations.

Breast cancer is a highly heterogeneous disease, which can be manifested at different molecular levels. Breast cancer genes (BRCA1 and BRCA2) produce tumor suppressor proteins, which help repair damaged DNA, ensuring the stability of the cells genetic materials. About 0.25% of the population carry mutated BRCA1 or BRCA2 genes [56], which no longer are capable of repairing broken DNA and preventing breast cancer. Individuals with BRCA mutations are more likely to develop breast cancer, and have a higher probability of recurrence once the primary cancer has been cured [57]. These mutations increase the lifetime risk for developing breast cancer, and 55-65% and 45% of women with BRCA1 mutation, or BRCA2 mutation, respectively, will develop breast cancer before the age of 70.

The second level of the omics cascade is transcriptomics, which is the study of gene expressions through measuring the transcripts of DNA, called RNA. Based on the gene expression profiles, five intrinsic subtypes of breast cancer have been established: luminal A, luminal B, HER2 enriched, normal-like, and basal-like [58]. The characteristic differences in the gene expression patterns of these subtypes correlate with tumor characteristics and clinical outcome. Luminal A breast cancers are most often ER / PgR positive and are associated with the best prognosis. Basal-like has the worst prognosis, is often associated with the BRCA1 mutated gene, and is primarily ER / PgR negative [59]. Luminal B is often associated with the BRCA2 mutated gene [60]. Gene expression profiling, for identifying the intrinsic subtype, has recently been approved for use in the clinic in Norway, through the Prosigna test [61]. The aim of the Prosigna test is to assess the expected benefit from chemotherapy for breast cancer patients, and will be used for patients with HR+ / HER2- tumors, without spread to lymph nodes. The objective is that for patients with a low risk of recurrence, the negative side effects of chemotherapy will outweigh the benefit of treatment.

Proteins are the functional products of genes, and do most of the work in cells and are required for the structure, function, and regulation of the body's tissues and organs [62]. Protein activity is however also affected by several ongoing processes, such as post-transcription modifications. Also the proteomic level plays an important role in current breast cancer clinical decision making in terms of optimal treatment plan, based on the differences in the expression of estrogen, progesterone and

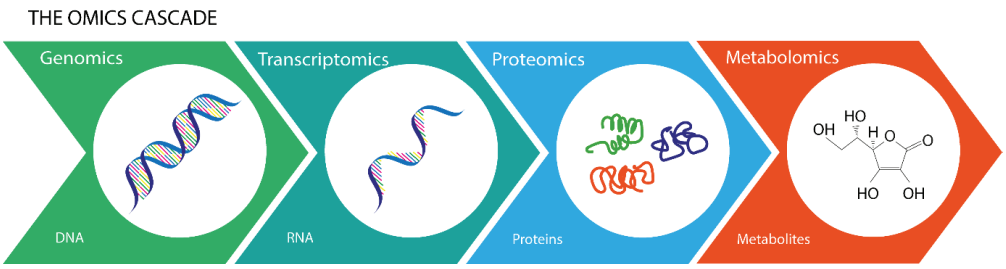


Figure 1.3 The omics cascade. All levels of the omics cascade interact with each other.

human epidermal growth factor receptors. Patients with an ER and/or PgR positive breast cancer will often benefit from hormone therapy, and HER2 positive breast cancers are often treated with anti-HER2 drugs, as described in chapter 1.1.3. However, triple negative breast cancer (TNBC) exhibit the greatest overlap with basal-like breast cancer, and has the worst prognosis. TNBC is ER, PgR and HER2 negative, and the tumors are thus unresponsive to hormone and anti-HER2 therapy [63]. TNBC is therefore treated solely with chemotherapy, except for metastatic cases, for which immunotherapy has been approved in March 2020 [64].

Six subtypes of breast cancer have been proposed based on the expression of proteins: basal, HER2, luminal A, luminal A/B, reactive I and reactive II [65]. These reverse phase protein array (RPPA) subgroups display considerable overlap with the gene intrinsic genetic subtypes, and have provided information about existing differences at the protein expression level. The reactive I and II protein subtypes are subsets of the luminal A intrinsic subtype and a combination of the intrinsic subtypes, respectively. The name reactive refers to the hypothesis, that many of the characteristic proteins are produced by the tumor microenvironment.

Metabolomics is the last level of the omics cascade, lying closest to the phenotype. Metabolites are end points or intermediates of chemical processes needed for cell viability. The metabolic profile of biological sample depends on the preceding omics levels as well as environmental factors. Metabolomics will be discussed in more detail in section 1.2.1.

1.2.1 Metabolomics

Metabolomics is the analysis of metabolites within a biological sample [66], and the metabolome represents the complete set of metabolites in the sample. The metabolic profile of a sample refers to a set of metabolites in the sample, as there is no analytical tool which simultaneously can measure all metabolites to date. Metabolites are small molecules (50-1500 Da), which are intermediates and downstream products of metabolism. The main groups of metabolites are sugars, amino acids, lipids, nucleotides and vitamins. The metabolome is comparable to the terms genome, transcriptome and proteome. Metabolites provide information closer to the phenotype, or the final observable endpoints of biological pathways, as they are the last level in the omics cascade [54]. The metabolome is a dynamic system, which in addition to the preceding levels in the omics cascade is influenced by environmental factors such as dietary intake, medication usage, gut microbiota and exercise [67-69]. The metabolic profile thus reflects the biological condition, giving an accurate snapshot of the current state of the system. Metabolomics has a wide range of common application, including plant biology, environmental studies, medicine and pharmacology, and has experienced an exponential growth during the last years [70, 71].

Metabolomics can be *targeted* meaning that only predefined metabolites are of interest, or *untargeted*, in which case all metabolites within a certain range are measured. Unlike genes, the total number of metabolites is undefined, and with the current analytical platforms we usually measure just a fraction of the whole metabolome.

There are two main analytical platforms for gaining insight into the metabolic profile of a sample: nuclear magnetic resonance spectroscopy (*NMR*) and mass spectroscopy (*MS*). These methods are complementary and have different strengths and limitations [72, 73]. The main difference is that MS has a higher sensitivity and can detect a higher range of metabolites, while NMR can provide information on chemical structure, is non-destructive and has a less extensive sample preparation.

1.2.2 Tumor metabolism

Cancer cells have an altered metabolism compared to normal cells and thus a reprogrammed energy metabolism for tumor survival, growth and proliferation [1, 74]. As metabolites are downstream products of the proceeding omics levels, small alterations in one of the preceding levels, such as alterations in the gene expression level, can be seen as amplified output of ongoing cellular activity, and can have an effect on metabolite concentrations. Cancer cells have three basic needs: 1) rapid generation of adenosine triphosphate (ATP) as a source of energy, 2) increased synthesis of lipids, carbohydrates, proteins and nucleic acids, and 3) proper redox stability, which are reflected in metabolic dysregulation of cancer cells [75].

Glycolysis is a linear metabolic pathway where glucose is broken down to pyruvate and a hydrogen ion. During this process the high-energy molecules ATP and reduced nicotinamide adenine dinucleotide (NADH) are formed. Depending on the presence or absence of oxygen, referred to as aerobic or anaerobic conditions, respectively, pyruvate can follow one of two possible pathways. If oxygen is available, pyruvate can be oxidized in the tricarboxylic acid (TCA) cycle, followed by phosphorylation to produce ATP. During this process about 23-30 ATP molecules are made per one oxidized glucose molecule, thus glucose is considered the main energy source of human cells. Under anaerobic conditions, however, pyruvate is broken down to lactate, yielding 2 ATP molecules, through a process called lactic acid fermentation. A specific characteristic of cancer cells is that most of the pyruvate is converted to lactate, independently on the presence of oxygen. This characteristic is called the Warburg effect, first described by Otto Warburg in 1930 [76]. Most tumors have an increase rate of glucose uptake, and perform glycolysis at a rate that is ten times faster than noncancerous tissues, to compensate for the inefficient ATP production [77, 78] . As cancer cells often experience hypoxia, they are dependent on the production of ATP by breaking down pyruvate to lactate, and elevated lactate levels have been observed in cancerous tissues [79]. It is thought that this effect is an

adaptation of cancer cells to facilitate the update of nutrients needed to produce new cells. Other functions of the Warburg effect have also been proposed, related to biosynthetic pathways, tumor microenvironment and cell signaling [78].

Amino acids are a group of organic compounds, characterized by the presence of amine and carboxyl functional groups, and a side chain specific to each amino acid. Amino acids serve as building blocks for proteins and play important roles as regulators or intermediates in several metabolic pathways for cell growth and maintenance. There are about 500 naturally occurring amino acids, though only 20 are present in the genetic code, of which nine are classified as essential. These are histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan and valine [80]. Essential amino acids are amino acids that cannot adequately be synthesized *de novo* by the organism and must be supplied through the diet, as opposed to non-essential amino acids. In tumor metabolism, however, non-essential amino acids play important roles in numerous aspects of tumor metabolism [81]. Their functions include providing precursors for biosynthesis of macromolecules, controlling redox status and antioxidant systems, and serving as substrates for post-translational and epigenetic modifications [82]. Glutamine is a non-essential amino acid, which has been found to be essential for rapidly dividing cells [83]. Glutamine can also be converted by glutaminase to glutamate which can be used for production of other amino acids, such as alanine, aspartate, serine and glycine. Glutamine is responsible for redox homeostatis and cancer signaling, and some cancer cell lines have shown glutamine addiction. Glutamate can also be utilized to produce ATP and thus help rapidly proliferating cells meet the increased demand for ATP, by replenishing TCA cycle intermediates. Glycine can also be produced from choline and has been associated with large tumors and poor prognosis [84, 85].

Lipid metabolism

Lipids refers to a class of large and diverse macromolecules, with multiple biochemical functions including energy storage, cell signaling and acting as structural components of cell membranes. They can be obtained from food (exogenous uptake) or can be synthesized by the liver (endogenous synthesis). A third pathway, called the reverse transport pathway, is a mechanism by which the body removes excess cholesterol from peripheral cells and transports it to the liver [86]. Figure 1.4 shows a simplified flowchart of lipid metabolism. Lipid metabolism involves lipid degradation and synthesis in cells, and abnormal lipid metabolism has been associated with numerous diseases, including type 2 diabetes, coronary artery disease, sleep apnea and cancer [87].

Lipid metabolism pathways in cancer cells are dysregulated by a number of cancer-cell intrinsic processes, and extensive studies have provided strong evidence for reprogramming of lipid metabolism in cancer [88]. Due to the metabolically challenging environment of cancer cells, with scare availability of oxygen and nutrients, the balance between the endogenous synthesis and

exogenous uptake of fatty acids is altered, and in general, cancerous tissues have an increased rate of lipid synthesis as part of the reprogrammed metabolism of cancer cells [89]. Depending on the tumor type, tumor cells can synthesize fatty acids de novo in spite of sufficient dietary lipid supply [90]. Activation of fatty acid synthesis is thought to be required for carcinogenesis and tumor cell survival. Lipid metabolism in cancer cells is differently regulated depending on environmental factors, in particular nutrient and oxygen availability.

Cancer cells with a sufficient supply of nutrients and oxygen mainly use glucose-derived acetyl-CoA for fatty acid synthesis for rapid cell proliferation [89, 91]. They can also acquire fatty acids from the environment, through utilizing both the lipogenic and lipolytic pathways [92, 93]. Under insufficient nutrient supply normoxic cancer cells mainly rely on endogenous fatty acid desaturation through acetate metabolism. The tumor microenvironment is however mostly hypoxic, meaning that it is deprived of adequate oxygen supply, and cancer cells either switch to alternative carbon sources (glutamine or acetate) or increase their fatty acid uptake. If the cells in addition are nutrient deprived, de novo fatty acid synthesis will be upregulated and the cancer cells will be fully dependent on glutamine or acetate for fatty acid synthesis.

The main two forms of circulating lipids in the body are triglycerides and cholesterol. These are insoluble in water and can be transported through the bloodstream as part of lipoproteins.

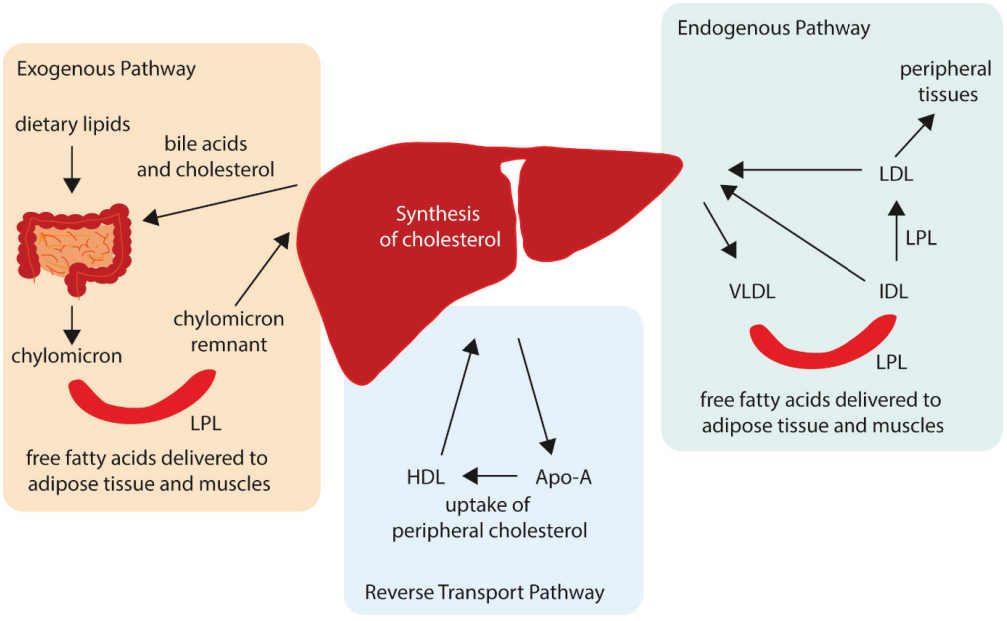


Figure 1.4 Simplified flowchart of lipoprotein metabolism, showing the main steps of the Exogenous, Endogenous and Reverse transport pathways. FFA: Free fatty acids; LPL: Lipoprotein lipase; HDL: High density lipoprotein; VLDL: Very low density lipoprotein; IDL: Intermediate density lipoprotein; LDL: Low density lipoprotein.

Lipoproteins are complex particles. They have an inner core, composed mainly of triglycerides and cholesteryl esters, surrounded by an outer core, which is a hydrophilic membrane consisting of free cholesterol, phospholipids and apolipoproteins. There are five main fractions of circulating lipoprototeins, each with its own characteristic protein and lipid composition: very low density lipoproteins (VLDL), intermediate density lipoproteins (IDL), low-density lipoproteins (LDL), high-density lipoproteins (HDL) and chylomicrons (CM) which can be further subdivided into subfractions based on their density [94, 95]. As lipid molecules are less dense than proteins, the most distinguishing feature of the main classes is the relative amounts of lipid and proteins, which is reflected in the density forming the basis for the definition of lipoprotein subfractions. Figure 1.5 shows the relationship between lipoprotein sizes and density, and the different parts of the lipoproteins. IDLs have a density and size in-between LDL and VLDLs, while CMs are much larger and less dense than VLDLs.

Chylomicrons are large particles rich in triglycerides, produced by the intestine. Their function is to transport dietary triglycerides and cholesterol to peripheral tissues, and their size depends on the amount of dietary fat. VLDLs are produces by the liver and are rich in triglycerides. They take part in the endogeneous pathway, where they transport lipids to the capillaries, in which triglycerides and cholesterol are taken up by muscles and adipose tissue, and their remnants (IDLs) are transported back to the liver or are synthesized to LDLs (which are enriched in cholesterol) and transported to peripheral tissues. HDLs are synthesized by the liver, and take part in the reverse transport pathway, where they pick up cholesterol in peripheral tissues and deliver it to tissues that need it, to other lipoproteins or back to the liver [95]. LDLs are pro-atherogenic (lead to buildup of cholesterol in the arteries), while HDL is anti-atherogenic, and for that reason HDL is commonly referred to as a good cholesterol, while LDL is sometimes called a bad cholesterol. Apolipoproteins are distributed over all lipoprotein main fractions, however with a varying proportion. They act as ligands for lipoprotein receptors, provide structure to the lipoprotein, guide the formation of lipoproteins and serve as activators involved in the metabolism of lipoproteins [95]. They may be classified as peripheral, or integral, depending on whether they are connected on the outside of the cell membrane, or are permanently embedded in the membrane. Apo-B is the major structural component of VLDLs, IDLs and LDLs. HDLs contain mostly Apo-A1 and Apo-A2, which are peripheral apolipoproteins, synthesized by the liver. Apo-A2 also activates lecithin cholesterolacyltransferase (LCAT) which is responsible for the formation of cholesteryl esters, while Apo-A1 activates hepatic lipase. There has been increased awareness that the lipid picture is more complicated, and that focusing on the subfractions instead of the main fractions gives important additional biological information. For example, Madssen et al. showed that chemotherapy induced an increase in LDL parameters, except for LDL2, which decreased

during the same period, in the serum of breast cancer patients [96]. Small LDLs have been found to be important biomarkers for atherosclerotic diseases [97] and it has been shown that LDL particle size and number provide are strong predictors of cardiovascular diseases [98]. The diverse characteristics of lipoproteins may aid the development of new therapeutic strategies for metabolic diseases [99-101].

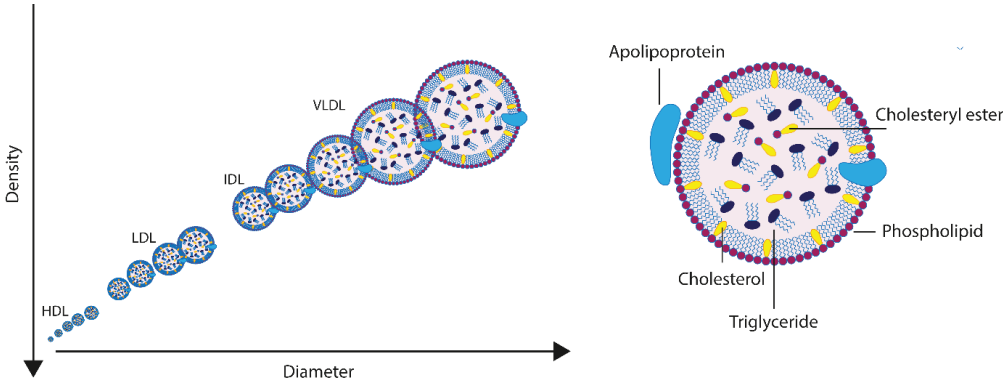


Figure 1.5 The relationship between lipoprotein sizes and density. IDLs have a density and size in-between LDL and VLDLs, while CMs are much larger and less dense than VLDLs. CMs are left out in this illustrations for simplicity. Apolipoproteins can be peripheral or integral, as shown in the figure to the right, where they are either connected to the outer cell membrane, or are embedded permanently in the membrane.

Serum metabolomics

In the context of cancer metabolism, the serum metabolome contains metabolic signals from both the tumor itself and the host organism [102-104]. The metabolic activity is influenced by several endogenous and exogenous factors, as depicted in Figure 1.6 [67-69]. Examples of endogenous factors are genetics, body composition, physical activity, microbiome, endocrine response, mental stress, inflammation and circadian rhythm. Examples of exogenous factors affecting the metabolism are diet, medication and smoking. This further implies that the serum metabolome varies highly across individuals. However, being minimally invasive, serum metabolomics can potentially be applied directly in the clinic for early diagnosis and treatment monitoring. Recent studies have reported associations between circulating metabolite concentrations and breast cancer risk in a prospective approach [105-108]. A study by Kuhn et al. found higher plasma levels of lysophosphatidylcholine 18:0 to be related to a lower risk of common cancers, including breast cancer [108]. His et al. performed prospective analysis of plasma metabolites and breast cancer risk on a cohort consisting of 1624 first primary incident invasive breast cancers and 1624 matched controls, where they concluded that the acetylcarnitine C2 and PC ae C36:3 were associated with risk of breast cancer [105]. Similarly, Léucyer et al. performed a prospective nested case-control study, which revealed plasma metabolites

associated with a risk of developing breast cancer within the following decade [107]. A study by Bro et al. published a model that can predict an increased risk for developing breast cancer 2-5 years after the sample has been taken [106].

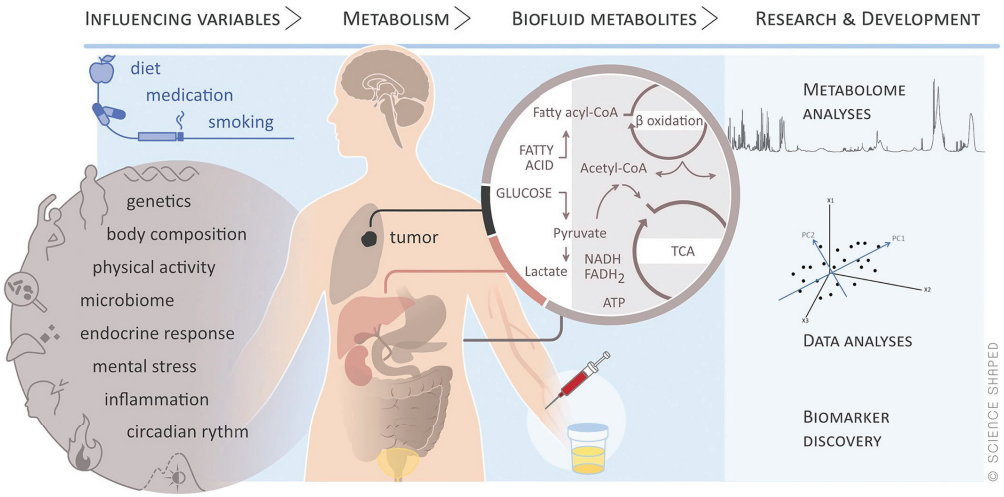


Figure 1.6 The serum metabolome is a highly dynamic system, affected by the preceding levels of the omics cascade, but also external factors such as physical activity, medication usage, age, and body composition. Reproduced with permission from [109].

1.3 Nuclear magnetic resonance (NMR) spectroscopy

1.3.1 Principles of NMR

The basis of nuclear magnetic resonance spectroscopy (NMR) is the concept of spin (I). Spin is a quantum mechanical property that atomic nuclei can possess. The spin depends on the number of protons and neutrons in the nucleus, thus distinct spin configurations will arise for different combinations of these particles. The overall spin of a nuclei is only present for uneven number of protons and neutrons, as even numbers of these particles will experience antiparallel spin pairs which will cancel each other out, giving $I = 0$. Nuclei having a non-zero spin generate their own magnetic moment (μ) proportional to the spin, giving rise to an NMR signal. This includes ^{1}H, ^{13}C, ^{14}N, ^{15}N, ^{19}F and ^{31}P, of which ^{1}H is most commonly used in biomedical applications as it the highest natural abundance and sensitivity. Quantum mechanics states that a nuclei has $2I + 1$ possible orientations, and thus energy states. ^{1}H has spin $I = ½$, and thus two possible energy states.

The magnetic moment μ of a nuclei is usually oriented at random. If placed in an external magnetic field, B_0, μ will align parallel or antiparallel to B_0, in a low ($I = ½$) or high ($I = -½$) energy state,

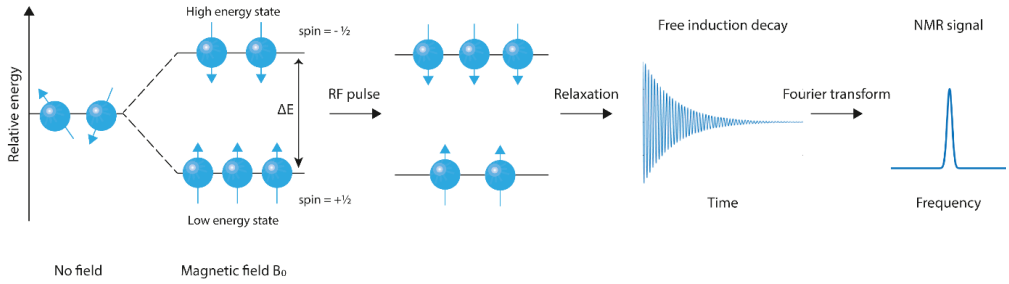


Figure 1.7 The basic principles of NMR. An atomic nuclei will orient in a random direction in the absence of a magnetic field. If an external magnetic field B_0 is applied, nuclei with spin number ½, will either align parallel or anti-parallel with the magnetic field, at a higher or lower energy state, respectively. A slight excess of nuclei will align in the low energy state, causing a net magnetization in the direction of B_0. The energy difference between these two spin states is dependent on the strength of B_0. A radio frequency pulse can excite nuclei to a higher energy state. These nuclei will thereafter return back to the original energy state through a process called relaxation. Energy released in this process can be detected as a signal called free induction decay, which after a Fourier transformation, will result in a spectrum in the frequency domain. RF: Radio frequency; NMR: Nuclear magnetic resonance.

respectively [110]. Given a constant temperature, a small excess of protons will be present at the lower energy than the higher energy state, producing a net magnetization (M_0) along B_0, from the sum of magnetic moments of all protons. The resonance frequency of the nuclei in an external magnetic field, called the Larmor frequency (ω_0), corresponds in the energy difference between the energy states and is given by $\omega_0 = \gamma B_0 / 2\pi$, where γ is the gyromagnetic ratio [111].

If a radio frequency (RF) pulse is applied at the same frequency as the Larmor frequency of the nuclei of interest, nuclei in the low energy state will excite to a higher energy state, disturbing the equilibrium and tilting the magnetization vector M_0 away from B_0. This tilt is dependent on the pulse magnitude and duration and a 90° pulse will flip M_0 with an angle of 90° from the z-axis to the xy-plane. Once the RF pulse is switched off, the excited nuclei gradually return to the equilibrium state with longitudinal (T_1) and transverse (T_2) relaxation times, releasing the absorbed energies. A signal, called the free induction decay (FID), can be detected during this process. The FID can be converted from the time domain to the frequency domain via a Fourier transformation, giving the NMR spectrum [112].

Nuclei of the same type that are in different magnetic environments will experience slightly different magnetic fields due to shielding from surrounding electrons, and will resonate at different frequencies. This frequency deviation due to electron shielding is known as the chemical shift, δ, and nuclei from different molecules appear as peaks at different positions of the spectrum. The chemical

shift is independent of the magnetic field strength, however other factors such as pH and temperature alter the chemical shift. The chemical shift is commonly expressed in terms of parts per million (ppm).

Due to spin-spin interactions, which is the influence by spins of closely located nuclei on the nuclei of interest, peaks may be split into singlets or multiplets. The chemical shift together with the splitting pattern provide information about the molecular structure allowing to identify the compounds. Signal intensity in an NMR spectrum is proportional to the concentration of the nuclei producing the signal, making it possible to quantify detected compounds, after necessary preprocessing. Figure 1.8 shows a representative CPMG spectrum of a serum sample with annotated metabolite peaks.

Transverse magnetization decays exponentially at a rate determined by the transverse relaxation rate constant R_{xy} given by:

$$M_x(t) = M_x(0)e^{-R_{xy}t},$$

where $M_x(t)$ is the x-magnetization at time t and $M_x(0)$ is the initial value [111]. The higher the rate constant, the faster the decay. The reciprocal of the rate constant R_{xy} is the time constant for the decay of transverse magnetization $T_2 = 1/R_{xy}$ called the T_2 relaxation. T_2 may be measured with a spin-echo sequence, such as the CPMG sequence. By running an experiment using a short pulse to form multiple echoes, the decay can be observed and used to estimate T_2 values [112]. Spectra corresponding to each echo can be used to determine the height of the peaks of individual

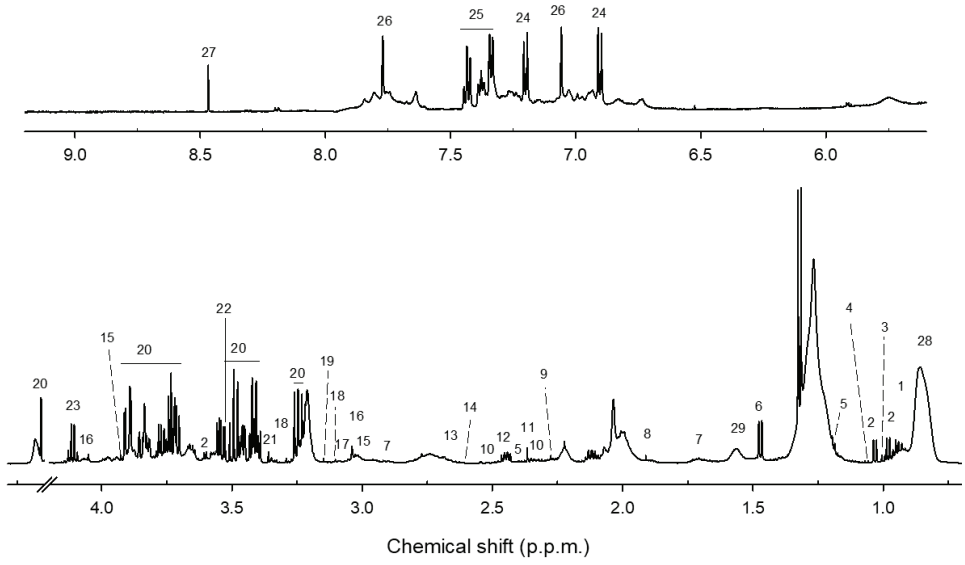


Figure 1.8 A representative CPMG spectrum of a serum sample with annotated metabolite peaks. 1: leucine; 2: valine; 3: isoleucine; 4: dimethylglutarate; 5: tri-hydroxybutyrate; 6: alanine; 7: lysine; 8: acetate; 9: acetoacetate; 10: glutamate; 11: pyruvate; 12: glutamine; 13: citrate; 14: methionine; 15: creatine; 16: creatinine; 17: ornithine; 18: proline-betaine; 19: dimethylsulfone; 20: glucose; 21: methanol; 22: glycine; 23: lactate; 24: tyrosine; 25: phenyllanine; 26: histidine; 27: formate; 28: lipid1; 29: lipid2.

metabolites. These values, and the corresponding times (t) can be used to fit an exponential function to model the decay and thus estimate R_{xy}. The percentage of signal present after a full CPMG experiment is given by $e^{-\tau/T_2}$, where τ is the length of the T_2 filter in the CPMG experiment.

1.3.2 Preprocessing of NMR metabolomics data

Raw NMR spectra are inadequate for statistical analysis as there may be variations in the spectra not related to the biological traits of interest [113]. Experimental inaccuracy can lead to differences in the sample weights of tissues or fluid volumes, which again can lead to higher spectral peaks. Also the composition of the tissue varies across samples, while for urine in specific, different concentrations of metabolites are strongly influenced by the amount of water (dilution) in a given sample. Further, peaks may shift during NMR acquisition due to instability in temperature or pH conditions during the experiments. Peaks appearing at slightly different positions in the ppm scale make comparisons between samples impossible. To remove these unwanted effects, preprocessing of the raw spectra is necessary prior to data analysis [114]. Preprocessing decreases the probability of inaccurate biological interpretations emerging from unwanted error sources, and common preprocessing steps together with their purposes are summarized in Table 1.2.

Table 1.2 Different steps in preprocessing of NMR spectra with their purposes.

Method	**Goal**
Baseline correction	Remove baseline distortions
Removal of water signal and contaminants	Avoid interference with signal of interest
Peak alignment	Correct for differences in chemical shift
Normalization	Correct for differences in concentrations of metabolites in samples
Centering and scaling	Correct for differences in average abundance of metabolites

Baseline correction

Acquired spectra may have a distorted baseline. To correct for this, baseline correction may be applied [115]. A simple method for correcting baseline offset is to subtract the minimum value of each spectra from the spectra. This method shifts the spectra vertically, removing the baseline offset, at the same time keeping the shape of the spectra unchanged. This method may however lead to problems if the spectra includes negative peaks. Another baseline correction method is asymmetric least squares method [116]. This method works by subtracting a spline function following the raw spectra baseline. One disadvantage of this method is, however, that the spline function is highly affected by broad peaks. Small neighboring peaks may thus be highly affected from this baseline correction method.

Peak alignment

Peaks can be shifted from their expected chemical shift due to changes in pH or temperature, inhomogeneous magnetic field or molecular interactions during acquisition. Shifted spectral peaks may be aligned using an alignment method. The icoshift algorithm is an approach based on correlation shifting of spectral intervals [117]. In this algorithm, the spectra are divided into segments, which may be equally distributed along the spectra, or user-defined regions. Each segment is treated individually and the spectra in each segment is aligned to a reference, referred to as the target, such that the correlation between the spectra segment and the reference is maximized, without distorting the shape of the signal. The reference may for example be the spectrum with the highest correlation to all other spectra in a given batch, or the mean of the spectra.

Removal of water signal and contaminants

Water signal appears as a big peak in the spectrum, and varies largely from sample to sample due to varying success of water suppression [118]. Contamination may also occur, which means that a compound, which does not occur naturally in a biological sample, is present. Different sources for a contamination exist, such as poor cleaning of the equipment, medication usage or long-term storage in cryotubes, and the origin might be difficult to identify. A contamination may give rise to one or multiple peaks in the spectra, which may overlap with signals of biologically relevant metabolites. Signals from water and any contaminations are not of interest in metabolomics studies, and should be removed from the spectra prior to normalization, if this is possible. If a signal from contamination is overlapping with signals of interest, correcting for it might be feasible, however it should be performed with caution.

Normalization

As metabolic responses are reflected in differences in concentration of specific metabolites, variations in dilution factors or sample weights should be removed making spectra comparable [119]. Different normalization approaches have been developed, which aim to remove this effect. Mean normalization, also referred to as area normalization, divides each data point by an equal total area of the spectrum. This eliminates variance related to the amount of sample analyzed. Another commonly used normalization method is probabilistic quotient normalization (PQN) [120]. In this approach, the most probable dilution factor is calculated based on the distribution of quotients. The quotients are derived from dividing the spectra to be normalized by a reference spectrum. When analyzing tissue biopsies, lipid peaks arising from adipose tissue, whose intensity may interfere with signals that are related to the biological effects of interest, should be removed [113].

Centering and scaling

Metabolites that are more abundant will generally display larger differences among samples in a batch, than metabolites that are of low abundance. The highly abundant metabolites will thus mask changes in low abundant metabolites, which may be biologically important. Scaling is an operation that aims to balance signal intensity variances that originate from difference in average abundance of metabolites. It is performed variable-wise, unlike normalization, which is performed independently for each spectra [121]. Prior to scaling mean centering is typically performed, which transforms all values so that they vary around zero instead of varying around the mean value. Autoscaling divides each variable by the standard deviation of the variable after mean centering, converting all metabolites to have unit variance. Noise will be given higher influence after applying autoscaling, thus it less suitable for spectra than for quantified metabolites.

Quantification of metabolites and lipoproteins

Metabolites may be quantified from the NMR spectra as the area under each peak is proportional to the concentration of the corresponding metabolite in the sample. The most common approach is to integrate the area under each peak, however recently methods for deconvolution of the NMR spectrum are increasingly used [122, 123].

Metabolite concentrations may be normalized by mean-normalization or PQN normalization, which yields relative concentrations in a batch. These relative concentration changes are informative, however, a direct comparison of significant metabolic changes across studies may be difficult. Further, the presence of lipids highly affects the relative metabolic concentration and thus the correlation between the variables in the data [124]. The advantage of absolute quantified concentrations is that comparison of findings across studies may be performed in a more accurate manner, keeping the original correlations of the biological variables unaltered. Absolute quantification is also more relevant for a clinical application. To perform absolute metabolite concentrations, each peak must be adjusted for the number of protons giving rise to the peak.

NMR spectroscopy is also well suited for the identification and quantification of lipoproteins in serum. Different lipoprotein subfractions have different chemical compositions and sizes [125, 126] giving rise to distinctive NMR signals. Lipoproteins contain triglycerides and cholesterol esters, which give rise to broad peaks at 0.8 and 1.2 ppm, arising from methyl ($-CH_3$) and methylene ($-CH_2-$) groups. The shape of these peaks is determined by the complexity of the composition of lipoproteins in the sample, and thus the envelopes of these peaks may be used for lipoprotein quantification for example through PLS regression [126].

1.4 Data analysis

1.4.1 Descriptive statistics

A variety of descriptive statistics have been used in this thesis for different purposes.

Median percentage change

Often it is the case that the metabolomics data are not normally distributed but rather follow a skewed (unsymmetric) distribution. Comparing the mean values of metabolite levels across groups may thus not be optimal, as the mean is highly influenced by extreme values. The median value is more appropriate as its value is less affected by the extremes. Assume we have a set of observed values. To calculate the median value, the values must first be ordered in an increasing direction. Median is the value which divides the observed values into two equal halves.

The median percentage change can be calculated to compare if a variable has undergone an increase or decrease between two measurements. Given repeated measurements of a variable at two distinct time points, with median values m_{t1} and m_{t2}, at time point t_1 and t_2, respectively. The median percentage change from time point 1 to time point 2 is given by

$$median\ percentage\ change = \frac{m_{t2}-m_{t1}}{|m_{t1}|} \cdot 100\%\,,$$

where || denotes the absolute value.

Coefficient of variation

The coefficient of variation (CV) is a measure of the dispersion in the data in relation to the mean [127]. Given a set of observed values, with a standard deviation, s and a mean value $\bar{x}$, the coefficient of variation is given by $\mathrm{CV} = \frac{s}{\bar{x}} \cdot 100\%$.

Thus, low CVs indicate little variation within the samples, whilst high CVs indicate high variation within the samples. CV is a useful descriptive statistic as, in contrast to the standard deviation, it is independent of the unit in which the measurement has been taken. It thus allows for comparisons across data sets with different units or with widely different means. CV is also a useful tool for comparing the reproducibility of repeated measurements.

Pearson correlation

The Pearson correlation coefficient $\hat{\rho}$ is a measure of how two variables x and y covary.

$$\hat{\rho} = \frac{\sum_{i=n}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{j=1}^{n}(y_j - \bar{y})^2}}$$

The range of values for the Pearson correlation coefficient is from +1 to -1. A value > 0 indicates that there is a positive association between the two variables, whereas a value < 0 indicates that there is a

negative association. A value of 0 indicates no association between the variables. Thus, the strength of the association between the two variables increases with the magnitude of the Pearson correlation coefficient, as shown in Figure 1.9.

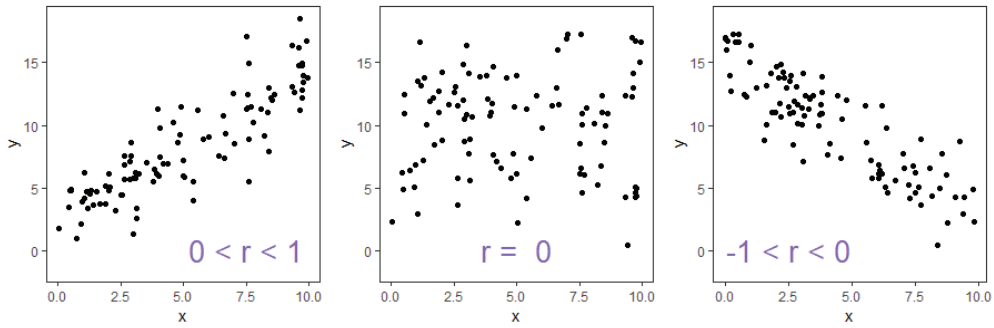


Figure 1.9 Pearson correlation. The magnitude of the Pearson correlation coefficient indicates how strong the association between two variables, while the sign indicates if the association in positive (Figure to the left) or negative (Figure to the right). A Pearson correlation coefficient close to 0 indicates no association (the middle Figure). r: Pearson correlation coefficient ($\hat{\rho}$).

Intra class correlation

The intra class correlation coefficient (ICC) is the ratio of the between-group variation, divided by the total variation in the data (*between*-group and *within*-group variations), and ranges between 0 and 1. A value close to 1 indicates high similarity between values from the same group, while a value close to 0 means that values from the same group are not similar [127] (Figure 1.10).

There is more than one possible method for calculating the ICC. Here, we will concentrate on the method involving a random effect model. Assume we have a random effect model:

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij},$$

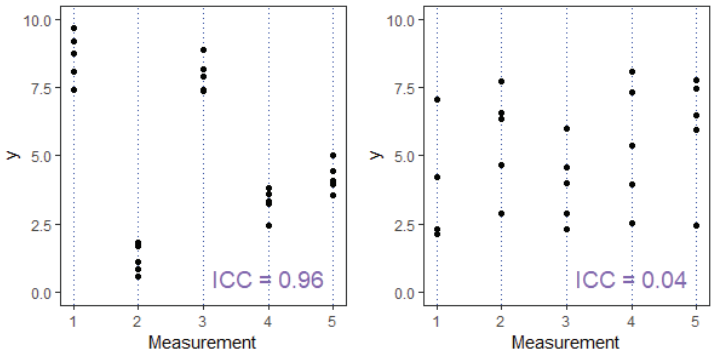


Figure 1.10 Intra class correlation coefficient. High (figure to the left) and low ICC (figure to the right) indicate high and low between-group variation, respectively.

where Y_{ij} is the i-th observation of the j-th group, μ is the overall mean, α_j is random effect specific for group j, and ε_{ij} is a noise term. Further, assume that $\varepsilon_{ij} \overset{i.i.d.}{\sim} N(0, \sigma^2)$, $\alpha_j \overset{i.i.d.}{\sim} N(0, \tau^2)$ and that the α_j's and ε_{ij}'s are mutually independent. The correlation between two observations in the same group, called the *within*-group correlation coefficient is given by

$$Corr(Y_{ij}, Y_{il}) = \frac{Cov(Y_{ij}, Y_{il})}{\sqrt{Var(Y_{ij})Var(Y_{il})}} = \frac{\tau^2}{\tau^2 + \sigma^2}$$

for j≠ l. This becomes the intra class correlation coefficient for inserted parameter estimates of the random intercept model.

1.4.2 Univariate analyses

The significance of differences in individual metabolite levels between two sampling points or groups may be assessed by performing univariate hypothesis testing on the quantified amounts. The null hypothesis, that there is no difference between the two groups, against the hypothesis that there is a significant difference between the groups is tested. Statistical tests can be either *parametric* or *non-parametric*, where the latter has lower statistical power. The choice of the statistical test must be dependent on the nature of the data. To apply a parametric test, underlying model assumptions should be checked. These are normality, homogeneity of variance (homoscedasticity) and independence. The requirement of normality is fulfilled if the data follows a normal distribution, i.e. a symmetric bell-shaped curve. Normality can be assessed by making a quantile-quantile (QQ) -plot of the data or by a normality test, such as the Shapiro-Wilk or Kolmogorov-Smirnov tests. The assumption of homoscedasticity implies that the variance in the data should be stable with no increasing or decreasing trend. Homoscedasticity can either be evaluated by a graphical inspection of the data or by a statistical test, e.g. the Levene's test. Data, which is not normal or homoscedastic can sometimes meet these requirements after a logarithmic transformation. Independency in the context of metabolomics implies that a sample is unrelated to the other sample.

Parametric tests

If the assumptions of normality and homoscedasticity are valid, the parametric t-test can be applied to the data. The t-test assumes that the two groups which are compared follow a normal distribution and compares the difference in the mean values of the two groups. If the samples are independent, an unpaired t-test is performed, otherwise, the paired t-test should be chosen, for example if one has repeated measurements of a variable.

Nonparametric tests

If the data is non-normal, a nonparametric test can be applied. Nonparametric tests do not rely on the data to belong to any parametric family of distributions. Depending whether or not the two groups compared are independent or not, the Mann-Whitney U test (also called the Wilcoxon rank-sum test) or the Wilcoxon signed-rank can be used.

Linear mixed-models

For analyses where there are repeated measurements the responses may be correlated. The repeated measurements can either be clustered (grouped) or longitudinal. Grouped data refers to the case that the data are nested and there is no natural ordering of the units within each group, like patients in different hospitals. Longitudinal data is when data for each individual are observed at multiple time points, like repeated sampling from patients undergoing treatment. Linear mixed models (LMM) are extensions of the linear regression model, which allow for the inclusion of both fixed and random effects [128]. A fixed effect is a parameter that is somewhat systematic, while random effects are parameters that are unsystematic, and which we want to account for, but are not interested in the estimated parameters. In a medical context, the fixed effect might be treatment type, while the random effect might be the patient id.

Mathematically, the response for group i can be expressed as:

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i,$$
$$\boldsymbol{\gamma}_i \sim \mathrm{N}(0, \mathbf{Q}), \quad \boldsymbol{\varepsilon}_i \sim \mathrm{N}(0, \sigma^2\mathbf{I})$$

where $\boldsymbol{Y}_i$ is a random vector of length n_i with the responses in group i, $\boldsymbol{X}_i$ $(n_i \times p)$ and $\boldsymbol{Z}_i$ $(n_i \times q)$ are known fixed-effects and random-effects design matrices (also called regressor or model matrices), $\boldsymbol{\beta}$ is a p-dimensional vector of fixed effects, which is common for all groups, $\boldsymbol{\gamma}_i$ is a q- dimensional vector of random effects used to model correlated responses, and $\boldsymbol{\varepsilon}_i$ is the random error vector [128, 129].

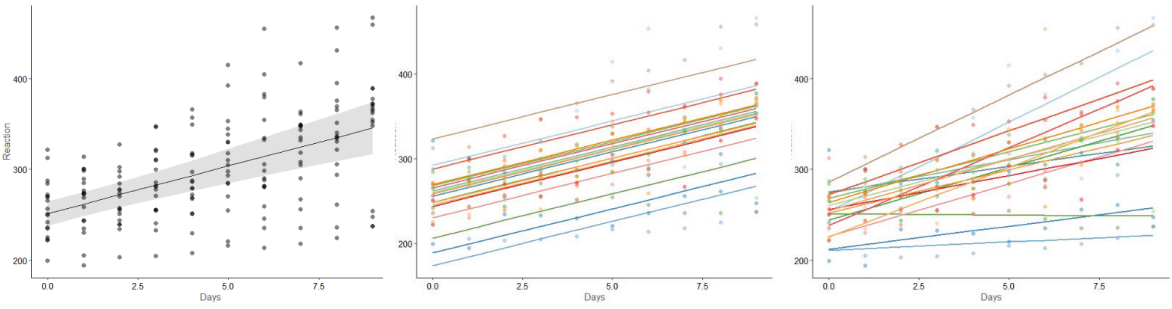


Figure 1.11 An linear regression model (left) a linear mixed model with a random intercept (middle) and a linear mixed model with a random intercept and random slope (right).

A random slope should be included if the relationship between the response variable and a covariate is different for each group i of the random effect matrix, and is modelled by adding an interaction term between the random and fixed-effect matrices, which modifies the $\boldsymbol{Z_i}$ matrix in the equation [129]. Figure 1.11 shows the difference between a linear regression model, a linear mixed model with a random intercept and a linear mixed model with a random intercept and a random slope. In particular, the linear regression model assumes that the mean response is equal for all groups.

Another advantage of linear mixed models, is that it can handle missing observations, meaning that if the response variable of a patient is missing at one time point, that patient is not left out from the analysis.

Logistic regression (LR)

Logistic regression is a regression analysis for binary response variables, i.e. variables that can take on one of two possible outcomes. It is a predictive analysis, and the task is to predict the outcome based on one or multiple independent variables. Let Y_i and $\boldsymbol{x_i} = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{ip})$ denote the response variable and the p independent variables for individual i, respectively. The response variable is typically coded as 0/1, where $Y_i = 1$ corresponds to class 1, success, and $Y_i = 0$ to class 2, failure. The probability that an observation comes from class 1, the probability of success, is $P(Y_i = 1 \,|\boldsymbol{x_i}) = \pi(\boldsymbol{x_i})$, and the probability of failure is $1 - \pi(\boldsymbol{x_i})$.

The model assumes a linear relationship between the covariates and the log-odds of the event $Y_i = 1$. Mathematically, given *p* covariates, logistic regression estimates a multiple linear regression function defined as:

$$logit(\pi(\boldsymbol{x_i})) = log\left(\frac{\pi(\boldsymbol{x_i})}{1-\pi(\boldsymbol{x_i})}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip},$$

where $\boldsymbol{x_i} = (x_{i1}, x_{i2}, \dots., x_{ip})$ are covariates of the i-th observation [130]. This assures that the probability lies in the interval between zero and one as the logistic curve is s-shaped. The parameter β_k for a covariate k determines the rate of increase or decrease of the s-shaped curve, and the sign of β_k indicates whether the curve ascends or descends [131].

An advantage of the logistic regression model is the interpretability associated with the odds, which is the probability of success divided by the probability of failure. The odds of $Y_i = 1$ for observation i is:

$$odds = \frac{\pi(\boldsymbol{x_i})}{1 - \pi(\boldsymbol{x_i})} = e^{\beta^T \boldsymbol{x_i}} = e^{\beta_0} \cdot (e^{\beta_1})^{x_{i1}} \cdot (e^{\beta_2})^{x_{i2}} \cdots (e^{\beta_p})^{x_{ip}}.$$

The odds increase multiplicatively by e^{β_k} for every one-unit increase in x_{ik}, thus the odds at level $x_{ik} + 1$ is equal to odds at x_{ik} multiplied by e^{β_k}, when all other variables are kept constant. When

$\beta_k = 0, e^{\beta_k} = 1$ the odds does not change as x_{ik} changes, and $\beta_k > 0$ and $\beta_k < 0$ correspond to an increased or decreased odds, respectively. For each coefficient in the model, a p-value is returned for testing the hypothesis H_0: there is no association between the covariate and the outcome (odds ratio = 1) vs. H_a: there is an association between the covariate and the outcome. Thus a significant p-value indicates that there is an association present.

When fitting a multivariate logistic regression model to a dataset, often not all of the covariates included will be significant. Keeping all of the variables in the model makes the interpretability of the model more challenging, and might lead to overfitting, especially for small sample sizes. There are various approaches for variable subset selection, aiming at retaining only a subset of the variables that are sufficient for explaining their joint effect on the response, such as forward-, backward-, and best subset selection. These approaches search through multiple candidate models by either adding or removing one and one covariate, and comparing the model fit [132]. Alternatively, a shrinkage method can be applied, such as the lasso, Rigde or elastic net regularization which will retain all variables, but shrink model coefficients which are of less importance [133]. Lasso has the ability to shrink some variables to zero, thus is suitable for subset selection [134]. It has however some limitations: 1) if there is a higher number of variables then observations ($p > n$) it selects at most n variables before it is saturated, and it therefore not well suited; 2) in the presence of high pairwise correlations, the lasso tends to select only one of the variables; and 3) when there is a higher number of observations then variables ($n > p$), if the variables are highly correlated, Ridge regression tends to perform better [135]. Elastic net regularization is a shrinkage method that linearly combines the penalties associated with the lasso and ridge methods [135].

1.4.3 Machine learning methods

Machine learning is a field within the artificial intelligence field, where a computer is given the ability to learn without being explicitly programmed [136]. By feeding a machine learning method with multiple examples (training data) it is able to learn from the data, and generate rules that can be applied to predict outcomes for new samples [137].

For the purpose of multivariate analysis, with n samples over p variables, it is common to condense the data into an $n \times p$ dimensional matrix $\boldsymbol{X}$, where each row and column represents one sample and variable (metabolite), respectively. Multivariate statistical methods can be subdivided into two types: *unsupervised* and *supervised* methods. Unsupervised methods aim to describe naturally occurring patterns in the data, without any knowledge on the response variable. Supervised methods make use of the labeling to create models, which can be used for predicting the outcome for unobserved data.

NMR data are highly collinear, making standard statistical methods unsuitable. In addition, the number of variables (p) often exceeds the number of samples (n). Multivariate methods utilizing the use of latent variables are thus commonly used for analyzing NMR data.

Principal component analysis (PCA)

Principal component analysis is an unsupervised statistical procedure, which uses orthogonal transformation to convert a set of observations of correlated variables into a set of linearly uncorrelated variables, called principal components (PCs). PCs are linear combinations of the original variables, where the first principal component, PC1, is chosen in the direction along which the samples show the largest variation [138], as illustrated in Figure 1.12. The second principal component is orthogonal (thus uncorrelated) to the first principal component in the direction of the greatest remaining variation. In this way, the dimension of the original data set is reduced, and most of the variation in the data is explained by few new variables (PCs). When the number of variables is larger than the number of samples and there is a high degree of collinearity, PCA may greatly reduce the dimensionality of the data without loss of information.

Mathematically, the decomposition of $\boldsymbol{X}$ by PCA results in $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^T + \boldsymbol{E}$, where $\boldsymbol{T}$ and $\boldsymbol{P}$ are the scores and loadings matrices, respectively, and $\boldsymbol{E}$ is a matrix of residual variance which is not a part of the model [138]. Here, ^{T} indicates the matrix transpose. PCA has only one tuning parameter, which is the number of principal components. This is directly related to the residual variance matrix $\boldsymbol{E}$ and as the number of included PCs in the model increase, the residual variance decreases. The optimal number of PCs can be determined from a scree plot, where the variance explained is plotted against the number of principal components in the model. Typically, the curve should be steep at first, then have a bend, followed by a flat part. The number of PCs that are associated with the bend should be chosen. PCA is an unsupervised method, it is often used as an explanatory tool for detecting underlying patterns and outliers.

The results of a PCA are displayed in a score plot in the new coordinate system defined by the PCs. This is a scatter plot, where each sample is projected down onto the new coordinate system, represented by a point. By coloring the points by groups of interest, the amount of separation between the groups can easily be observed. Also, any underlying patterns in the data can be revealed. Complementary to the score plot is the loadings-plot, where the contribution of each original variable, its weight, on the PCs is shown. The scores and loadings can alternatively be plotted together in a biplot, to rapidly overview the correlation between sample scores and variable loadings.

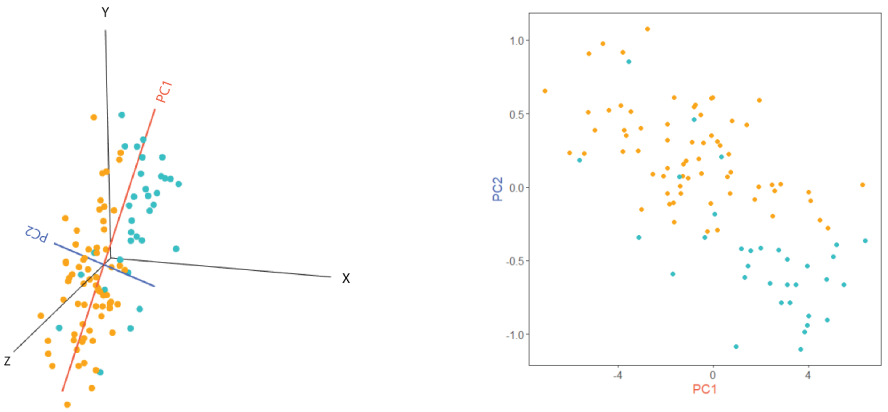


Figure 1.12 Principal component analysis. The figure to the left shows data in the original 3-dimensional space, where points corresponding to two different groups are colored in different colors. The figure to the right shows the same points propagated down to a 2-dimensional space, where the highest variance along PC1.

Partial least squares discriminant analysis (PLS-DA)

Partial least squares discriminant analysis (PLS-DA) is a supervised statistical procedure, with many similarities to PCA. PLS-DA however is a combination of dimensionality reduction and discriminant analysis, in one algorithm [139]. The original data matrix $\boldsymbol{X}$ is decomposed into a set of new orthogonal variables, called latent variables (LV).

The general underlying model for a PLS-DA can be written as:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^T + \boldsymbol{E}$$
$$\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{Q}^T + \boldsymbol{F}$$

Where $\boldsymbol{X}$ is the original data matrix and $\boldsymbol{Y}$ is the response matrix with class memberships. Further, $\boldsymbol{P}$ and $\boldsymbol{U}$ are the score matrices for $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. Similarly, $\boldsymbol{P}$ and $\boldsymbol{Q}$ are the loading matrices for

$\boldsymbol{X}$ and $\boldsymbol{Y}$, and $\boldsymbol{E}$ and $\boldsymbol{F}$ are the error matrices. $\boldsymbol{X}$ and $\boldsymbol{Y}$ are decomposed in such a way that the covariance between $\boldsymbol{T}$ and $\boldsymbol{U}$ is maximized. Thus, the first LV will place the highest weights on the variables most strongly related to the response. There exist a number of algorithms for performing PLS-DA, of which SIMPLS [140] is most commonly implemented in statistical software, as it performs faster than the traditional non-linear iterative partial least squares (NIPALS) algorithm.

Model tuning and interpretation is similar to that of PCA. As PLS-DA is a supervised method, the number of LVs should be chosen based on the prediction accuracy through cross validation. As for PCA, the results of a PLS-DA model can be visualized by score and loading plots. The score plot shows

each sample projected to the new coordinate system spanned by the principal components, and the loading plot shows the weight of each original variable on the score.

PLS-DA can be extended through orthogonal projections to latent structures (OPLS-DA). This results in separating the predictive from non-predictive variation, in cases where there are more than one LV in the model. This extension improves the interpretability of the model, but not the predictive performance [141].

Multilevel PLS-DA

Multilevel PLS-DA is an extension of PLS-DA, effective for paired data structures in a multilevel study [142]. The multilevel approach consists of two steps. First, the variation between individuals is separated from the variation within the samples. Secondly, PLS-DA analysis is performed on the within subject variation.

Let $\boldsymbol{T1}$ be a matrix containing all observations from sampling point 1, and similarly $\boldsymbol{T2}$ a matrix containing all observations from sampling point 2. Then the *between* subject variation, $\boldsymbol{M}$, in the multilevel model is defined by

$$\boldsymbol{M} = \frac{1}{2}\begin{bmatrix}\boldsymbol{T1} + \boldsymbol{T2} \\ \boldsymbol{T1} + \boldsymbol{T2}\end{bmatrix},$$

and is thus the average of the two sampling time points.
The within subject variation, $\boldsymbol{W}$, is defined by

$$\boldsymbol{W} = \begin{bmatrix}\boldsymbol{T1} - \boldsymbol{T2} \\ \boldsymbol{T2} - \boldsymbol{T1}\end{bmatrix},$$

And is the net difference between the two sampling time points.

The second step in the multilevel approach is to analyze the within and between subject variations separately using PLS-DA. This is an effective tool for longitudinal data, where there are two or more multivariate measurements per subject. This approach can also be used for evaluating a treatment intervention, where $\boldsymbol{T1}$ would then contain observations from treatment group, and $\boldsymbol{T2}$ from control group. Exploiting the paired data structure provides complementary information about the diversity and abundance of the treatment effect within different subgroups across the study population as well as the intrinsic differences between the individuals in the study.

Random forest

A decision tree is made up by binary splitting of the decision space following a top-down approach. At each step of the algorithm a split of the decision is made to maximize the homogeneity of the two new parts. This is best illustrated by an example. Suppose we want to predict the body mass index (BMI) of an individual, given only the waist and the hip circumferences of the individual. The data is

shown in Figure 1.13, where the points are colored according to the associated BMI category, as defined by the world health organization (WHO) [143].

Figure 1.14 shows the process of fitting a classification tree to the data. The top split assigns individuals having hip circumference < 97.5 cm to the left branch. The predicted BMI category for these individuals is the majority vote of the BMIs of all points enclosed in this space, namely underweight. Thus if the algorithm was stopped here, all individuals with a hip circumference < 97.5 cm would be classified as underweight, while all above as having a normal weight. The second split is at hip circumference = 112.5 cm. Now, in addition to the previous split, all individuals with a hip circumference above 112.5 cm are classified as pre-obese. This splitting is repeated, until a stopping criteria is met. The bottom figure of Figure 1.14 (bottom-left) shows the final classification tree fit to the data, and the associated partitions of the predictor space. Each rectangle in the predictor space is associated with a terminal node, and the points along the tree where the predictor space is split, are referred to internal nodes. The lines connecting the nodes are named branches. To make a new prediction, one starts at the top of the tree, and follows the criteria met at each internal node, until a terminal in met. This example showed a classification tree. However, it could easily be transformed to a regression tree, by replacing the BMI categories with the exact BMI values. In such a case, the predicted value at each terminal node would simply be the mean value of all observations in the corresponding space. A decision tree is intuitive and easily interpretable, however its performance is often worse than for other regression and classification approaches. Trees are also non-robust, meaning that a small change in the data can cause a large change in the final estimated tree.

Figure 1.13 The decision space for BMI as a function of the variables WaistCirc and HipCirc, which are waist and hip circumferences of some individuals, respectively.

By combining multiple decision trees the performance can be much improved. A random forest (RF) is an ensemble of B decision trees, where each tree in the forest is constructed on a bootstrapped sample of the original data [133, 144, 145]. The bootstrapped sample is created by sampling with replacement from the original data [146]. This approach largely reduces the variance and increases the robustness of the model. Another difference between a decision tree and a random forest, is that when growing a tree in the random forest, only a random subset, m, of the p predictor variables can be considered at each split of the tree.

The objective for this is to break up the correlation between the trees in the forest, and to hinder strong predictors to dominate the model. When a random forest is grown, the final prediction for a new observation is the majority vote, or the mean value, of the individual trees, for classification and regression problems, respectively. A random forest has two tuning parameters. The size of the tree ensemble, B, must be sufficiently large, for the error rate to reach a stable minimum, however it will not overfit if B is increased [133]. The number of predictor variables, which can be considered at each split, must also be chosen. It is typically set to m= $\sqrt{p}$ or $p/3$ when building a random forest of classification or regression trees, respectively. However, the best choice will depend on the problem, and a small value of m will typically be beneficial when there is a large number of correlated predictors in the data [145]. A disadvantage of RFs is that the interpretability is lost at the expense of an increased prediction accuracy, compared to decision trees. The overall summary of importance of each predictor can however be obtained using the Gini index or root-mean squared error (RSS) for classification and regression problems, respectively, and is often represented in a bar plot. The total amount that the Gini index or RSS is decreased by splits over a given predictor variable, averaged over all B trees can be recorded. A high value will indicate an important predictor.

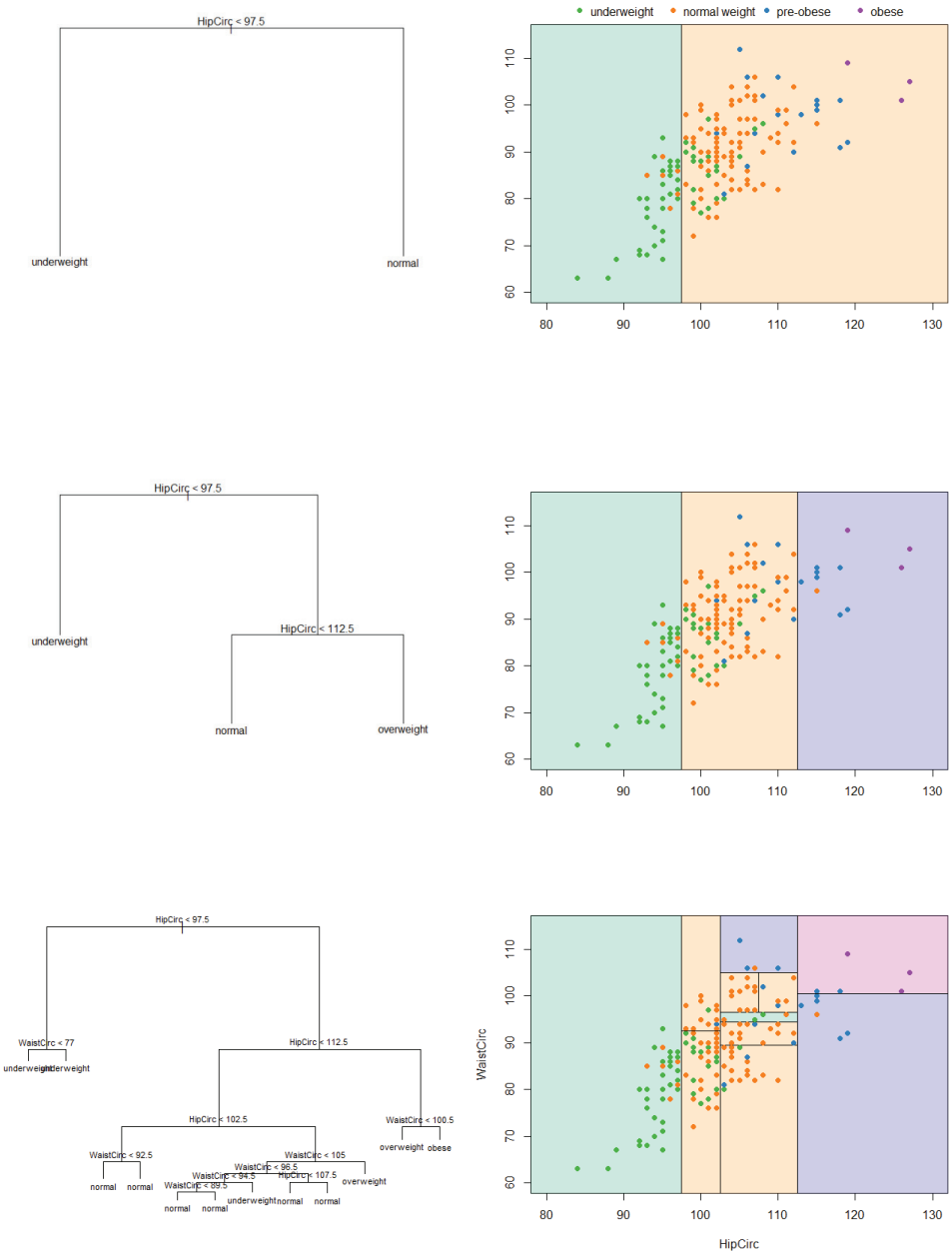


Figure 1.14 The process of fitting a classification tree to the decision space in Figure 1.13.

Gradient boosting machines

Similarly to random forests, gradient boosting machine is an ensemble of decision trees. However, boosting does not involve bootstrap sampling, instead trees are grown sequentially, where each tree is fit on a modified version of the original data set [133, 145, 147, 148]. Given a current model, a decision tree is fit to the residuals from the model. This new tree is added to the existing tree ensemble, and the residuals are updated. The model is thus improved in areas where it performs poorly. The boosting approach is thus to learn slowly, where each individual tree is rather small. There are different variants of boosting algorithms, which will not be discussed in detail here [136]. Depending on the exact implementation of the algorithm, there can be different tuning parameters. Importantly, the distribution function must be chosen to be Gaussian or Binomial, depending on if it is a regression or classification problem, respectively. The size of the individual trees must also be set, governed by the interaction depth, d. A model with only tree stumps ($d = 1$) often outperforms models with higher interaction depths, if the number of trees is sufficiently large. To make the algorithm learn even more slowly, a shrinkage parameter, λ, can be set [148]. By setting λ to be below one, at iteration of the algorithm a shrunk version of a new tree is added to the tree ensemble. It is typically set to 0.01 or 0.001, depending on the problem. A very small λ might require using a very large ensemble size for achieving good performance. Another important tuning parameter is the number of trees in the ensemble, B. This number is typically considerably higher than for a random forest, as each individual tree is here very small. These tuning parameters can be chosen during the training process by cross validation. Similarly as to RFs, relative variable influence in the model is given as an output.

1.4.4 Deep learning

An artificial neural network (ANN) is a ML algorithm inspired by the human brain. It consists of layers of interconnected neurons [137]. These neurons are computational units, which can receive some input, process it, and propagate some output downstream, as visualized in Figure 1.15. The input layer takes in the raw data, which is then propagated to the nodes of the hidden layer. At each node, a weighted sum of the input values is calculated, biased and passed through an activation function, before being propagated further through the network [149]. For neurons in the first layer, the input will be the raw data. For neurons further downstream in the network, the input will be the output of the previous layer. The type of activation function and the number of nodes in the hidden layer can be varied. Some commonly used activation functions are rectifying linear units (ReLU), hyperbolic tangent, the unit step function and identity.

Deep Learning (DL) is a class of ML methods based on artificial neural networks, with more than one hidden layer. In general, the more number of hidden layers and neurons per layer, the more complex

models can be built. Opposed to other ML algorithms, a DL system is self-teaching, meaning that it learns by filtering information through multiple hidden layers [150]. This is achieved through back-propagation, where during training of the network, the output of the network (predicted value) is compared to the truth, and the weights are adjusted [149]. The mismatch between the truth and the predicted value is calculated through the loss function, which takes the prediction of the network, and the true target, and computes a distance score. There exist different methods for the optimization of neural networks, where stochastic gradient descent (SGD) is one of the most commonly used. This algorithm searches for the minimum point of the loss function, the point for which the gradient of the loss function is 0 [137]. This gives the analytical combination of the weight values that yield the smallest possible loss. The algorithm is as follows: 1. Run the network with a set of initial weights to obtain predicted values 2. Compute the loss of the network 3. Compute the gradient of the loss with regard to the network's parameters 4. Move the parameters in the opposite direction from the gradient: new weights = old weights – step*gradient, thus reducing the loss slightly [137]. The steps 1-4 are repeated until a further decrease in the loss cannot be obtained. The backpropagation algorithm applies the chain rule to compute the gradient values. It starts with the final loss value, and iterates backward from last to first layer, computing the gradient one layer at a time.

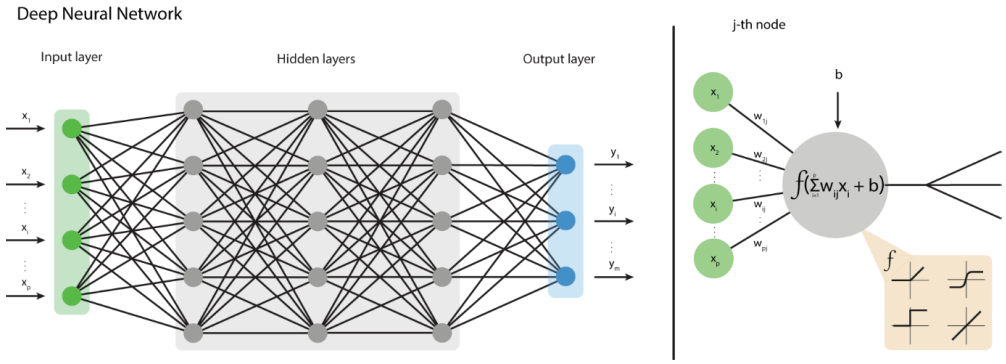


Figure 1.15 A schematic representation of the architecture of a deep neural network consisting of an input layer, multiple hidden layers and an output layer. The input is passed through the first hidden layer, and a weighted sum is passed through an activation function before it is down-propagated to the next layer. When training, the output is compared to the true value, and the weights are adjusted by back-propagation. The figure to the right depicts one node of the neural network, where a weighted sum of the values of the nodes in the proceeding layer is passed through an activation function (in the yellow box).

Table 1.3 Activation and loss functions for different types of problems. This table is reproduced from [137].

Problem type	**Last-layer activation**	**Loss function**
Binary classification	Sigmoid	Binary cross-entropy
Multiclass, single-labeled classification	Softmax	Categorical cross-entropy
Multiclass, multilabel classification	Sigmoid	Binary cross-entropy
Regression to arbitrary values	None	Mean-squared error (MSE)
Regression to values between 0 and 1	Sigmoid	MSE or binary cross-entropy

Data should be variable-wise normalized (auto-scaled) before feeding into a neural network, as it makes the learning of the network easier. The architecture of a neural network will depend on the problem at hand. There are different types of activation functions and loss functions that should be used depending on the problem type, as summarized in Table 1.3. In addition, the number of layers, the number of nodes in the each layer, and the number of epochs must be chosen. An epoch is a full iteration over all of the training data. If the task is to classify data into a large number of categories, any intermediate layer should not be smaller than the size of the final layer, to avoid information bottlenecks which may drop out relevant information. The model can be regularized further by adding dropout, weight regularization or by adding or removing variables. Dropout consist of randomly setting to 0 a number of outputs (governed by the dropout rate, which is the fraction of features that are zeroed out) of a layer during training [136]. Adding dropout will make the network more generalizable and reduce overfitting. Weight regularization is a way of forcing the weights of the network to take on only small values, thus putting a constraint on the complexity of the network, and is done by adding a cost associated with having large weights to the loss function.

Although all deep neural networks have the basic structure of an artificial neuron network, several types of networks have evolved, enabling a wide range of potential applications, such as speech recognition, medical image analysis and natural language processing. Some of the most commonly used DL algorithms are Feed-forward Neural Networks (FFNN), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN) and Probabilistic Neural Networks (PNN) [151, 152].

A FFNN is a type of neural network where connections between nodes do not form a cycle. For making a prediction given a new set of input variables, these variables are passed through the hidden layers to the output layer, thus they are fed forward through the network. A FFNN is commonly trained using backpropagation and is then referred to as a multi-layer perceptron.

CNNs are most commonly applied to analyzing or classifying images, but are also widely applied to natural language processing (NLP) [153]. The structure of a CNN is one input layer (typically an 2D

image, or a matrix of words for NLP) and one output layer, as well as multiple hidden layers [154]. The hidden layers in a CNN consist of a series of convolutional layers. The convolution is a sliding dot product, which works like a filter. Often it is a 2D matrix, of much smaller dimensions than the image (input layer). Initially it starts at the top left corner of the image, and the dot product of the image and the convolution is calculated. Then the convolutional kernel is moved to the right by one pixel, and the same action is performed. In this way, only important information is filtered through this convolutional layer. This filtering depends on the features one wants to extract and highlight. Following the convolutional layers, pooling is performed, which is responsible for reducing the dimensions of the convolved feature, thus reducing the computational power. Two common types of pooling are max pooling and average pooling. In both types, the convolved layer is divided into non-overlapping squares, and the operation is performed on each of these. Max pooling returns the maximum value of the values in each square, while average pooling returns the average value of this square. The final step of the CNN is to flatten the output, which is fed into a regular NN. In this way, non-linear combinations of the extracted features of interest can be modelled [155].

1.4.5 Model selection and validation

A central challenge in machine learning is that a model must perform well on new, previously unseen data, not only on the training data [156]. An optimal multivariate model is one that has a high prediction accuracy, but that also has the ability to generalize well to new data. *Overfitting* is the case when the model is too closely fit to the data on which it is trained. The model includes details and noise, making its prediction accuracy for new, unseen, data low. *Underfitting* is the case when the model is too general and the model fails to capture necessary information in the data, also resulting in a model with poor performance [133]. To avoid over- or underfitting, multivariate models need to be carefully selected and validated. The optimal way to achieve this is to divide the original data set into three parts: a training set, a test set and a validation set [149]. The training set is used to build an initial model, using a supervised learning method. The fitted model is then employed to the test set, where its performance is evaluated. This step includes refining of the model by tuning model parameters, and is the crucial step for avoiding over- and underfitting. The final model is then employed on the validation set, to get a correct prediction or classification accuracy. In medicine, however, the data sets are often too small to allow for such a procedure.

Cross validation

Cross validation is a model validation procedure, in which the original data set is divided into k parts, where each part is in turn hold out in the model building and is referred to as a test set. The remaining $k - 1$ parts of the original data set together make up the training set, and are used for model building [133]. Thus, every part of the original data set gets to be in a test set exactly once and in a training set

$k-1$ times. The overall model performance is then an average of the performance of the k models. Cross validation can also be used to determine optimal parameters for a multivariate model, when applied on the training set.

A double CV procedure can be employed when the number of observations is large enough. In double CV, an inner loop, in which parameter tuning is performed, is nested in an outer loop for prediction performance assessment [157]. The outer loop makes its assessment using a test set, which has been completely left out during the model tuning, thus giving a more accurate estimate of the prediction performance, as illustrated in Figure 1.17.

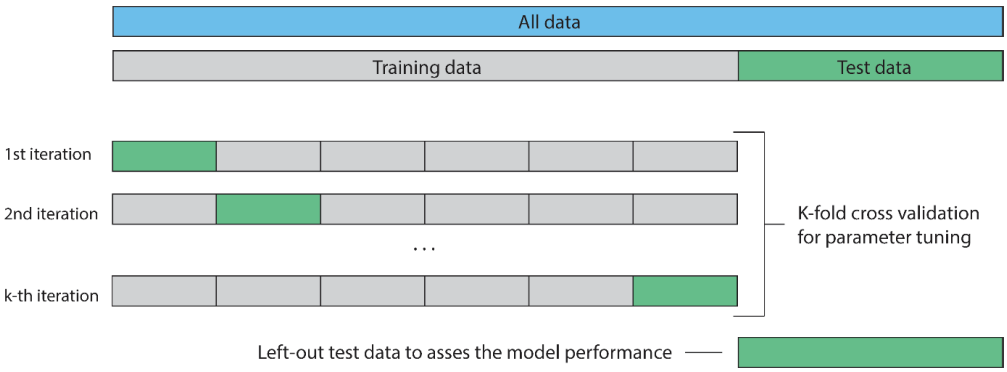


Figure 1.16 Schematic representation of a double cross-validation procedure. The inner loop consists of a K-fold CV, for parameter tuning. The model giving the best performance is chosen and tested on a test set to assess the predictive power of the final model, in the outer loop. The test set samples are unseen in the model optimization procedure.

Permutation testing

Permutation testing can be employed to verify the statistical significance of a multivariate model. The idea behind permutation testing is to test if a given classification model is better than random. Suppose we have a data matrix $\boldsymbol{X}$, with observations from each individual in one row, a corresponding response vector $\boldsymbol{y}$, and a multivariate prediction model with classification error E. In permutation testing, the original class labels ($\boldsymbol{y}$) of the samples in a data set ($\boldsymbol{X}$) are shuffled among the individuals [158]. In the permuted data set, the original class labels are thus assigned to the individuals in a random order. On this permuted data set, a new multivariate model is fit in the exact same way as previously, and the classification error E_i is calculated. This procedure is repeated N times, where the number of permutations should be set to at least 1000. The proportion of classifications equal to or better than the original classification are used to calculate the p-value assessing the statistical significance of the original model:

$$p = \frac{\sum_{i=1}^{N} I(E_i \geq E)}{N},$$

where I is the indicator function, which takes the values 1 or 0, depending on weather the condition in the parenthesis is true or not, respectively. A p-value less than or equal to 0.05 or 0.001 indicates a significant model with 95% or 99% confidence, respectively.

Multiple testing correction

In metabolomics, statistical tests are often performed on each metabolite in parallel. When performing multiple statistical tests a fraction of the tests will result in false positives, meaning that significant p-values will be obtained merely by chance. Increasing the number of tests increases the probability of obtaining false positive findings. There is a number of methods for correcting the p-values from multiple tests, among them the Bonferroni [159, 160] and Benjamini-Hochberg adjustments [161]. The Bonferroni adjustment implies multiplying each p-value by the number of tests performed. This adjustment controls the familywise error rate, which is the probability of obtaining one or more false discovery. The Bonferroni adjustment is thought to be too strict for metabolomics studies, thus the alternative, Benjamini-Hochberg adjustment, is thus commonly used. The Benjamini-Hochberg adjustment provides a balance between false positives and false negatives by controlling for the false discovery rate.

1.4.6 Study designs

Case-control studies are designed to measure disease occurrence and its association with an exposure in a retrospective manner (Figure 1.18). In a case-control study, two groups with a different outcome (e.g. healthy/ill, dead/alive) are drawn from a larger cohort and compared. Cases are identified and are given a response value $y = 1$. Suppose we have n cases that together make up the case group. A control group is created, by randomly selecting a group of individuals with a different outcome than the cases. These are given the response value $y = 0$. The cases can be matched to the controls on different levels (e.g. age or sex), if the disease is highly dependent on a variable. However, matching should be restricted, so that the control group is representative for a random population. Increasing the number of controls over the number of cases, up to a ratio of about 4 to 1, increases the statistical power of the analysis. Once a group of cases and controls is created, differences in conditions prior to the event (outcome) can be compared to identify predictive biomarkers of the disease. Case-control studies have several advantages: they can study rare diseases or diseases with a long latency, they are relatively inexpensive, existing records can be used, multiple risk factors can simultaneously be examined and are relatively quick to conduct. The main disadvantage of case-control studies is that it is susceptible to selection biases and confounding. Selection bias occurs when cases or controls are selected on criteria related to the exposure of interest [162]. Confounding may be present if other risk factors are present that were not measured. In addition, the rate of disease in exposed and unexposed individuals cannot be determined.

Cohort studies provide an alternative study design and can be either retrospective (looking into the past) or prospective (looking in to the future) [163]. In a *retrospective* cohort study subjects are selected from a homogeneous population, based on exposure to a specific risk factor. The outcome of interest (typically disease incidence rate) is compared between the exposed and unexposed group. A *prospective* cohort study is an approach that recruit and follow participants who share a common characteristic over a period of time. A part of the cohort will be exposed to a specific risk factor during the follow-up period. By measuring the outcome of interest at the end of the follow-up, the impact of this variable can be assessed. The disadvantage of this study design is that some diseases will require a long follow-up period for the event to occur, and is vulnerable to a high loss to follow-up rate. *Cross-sectional* studies examine the data (exposures to risk factors and disease status) at one particular time point in a defined population.

The study designs discussed above are all observational studies, which mean that the study participants are only observed, and with no action from the researcher. Study designs including an intervention are called experimental studies, of which randomized controlled trials (RCT) are the most commonly used. An RCT aims to reduce bias when testing the effectiveness of new treatment strategies by randomly allocating individuals to receive one of several interventions. One of these interventions is either a standard practice, placebo or no intervention at all, and individuals receiving this intervention make up the control group. The efficacy of the other interventions is assessed in comparison to the control.

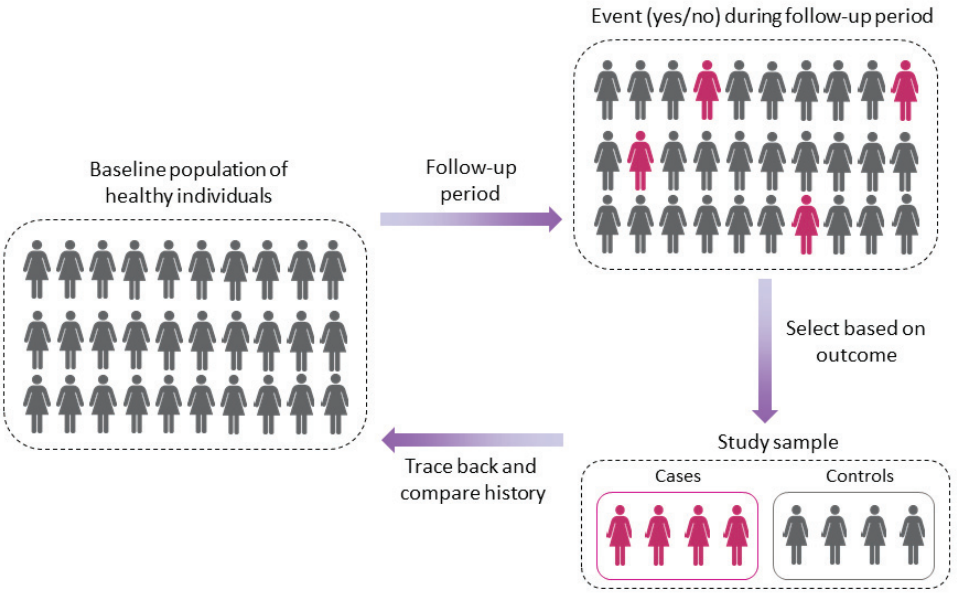


Figure 1.17 A case-control study design. Study participants which have developed the outcome variable of interest are identified, and a suitable control group is selected. Their past exposure to suspected risk factors is then compared to investigate associations between the risk factor and the outcome.

2 Aims of thesis

The overall aim of this thesis was to use metabolic profiling to predict individual breast cancer risk, and to investigate treatment-induced metabolic changes in breast cancer patients undergoing treatment. This overall aim was accomplished by individual aims for each of the studies included in this thesis:

1. Assess the effect of repeated freeze and thaw cycles on the NMR measured serum and urine metabolome to ensure the quality of metabolomics findings from samples which have been thawed more than once (Paper II).

2. Determine the systematic metabolic effect of neoadjuvant therapy in breast cancer patients, and establish differences in serum and tissue metabolism between treatment responders and non-responders (Paper I).

3. Identify predictive biomarkers for individual breast cancer risk, by metabolic characterization of serum samples of healthy women who later developed breast cancer (Paper III).

3 Materials and methods

This thesis includes three papers related to metabolomics of biofluids and tissue samples. Paper II is a methodological article evaluating the effects of repeated freeze and thaw cycles on NMR-measured lipoproteins and metabolites in serum and urine samples. Paper I is a scientific article assessing the serum metabolic response of breast cancer patients undergoing treatment. Paper III presents the predictive potential of circulating serum biomarkers for future breast cancer. Table 1.3 summarizes the material and methods used in Papers I-III of this thesis.

Table 3.1 Material and methods used in Papers I-III of this thesis.

	Paper I NeoAva cohort	Paper II Healthy volunteers	Paper III HUNT2 biobank
Materials	Serum and tissue samples of 132 women measured before, during and after treatment.	20 serum and 20 urine samples from donors at the St. Olav's hospital.	Blood samples from over 1153 patients who later developed breast cancer and a control group of equal size.
Methods	NMR spectroscopy	NMR spectroscopy	NMR spectroscopy
Quantification	Spectral integration in Matlab	Absolute quantifications (Bruker Biospin)	Absolute quantifications (Bruker) and spectral integration in Matlab
Data analysis	PCA; PLS-DA; multilevel PLS-DA; univariate tests	PCA; CV; ICC; univariate tests	PLS-DA; RF; LR; GBM; DL

3.1 Study cohorts

3.1.1 NeoAva cohort

For Paper I, serum and tissue samples obtained within the Neoadjuvant Avastin in Breast Cancer (NeoAva) cohort were included. The NeoAva study is a multicenter randomized phase II clinical trial to evaluate the effect of Avastin (bevacizumab) in combination with neoadjuvant treatment regimes. Women with non-metastatic, large (diameter ≥ 2.5 cm), HER2 positive- breast tumors were recruited in the years 2008- 2012 at St. Olav University Hospital in Trondheim and Oslo University Hospital (Radium Hospital and Ullevål Hospital). Nonfasting serum was collected at four sampling time points: prior to treatment (TP1), 12 weeks into treatment (TP2), at surgery (TP3) and six weeks after surgery (TP4). Tissue samples were in addition obtained at TP1-TP3, by ultrasound-guided needle biopsies

(TP1 and TP2) or from the surgically removed tumor (TP3). All patients had signed a written informed consent and the study has been approved by the Regional Ethics Committee (REK) and the Norwegian Medical Agency.

Treatment protocol

Patients in Paper I were administered neoadjuvant chemotherapy according to Norwegian guidelines, and were randomized to additionally receive bevacizumab. Chemotherapy consisted of four cycles of anthracyclines in the form of FEC100 (5-fluorouracil 600 mg/m^2, epirubicin 100 mg/m^2, cyclophosphamide 600 mg/m^2) every three weeks, followed by 12 weeks of taxane-based therapy (weekly infusion of paclitaxel 80 mg/m^2 or four cycles of docetaxel 100 mg/m^2). Bevacizumab was administered once every three weeks (15 mg/kg) in parallel with chemotherapy for half of the patients. Docetaxel treatment was changed to paclitaxel for the majority of patients due to toxicity. For these patients, the dose of bevacizumab was changed to 10 mg/kg every two weeks. Figure 3.1 shows the treatment regime and experimental set up of the study.

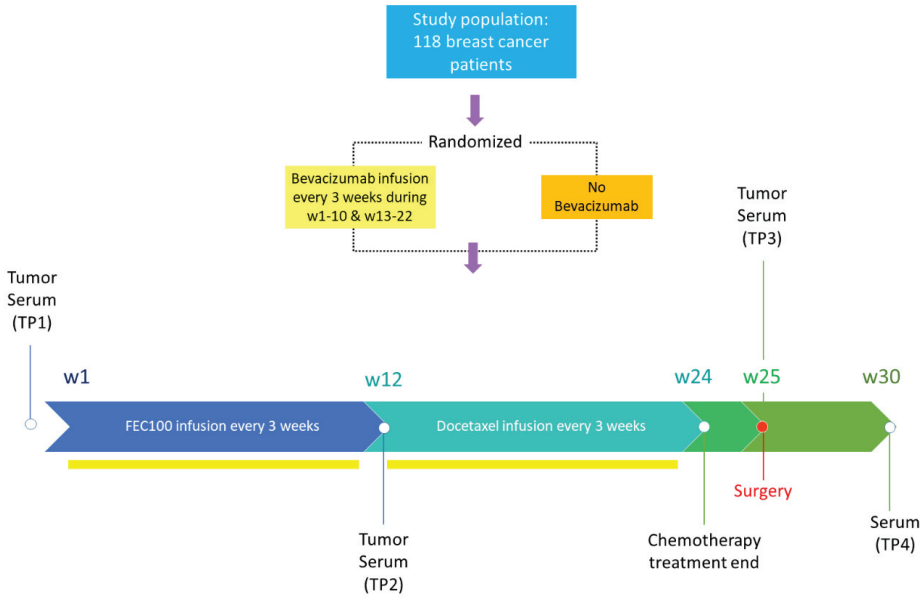


Figure 3.1 Diagram showing the treatment regime and experimental set up of the NeoAva study.

Prognostic measures and survival evaluation

For Paper I, the residual cancer burden (RCB) was used as a response measure. It is a continuous index, combining the pathological measurement of the primary tumor with nodal metastases, thus taking into account the size and cellularity of the tumor, as well as number and size of lymph node

metastases. RCB can be divided into four classes, where class 0 is equivalent to pathological complete response (pCR), meaning that no cancer or lymph node metastasis is present. Due to a limited patient cohort, patients having an RCB class 0 or I were classified as having a good response, while patients having a RCB class II or III were classified as having a poor response to treatment.

Patients deceased within 5 years after diagnosis were classified as non-survivors, whereas patients surviving ≥ 5 years were classified as survivors.

3.1.2 Freeze/Thaw cohort

For paper II, non-fasting serum (n = 20) and spot-urine (n = 20) samples were obtained from two sets of anonymized healthy female and male adult donors. Blood samples were obtained from the Blood bank at St Olav's Hospital (Trondheim, Norway) and urine samples were obtained from anonymized donors. Each serum and urine sample was divided into five aliquots before following the work flow illustrated in Figure 3.2. The five aliquots of each sample were subjected to 1-5 freeze-and thaw cycles before NMR analysis. The need for a formal ethical approval was discussed with REK Central Norway, and found to be unnecessary as this was a quality control study using completely anonymized samples from healthy volunteers.

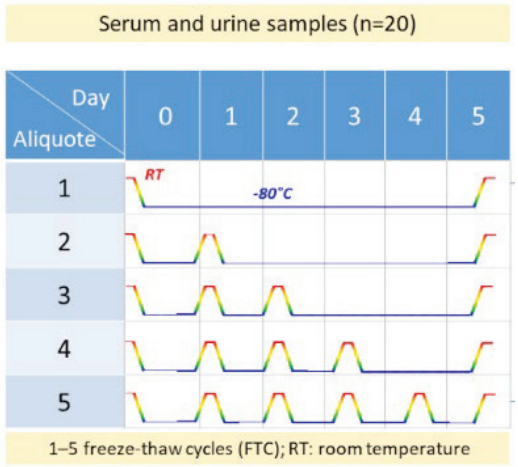


Figure 3.2 Study design for evaluating the effect of freeze and thaw cycles on samples in Paper II.

3.1.3 HUNT2 biobank

For Paper III, serum samples (n = 2306) from the second wave of the Trøndelag Health Study (The HUNT Study) were included. The HUNT study is a database of questionnaire data, clinical measurements and biological samples from inhabitants in the Northern part of Trøndelag collected over four waves, from 1984 onwards. HUNT2 was the second wave and data was collected in the years

1995-1997. All participants at age 20 or older donated a blood sample. For this paper, serum samples were collected through a nested case-control design, from healthy female participants. By matching the data from HUNT2 with the Cancer Registry of Norway, all female participants which later developed breast cancer were identified and chosen as cases for this study. For each case, an age-matched (using intervals of 5 years) control was selected at random. The controls remained breast-cancer free until data collection in 2018. All participants of the HUNT study have signed a written informed consent, and the study has been approved by the REK.

3.2 NMR protocol

3.2.1 Sample preparation

All serum samples were stored at -80 °C until NMR analysis. Thawing took place at 4°C (Paper I) or at room temperature (Paper II and III), samples were then turned upside down a couple of times to remove the freezing gradient. Equal amounts (150 µL) of serum and buffer (D20 with 0.075 mM Na_2HPO_4, 5 mM NaN_3, 3.5 mM trimethylsilylpropanoic acid (TSP), pH 7.4) were mixed and transferred to 3 mm NMR tubes for analysis.

Urine samples were stored at -80 °C until NMR analysis. Urine samples were thawed at room temperature. Thawed samples were centrifuged at 12121 g for 5 minutes at 4°C. 540 µL of supernatant was mixed with 60 µL buffer (1.5 mM KH_2PO_4 in D_20, 0.1% TSP, pH 7.4). Urine and buffer mixtures were transferred to 5 mm NMR tubes.

Spectral acquisition from tissue samples was completed prior to the start of work described in this thesis [164]. In short, biopsies were snap-frozen after collection and stored at -80 °C. Sample preparation took place on a dedicated work station ensuring that the samples were kept frozen during the preparation. Approximately 4 mg of tissue were cut to fit into 30 µL inserts containing 3.0 µL of 24.29 mM sodium formate in D_2O. Each insert was set into a MAS zirconium rotor with a diameter of 4 mm.

3.2.2 Spectral acquisition

All NMR analyses of biofluids performed locally at NTNU were carried out on a Bruker Avance III spectrometer operating at 600 MHz (Bruker BioSpin GmbH, Rheinstetten, Germany) equipped with a 5 mm QCI CryoProbe. Sample handling and data acquisition were automatically performed using SampleJet sample changer and Icon-NMR on Topspin 3.5 (Bruker BioSpin). For serum and urine samples, NMR spectra were recorded using a one-dimensional nuclear overhauser effect spectroscopy pulse sequence (noesygppr1d), using 96k data points and 30 ppm spectral width, and two-dimensional JRES spectra (jresqpprqf). For serum samples, Carr Purcell Meiboom Gill (cpmgpr1d) spectra with

water presaturation were acquired at a temperature of 37 °C in addition. To aid in peak annotation, 2D HSQC and HMBC were recorded from selected samples.

HR MAS MR spectra were acquired from the tissue samples on a Bruker Avance Avance III spectrometer (Bruker Biospin GmbH, Germany) equipped with a $^{1}H/^{13}C$ MAS probe. For all samples, one-dimensional Carr Purcell Meiboom Gill (cpmgpr1) spectra with presaturation were acquired at a temperature of 5 °C to minimize tissue degradation.

3.2.3 Spectral preprocessing and metabolite quantification

For Paper I, preprocessing of the tissue NMR spectra had been completed prior to the work described in this thesis, and the details have been described by Euceda et al, 2017 [164]. In short, spectra were baseline corrected, peak aligned using the icoshift algorithm and normalized by probabilistic quotient normalization (PQN) [120] after the removal of lipid signals. Quantified metabolites were normalized by PQN. For serum NMR spectra, the free induction decays (FIDs) were Fourier-transformed to 128 K real data points after modification by an exponential line broadening factor of 0.3 Hz. Spectral data was further preprocessed in Matlab R2017b (The Mathworks, Inc. Natick, USA) [165]. The left peak of the alanine doublet at 1.47 ppm was used as a chemical shift reference, and the region of interest was defined to be between 0.2 and 9.2 ppm, excluding the water region (4.3-5.0 ppm). Baseline correction by asymmetric least squares (AsLS) was attempted, but did not show an improvement, thus uncorrected baseline was used for further analysis. The spectra were peak aligned using the icoshift algorithm [117] and normalized by mean normalization. Metabolite peaks were identified using the human metabolome database, published literature and an in-house overview over previously assigned spectral peaks in serum based on 2D HSQC acquisitions, and the STOSCY algorithm [166]. Quantification was performed by integrating fixed spectral regions corresponding to each peak. For metabolites with more than one identifiable peak, the mean value of the multiple peaks were calculated and used for further analysis. In total, 29 distinct peaks were quantified (27 metabolites and two lipid signals).

For Paper II, 32 scans were recorded and the free induction decays (FIDs) were Fourier-transformed to 128k real data points after modification by an exponential line broadening factor of 0.3 Hz. Serum and urine metabolite concentrations were automatically quantified using Bruker B.I. Quant-PS™, and Bruker B.I. Quant-UR™ methods, respectively. These methods are based on algorithms developed for fitting predefined proton signals [167]. In total, 26 serum and 50 urine metabolite concentrations were quantified. Serum lipoproteins were quantified using the commercial Bruker IVDr Lipoprotein Subclass Analysis (B.I.LISA™) method from Bruker BioSpin. This method yields 112 quantitative lipoprotein parameters: the concentrations of lipids [cholesterol (CH), free cholesterol (FC), triglycerides (TG), and

phospholipids (PL)] in serum, and in four main lipoprotein classes: VLDL, IDL, LDL, and HDL, as well as 15 subclasses (VLDL 1−5, LDL 1−6, and HDL 1−4). Simultaneously, the concentrations of apolipoproteins (Apo-A1, Apo-A2, and Apo-B) in serum and two main classes (HDL and LDL) and 10 subclasses (HDL 1−4 and LDL 1−6) are quantified. The model also returns 12 calculated parameters, including ratios of LDL-CH/HDL-CH and Apo-B/Apo-A1, and 10 particle numbers (particle numbers of total serum, VLDL, IDL, LDL, and LDL 1−6). The density range of lipoprotein subfractions is continuous, and the subfractions referred to in this thesis are in accordance with the subfractions, with the corresponding density ranges, as defined in protocols from Bruker BioSpin: LDL1: 1.019-1.031 kg/L, LDL2: 1.031-1.034 kg/L, LDL3: 1.034-1.037 kg/L, LDL4: 1.037-1.040 kg/L, LDL5: 1.040-1.044 kg/L, and LDL6: 1.044-1.063 kg/L. HDL1: 1.063-1.100 kg/L, HDL2: 1.100-1.112 kg/L, HDL3: 1.112-1.125 kg/L, and HDL4: 1.125-1.210 kg/L.

For Paper I, measurements of lipoprotein parameters and methods for absolute quantification were not available, thus relative metabolite concentrations were used. Newly developed software for absolute metabolite quantification by Bruker enabled the use of absolute metabolite and lipoprotein concentrations for Paper II. For Paper III, absolute metabolite concentrations were used, by combining automatic quantification using the Bruker developed software in addition to our own in-house routine for quantification for metabolites.

Approximately half of the samples for Paper III were analyzed at MR Core facility, NTNU, while the second half was shipped to Bruker BioSpin GmbH, Germany, for analyses. In order to achieve absolute metabolite concentrations, raw CPMG spectra were imported into Matlab, and areas under the spectral regions corresponding to distinct metabolites were integrated. These integrals were corrected for the number of protons giving rise to the signals, and were adjusted for T2 relaxation times.

To obtain T2 relaxation times, we performed experiments on three separate serum samples, and modelled the exponential decay for each signal separately, based on the area under the signal. The details of the cpmg sequences are: D20 = 0.0003 and VClist = [2; 4; 8; 16; 50; 100; 200; 350; 500; 650; 800; 1000; 2000; 4000; 6000; 10000]. The exponential decay was modelled using a two-component exponential function, yielding a separate component for the decay of the lipid signals and the metabolite decay. To obtain a better fit of the exponential function to the data, which tended to overestimate the decay, every second point at the beginning of the decay (points 2, 8, 50, 200, 500 and 800 of the VClist variable) was left out when fitting the line, giving more weight to the smaller value closer to zero. The T2 values varied across these three samples, due to their slightly different metabolic compositions. Therefore, the signal (spectral integral) was corrected based on the mean value of the T2 values of the three serum samples.

Signals arising from the same metabolite were thereafter averaged, giving a total of 28 quantified metabolites. The concentration of glucose was set equal to the automatically quantified glucose concentration (Bruker B.I. Quant-PS^{TM}) and the remaining metabolite concentrations were scaled accordingly using the same factor, thus giving absolute metabolite concentrations. The analyses of lipoprotein parameters was performed as described for Paper II, after correcting for the signal arising from the contamination, as described in detail in Chapter 4 of this thesis.

3.2.4 Quality control samples

For Paper II and III, quality control (QC) samples were prepared and run in parallel with serum and urine samples of interest. These QC samples were prepared by pooled serum or urine samples of anonymous donors. The main purpose of the QC samples was to assess the reproducibility of the NMR method, and to be able to detect an instrument malfunction as soon as possible.

For Paper II, the QC samples were in addition used to compare the variability associated with preparing and running an identical sample several times, with the variability of the FTC samples. For Paper III, the QC samples were used to evaluate batch differences of metabolite concentrations between samples run in our local lab and those which were shipped to and analyzed in Germany.

3.3 Data analysis

3.3.1 Imputation of missing data

For Paper II, metabolite concentrations below a threshold for quantification, the limit of detection (LOD), were not reported, thus the automatically quantified metabolite data had missing values. Metabolites with concentrations > 0 for more than 30% of serum or urine samples were classified as quantifiable, while the rest were excluded from the analysis. For the included metabolites, zero values were replaced by half of the lowest detected value of the corresponding metabolite.

3.3.2 Univariate data analysis

For Papers I and II, due to non-normality of the serum metabolites, the nonparametric Wilcoxon-signed-rank test was used to test the significance of the changes in serum metabolite levels between the different sampling points. P-values were adjusted using the Benjamini-Hochberg procedure to correct for multiple comparisons. Statistical significance was considered for adjusted p-values ≤ 0.05. The univariate tests were performed in R 3.5.0 (R Foundation for Statistical Computing) [168].

For Paper II, the coefficients of variation and intra-class correlations were calculated to assess the degree of variation within samples subjected to multiple freeze-and-thaw cycles, from individual patients. These analyses were performed in R.

In Paper III Student t-tests were performed to compare the baseline characteristics for the study cohort, i.e. to assess if there are variables which at baseline were significantly different between the cases and controls. These tests were performed in R.

3.3.3 Multilevel and multivariate analysis

For Paper I and II, PCA was used to explore naturally occurring groupings. The number of PCs was selected using residual explained variance plots. PCAs were performed in Matlab 2017b [165] using the PLS Toolbox 8.6.2 (Eigenvector Research Inc., U.S.A.) [169].

For Paper I, PLS-DA was employed to build classification models for different clinical variables. The models were fitted and validated using 10-fold cross-validation, repeated 20 times. The optimal number of LVs was chosen to be the number of LVs corresponding to the first minima in the cross-validated classification errors. Furthermore, permutation testing (with 1000 repetitions) was employed to verify the statistical significance of the PLS-DA models.

Multilevel PLS-DA was performed in Paper I, to assess the treatment effect on the serum metabolome, after the removal of the between-patient variation. These analyses were performed in Matlab 2017b using the PLS Toolbox 8.6.2.

For Paper III, multiple machine learning methods were employed. PLS-DA was performed in Matlab 2020 using the PLS Toolbox 8.6.2. Further, RF, GBM and LR was performed in R 4.0.0 using the randomForest v4.6-14 [144], gbm v2.1.8 [170] and the stats v3.6.2 [168] packages, respectively. For LR, also glmnet was used for the adaptive lasso procedure [171]. For DL, Python 3.8 was used, making use of the Keras library [172] and numpy [173], pandas [174, 175] and scikit-learn [176] packages.

4 Summary of papers and additional results

4.1 Paper I

Assessing treatment response and prognosis by serum and tissue metabolomics in breast cancer patients.

Julia Debik, Leslie R. Euceda, Steinar Lundgren, Olav Engebraaten, Øystein Garred, Elin Borgen, Hedda von der Lippe Gythfeldt, Tone F. Bathen and Guro F. Giskeødegård

Journal of Proteome Research: 2019 Oct 4; 18(10):3649-3660. doi: 10.1021/acs.jproteome.9b00316

Patients with locally advanced breast cancer have a worse prognosis compared to patients with localized tumors and require neoadjuvant treatment before surgery. The aim of this study was to characterize the systemic metabolic effect of neoadjuvant chemotherapy in patients with large primary breast cancers and to relate these changes to treatment response and long-term survival.

This study included 132 patients with large primary breast tumors randomized to receive neoadjuvant chemotherapy with or without the addition of the antiangiogenic drug Bevacizumab. Tumor biopsies and serum were collected before and during treatment and, serum additionally 6 weeks after surgery. Samples were analyzed by nuclear magnetic resonance spectroscopy (NMR).

Correlation analysis showed low correlations between metabolites measured in cancer tissue and serum. Multilevel partial least squares discriminant analysis (PLS-DA) showed clear changes in serum metabolite levels during treatment (p-values ≤ 0.001), including unfavorable changes in lipid levels. PLS-DA revealed metabolic differences between tissue samples from survivors and nonsurvivors collected 12 weeks into treatment with an accuracy of 72% (p-value = 0.005); however, this was not evident in serum samples.

Our results demonstrate a potential clinical application for serum-metabolomics for patient monitoring during and after treatment, and indicate potential for tissue NMR spectroscopy for predicting patient survival.

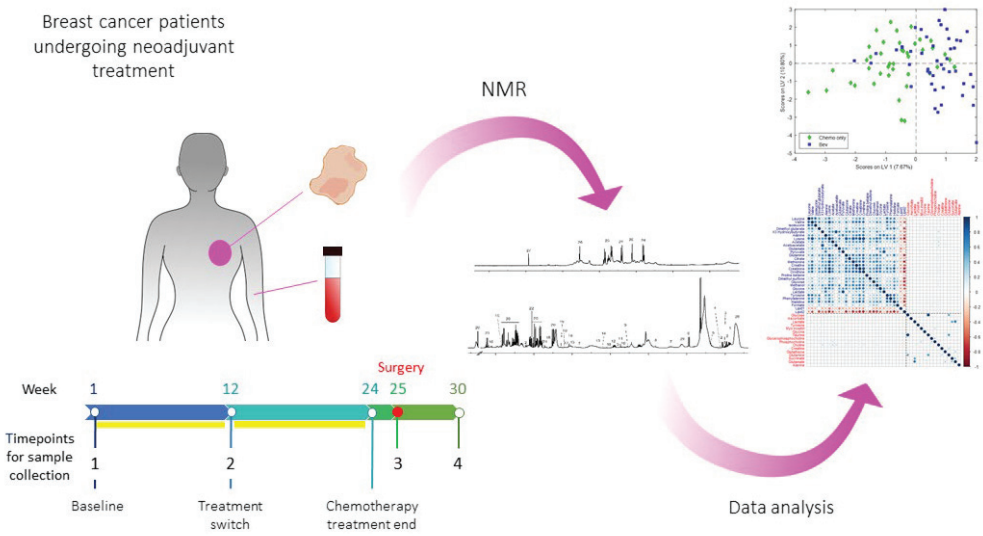


Figure 4.1 Graphical abstract. This figure summarizes the study design of the NeoAva study and data analysis.

4.2 Paper II

Effect of repeated freeze-thaw cycles on NMR measured lipoproteins and metabolites in biofluids.

Feng Wang[*], Julia Debik[*], Trygve Andreassen, Leslie R. Euceda, Tonje H. Haukaas, Claire Cannet, Hartmut Schäfer, Tone F. Bathen[#], Guro F. Giskeødegård[#]; [*]shared first authorship; [#]shared last authorship

Journal of Proteome Research: 2019 Oct 4; 18(10): 3681-3688. doi: 10.1021/acs.jproteome.9b00343

Metabolic profiling of biofluids by Nuclear Magnetic Resonance (NMR) spectroscopy serves as an important tool in disease characterization, and its accuracy largely depends on the quality of samples. We aimed to explore possible effects of repeated freeze-thaw cycles (FTCs) on concentrations of lipoprotein parameters in serum and metabolite concentrations in serum and urine samples. After one to five FTCs, serum and urine samples (n=20) were analyzed by NMR spectroscopy and 112 lipoprotein parameters, 20 serum and 35 urine metabolites were quantified by a commercial analytical platform.

Principal component analysis showed no systematic changes related to FTCs, and samples from the same donor were closely clustered, showing a higher between-subject variation than within-subject variation. The coefficients of variation were small (medians of 4.3%, 11.0% and 4.9% for lipoprotein parameters and serum and urine metabolites, respectively). Minor, but significant accumulated freeze-thaw effects were observed for 32 lipoprotein parameters and one serum metabolite (acetic acid) when comparing FTC1 to further FTCs. Remaining lipoprotein and metabolite concentrations showed no significant change.

In conclusion, five FTCs did not significantly alter the concentrations of urine metabolites and introduced only minor changes to serum lipoprotein parameters and metabolites evaluated by the NMR-based platform.

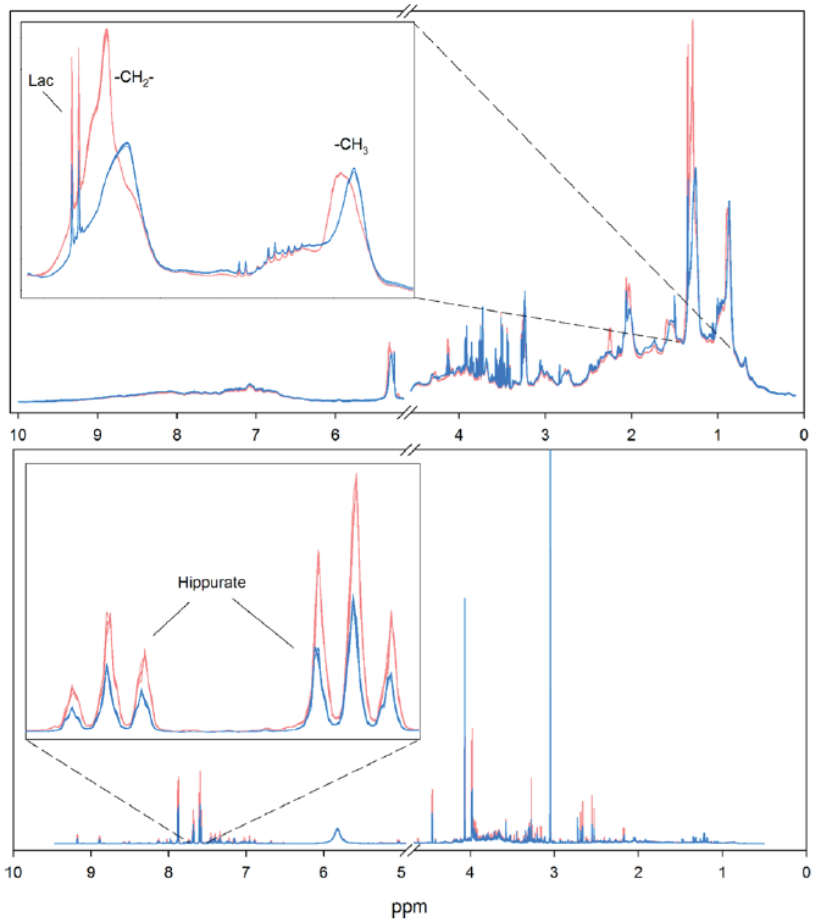


Figure 4.2 NMR spectra of serum and urine samples. Serum (top) and urine (bottom) spectra from two representative donors, colored in orange and blue, respectively. All five spectra, one from each FTC are plotted. For the serum spectra, the area in focus shows part of the spectral region where the lipoprotein signals appear. The spectral positions of the –CH_2- and –CH_3 signals reflect the lipoprotein particle size. The area in focus for the urine spectra shows signals from hippurate as an example.

4.3 Paper III

Serum metabolic profiling for assessment of breast cancer risk in women participating in the HUNT2 study.

Julia Debik, Hartmut Schaefer, Trygve Andreassen, Feng Wang, Fang Fang, Claire Cannet, Manfred Spraul, Tone F. Bathen, Guro F. Giskeødegård

Background

Breast cancer is the most common cancer in women worldwide and early diagnosis is of vital importance. The aim of the present study was to investigate if serum metabolic profiles of healthy women could contribute to predict the risk of developing breast cancer in the future and to gain a better understanding of the etiology of the disease.

Methods

A nested case-control study within the Trøndelag Health Study (HUNT study) was performed, including 1153 participants diagnosed with breast cancer after inclusion into HUNT and 1153 age-matched controls. Using nuclear magnetic resonance spectroscopy (NMR), 28 metabolites and 112 lipoprotein subfractions were quantified from pre-diagnostic serum samples.

Results

Logistic regression identified significant associations between multiple circulating molecules and a future breast cancer 0-22 years after serum collection. However, multivariate prediction models could not accurately distinguish between breast cancer cases and controls.

Conclusions

There were significant positive associations between triglycerides in VLDLs, free cholesterol and Apo-A2 in HDLs, acetate, and valine with long-term breast cancer risk, while there were significant inverse associations between total amount of Apo-A1, free and esterified cholesterol in VLDLs, phospholipids in HDLs and glycine with long-term breast cancer risk.

4.4 Additional results related to Paper I

Paper I gives insight into the serum metabolic changes during neoadjuvant treatment of breast cancer patients, based on a cohort of 132 patients and four sampling time points. Most of the results published in this paper are based on PLS-DA models. Attempts were made to increase the prediction accuracies for the models in Paper I, by applying RFs instead of PLS-DA.

The results of the RFs for predicting the same outcomes as the ones presented in Paper I are summarized in Table 4.1. Details about the implementation of the PLS-DA models can be found in the published article. For RFs, the number of trees (ntree) was set to 500 to ensure that the error has reached its minimum, while the optimal number of variables tried at each split (mtry) in the trees was found by comparing the out-of-bag error rate for all possibilities of mtry from 1 to 29 (the number of metabolites). The out-of-bag error rate is calculated by averaging the error of the predictions of left-out samples from each bag (bootstrapped data) and is equivalent to cross validation. The sensitivity and specificity of the RF classifiers were calculated in a similar way, based on the predictions of the left-out samples from each bag. The model metrics are averaged over 20 repetitions of model fitting and evaluation. The PLS-DA analysis were performed in Matlab 2017b [165] using the PLS toolbox [169], while RFs were performed in R 3.5.0 [168] using the randomForest package [144].

Table 4.1 Summary of prediction models, comparing the performance of PLS-DA models and RFs. The accuracies are averages over 20 cross-validated classification errors.

			PLS-DA		RF	
		N (class1 / class2)	Sens./Spec. (%)	Accuracy	Sens./Spec. (%)	Accuracy
Bev. Treat / Chemo treat. only	TP2	89 (46 / 43)	58 / 70	64 %	62 / 63	63 %
	TP3	93 (46 / 47)	60 / 57	59 %	50 / 52	51 %
	TP4	86 (44 / 42)	67 / 47	57 %	55 / 54	52 %
RCB class 0+I / RCB class II+III	TP1	89 (30 / 59)	27 / 44	36 %	22 / 64	60 %
	TP2	89 (28 / 61)	33 / 63	48 %	NaN / 68	67 %
	TP3	93 (31 / 62)	58 / 57	58 %	37 / 67	62 %
	TP4	86 (28 / 60)	65 / 73	69 %	0 / 69	65 %
5 year survival	TP1	89 (81 / 8)	5 / 70	37 %	91 / NaN	91 %
	TP2	89 (81 / 8)	48 / 81	64 %	91 / NaN	90 %
	TP3	93 (85 / 8)	43 / 79	61 %	92 / NaN	91 %
	TP4	86 (79 / 7)	23 / 73	48 %	92 /NaN	92 %

4.5 Contamination problem related to Paper III

For paper III, 2306 serum samples from HUNT2 study were analyzed. The spectral acquisition revealed a contamination present in all of these samples, which interfered with one of the lipid signals. This lipid signal arises from methyl (-CH_3) groups at approximately 0.88 ppm, mainly from triglycerides and esterified cholesterol within the lipoprotein particles. From this lipid signal detailed information about lipoprotein subfractions is extracted through the B.I.LISA^{TM}. Figure 4.3 shows the raw spectra of the first 1200 HUNT samples, in the area in which this contamination peak appears.

Because of this contamination, lipoprotein subfractions as reported during spectral acquisition of the serum samples, could not be used directly. Exclusion of these variables during data analysis is associated with a reduction of 112 variables, which are valuable for giving a precise picture of the lipid metabolism.

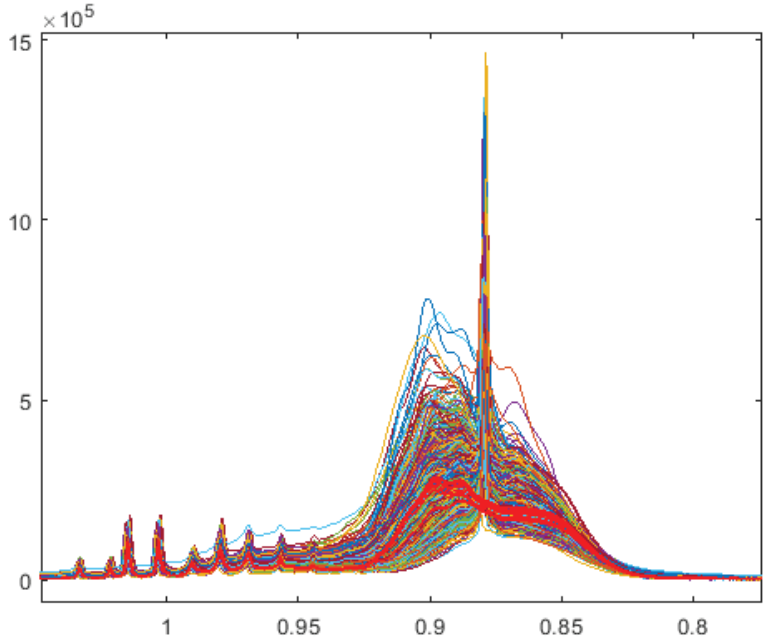


Figure 4.3 Raw spectra from approximately 1200 serum samples from the HUNT2 cohort, and 20 QC samples run in parallel (in red), in the region where the lipid peak arising from –CH3 appears in the spectra. The peak at approximately 0.88 ppm comes from neopentyl glycol and interferes with the signal from the lipid. The QC samples have no such peak.

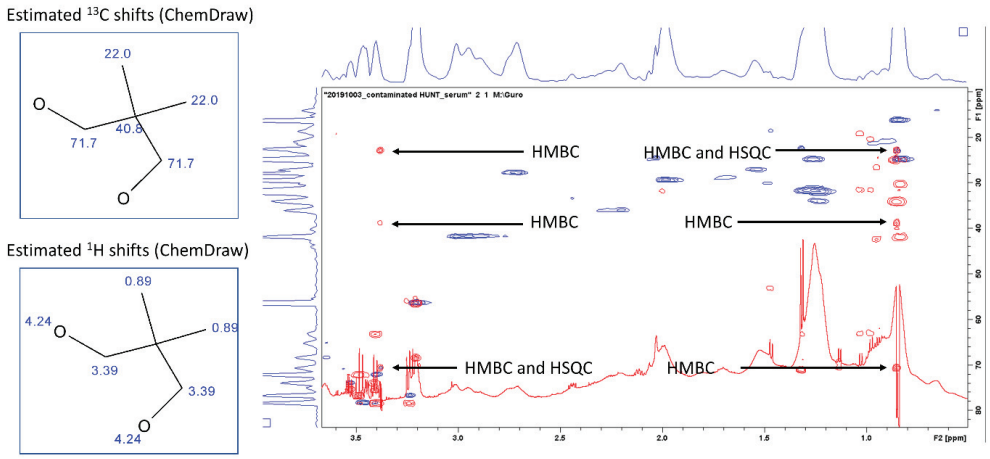


Figure 4.4 Results from 2D NMR experiments, verifying that the contamination comes from the compound neopentyl glycol, verified by comparing to the molecular structure of the compound. HMBC: Heteronuclear multiple-bond correlation spectroscopy; HSQC: Heteronuclear single-quantum correlation spectroscopy.

Through 2D experiments (Figure 4.4) we were able to identify that the peak was arising from the compound neopentyl glycol. Both HSQC and HMBC experiments showed crosspeaks at 0.85 / 23 and 3.38 / 71 ppm (^{1}H/^{13}C), indicating that protons at 0.85 and 3.38 ppm are coupled to carbons at 23 and 71 ppm through both one-bond and multiple bonds. This points to a highly symmetric molecule like neopentyl glycol. Further evidence were found through high agreement with calculated chemical shifts (ChemDraw, Alfasoft). Effort was made to find out the source of this contamination, however without luck. We believe the contamination is either from the blood collection tubes or from the tubes in which the blood has been stored. Identification of the compound corresponding to the contamination peak made it possible to perform spiking experiments, where known amounts of the contaminant were added to clean serum samples (from anonymous blood donors at the St.Olav's University hospital).

The results of the spiking experiment showed that, although the peak arising from neopentyl glycol consistently appears in the same position at the ppm scale, the majority of the quantified lipoprotein subfractions quantified by B.I.LISA^{TM} were largely affected. The subfractions associated with the low density lipoproteins were mostly affected, with a high overestimation of the particle numbers. However, as these lipoprotein subfractions are not calculated independently from one another, for example overestimated values for the subfractions L1CH, L2CH and L3CH resulted in underestimates of L4CH, L5CH and L6CH. In addition, the shape of the lipid signals varies across individuals, and thus the contamination affects different lipoprotein variables for different lipid profiles, making the correction for the contamination peak not straight forward. This is shown in Figure 4.5, where clean (in blue) and spiked spectra (in red) are shown for four of the donors.

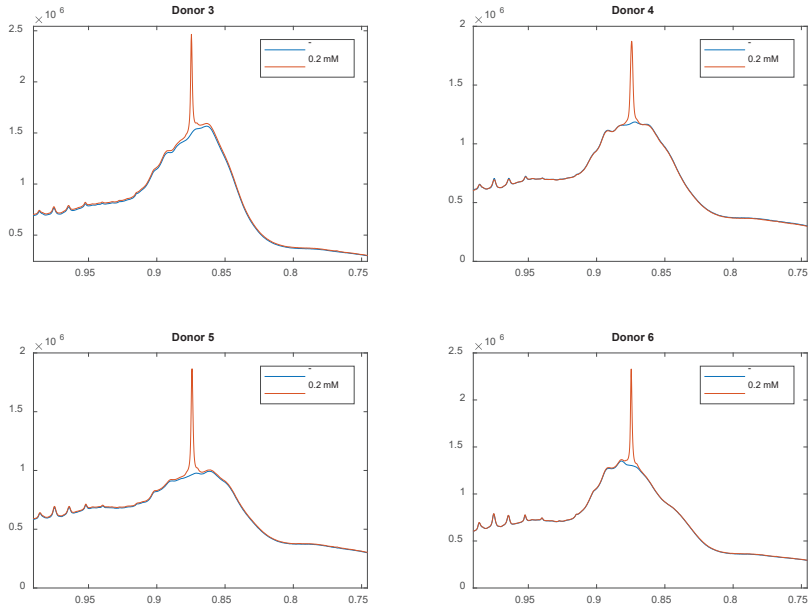


Figure 4.5 Results from the spike-in experiment, for four different donors, comparing a clean serum sample, with a serum sample spiked by a 0.2 nM concentration of neopentyl glycol, in the region with the lipid peak arising from –CH_3.

We proceeded with an experiment, where three aliquots of a clean serum sample were spiked with different concentrations of neopentyl glycol, while one remained clean, to investigate if there is a linear trend relating the concentration of neopentyl glycol to the lipoprotein variables. This experiment showed us that the lipoprotein variables were differently affected by increasing the concentration. As some of the variables indeed showed a linear trend with increased concentration, other variables did not increase further after reaching a maximum or minimum value. Figure 4.6 shows the results, where the different concentrations of the lipoprotein subfractions associated with LDL from the four samples with different concentrations of neopentyl glycol are shown.

As collaborators in the project, Bruker Biospin assisted with developing a routine to model the peak arising from the contamination and subtract it from the raw spectra. The processed raw spectra was then used to re-calculate the concentrations of the lipoprotein subfractions using B.I.LISA^{TM} protocols. To evaluate the accuracy of the concentrations calculated from the processed spectra, this algorithm was tested on our spiked samples, and compared with the corresponding clean samples. The fit of the modelled peak varied across donors, see Figure 4.7 for examples of a good (4.7A) and a poor fit (4.7B). Looking at the percentage differences between lipoprotein concentrations calculated from the clean spectra and the processed contaminated spectra, it was apparent that some of the variables were not reliable. Based on the spike-in experiment, which included 10 donors, 3 of the variables had a percentage difference exceeding +/- 10% for 9 of the donors. 4 of the variables had a percentage difference exceeding +/- 10% for 8 of the donors, 4 variables for 7 of the donors, and 7 variables for 6 and 5 of the donors. As there are no clean HUNT2 samples to compare with, the exclusion of some variables from the data analysis in this project has been done on the basis of the spiking experiment.

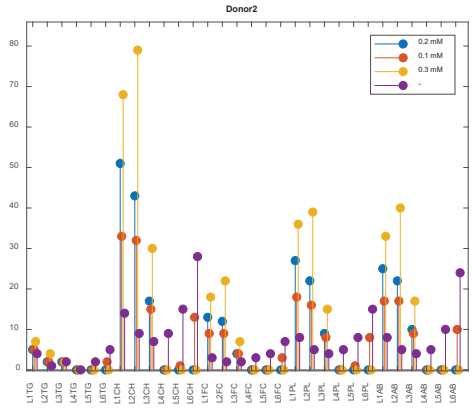


Figure 4.6 Concentrations of different lipoprotein subfractions associated with LDL from four different aliquots from the same donor spiked with different concentrations of neopentyl glycol (0, 0.2 mM, 0.1 mM and 0.3 mM).

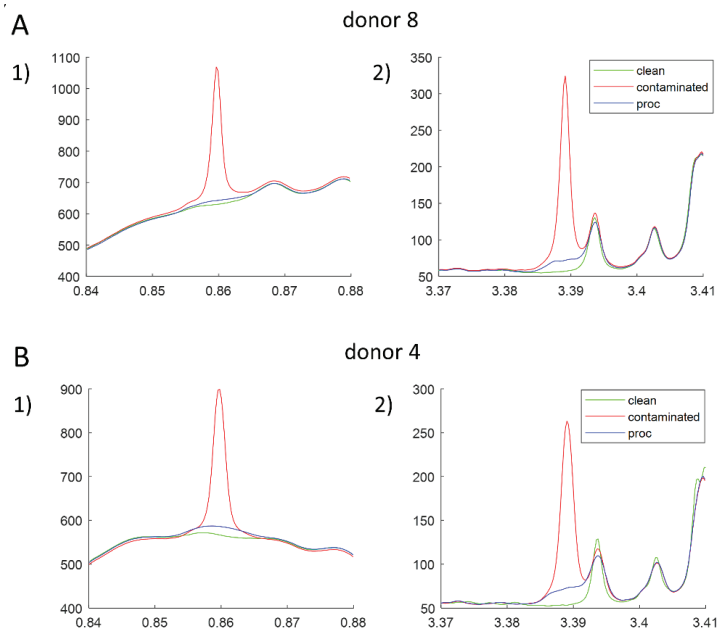


Figure 4.7 Results from the algorithm that models the peaks arising from the contamination and removes it from the raw spectra, when tested on the spike-in samples. The green line shows the spectra of the clean sample, the red line shows the spectra of the corresponding sample, with added neopentyl glycol, and the blue line shows the processed spectra, where the peaks arising from the contamination have been modelled and removed. The top panel (A) shows an example of a good fit, and thus small differences between the clean and the processed spectra. The bottom panel (B) shows an example of a poor fit, where there is more discrepancy between the clean and the processed spectra. 1) and 2) show the two regions of the spectra where signals from neopentyl glycol appear.

Table 4.2 Results of the algorithm that models the peaks arising from the contamination and removes it from the raw NMR spectra, tested on the samples from the spike-in experiment.

Variable	Total no. of donors with concentration exceeding a percentage change of +/- 10 %	Variable	Total no. of donors with concentration exceeding a percentage change of +/- 10 %	Variable	Total no. of donors with concentration exceeding a percentage change of +/- 10 %
TPTG	0	IDPL	2	L5TG	9
TPCH	0	LDPL	0	L6TG	4
LDCH	0	HDPL	0	L1CH	6
HDCH	0	HDA1	1	L2CH	8
TPA1	1	HDA2	1	L3CH	5
TPA2	1	VLAB	0	L4CH	6
TPAB	0	IDAB	3	L5CH	4
LDHD	0	LDAB	0	L6CH	2

ABA1	1	V1TG	1	L1FC	6
TBPN	0	V2TG	0	L2FC	9
VLPN	0	V3TG	2	L3FC	1
IDPN	3	V4TG	1	L4FC	7
LDPN	0	V5TG	0	L5FC	3
L1PN	6	V1CH	3	L6FC	2
L2PN	8	V2CH	2	L1PL	6
L3PN	4	V3CH	5	L2PL	8
L4PN	7	V4CH	3	L3PL	3
L5PN	4	V5CH	2	L4PL	7
L6PN	2	V1FC	2	L5PL	4
VLTG	1	V2FC	5	L6PL	2
IDTG	1	V3FC	3	L1AB	6
LDTG	3	V4FC	4	L2AB	8
HDTG	1	V5FC	5	L3AB	4
VLCH	1	V1PL	3	L4AB	7
IDCH	5	V2PL	0	L5AB	4
VLFC	0	V3PL	3	L6AB	2
IDFC	3	V4PL	2	H1TG	3
LDFC	0	V5PL	2	H2TG	1
HDFC	0	L1TG	2	H1CH	4
VLPL	0	L2TG	5	H2CH	2
H3TG	1	L3TG	1	H3CH	1
H4TG	3	L4TG	9	H4CH	1
H1FC	1	H2PL	1	H3A1	1
H2FC	2	H3PL	1	H4A1	1
H3FC	1	H4PL	1	H1A2	2
H4FC	1	H1A1	6	H2A2	2
H1PL	5	H2A1	0	H3A2	0

5 Discussion

The work in this thesis describes metabolic characterization of breast cancer. The main objective was to investigate the potential of metabolic profiling for treatment monitoring during breast cancer treatment and for early risk stratification of a future breast cancer. This thesis also evaluated the metabolic effects of sample handling prior to NMR experiments of biofluids, specifically the effect of repeated freeze and thaw cycles on quantified concentrations of metabolites and lipoprotein particles.

In Paper I, the metabolic effects from neoadjuvant treatment in patients diagnosed with locally advanced breast cancer were examined. Survivors and non-survivors were successfully discriminated based on the NMR spectra from tissue samples, while clear serum metabolic changes were observed in patients undergoing treatment. Also the type of treatment could be discriminated based on the serum metabolic profiles. This paper showed the complementary nature of studying the tissue and serum metabolomes.

In Paper II, the metabolic effects of sample handling, with a focus on the effects of repeated freezing and thaw cycles on the concentrations of metabolites in serum and urine, and lipoprotein particles in serum, were evaluated. We found that up to five freeze and thaw cycles did not induce systematic changes in the concentrations measured by NMR. However, minor, but significant accumulated freeze-thaw effects were observed for 32 lipoprotein parameters and acetic acid in serum, when comparing one freeze and thaw cycle to further cycles. This paper constitutes an important basis for Paper III, in which the sample material was from the HUNT2 biobank. Samples stored in biobanks often undergo several freeze and thaw cycles, and it was hypothesized that several freeze and thaw cycles may influence the sample composition due to chemical degradation processes.

Paper III included a large collection of serum samples from healthy participants in the HUNT2 biobank, of which half later developed breast cancer, while the second half remained breast cancer free during the follow-up period. The aim of this paper was to investigate associations between metabolites and lipoproteins in serum with breast cancer risk. This paper revealed several significant associations, however, these were not strong enough to be used for predicting a future breast cancer case.

5.1 Potential clinical applications of metabolomics findings

Metabolomics is the final level of the omics cascade, and metabolites are downstream products of the proceeding levels. Detailed information on many thousands of metabolites have been described in metabolome databases [177, 178] and metabolomics has in the recent years gained interest as a promising tool for disease detection and characterization, as it has the ability to detect subtle differences in metabolism. Biological findings in Paper I are related to treatment response among a

specific group of breast cancer patients, while biological findings in Paper III are related to the very early disease development. Paper I included small-molecular metabolites quantified in tumor and serum samples, while Paper III included serum levels of both metabolites and lipoprotein parameters.

5.1.1 Tissue versus biofluids

As described previously, there are mainly two commonly used analytical platforms for metabolomics studies, NMR and MS, of which NMR has the advantages of being less labor-intensive and more reproducible, and in that aspect more promising for clinical applications. Tissue metabolic profiles provide a detailed picture of the biological processes inside the tumor and have been found to be correlated with cancer-specific variables [179-181]. In comparison to tissue metabolomics, the metabolic profile of a biofluid reflects the metabolic state of the entire organism. Blood circulates through the whole body, and is responsible for delivering nutrients to all organs and tissues. Urine is a by-product of metabolism, containing compounds that have been excreted, and flows through from the kidneys to the bladder. Metabolomics of biofluids are minimally-invasive, and thus appealing for use in clinical applications. As illustrated in Figure 5.1, NMR metabolic profiling of biofluids may be applied along the entire cancer timeline; for risk stratification of healthy individuals, for diagnosis, for diseases subtyping leading to treatment stratification, for monitoring treatment response and also for detecting a disease relapse. Serum and urine metabolism is influenced by several endogenous processes as well as environmental and lifestyle factors [67, 182-184], including the circadian rhythm [185]. Because of this high variability, identification of robust biomarkers is more challenging, and effort should be made to minimize variation not related to the biological trait of interest.

Figure 5.1 NMR metabolic profiling of biofluids has a wide range of applications, from risk assessment to detecting a disease relapse.

The main objective of this thesis was to perform metabolic characterization of breast cancer for improved management of breast cancer. Paper I focused on tissue and serum metabolic changes during neoadjuvant treatment of locally advanced breast cancer patients, while Paper III investigated the associations between circulating metabolites and lipoprotein parameters with a long-term breast cancer risk. As the work in Paper I was a continuation of a previously published work, NMR analyzes

of tissue samples had been completed before the work on this thesis began. Thus, although this thesis includes metabolomics analyses of tissue, metabolomics of biofluids has been the main focus.

In Paper I we had a unique opportunity to compare treatment response of breast cancer patients undergoing treatment in the tumor tissue and serum metabolome. In this paper, also the correlations of metabolites from tissue biopsies and serum samples from the same patients were evaluated. The correlations were also calculated for each medium alone. The results showed that there is little correlation between serum and tissue metabolic profiles. For example, tumors are often characterized by a high lactate production, however this study showed a lack of correlation between tissue and serum lactate levels. Thus, despite possible leakage of metabolites from the tumor into the bloodstream, tumor specific metabolites will in general be masked by the overall high variation of the serum metabolome.

5.1.2 Treatment response

Paper I focused on treatment-induced serum and tissue metabolic changes in breast cancer patients, and included up to four sampling time points (prior to treatment, during treatment, and post-treatment (serum only)). This study was quite unique due to its longitudinal study design and the availability of two types of biological samples; tumor tissue and serum. In this paper, changes in the tissue metabolic profiles as an effect of chemotherapy were observed, which could be linked to patient survival (classification accuracy for separating survivors and non-survivors = 72%). For the same patient cohort, based on the tissue metabolic profiles, patients with a pathological minimal residual disease could be successfully discriminated from pathological non-responders from samples taken after treatment completion (accuracy = 77%) [186]. Patients with a pathological minimal residual disease showed elevated levels of glucose, compared to pathological non-responders, indicating a decline in the consumption of glucose in the former group. Similarly, decreased levels of glycerophosphocholine, phosphoscholine and choline was found comparing these groups. These metabolites are involved in the metabolism of phosphatidylcholine, which is a phospholipid highly abundant in eukaryote cell membranes. Also in other studies, tissue metabolic profiles have shown promise for clinical application prior to and during breast cancer treatment for treatment stratification. For example, tissue metabolic profiles have shown to have predictive value for determining the ER status of breast cancer patients [179], and for discriminating between triple negative and triple positive breast cancers [180]. However, in Paper I, patients receiving different treatment regimens (chemotherapy combined with Bevacizumab versus chemotherapy only) could not be discriminated at any time points based on tissue metabolic profiles. In contrast, they were successfully discriminated by serum metabolic profiles obtained at the early stage of treatment (accuracy = 64 %), which again demonstrates the complementary nature of these two biological

matrices. The group of patients treated with Bevacizumab had higher serum levels of leucine, acetoacetate and tri-hydroxybutyrate, and lower levels of formate, compared to patients treated with chemotherapy only. Interestingly, tissue metabolic profiles revealed a significant interaction between time and bevacizumab for glutathione [186]. This finding indicates that bevacizumab might play a redox destabilizing role in cancer cells, inducing oxidative stress to promote apoptosis.

From serum metabolic profiles taken after completed treatment, the residual cancer burden could be predicted with an accuracy of 69 %. Moreover, serum metabolic changes during treatment revealed increased levels of circulating lipids during treatment. This is an important finding, pinpointing the increased risk of cardiovascular diseases for breast cancer survivors [187-189]. Similar treatment-induced increases in lipid levels were observed in a different patient cohort, for breast cancer patients undergoing adjuvant treatment [96]. For the same patient cohort as examined in Paper I, by combing serum metabolites with inflammation related markers (measured 12 weeks into treatment, TP2), a set of 10 molecules were found to be associated with treatment response [190]. Analyses were performed by applying a Lasso logistic regression model, and serum metabolites included in the lasso were leucine, creatinine, proline-betaine and dimethyl-sulfone. The serum metabolic profiles alone of samples taken at TP2, could however not successfully discriminate patients with a good (RCB class 0 or I) or poor (RCB class II or III) response.

5.1.3 Assessing biomarkers for future development of breast cancer

Paper III presents one of the largest prospective analyses of serum metabolic profiles and breast cancer risk to date [105, 107, 191], and associations between several lipoprotein parameters and metabolites and long-term risk of breast cancer were identified. Significant positive associations between triglycerides in VLDLs, free cholesterol and Apo-A2 in HDLs, acetate, and valine with long-term breast cancer risk were found, and inverse associations between total amount of Apo-A1, free and esterified cholesterol in VLDLs, phospholipids in HDLs and glycine and long-term breast cancer risk. This study partially confirmed previous findings, linking an increased breast cancer risk with an increased risk of cardiovascular diseases [192], but provided a more detailed picture of the lipidomic profile. For example, we revealed additional biological insight into the relationship between HDL metabolism and breast cancer risk, by showing that the esterified cholesterol, which is within the lipoprotein, and the free cholesterol in the micellar membrane of the lipoprotein have opposite associations with breast cancer risk [100]. Most similar prospective studies on breast cancer risk have been performed based on metabolic profiles obtained by mass spectrometry [105, 108, 184, 191], which makes the comparison of significant findings across studies difficult due to a different panel of metabolites being measured. However, in this study we were able to confirm some of the findings reported by Lécuyer et al., such as positive associations between valine, which has been suggested as

a marker for poor cardiovascular health, and breast cancer risk, in addition to positive associations of creatinine. The association of creatinine was however conflicting to results presented by Bro et al., which observed an inverse association between a NMR signal corresponding to a spectral overlap of creatine and creatinine and short-term breast cancer risk [106]. Our study also found inverse associations between the non-essential amino-acid glycine and breast cancer risk. This metabolite is involved in a wide range of metabolic pathways, and circulating levels of glycine have previously been associated with cardiovascular diseases (CVD) [193], insulin resistance [194], metabolic syndrome [195] and type 2 diabetes [193, 196]. Glycine has also a key role in rapid cancer cell proliferation [197] and differential expression of enzymes associated with serine/glycine metabolism differ across breast cancer subtypes [198]. Glycine levels in Paper I remained relatively stable during treatment, with a significant decrease after treatment completion, possibly indicating a negative side-effect of treatment. A more systematic study comparing the etiologies of CVD, diabetes and breast cancer would be very interesting.

The large sample size and long follow-up period together with the availability of numerous lifestyle factors allow for further stratified analysis of the HUNT2 cohort for assessment of breast cancer risk in future investigations. Also, if the associations found are merely mediators of already known risk factors (such as alcohol intake or overweight) should be investigated, by performing analyses corrected for known breast cancer risk factors. Further work should also include analyzes on more homogenous subgroups of individuals, stratified on breast cancer specific variables (e.g. ER status), to investigate how associations differ among the different breast cancer subtypes. Other studies have shown that there exists a heterogeneity in the serum metabolic profile by breast cancer subtype [199, 200]. Analyzes stratified on pre- or post-menopausal status or the use of hormone therapy should also be performed, as these variables are likely to influence the associations of metabolites with breast cancer risk [105, 200].

Paper III includes pre-diagnostic serum metabolic profiles, while Paper I includes serum metabolic profiles of women prior to, during and after treatment. The presence of linear trends in the development of breast cancer could be investigated by comparing the serum metabolic profiles of individuals which later developed breast cancer included in Paper III with serum metabolic profiles of newly diagnosed breast cancer patients (samples at TP1 in Paper I). Paper I however includes relative serum metabolic concentrations, while Paper II and III include absolute serum metabolic concentrations, making direct comparison of the values challenging.

5.1.4 Recent developments in serum metabolomics

Major obstacles for the translation of metabolic findings to clinical application have been the lack of standardization and relatively small patient cohorts, resulting in a large heterogeneity in reported metabolites which have been found significant in cancer research [201]. This makes findings across studies difficult to compare.

Several commercial actors for automatic metabolic quantification have been established, such as Metabolon (https://www.metabolon.com/) and Nightingale (https://nightingalehealth.com/) for MS and NMR metabolomics, respectively. These platforms are convenient for large-scale metabolomics studies, allowing for comparing metabolic profiles of thousands of individuals. Studies with impressive cohort sizes have already been performed utilizing the Nightingale platform, where significant metabolic risk factors have been described [202-204]. Similar to the B.I.LISA™ lipoprotein subclass analysis developed by Bruker BioSpin, Nightingale provide information on the lipid composition of serum samples, however their method is based on electrophoresis, while the algorithms developed by Bruker BioSpin are based on ultracentrifugation, making comparison of results challenging. The downside of such commercial actors is however that the researcher is left with little control over the sampling handling and data acquisition. These companies receive biological samples and return metabolic concentrations, but generally not the raw data, which is a big disadvantage. Importantly, the platform used in the current thesis (Bruker BioSpin's quantification methods) allow for automatic metabolite and lipoprotein parameter quantification, returning the raw NMR spectra in addition. Section 5.2 of this thesis illustrates some of the challenges that might be present during metabolomics analyses, such as the presence of a contamination or variability associated with algorithms for automatic quantification. In addition, several challenges common in MR based metabolomics analyses were also present in the current work, such as NMR spectra with poor shim and varying success of water suppression. How to handle missing data due to metabolic concentrations below the limits of detection may also be important for correct interpretation of data analysis. These large platforms have so far been reluctant to sharing exact procedures for handling these problems. During the recent years there has also been improvements for automatic spectral preprocessing and metabolite quantification (Chenomx, B.I.QUANT-PS2™, BATMAN [122], Mnova [205]), which is necessary for large-scale metabolomics studies.

5.2 Metabolite and lipoprotein quantification and reliability

5.2.1 Absolute versus relative metabolite quantification

For Paper I, there was no procedure available for absolute metabolite quantification. In addition, spectra were run using an older protocol, which has not been developed for lipoprotein parameter

quantification. For the quantification of metabolites and lipoprotein parameters in Paper II, we used commercially available analytical platforms developed by Bruker BioSpin: B.I.LISA^{TM}, Bruker B.I. Quant-PS^{TM} and Bruker B.I.Quant-UR^{TM} for lipoprotein parameters, serum and urine metabolites, respectively [167]. In Paper II we observed that the procedure for automatic metabolite quantifications in serum is associated with a high variability in the concentrations, as the CVs for the serum metabolites of the QC samples were high, compared to the CVs of the lipoprotein parameters and urine metabolites. We therefore decided to quantify serum metabolites manually for Paper III, however making use of the absolute concentrations of glucose obtained by B.I. Quant-PS^{TM} in order to convert spectral integrals to absolute concentrations.

Metabolites in Paper I were quantified manually, by importing raw NMR spectra into Matlab, and integrating the area under each peak, after necessary preprocessing. Mean normalization was then applied to the spectra prior to metabolite integration, which provided relative metabolic concentrations. The disadvantage of relative metabolite concentrations is that metabolic findings across studies are more difficult to compare. Also, even though the metabolic response of different groups can be compared, it is difficult to say something about the magnitude of these differences. As mentioned in Section 5.1.4 one of the main challenges of metabolomics is that there is no standardized pipeline to guide researchers trough sample handling, spectral acquisition, preprocessing and to quantification. Different approaches of preprocessing are commonly used. Moreover, different sets of metabolites may be detected by different analytical platforms [206]. Non-standardized analytical procedure hinders the translation of metabolic results to clinical practice [109]. Different types of normalization emphasize different metabolites [207] and in Paper I we saw how the inclusion or exclusion of different spectral regions (in specific the broad lipid peaks, which make up a substantial part of the spectra and have a high between-individual variation) influences the results, and thus the biological interpretation. Absolute metabolic concentrations allow for easier comparisons between studies, and possibly also across different analytical platforms and pinpoints why absolute concentrations are preferable over relative concentrations. We therefore strived to obtain absolute concentrations for Papers II and Papers III.

We observed in Paper II that the reproducibility of the automatically quantified serum metabolites was limited due to the protocol for automatic quantification. This was also evident in Paper III (the median CV (calculated from the QC samples) for all samples combined was CV_C = 55.4 %, and only alanine, isoleucine, valine, lactate and pyruvate had a CV < 15 %) and motivated us to performed manual quantification to obtain as low CVs as possible. Quantification was performed in the cpmg spectra as suppression of lipids and macromolecules makes detection of the metabolite peaks easier. Cmpg spectra are however affected by T2 relaxation, thus when all metabolites were identified and

quantified, they were adjusted by the T2 relaxation times. T2 values were estimated based on three separate serum samples. These T2 values varied across these three samples, which had slightly different metabolic compositions. The variation was in specific high for small peaks or peaks lying close to a lipid signal and this variation demonstrates a need for performing a more systematic study of how the T2 values are affected by factors such as the amount of lipid in the serum. For the metabolites for which the T2 values had a high variation (proline-betaine, lysine, ornithine, citrate, acetoacetate, 3-hydroxybutyrate, glutamate, ethanol, isoleucine and dimethyl-glutarate), there is a higher uncertainty in the exact metabolic concentrations in Paper III. In order to get absolute metabolite concentrations, the concentrations of glucose were set equal to the absolutely quantified concentrations of glucose by B.I.Quant-PS (as this metabolite had the lowest coefficient of variation, CV_C = 8.37 % for all samples combined, while CV_T = 2.55 %, CV_G = 2.92% calculated from samples in the two labs separately, where all CVs are calculated from the QC samples), while the remaining metabolites were scaled accordingly by the same factor. The concentrations quantified manually had lower CVs than automatically quantified concentrations (median CV_C = 9.95 % of which 23 had a CV < 15 %), thus we were able to substantially reduce variability caused by the quantification protocol. This reduction in the variation resulted in CVs smaller than those reported in Paper II, where the protocol for absolute metabolite quantification was employed, and for which the median CV was 11.0%.

Absolute quantification is a big step towards standardization and better tools for automated metabolite quantification are necessary. Moreover, data management according to the FAIR principles (findable, accessible, interoperable and reusable) [208] and multicenter studies where the same protocol is used are essential for validation of biomarkers and for further advances in this field.

5.2.2 Reproducibility of NMR

In metabolomics studies there is variability associated with each step from sample acquisition to measuring metabolic concentrations. Metabolic profiles, when measured by NMR, have a high reproducibility [209]. However, sample handling and storage prior to analyses may influence the metabolic composition of the biological medium [210, 211]. A proportion of variability will always be associated with lab work, such as slight variations in the proportions between the biological medium of interest and buffer, due to pipetting. There is also some variability associated with the NMR acquisition, due to differences in for example shim and the success of water suppression. As stated by Dunn et al. the use of QC samples provide the ability to perform signal correction in order to reduce analytical variation and to quantitatively determine analytical precision [212].

In Paper II and III we ran QC samples in parallel with the samples of interest. In Paper II these were used to compare the variability due to repeated freeze and thaw cycles with variability from other

sources. For Paper III, by mixing serum from 10 healthy donors we were able to prepare a set of QC samples large enough to be run throughout the whole project period, and on average one QC sample was run together with 60 serum samples. The QC samples in Paper III turned out to be particularly useful for comparing the metabolic levels of samples run at the two different labs. Another advantage of QC samples was that the reproducibility of specific metabolites and lipoprotein parameters could be assessed. The method for quantifying lipoprotein parameters was released in 2016 [213] and has previously shown robust results [167], reflected in overall low CVs. The software for automatic quantification of serum metabolites was however newly released [214], and CVs were in general higher indicating a lower reproducibility for some of the serum metabolites, such as trimethylamine-N-oxide and acetoacetic acid.

Studies for which a long follow-up period is desired, will always require sample material that has been stored for a long time. Paper III was based on serum samples from the HUNT2 biobank, which have been collected in the years 1994-5 and have been used for several research project throughout the years. It is therefore likely that the samples have undergone multiple freeze and thaw cycles due to aliquoting [215]. Research on how this might affect the metabolic composition of biological samples, when measured by NMR, was limited, thus we performed a study evaluating this effect. In this study we found no systematic changes related to repeated freeze and thaw cycles in neither serum nor urine. Minor, but significant accumulated effects were however observed for one serum metabolite (acetic acid), and 32 lipoprotein parameters in serum. Paper II thus justified the use of biobank samples for Paper III. From the significantly associated metabolites with BC risk found in Paper III, none were affected by multiple freeze and thaw cycles, making them reliable biomarkers, while for the lipoprotein parameters, V3FC, VLCH, V2FC, V4CH and H1A2 had some accumulated effects, with slight increases in their concentrations. This should however not influence the associations with breast cancer in paper III, as all samples have been stored in the HUNT2 biobank for the same amount of time, where they have been treated similarly, and all samples have been thawed only once after sample retrieval for this project. Nevertheless, this adds extra variability to the data which may mask significant findings.

In Paper III we also detected a contamination of neopentyl glycol in the samples. The degree of success of the method developed by Bruker BioSpin for correction of this contamination signal varied across samples, inducing extra variability in the data of this project. All such additional variability makes the discovery of biomarkers more challenging as they may contribute to mask the biological information of main interest.

5.2.3 Batch effect observed in Paper III

Samples that made up the basis for the analysis in Paper III were analyzed at two different labs: locally at NTNU, Trondheim and at Bruker BioSpin in Germany. As previously described, QC samples from the same serum mix were analyzed at the lab in Trondheim and in Germany, allowing us to investigate the presence of a batch effect. A study performed by Jiménez et al. reported excellent reproducibility for lipoprotein parameters and small molecular weight metabolites, when analyzed at 11 different spectrometers, distributed across four different geographical locations [167]. In their study, nine 600 MHz Bruker Avance III HD spectrometers and two 600 MHz Bruker Avance III spectrometers were used for NMR acquisition. Even though the two labs analyzing samples for paper III followed the same protocol for sample handling, there were apparent differences in the metabolic profiles across the two sites. Only five serum metabolites had CVs < 15 %, when calculated using QC samples analyzed at the two labs, but when calculated separately for each lab, the median was CV_T = 21.6 % and CV_G = 56.4 %, and 14 and 6 metabolites had CV < 15 % from the Trondheim and Germany batches, respectively. Effort has been made to understand what caused this batch effect, however no certain explanation has been found. Some of our hypothesis for this batch effect include pipetting procedures and different pipettes. Differences in the shim may also influence the B.I.Quant models, which are based on the lineshapes of the metabolite signals. Also, NMR analyses were carried out on a Bruker Avance III HD Ultrashield Plus 600 MHz spectrometer equipped with a 5 mm TCI probe in Trondheim, while at the lab in Germany they were carried out on an Avance-IVDr spectrometer. Even though the equipment at both labs has been calibrated for use of the same protocol, it is possible that the difference in the type of probe used is the origin of the batch effect [216].

The presence of a batch effect was clearly evident when the concentrations of the QC samples were plotted alone, showing variability around two different concentrations. Interestingly, the batch effect was not systematic, meaning that concentrations from samples analyzed at one lab were not consistently lower or higher than from the lab, which rules out that the batch effect is due to systematic variation from pipetting or the types of pipettes. Figure 5.2 shows the raw concentrations of a number of lipoprotein parameters colored by the lab at which they were analyzed, together with the corresponding concentrations of QC samples. In this figure the concentrations of L6PN are clearly higher for samples analyzed in Trondheim compared to the ones in Germany, while the opposite is true for L3PN. Figure 5.3 shows PCA scores plots, for the first three PCs when applied to the concentrations of the lipoprotein parameters, clearly showing a batch effect.

As mentioned earlier, in an attempt to decrease the batch effect and the CVs of the serum metabolites, NMR spectra were imported into Matlab for manual quantification. The batch effect was also observed on the raw NMR spectra, thus quantification had to be performed on the two batches

separately. In particular, there was a difference in signal intensity and metabolite peaks experienced shifts at the spectral ppm scale which were not systematic (not in the same direction for the different metabolites). As described in Section 5.2.1, we were able to lower the CVs in Paper III substantially by quantifying manually. The resulting CVs were smaller than those reported in the multicenter study mentioned previously, in which CVs < 15 % are reported for 20 out of 24 detected metabolites [167]. These values have however been calculated using the mean concentration from each lab, while looking at all samples combined, the reported average CV is 40%, which is substantially higher than what we were able to obtain in Paper III. The panel of serum metabolites differs slightly across the studies mentioned in this section, of which the majority is overlapping.

Furthermore, since the automatically quantified metabolites had a batch effect, and we used the automatically quantified values of glucose to convert spectral integrals to absolute concentrations, the batch effect was necessarily propagated on the metabolite concentrations obtained by manual quantification. As the batch effect was not systematic, correcting for it could cause unwanted alterations in the metabolic profiles. For that reason, we decided not to correct for the batch effect, so that the naturally occurring composition of metabolites and lipoprotein parameters in the serum samples would not be altered. Lab was however included as a fixed effect in the LR models. Figure 5.4 (left) shows a PCA score plot of absolute metabolic concentrations, after a thorough manual quantification in Matlab, where points are colored according to the lab at which they have been analyzed, clearly visualizing a batch effect. Figure 5.4 (right) shows the PC1 loadings for each variables, showing that in specific dimethyl-glutarate, leucine, valine and lysine cause this effect. This is in good correspondence with the CV values for these metabolites, which were small when calculated separately for the two labs, but larger when combined, and are: for dimethyl-glutarate CV_T = 6.2%, CV_G = 6.2%, while combined CV_C = 14.3 %; leucine: CV_T = 2.9 %, CV_G = 3.6 % & CV_C = 11.5%; valine: CV_T = 2.4 %, CV_G = 2.9 % & CV_C = 8.5 %; lysine: CV_T = 2.4 %, CV_G = 2.9 % & CV_C = 4.6 %.

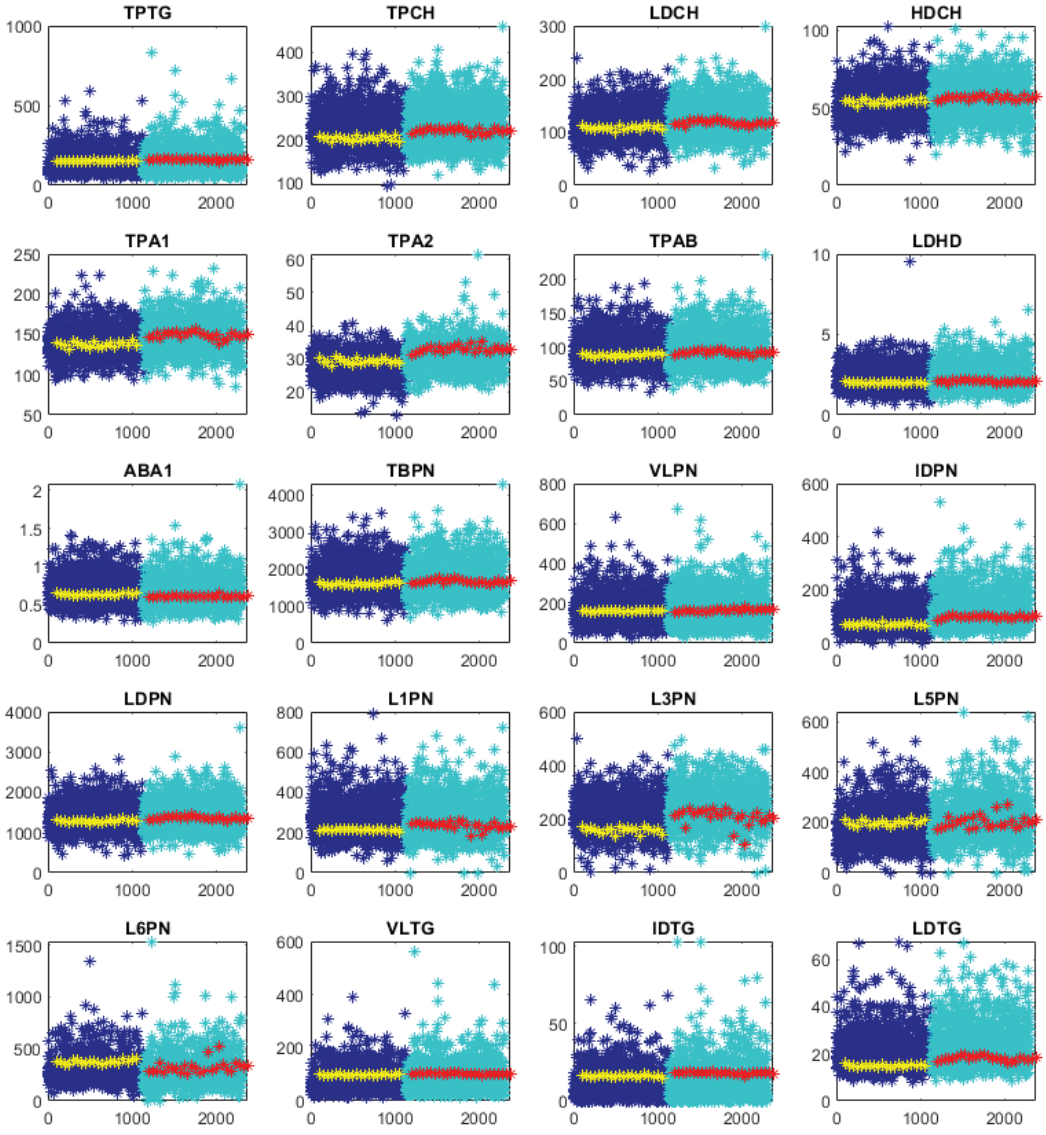


Figure 5.2 A selection of lipoprotein parameters from the HUNT2 cohort. Navy: samples analyzed in Trondheim; Turquoise: samples analyzed in Germany; Yellow: QC samples analyzed in Trondheim; Red: QC samples analyzed in Germany.

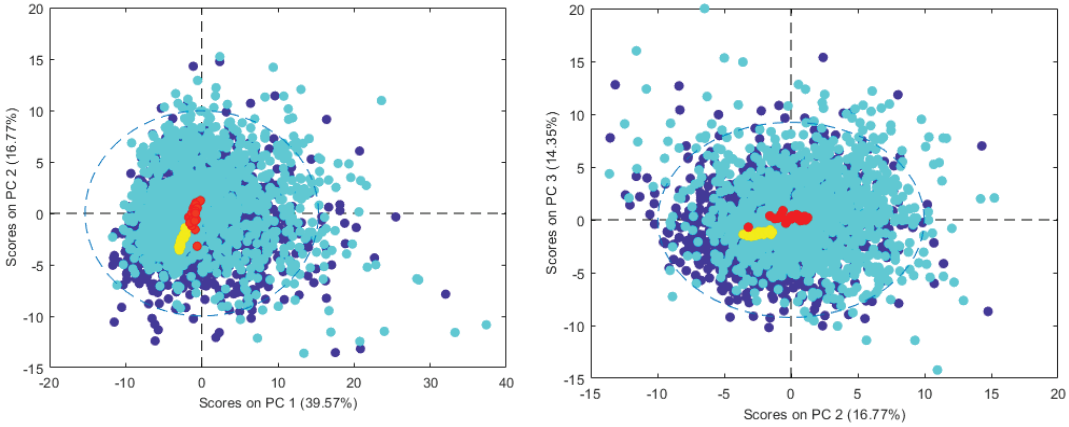


Figure 5.3 PCA plots of the lipoprotein parameters of samples in the HUNT2 cohort, colored according to the lab at which they have been analyzed. Left: scores on PC1 and PC2; Right: scores on PC2 and PC3. The batch-effect on the serum metabolic profiles is seen as a shift in the PC3 scores. Navy: Samples analyzed in Trondheim; Turquoise: samples analyzed in Germany; Yellow: QC samples run in Trondheim; Red: QC samples run in Germany.

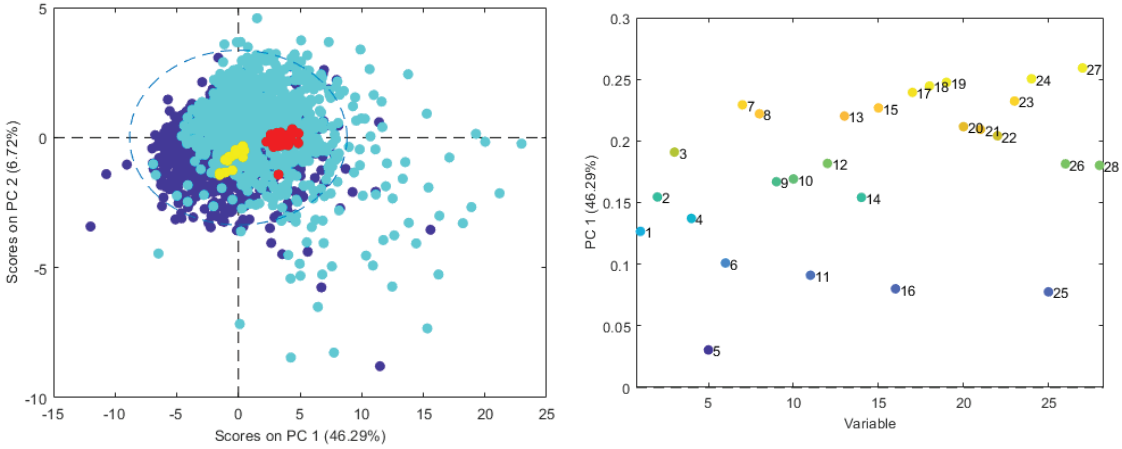


Figure 5.4 Scores on PC1 and PC2 when performing PCA on the serum metabolic profiles of samples in the HUNT2 cohort, colored according to the lab at which they have been analyzed. Left: scores on PC1 and PC2; Navy: Samples analyzed in Trondheim; Turquoise: samples analyzed in Germany; Yellow: QC samples run in Trondheim; Red: QC samples run in Germany; 1: formate; 2: creatine; 3: lactate; 4: glycine; 5: methanol; 6: dimethyl-sulfone; 7: ornithine; 8: methionine; 9: glutamine; 10: citrate; 11: acetate; 12: acetoacetate; 13: glutamate; 14: pyruvate; 15: alanine; 16: ethanol; 17: isoleucine; 18: 2-methylglutarate; 19: leucine; 20: phenylalanine; 21: glucose; 22: tyrosine; 23: creatinine; 24: valine; 25: proline-betaine; 26: histidine; 27; lysine; 28: 3-hydroxybutyrate.

5.3 Data analysis

For examining disease characteristics in a metabolomics study, multivariate analysis methods are frequently applied, largely due to the nature of the chemical signals [217]. The main advantage of these methods is their ability to identify patterns of several metabolites simultaneously. Especially PLS-DA is commonly used for creating discriminatory models [218, 219] and PCA for identifying groups in data. This is because the data often contains few samples compared to the number of variables and because these methods are able to overcome the multicolinearity problem. These methods utilize the correlation structure of the data to extract the principal components [138] and project the multivariate data onto a lower-dimensional space, and were used extensively in this thesis (PCA was used in all papers, while PLS-DA was used in Papers I and III).

5.3.1 Analyses of repeated measurements

In Paper I we had repeated measurements from individuals undergoing treatment, thus Paper I included the use of paired multivariate data analysis, taking advantage of the multilevel structure of the data. As the serum metabolome is highly dynamic, variability across different individuals is high. This was also shown in Paper II, where samples from the same donors were clustered closely together in the PCA plot, and the distance to the samples from different donors was much longer. In multilevel analyses the total variation is split into the within- and between- sample variation, and the net differences pre and post treatment is used as input for the discriminant model. An enhancement of the variation resulting from the treatment is obtained [220]. This approach was beneficial, and the number of correctly classified samples increased to 90%, compared to the regular PLS-DA approach which gave an accuracy of 76%, when discriminating between the serum metabolic profiles of breast cancer patients at different time points of treatment as shown in Table 5.1.

Table 5.1 Prediction accuracies for discriminating between serum metabolic profiles at the different sampling points in the NeoAva study (Paper I), when applying multilevel PLS-DA and ordinary PLS-DA.

	PLS-DA	Multilevel PLS-DA
TP1 vs TP2	76 %	90 %
TP2 vs TP3	61 %	77 %
TP3 vs TP4	70 %	87 %

Also linear mixed models would provide useful information about any univariate trends in the metabolic changes over the multiple time points in Paper I. This was attempted, but due to high deviations from normality the models experienced poor fits. This was also the case after logarithmic transformation of the data, thus LMMs were left out from the paper, and the trends in metabolite concentrations were instead assessed based on median percentage changes between the different time points. LMM was however employed in Paper II to calculate the intra-class correlation coefficients, as it calculates the between cluster (here subject) variance and the total variance. Calculating the ICCs allowed us to investigate the proportion of the overall variation explained by clustering. The ICCs were generally high, showing a good reproducibility of the NMR analyzes, in accordance with what we observed in the PCA. In this study we also observed that sometimes variables which had a low CV, thus showing a good reproducibility, also had a low ICC value, or vice versa, which was at first surprising. After investigating this issue, we found that this was the case when the metabolite levels were close to the limit of detection. The CV is calculated by dividing the standard deviation by the mean, thus very small mean values will give high CVs. Also, the CVs were calculated based on five samples, thus increasing the sample size would provide more accurate estimates.

Paper II consisted of a relatively small data set, where serum and urine samples from 20 donors were compared over several freeze and thaw cycles. Because the variability of the variables was of main concern, PCA was chosen for answering the research question. It is a dimension reduction technique which allows for observing naturally occurring clusters on a lower-dimensional space, thus it allowed us to evaluate the clustering of the samples colored by the corresponding donor. PCA was also very useful for the exploratory analysis in Paper III. By plotting the raw concentrations of metabolites and lipoprotein parameters from samples analyzed in Trondheim and Germany in the same figure, and also including the concentrations of the QC samples, batch effects were detected as previously discussed. PCA allowed us to evaluate the combined impact of these differences on the whole metabolic profiles.

In Paper II, for univariate testing the Wilcoxon signed-rank test was chosen as the metabolic concentrations did not follow a normal distribution. The student's t test could probably have been used instead, as it quite robust to deviations from normality [221], and is more powerful than the non-parametric Wilcoxon test for small samples [222], however we chose to use Wilcoxon as that is more accepted. The p-values were adjusted for multiple testing using the Benjamini-Hochberg procedure, due to the high number of comparisons, and to minimize the probability of incidental findings. Correcting for multiple testing in metabolomics studies, in cases when the number of variables is much higher than the number of samples, may often reduce most of the p-values below a significance

threshold. Multivariate methods allow for looking at the variables together, thus allowing to identify patterns which are easily seen when through univariate testing.

5.3.2 Use of machine learning in metabolomics

We wanted to investigate if there may be other machine learning approaches (and in specific deep learning approaches) which could achieve a better performance than PLS-DA, inspired by several other studies [179, 223, 224]. A study comparing the predictive performance of a wide range of machine learning approaches on publically available metabolomics data sets showed that support vector machines (SVMs) with a radial kernel basis function and artificial neural networks often showed a better performance than PLS-DA, although the differences in performance were small [223]. Another study showed the superiority of a DL model when predicting the ER status from metabolic profiles of breast cancer tissue [179] compared to other ML algorithms.

An effort was made to try to increase the prediction accuracies for the models in Paper I by applying RFs instead of PLS-DA. DL methods are in general dependent on enough training data to reach a good performance [225] and thus the application of DL for this cohort was not attempted due to the limited sample size. The problem encountered with RF in this study was that it performed poorly on highly imbalanced data. Imbalanced data means that there is a large imbalance between the number of observations (here patients) in different classes (e.g. survivor, non-survivor). The prediction accuracy was in these cases very high, but only because almost all new prediction were assigned the majority class, resulting in a very high number of false positives or false negatives, as seen in Table 4.1. This underpins the importance of looking at sensitivity (true positives/(true positives + false negatives)) and specificity (true negatives/(true negatives + false positives)) in addition to overall accuracy. The imbalanced classification results is a direct result of how a tree, which is the basic unit of a RF, is constructed. A training set consisting of different number of representatives from either class may result in a classifier that is biased towards the majority group. As each tree in a RF is built on a bag (bootstrapped sample of the original data), each tree will in average be biased in the same direction and magnitude by class imbalance. As this is a common problem for binary classifiers, several techniques exist for reducing the class imbalance, such as under-sampling of the majority class or over-sampling of the small class [226]. Alternatively, the way in which the accuracy of the model is calculated can be modified [227]. However, since the RF classifiers showed a very similar performance on the data as the PLS-DA, in cases where class imbalance was not an issue, we chose to present the results of PLS-DA in Paper III.

In Paper III, we found multiple metabolites and lipoprotein parameters significantly associated with a long-term risk of developing breast cancer. The discriminative ability of these associations were

however modest. The data material consisted of both a high number of samples and variables, which provided a unique opportunity to test the performance of different machine learning approaches for data analysis. The ML models tested included RF, GBM, PLS-DA and a neural network. However, although several variables had significant odds ratios, the metabolic information contained in the serum samples of healthy females was not capable of predicting a future breast cancer case. This is probably due to a high heterogeneity of the study population and needs to be investigated in further work on this cohort. By taking advantage of the lifestyle and cancer-specific variables, we can select more homogeneous subgroups of the full cohort. This will reduce the variability which is not linked to the outcome of interest, and thus might increase predictive performance.

The ML models in Paper III were assessed by a double cross validation routine, where the data was split into an 80/20 training and test set. Inside the training set was a separate cross-validation loop for tuning the model parameters, and the performance of the optimal model was tested on the test set. This made it possible to calculate the performance metrics prediction accuracy and AUC score, and also the standard deviation of the metrics. Based on the AUC score, the RF classifier showed the overall best performance, with a mean AUC score of 0.543, while based on the prediction accuracy, PLS-DA models showed the best performance with a mean accuracy of 51.6%. A permutation test, based on 1000 random permutations, on the PLS-DA model for predicting breast cancer within 15 years of follow up, gave a p-value of 0.082, showing that this model is significant at a 0.1 significance level. The PLS-DA models for predicting breast cancer within 5 years and 10 years were however not significant (p-values = 0.144 and 0.480 for 5 and 10 years, respectively). Shrinking the follow-up window did not affect the mean model performances, however an increase in the standard deviation of the metrics was observed indicating less robust models. In terms of the robustness of the ML algorithms, the DNN model had slightly higher standard deviation of the performance metrics, in particular for predicting a short-term BC risk. This could indicate that the model had insufficient training data to reach stability. Across all models and the two types of performance metrics, the standard deviations were in the rage 0.035-0.055, when excluding the standard deviation of the accuracy for the DNN model for short-term risk prediction. The GBM models had an overall lowest standard deviation in measuring the accuracy, while RFs had the lowest standard deviation in estimating the AUC score.

As mentioned earlier, data analysis in metabolomics studies is dominated by partial least squares methods, which project data into latent variables. DL has seldom been applied in metabolomics studies for many reasons: lack of community acceptance, limited sizes of data sets and limitations in computational powers [218]. In the recent years, these issues have become smaller [152, 153, 228]. DL has become much more accessible [229] and has shown promise for metabolomics data, even for limited data sizes [179, 230, 231]. DL is more commonly accepted in many research fields, the

computational powers have increased and the trend is increasing study cohorts. The however main disadvantage of DL is that these models do not return variable contributions to the model, and thus information about the underlying biology is lost. PLS based methods on the other hand are easily interpretable, and return loadings or VIP scores which link metabolite abundance to the outcome. The rationale for applying a wide range of ML algorithms on the cohort in Paper III was the size of the data and was driven largely by a personal interest in data analysis. If one of the ML models, which is a more "black-box" in nature (e.g. the RF, GBM or DNN) would yield a good prediction model, it would be difficult to gain a deeper biological understanding of the formation of cancer based solely on the model. However, it would reveal the presence of a predictive potential in the data set, and other models could be used to investigate which variables were important for making a correct prediction. Also, by applying a range of ML methods, variables important across the models could be compared, thus the models could both reveal robust biomarkers and also give complementary biological information. A data set with stronger biological correlations with the outcome variable would be more appropriate for comparing the model performances.

For the data set in Paper III, a non-targeted approach, where the whole NMR spectra are used as inputs to a DL model might give better predictions, as the variability associated with spectra preprocessing and metabolite quantification is avoided. This could be accomplished by retraining an existing DL model developed for image classification through transfer learning.

5.3.3 Statistical inference

In paper III, fitting the logistic regression model on the whole data caused problems due to the highly multi-collinear nature of the variables, in specific the lipoprotein parameters. For this data set, for pairs of variables with a correlation exceeding approximately 0.98, the model overestimated these parameters and forced them to cancel each other out. This is due to the nature of estimation by least squares: when two strongly correlated variables are included in the model, overestimation of one of the parameters is accompanied by underestimation of the other parameter [232]. This resulted in odds ratios which were highly implausible. For that reason, groups of highly correlated variables were in turn left out from the analyses, and the models were compared for the different subsets of variables. This approach resulted in more stable models with more reasonable results. To find a best subset of variables associated with an increased risk of breast cancer, the best subset selection procedure would be unfeasible due to the high number of model comparisons (because of a high number of variables). Therefore, both forward and backward stepwise selection procedures were tried. Backward selection resulted in a high number of variables included in the model, while forward selection resulted in a more modest number of variables. Because stepwise selection procedures are not optimal for highly correlated variables, also regularization by Lasso was attempted. It however

turned out that also Lasso performed badly shrinking almost all variables to zero. This is probably due to the high correlations among the variables [135]. To overcome this, an adjusted lasso approach was employed instead [233]. This approach gives the best subset of variables that together show the best prediction performance, but which are not necessarily statistically significant on their own. As the predictive performance was evaluated using the ML approaches, the aim of the logistic regression analyses were more statistical inference, and for that reason the discussion in Paper III was mostly based on the results from the forward stepwise subset selection.

There were some discrepancy in which variables were considered important for breast cancer risk assessment depending on the procedure for selecting an optimal subset of variables in the LR models. These discrepancies may reflect different biological mechanisms related to short- and long-term breast cancer risk, but may also be related to 1) differences in sample sizes 2) highly correlated variables and 3) the difference in the natures of subset selection procedures. Performing stratified analyses, where the follow-up time was reduced, resulted in a lower number of samples while the number of variables was kept unchanged. It is recommended that LR models should have 10 observations for one variable, for the smallest of the two outcome groups [222], a condition which was not satisfied in the analyses stratified on years until diagnosis. The high correlation of the variables affects the stepwise selection procedure. Even though the variables with highest correlations were split into different subsets in the LR models, the lipoprotein parameters in specific are highly mulitcollinear. Thus a variable selected in the stepwise selection procedure in one model, could easily be substituted by another variable, without altering the model performance. This might also be the reason that there is a variability in which variables have been selected by the stepwise procedure for models with a different follow-up window.

Figure 5.5 shows a Venn diagram, illustrating the number of overlapping variables chosen for the stratified analyses. Three variables had a significant association with breast cancer for all follow-up periods. Moreover, variables selected by the stepwise procedure and the variables selected by the adaptive lasso were not always overlapping. This reflects the different nature of these two procedures, where the former has a focus of statistical inference, while the later focus on prediction accuracy. Also, for selecting the optimal penalty and weights in the adaptive lasso, a 10-fold cross validation has been implemented, which necessarily will result in some variability in terms of selected variables. The biomarkers (metabolites and lipoprotein parameters) significantly associated with breast cancer risk in Paper III, although not being strong enough for predicting a future breast cancer case, can be considered robust findings as they were identified despite the variation in the data not related to the outcome of interest.

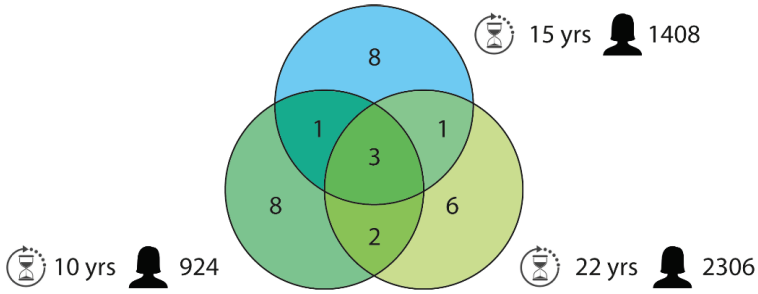


Figure 5.5 A Venn diagram showing overlapping variables which were significantly associated with a future breast cancer in Paper III, in LR models stratified on the length of the follow-up period.

6 Concluding remarks and future perspectives

This thesis illustrates the potential of metabolomics for a more personalized management of breast cancer patients, with a wide spectrum of clinical applications: from risk assessment, through diagnosis, disease subtyping and thus treatment stratification, treatment monitoring and to detection of relapse.

The reproducibility of NMR for metabolic profiling has been investigated in this thesis. Paper II has evaluated the effect of multiple freeze and thaw cycles on the concentrations serum and urine metabolites, and lipoprotein parameters in serum. This study showed an overall good reproducibility of the metabolites and lipoproteins, but with minor, though significant, accumulated effects for some of the lipoprotein parameters.

This thesis has also illustrated the complementary nature of tissue metabolomics, which focuses on the biological processes within the cancerous tissue, and serum metabolomics, which reflects the current state of the whole body. Paper I showed that treatment response of locally advanced breast cancer patients could be predicted from tissue metabolic profiles. Also treatment-induced serum metabolic changes were evaluated, and in specific, treatment-induced increases in the amount of lipids in the serum was observed. This side-effect of treatment has been observed also in other studies, and is likely linked with a higher risk of cardiovascular events for breast cancer survivors.

Paper III investigated associations of circulating metabolites and lipoprotein parameters for a long-term risk of developing breast cancer. The work carried out in this paper provides a better biological understanding of the etiology of breast cancer, and has shown some similarities with the etiology of cardiovascular diseases. In specific, significant positive associations between triglycerides in VLDLs, free cholesterol and Apo-A2 in HDLs, acetate, and valine and a long-term breast cancer risk have been revealed. Similarly, inverse associations between total amount of Apo-A1, free and esterified cholesterol in VLDLs, phospholipids in HDLs and glycine with long-term breast cancer risk were found.

The use of serum metabolomics for risk assessment for breast cancer needs further investigations. Analyzes on more homogeneous subcohorts (e.g. stratified on breast cancer specific variables, or use of hormone therapy) of the HUNT2 cohort might reveal additional insight into the etiology of breast cancer. The large sample size and long follow-up period together with the availability of numerous lifestyle factors allow for future stratified analysis. Further studies should also be performed to investigate if the associations found in Paper III are merely mediators of already known risk factors, such as alcohol intake or overweight. Moreover, using the whole NMR spectra as inputs instead of metabolite concentrations (hence minimizing the variability associated with spectra preprocessing and metabolite quantification) to a DL model developed for image classification, might give a model

with a better predictive performance. Relating serum metabolic profiles with lifestyle factors, and comparing with other diseases, such as cardiovascular diseases and diabetes, the influence of lifestyle and diet on the disease development should be exploited further.

This thesis has also demonstrated the recent developments in metabolomics: from relative to absolute metabolic concentrations, and a trend towards larger patient cohorts, with the advantages that follow. The lack of standardization in this field still makes comparison of metabolic findings across different studies difficult, and hinders the translation of findings to a clinical practice. An increased focus on standardization at all stages in metabolomics research, better tools for automated metabolite quantification, good systems for open publishing, data sharing (of raw data) and multicenter studies are important for the further advances in this field.

7 References

1. Hanahan D, W.R., *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
2. Cooper, C.M. and R.E. Hausman, *The cell: A Molecular Approach*. 7th Edition ed. 2015: Sinauer Associates Inc.,U.S.
3. Peto, J., *Cancer epidemiology in the last century and the next decade.* Nature, 2001. **411**(6835): p. 390-5.
4. Ferlay, J., et al., *Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012.* International Journal of Cancer, 2015. **136**(5): p. E359-E386.
5. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
6. Cancer Registry of Norway, *Cancer in Norway 2017 - Cancer incidence, mortality, survival and prevalence in Norway*. 2018.
7. Cancer Registry of Norway, *Cancer in Norway 2018 - Cancer incidence, mortality, survival and prevalence in Norway*. 2019.
8. American Cancer Society, *Global Cancer Facts & Figures 4th Edition*. 2018.
9. Yung, R.L. and J.A. Ligibel, *Obesity and breast cancer: risk, outcomes, and future considerations.* Clin Adv Hematol Oncol, 2016. **14**(10): p. 790-797.
10. Howell, A., et al., *Risk determination and prevention of breast cancer.* Breast Cancer Res, 2014. **16**(5): p. 446.
11. Haukaas, T.H., *Metabolic profiling of breast cancer using ex vivo MR spectroscopy*. 2016, Norwegian University of Science and Technology.
12. Sitter, B., et al., *Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy.* NMR in biomedicine, 2009. **23**(4).
13. Geddes, D.T., *Inside the lactating breast: The latest anatomy research.* Journal of Midwifery & Womens Health, 2007. **52**(6): p. 556-563.
14. Hynes, N.E. and C.J. Watson, *Mammary gland growth factors: roles in normal development and in cancer.* Cold Spring Harb Perspect Biol, 2010. **2**(8): p. a003186.
15. Pandya, S. and R.G. Moore, *Breast development and anatomy.* Clin Obstet Gynecol, 2011. **54**(1): p. 91-5.
16. Makki, J., *Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance.* Clin Med Insights Pathol, 2015. **8**: p. 23-31.
17. Hanby, A.M. and T.A. Hughes, *In situ and invasive lobular neoplasia of the breast.* Histopathology, 2008. **52**(1): p. 58-66.
18. Chang, J. and O. Chaudhuri, *Beyond proteases: Basement membrane mechanics and cancer invasion.* J Cell Biol, 2019. **218**(8): p. 2456-2469.
19. Kalluri, R., *Basement membranes: structure, assembly and role in tumour angiogenesis.* Nat Rev Cancer, 2003. **3**(6): p. 422-33.
20. Helsedirektoratet. *Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av pasienter med brystkreft*. 2020 [cited 2020 24.09.20]; Available from: https://www.helsebiblioteket.no/retningslinjer/brystkreft/kirurgisk-og-kurativ-behandling/lokalavansert-brystkreft.
21. Garg, P.K. and G. Prakash, *Current definition of locally advanced breast cancer.* Curr Oncol, 2015. **22**(5): p. e409-10.
22. National Breast Cancer Foundation, I. *Male Breast Cancer*. 2020 [cited 2020 24.08.20]; Available from: https://www.nationalbreastcancer.org/male-breast-cancer.
23. Gucalp, A., et al., *Male breast cancer: a disease distinct from female breast cancer.* Breast Cancer Res Treat, 2019. **173**(1): p. 37-48.
24. Rojas, K. and A. Stuckey, *Breast Cancer Epidemiology and Risk Factors.* Clin Obstet Gynecol, 2016. **59**(4): p. 651-672.

25. Chlebowski, R.T., *Nutrition and physical activity influence on breast cancer incidence and outcome.* Breast, 2013. **22 Suppl 2**: p. S30-7.
26. Simpson, E.R., *Sources of estrogen and their importance.* Journal of Steroid Biochemistry and Molecular Biology, 2003. **86**(3-5): p. 225-230.
27. Siiteri, P.K., *Adipose-Tissue as a Source of Hormones.* American Journal of Clinical Nutrition, 1987. **45**(1): p. 277-282.
28. Breastcancer.org. *Breast Cancer Risk Factors*. 2020 [cited 2020 24.08.2020]; Available from: https://www.breastcancer.org/risk/factors.
29. Naume, B. *Årsaker til brystkreft*. Oncolex 2015 [cited 2020 24.08.20]; Available from: http://oncolex.no/Bryst/Bakgrunn/Arsaker.
30. Jones, M.E., et al., *Smoking and risk of breast cancer in the Generations Study cohort.* Breast Cancer Res, 2017. **19**(1): p. 118.
31. Liu, Y., N. Nguyen, and G.A. Colditz, *Links between alcohol consumption and breast cancer: a look at the evidence.* Womens Health, 2015. **11**(1): p. 65-77.
32. McDonald, J.A., A. Goyal, and M.B. Terry, *Alcohol Intake and Breast Cancer Risk: Weighing the Overall Evidence.* Curr Breast Cancer Rep, 2013. **5**(3).
33. Gavaler, J.S., et al., *Alcohol and estrogen levels in postmenopausal women: the spectrum of effect.* Alcohol Clin Exp Res, 1993. **17**(4): p. 786-90.
34. Fabian, C.J., B.F. Kimler, and S.D. Hursting, *Omega-3 fatty acids for breast cancer prevention and survivorship.* Breast Cancer Res, 2015. **17**: p. 62.
35. Naume, B. *Stadier ved brystkreft*. 2015 [cited 2020 24.08.20]; Available from: http://oncolex.no/Bryst/Bakgrunn/Stadier.
36. Zahoor, S., et al., *Sentinel Lymph Node Biopsy in Breast Cancer: A Clinical Review and Update.* J Breast Cancer, 2017. **20**(3): p. 217-227.
37. Chew, H.K., *Adjuvant therapy for breast cancer: who should get what?* West J Med, 2001. **174**(4): p. 284-7.
38. Hammond, M.E., et al., *American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer.* Arch Pathol Lab Med, 2010. **134**(6): p. 907-22.
39. Allison, K.H., et al., *Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update.* J Clin Oncol, 2020. **38**(12): p. 1346-1366.
40. Dai, X., et al., *Cancer Hallmarks, Biomarkers and Breast Cancer Molecular Subtypes.* J Cancer, 2016. **7**(10): p. 1281-94.
41. Weigel, M.T. and M. Dowsett, *Current and emerging biomarkers in breast cancer: prognosis and prediction.* Endocr Relat Cancer, 2010. **17**(4): p. R245-62.
42. Engstrom, M.J., et al., *Molecular subtypes, histopathological grade and survival in a historic cohort of breast cancer patients.* Breast Cancer Res Treat, 2013. **140**(3): p. 463-73.
43. Figueroa-Magalhaes, M.C., et al., *Treatment of HER2-positive breast cancer.* Breast, 2014. **23**(2): p. 128-136.
44. Iqbal, N. and N. Iqbal, *Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications.* Mol Biol Int, 2014. **2014**: p. 852748.
45. National Cancer Institute. *Possible Side Effects*. [cited 2020 23.11.2020]; Available from: https://training.seer.cancer.gov/treatment/chemotherapy/sideeffects.html.
46. Folkman, J., *Tumor angiogenesis: therapeutic implications.* N Engl J Med, 1971. **285**(21): p. 1182-6.
47. Ranieri, G., et al., *Vascular endothelial growth factor (VEGF) as a target of bevacizumab in cancer: from the biology to the clinic.* Curr Med Chem, 2006. **13**(16): p. 1845-57.
48. Huang, H., et al., *An updated meta-analysis of fatal adverse events caused by bevacizumab therapy in cancer patients.* PLoS One, 2014. **9**(3): p. e89960.
49. Avdeling spesialisthelsetjenester, *Nasjonalt handlingsprofram med retningslinjer for diagnostikk, behandling og oppfølging av kreft i tykktarm og endetarm*. 2019: Oslo.

50. von Minckwitz, G., et al., *Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes.* J Clin Oncol, 2012. **30**(15): p. 1796-804.
51. Symmans, W.F., et al., *Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy.* J Clin Oncol, 2007. **25**(28): p. 4414-22.
52. Eisenhauer, E.A., et al., *New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1).* European Journal of Cancer, 2009. **45**(2): p. 228-247.
53. Renata Bujak, W.S.-L., Michal J. Markuszewski, Roman Kaliszan, *Metabolomics for laboratory dianogstics.* Journal of Pharmaceutical and Biomedical Analysis, 2015. **113**: p. 108-120.
54. Baharum, S.N. and K.A. Azizan, *Metabolomics in Systems Biology*, in *Omics Applications for Systems Biology*, W.M. Aizat, H.-H. Goh, and S.N. Baharum, Editors. 2018, Springer International Publishing: Cham. p. 51-68.
55. Oldiges, M., et al., *Metabolomics: current state and evolving methodologies and tools.* Appl Microbiol Biotechnol, 2007. **76**(3): p. 495-511.
56. King, M.C., et al., *Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2.* Science, 2003. **302**(5645): p. 643-6.
57. Mavaddat, N., et al., *Pathology of Breast and Ovarian Cancers among BRCA1 and BRCA2 Mutation Carriers: Results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA).* Cancer Epidemiology Biomarkers & Prevention, 2012. **21**(1): p. 134-147.
58. Sørlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **98**(19).
59. Dai, X.F., et al., *Breast cancer intrinsic subtype classification, clinical use and future trends.* American Journal of Cancer Research, 2015. **5**(10): p. 2929-2943.
60. Larsen, M.J., et al., *Classifications within Molecular Subtypes Enables Identification of BRCA1/BRCA2 Mutation Carriers by RNA Tumor Profiling.* Plos One, 2013. **8**(5).
61. Fagerlund, B.C., et al., *Single Technology Assessment: Prosigna Gene Signatures to Assess Expected Benefit from Chemotherapy in Breast Cnacer. Assessment of manufacturer's submission*. 2019, Norwegian Institute of Public Health: Oslo.
62. Snustad, D.P. and M.J. Simmons, *Principles of Genetics*. 2015: WILEY.
63. Lyons, T.G., *Targeted Therapies for Triple-Negative Breast Cancer.* Curr Treat Options Oncol, 2019. **20**(11): p. 82.
64. Moe, L., *Nå skal noen brystkreftpasienter behandles med immunterapi*, in *Dagens Medisin*. 2020: https://www.dagensmedisin.no/artikler/2020/06/29/na-tilbys-immunterapi-til-noen-brystkreftpasienter/.
65. Cancer Genome Atlas, N., *Comprehensive molecular portraits of human breast tumours.* Nature, 2012. **490**(7418): p. 61-70.
66. Clish, C.B., *Metabolomics: an emerging but powerful tool for precision medicine.* Cold Spring Harb Mol Case Stud, 2015. **1**(1): p. a000588.
67. Yu, Z., et al., *Human serum metabolic profiles are age dependent.* Aging Cell, 2012. **11**(6): p. 960-7.
68. Esko, T., et al., *Metabolomic profiles as reliable biomarkers of dietary composition.* Am J Clin Nutr, 2017. **105**(3): p. 547-554.
69. Sato, S., et al., *Human metabolomics reveal daily variations under nutritional challenges specific to serum and skeletal muscle.* Mol Metab, 2018. **16**: p. 1-11.
70. Troisi, J., et al., *A metabolomics-based approach for non-invasive screening of fetal central nervous system anomalies.* Metabolomics, 2018. **14**(6): p. 77.
71. Tenenboim, H. and Y. Brotman, *Omic Relief for the Biotically Stressed: Metabolomics of Plant Biotic Interactions.* Trends Plant Sci, 2016. **21**(9): p. 781-791.
72. Emwas, A.H., *The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research.* Methods Mol Biol, 2015. **1277**: p. 161-93.

73. Dettmer, K., P.A. Aronov, and B.D. Hammock, *Mass spectrometry-based metabolomics.* Mass Spectrom Rev, 2007. **26**(1): p. 51-78.
74. Gandhi, N. and G.M. Das, *Metabolic Reprogramming in Breast Cancer and Its Therapeutic Implications.* Cells, 2019. **8**(2).
75. Long, J.P., X.N. Li, and F. Zhang, *Targeting metabolism in breast cancer: How far we can go?* World J Clin Oncol, 2016. **7**(1): p. 122-30.
76. Warburg, O., *On the origin of cancer cells.* Science, 1956. **123**(3191): p. 309-14.
77. Vander Heiden, M.G., L.C. Cantley, and C.B. Thompson, *Understanding the Warburg effect: the metabolic requirements of cell proliferation.* Science, 2009. **324**(5930): p. 1029-33.
78. Liberti, M.V. and J.W. Locasale, *The Warburg Effect: How Does it Benefit Cancer Cells?* Trends Biochem Sci, 2016. **41**(3): p. 211-218.
79. Jiang, B., *Aerobic glycolysis and high level of lactate in cancer metabolism and microenvironment.* Genes Dis, 2017. **4**(1): p. 25-27.
80. Lopez, M.J. and S.S. Mohiuddin, *Biochemistry, Essential Amino Acids*, in *StatPearls*. 2020: Treasure Island (FL).
81. Lieu, E.L., et al., *Amino acids in cancer.* Exp Mol Med, 2020. **52**(1): p. 15-30.
82. Choi, B.H. and J.L. Coloff, *The Diverse Functions of Non-Essential Amino Acids in Cancer.* Cancers (Basel), 2019. **11**(5).
83. Yang, C., et al., *Glutamine oxidation maintains the TCA cycle and cell survival during impaired mitochondrial pyruvate transport.* Mol Cell, 2014. **56**(3): p. 414-24.
84. Sitter, B., et al., *Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy.* NMR Biomed, 2010. **23**(4): p. 424-31.
85. Sitter, B., et al., *Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters.* NMR Biomed, 2006. **19**(1): p. 30-40.
86. Marques, L.R., et al., *Reverse Cholesterol Transport: Molecular Mechanisms and the Non-medical Approach to Enhance HDL Cholesterol.* Front Physiol, 2018. **9**: p. 526.
87. Tang, Q.Q., *Lipid metabolism and diseases.* Science Bulletin, 2016. **61**(19): p. 1471-1472.
88. Cheng, C., et al., *Lipid metabolism reprogramming and its potential targets in cancer.* Cancer Commun (Lond), 2018. **38**(1): p. 27.
89. Munir, R., et al., *Lipid metabolism in cancer cells under metabolic stress.* Br J Cancer, 2019. **120**(12): p. 1090-1098.
90. Zaidi, N., et al., *Lipogenesis and lipolysis: the pathways exploited by the cancer cells to acquire fatty acids.* Prog Lipid Res, 2013. **52**(4): p. 585-9.
91. Santos, C.R. and A. Schulze, *Lipid metabolism in cancer.* FEBS J, 2012. **279**(15): p. 2610-23.
92. Nomura, D.K., et al., *Monoacylglycerol Lipase Regulates a Fatty Acid Network that Promotes Cancer Pathogenesis.* Cell, 2010. **140**(1): p. 49-61.
93. Kuemmerle, N.B., et al., *Lipoprotein lipase links dietary fat to solid tumor cell proliferation.* Mol Cancer Ther, 2011. **10**(3): p. 427-36.
94. Thompson, T.E. *Lipid*. 2020 21.02.20202 [cited 2020 01.09]; Available from: https://www.britannica.com/science/lipid.
95. Feingold, K.R. and C. Grunfeld, *Introduction to Lipids and Lipoproteins*, in *Endotext*, K.R. Feingold, et al., Editors. 2000: South Dartmouth (MA).
96. Madssen, T.S., et al., *Metabolite and lipoprotein responses and prediction of weight gain during breast cancer treatment.* Br J Cancer, 2018. **119**(9): p. 1144-1154.
97. Ivanova, E.A., et al., *Small Dense Low-Density Lipoprotein as Biomarker for Atherosclerotic Diseases.* Oxid Med Cell Longev, 2017. **2017**: p. 1273042.
98. Superko, H.R. and R.R. Gadesam, *Is It LDL Particle Size or Number that Correlates with Risk for Cardiovascular Disease?* Current Atherosclerosis Reports, 2008. **10**(5): p. 377-385.
99. Ren, L., et al., *Apolipoproteins and cancer.* Cancer Med, 2019. **8**(16): p. 7032-7043.
100. Borgquist, S., et al., *Apolipoproteins, lipids and risk of cancer.* Int J Cancer, 2016. **138**(11): p. 2648-56.

101. Katzke, V.A., et al., *Blood lipids and lipoproteins in relation to incidence and mortality risks for CVD and cancer in the prospective EPIC-Heidelberg cohort.* BMC Med, 2017. **15**(1): p. 218.
102. Jove, M., et al., *A plasma metabolomic signature discloses human breast cancer.* Oncotarget, 2017. **8**(12): p. 19522-19533.
103. Cala, M.P., et al., *Multiplatform plasma metabolic and lipid fingerprinting of breast cancer: A pilot control-case study in Colombian Hispanic women.* PLoS One, 2018. **13**(2): p. e0190958.
104. Xie, G., et al., *Lowered circulating aspartate is a metabolic feature of human breast cancer.* Oncotarget, 2015. **6**(32): p. 33369-81.
105. His, M., et al., *Prospective analysis of circulating metabolites and breast cancer in EPIC.* BMC Med, 2019. **17**(1): p. 178.
106. Bro, R., et al., *Forecasting individual breast cancer risk using plasma metabolomics and biocontours.* Metabolomics, 2015. **11**(5): p. 1376-1380.
107. Lecuyer, L., et al., *NMR metabolomic signatures reveal predictive plasma metabolites associated with long-term risk of developing breast cancer.* Int J Epidemiol, 2018. **47**(2): p. 484-494.
108. Kuhn, T., et al., *Higher plasma levels of lysophosphatidylcholine 18:0 are related to a lower risk of common cancers in a prospective metabolomics study.* BMC Med, 2016. **14**: p. 13.
109. Giskeødegård, G.F., et al., *NMR-based metabolomics of biofluids in cancer.* NMR Biomedicine, 2018(Special Issue Review Article).
110. Gil, R.R. and A. Navarro-Vázquez, *Modern NMR Approaches to the Structure Elucidation of Natural Products*. Vol. Volume 2: Data Acquisition and Applications to Compound Classes. 2016.
111. Keeler, J., *Understanding NMR Spectroscopy*. Second ed. 2010: Wiley.
112. Claridge, T.D.W., *High-Resolution NMR Techniques in Organic Chemistry*, ed. J.E. Baldwin. Vol. 19. 1999: Elsevier Science Ltd.
113. Euceda, L.R., G.F. Giskeødegård, and T.F. Bathen, *Preprocessing of NMR metabolomics data.* Scandinavian Journal of Clinical and Laboratory Investigation, 2015: p. 192-203.
114. Smolinska, A., et al., *NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review.* Anal Chim Acta, 2012. **750**: p. 82-97.
115. Xi, Y. and D.M. Rocke, *Baseline correction for NMR spectroscopic metabolomics data analysis.* BMC Bioinformatics, 2008. **9**: p. 324.
116. Engel, J., et al., *Breaking with trends in pre-processing?* Trac-Trends in Analytical Chemistry, 2013. **50**: p. 96-106.
117. Tomasi, G., F. Savorani, and S.B. Engelsen, *icoshift: An effective tool for the alignment of chromatographic data.* J Chromatogr A, 2011. **1218**(43): p. 7832-40.
118. Price, W.S., *Water signal suppression in NMR spectroscopy.* Annual Reports on Nmr Spectroscopy, Vol 38, 1999. **38**: p. 289-354.
119. Emwas, A.H.M., et al., *NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations.* Metabolomics, 2013. **9**(5): p. 1048-1072.
120. Dieterle, F., et al., *Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics.* Anal Chem, 2006. **78**(13): p. 4281-90.
121. Craig, A., et al., *Scaling and normalization effects in NMR spectroscopic metabonomic data sets.* Anal Chem, 2006. **78**(7): p. 2262-7.
122. Hao, J., et al., *Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN.* Nat Protoc, 2014. **9**(6): p. 1416-27.
123. Mercier, P., et al., *Towards automatic metabolomic profiling of high-resolution one-dimensional proton NMR spectra.* J Biomol NMR, 2011. **49**(3-4): p. 307-23.
124. Debik, J., et al., *Assessing Treatment Response and Prognosis by Serum and Tissue Metabolomics in Breast Cancer Patients.* J Proteome Res, 2019. **18**(10): p. 3649-3660.

125. Bathen, T.F., et al., *Quantification of plasma lipids and apolipoproteins by use of proton NMR spectroscopy, multivariate and neural network analysis.* NMR Biomed, 2000. **13**(5): p. 271-88.
126. Aru, V., et al., *Quantification of lipoprotein profiles by nuclear magnetic resonance spectroscopy and multivariate data analysis (vol 94, pg 210, 2017).* Trac-Trends in Analytical Chemistry, 2019. **119**.
127. Rosner, B., *Fundamentals of Biostatistics*. Eight ed. 2015, Boston: Cengage Learning.
128. Pinheiro, J.C. and D.M. Bates, *Mixed-Effects Models in S and S-PLUS*. Statistics and Computing. 2000.
129. Zuur, A.F., et al., *Mixed effects models and extensions in ecology with R*. Statistics for Biology and Health. 2009, New York: Springer.
130. Hosmer, D.W. and S. Lemeshow, *Applied Logistic Regression*. Second edition ed. 2000: John Wiley & Sons, Inc.
131. Agresti, A., *An introduction to categorical data analysis*. 1996: New York N.Y. : J. Wiley & Sons.
132. Steyerberg, E.W., et al., *Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets.* Statistics in Medicine, 2000. **19**(8): p. 1059-1079.
133. Hastie, T., R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning - Data Mining, Inference and Prediction*. 2nd ed. Springer Series in Statistics. 2009: Springer New York Inc.
134. Tibshirani, R., *Regression shrinkage and selection via the Lasso.* Journal of the Royal Statistical Society Series B-Methodological, 1996. **58**(1): p. 267-288.
135. Zou, H. and T. Hastie, *Regularization and variable selection via the elastic net (vol B 67, pg 301, 2005).* Journal of the Royal Statistical Society Series B-Statistical Methodology, 2005. **67**: p. 768-768.
136. Géron, A., *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 2nd ed. 2019: O'Reilly Media.
137. Chollet, F., *Deep Learning with Python*. 2018, Shelter Island, NY 11964: Manning Publications Co.
138. Wold, S., K. Esbensen, and P. Geladi, *Principal Component Analysis.* Chemometrics and Intelligent Laboratory Systems, 1987. **2**: p. 37-52.
139. Wold, S., M. Sjöström, and L. Eriksson, *PLS-regression: a basic tool of chemometrics.* Chemometrics and Intelligent Laboratory Systems, 2001. **58**(2): p. 109-130.
140. Dejong, S., *Simpls - an Alternative Approach to Partial Least-Squares Regression.* Chemometrics and Intelligent Laboratory Systems, 1993. **18**(3): p. 251-263.
141. Bylesjö, M., et al., *OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification.* Journal of Chemometrics, 2006. **20**: p. 341-351.
142. Westerhuis, J.A., et al., *Multivariate paired data analysis: multilevel PLSDA versus OPLSDA.* Metabolomics, 2010. **6**(1): p. 119-128.
143. World Health Organization - Regional Office for Europe. *Body mass index - BMI*. 2020 [cited 2020 24.08.20]; Available from: https://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi.
144. Liaw, A. and M. Wiener, *Classification and Regression by randomForest.* R News, 2002. **2**: p. 18-22.
145. James, G., et al., *An Introduction to Statistical Learning: with Applications in R*. 2014: Springer Publishing Company, Incorporated.
146. Efron, B., *1977 Rietz Lecture - Bootstrap Methods - Another Look at the Jackknife.* Annals of Statistics, 1979. **7**(1): p. 1-26.
147. Boehmke, B. and B.M. Greenwell, *Gradient Boosting*, in *Hands on machine learning with R*, C.a. Hall/CRC, Editor. 2019.
148. Ridgeway, G., *Generalized Boosted Models: a guide to the gbm package.* 2019.
149. Chollet, F. and J.J. Allaire, *Deep Learning with R*. 1st ed. 2018: Manning Publications Co.

150. Zemouri, R., N. Zerhouni, and D. Racoceanu, *Deep Learning in the Biomedical Applications: Recent and Future Status.* Applied Sciences-Basel, 2019. **9**(8).
151. Perez-Enciso, M. and L.M. Zingaretti, *A Guide for Using Deep Learning for Complex Trait Genomic Prediction.* Genes (Basel), 2019. **10**(7).
152. Mamoshina, P., et al., *Applications of Deep Learning in Biomedicine.* Mol Pharm, 2016. **13**(5): p. 1445-54.
153. Cao, C., et al., *Deep Learning and Its Applications in Biomedicine.* Genomics Proteomics Bioinformatics, 2018. **16**(1): p. 17-32.
154. Kone, C. *Introducing Convolutional Neural Networks in Deep Learning*. 2019 [cited 2020 18.08.2020]; Available from: https://towardsdatascience.com/introducing-convolutional-neural-networks-in-deep-learning-400f9c3ad5e9.
155. Saha, S. *A Comprehensive Guide to Convolutional Neural Networks - the ELI5 way*. 2018 [cited 2020 18.08.2020]; Available from: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.
156. Goodfellow, I., Y. Bengio, and A. Courville, *Deep learning*. 2016, Cambridge, Mass: MIT Press.
157. Krstajic, D., et al., *Cross-validation pitfalls when selecting and assessing regression and classification models.* Journal of Cheminformatics, 2014. **6**.
158. Westerhuis, J.A., et al., *Assessment of PLSDA cross validation.* Metabolomics, 2008. **4**(1): p. 81-89.
159. Bonferroni, C.E., *Teoria statistica delle classi e calcolo delle probabilità.* Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze, 1936. **8**: p. 3-62.
160. Bonferroni, C.E., *Il calcolo delle assicurazioni su gruppi di teste.* Studi in Onore del Professore Salvatore Ortu Carboni, 1935: p. 13-60.
161. Benjamini, Y. and Y. Hockberg, *Controlling the false discobery rate: A practical and powerful approach for multiple testing.* Journal of the Royal Statistical Society B, 1995. **57**: p. 289-300.
162. Lewallen, S. and P. Courtright, *Epidemiology in practice: case-control studies.* Community Eye Health, 1998. **11**(28): p. 57-8.
163. Song, J.W. and K.C. Chung, *Observational studies: cohort and case-control studies.* Plast Reconstr Surg, 2010. **126**(6): p. 2234-42.
164. Euceda, L.R., et al., *Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy.* Metabolomics, 2017. **13**(4): p. 37.
165. *MATLAB*. 2017, The MathWorks Inc.: Natick, Massachusetts.
166. Cloarec, O., et al., *Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets.* Anal Chem, 2005. **77**(5): p. 1282-9.
167. Jimenez, B., et al., *Quantitative Lipoprotein Subclass and Low Molecular Weight Metabolite Analysis in Human Serum and Plasma by (1)H NMR Spectroscopy in a Multilaboratory Trial.* Anal Chem, 2018. **90**(20): p. 11962-11971.
168. R Development Core Team, *R: A Language and Environment for Statistical Computing*. 2009: Vienna, Austria.
169. Eigenvector Research, I., *PLS Toolbox*. 2018: Manson, WA USA 98831.
170. Greenwell, B., et al., *gbm: Generalized Boosted Regression Models*. 2020.
171. Friedman, J., T. Hastie, and R. Tibshirani, *Regularization Paths for Generalized Linear Models via Coordinate Descent.* Journal of Statistical Software, 2010. **33**: p. 1-22.
172. Chollet, F., *Keras*. 2015, GitHub.
173. Harris, C.R., et al., *Array programming with NumPy.* Nature, 2020. **585**: p. 357-362.
174. The pandas development team, *pandas-dev/pandas: Pandas*. 2020.
175. McKinney, W. *Data Structure for Statistical Computing in Python*. 2010.
176. Pedregosa, F., et al., *Scikit-learn: Machine Learning in Python.* Journal of Machine Learning Research, 2011. **12**: p. 2825-2830.

177. Psychogios, N., et al., *The human serum metabolome.* PLoS One, 2011. **6**(2): p. e16957.
178. Bouatra, S., et al., *The human urine metabolome.* PLoS One, 2013. **8**(9): p. e73076.
179. Alakwaa, F.M., K. Chaudhary, and L.X. Garmire, *Deep Learning Accurately Predicts Estrogen Receptor Status in Breast Cancer Metabolomics Data.* J Proteome Res, 2018. **17**(1): p. 337-347.
180. Cao, M.D., et al., *Metabolic characterization of triple negative breast cancer.* BMC Cancer, 2014. **14**: p. 941.
181. Haukaas, T.H., et al., *Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes.* Cancer Metab, 2016. **4**: p. 12.
182. Beckmann, M., et al., *Dietary exposure biomarker-lead discovery based on metabolomics analysis of urine samples.* Proc Nutr Soc, 2013. **72**(3): p. 352-61.
183. Pechlivanis, A., et al., *1H NMR study on the short- and long-term impact of two training programs of sprint running on the metabolic fingerprint of human serum.* J Proteome Res, 2013. **12**(1): p. 470-80.
184. Carayol, M., et al., *Blood Metabolic Signatures of Body Mass Index: A Targeted Metabolomics Study in the EPIC Cohort.* J Proteome Res, 2017. **16**(9): p. 3137-3146.
185. Davies, S.K., et al., *Effect of sleep deprivation on the human metabolome.* Proc Natl Acad Sci U S A, 2014. **111**(29): p. 10761-6.
186. Wood, L.R.E., *Metabolic characterization of breast cancer heterogenity and response to treatment*. 2016, Norwegian University of Science and Technology.
187. Bradshaw, P.T., et al., *Cardiovascular Disease Mortality Among Breast Cancer Survivors.* Epidemiology, 2016. **27**(1): p. 6-13.
188. Hooning, M.J., et al., *Long-term risk of cardiovascular disease in 10-year survivors of breast cancer.* J Natl Cancer Inst, 2007. **99**(5): p. 365-75.
189. Aleman, B.M., et al., *Cardiovascular disease after cancer therapy.* EJC Suppl, 2014. **12**(1): p. 18-28.
190. Nome, M.E., et al., *Serum levels of inflammation-related markers and metabolites predict response to neoadjuvant chemotherapy with and without bevacizumab in breast cancers.* Int J Cancer, 2020. **146**(1): p. 223-235.
191. van Roekel, E.H., et al., *Circulating Metabolites Associated with Alcohol Intake in the European Prospective Investigation into Cancer and Nutrition Cohort.* Nutrients, 2018. **10**(5).
192. Mehta, L.S., et al., *Cardiovascular Disease and Breast Cancer: Where These Entities Intersect A Scientific Statement From the American Heart Association.* Circulation, 2018. **137**(8): p. E30-E66.
193. Wittemans, L.B.L., et al., *Assessing the causal association of glycine with risk of cardio-metabolic diseases.* Nat Commun, 2019. **10**(1): p. 1060.
194. Lustgarten, M.S., et al., *Serum Glycine Is Associated with Regional Body Fat and Insulin Resistance in Functionally-Limited Older Adults.* Plos One, 2013. **8**(12).
195. Li, X., et al., *Association of serum glycine levels with metabolic syndrome in an elderly Chinese population.* Nutr Metab (Lond), 2018. **15**: p. 89.
196. Yan-Do, R. and P.E. MacDonald, *Impaired "Glycine"-mia in Type 2 Diabetes and Potential Mechanisms Contributing to Glucose Homeostasis.* Endocrinology, 2017. **158**(5): p. 1064-1073.
197. Jain, M., et al., *Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation.* Science, 2012. **336**(6084): p. 1040-4.
198. Kim, S.K., W.H. Jung, and J.S. Koo, *Differential expression of enzymes associated with serine/glycine metabolism in different breast cancer subtypes.* PLoS One, 2014. **9**(6): p. e101004.
199. Playdon, M.C., et al., *Nutritional metabolomics and breast cancer risk in a prospective study.* Am J Clin Nutr, 2017. **106**(2): p. 637-649.
200. Moore, S.C., et al., *A Metabolomics Analysis of Body Mass Index and Postmenopausal Breast Cancer Risk.* J Natl Cancer Inst, 2018. **110**(6): p. 588-597.

201. Moore, S.C., *Metabolomics and breast cancer: scaling up for robust results.* BMC Med, 2020. **18**(1): p. 18.
202. Onderwater, G.L.J., et al., *Large-scale plasma metabolome analysis reveals alterations in HDL metabolism in migraine.* Neurology, 2019. **92**(16): p. e1899-e1911.
203. Julkunen, H., et al., *Blood biomarker score identifies individuals at high risk for severe COVID-19 a decade prior to diagnosis: metabolic profiling of 105,000 adults in the UK Biobank*. medRxiv, 2020: p. 2020.07.02.20143685.
204. Deelen, J., et al., *A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals.* Nat Commun, 2019. **10**(1): p. 3346.
205. Mestrelab Research. *Mnova*. 2020 [cited 2020 27.11.2020]; Available from: https://mestrelab.com/software/mnova/.
206. Markley, J.L., et al., *The future of NMR-based metabolomics.* Curr Opin Biotechnol, 2017. **43**: p. 34-40.
207. Saccenti, E., *Correlation Patterns in Experimental Data Are Affected by Normalization Procedures: Consequences for Data Analysis and Network Inference.* Journal of Proteome Research, 2017. **16**(2): p. 619-634.
208. Wilkinson, M.D., et al., *The FAIR Guiding Principles for scientific data management and stewardship.* Sci Data, 2016. **3**: p. 160018.
209. Emwas, A.H., et al., *NMR Spectroscopy for Metabolomics Research.* Metabolites, 2019. **9**(7).
210. Giskeodegard, G.F., et al., *The effect of sampling procedures and day-to-day variations in metabolomics studies of biofluids.* Anal Chim Acta, 2019. **1081**: p. 93-102.
211. Haukaas, T.H., et al., *Impact of Freezing Delay Time on Tissue Samples for Metabolomic Studies.* Front Oncol, 2016. **6**: p. 17.
212. Dunn, W.B., et al., *The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans.* Bioanalysis, 2012. **4**(18): p. 2249-64.
213. Bruker BioSpin. *Bruker Announces More Comprehensive and Efficient NMR Profiling Module for Detailed Lipoprotein Subclass Analysis in Cardiovascular Disease Research*. 2016 [cited 2020 27.11.2020]; Available from: https://ir.bruker.com/press-releases/press-release-details/2016/Bruker-Announces-More-Comprehensive-and-Efficient-NMR-Profiling-Module-for-Detailed-Lipoprotein-Subclass-Analysis-in-Cardiovascular-Disease-Research/default.aspx.
214. Bruker BioSpin. *Bruker Announces Unique New NMR and MS Tools at Metabolomics 2018*. 2018 27.11.2020]; Available from: https://www.bruker.com/news/bruker-announces-unique-new-nmr-and-ms-tools-at-metabolomics-2018.html.
215. Yin, P., R. Lehmann, and G. Xu, *Effects of pre-analytical processes on blood samples used in metabolomics studies.* Anal Bioanal Chem, 2015. **407**(17): p. 4879-92.
216. Lacy, P., et al., *Signal intensities derived from different NMR probes and parameters contribute to variations in quantification of metabolites.* PLoS One, 2014. **9**(1): p. e85732.
217. Pinto, R.C., *Chemometrics Methods and Strategies in Metabolomics.* Adv Exp Med Biol, 2017. **965**: p. 163-190.
218. Mendez, K.M., D.I. Broadhurst, and S.N. Reinke, *The application of artificial neural networks in metabolomics: a historical perspective.* Metabolomics, 2019. **15**(11): p. 142.
219. Gromski, P.S., et al., *A tutorial review: Metabolomics and partial least squares-discriminant analysis--a marriage of convenience or a shotgun wedding.* Anal Chim Acta, 2015. **879**: p. 10-23.
220. Cao, M.D., et al., *Prognostic value of metabolic response in breast cancer patients receiving neoadjuvant chemotherapy.* BMC Cancer, 2012. **12**: p. 39.
221. Skovlund, E. and G.U. Fenstad, *Should we always choose a nonparametric test when comparing two apparently nonnormal distributions?* Journal of Clinical Epidemiology, 2001. **54**(1): p. 86-92.
222. Lydersen, S., *Statistical review: frequently given comments.* Ann Rheum Dis, 2015. **74**(2): p. 323-5.

223. Mendez, K.M., S.N. Reinke, and D.I. Broadhurst, *A comparative evaluation of the generalised predictive ability of eight machine learning algorithms across ten clinical metabolomics data sets for binary classification.* Metabolomics, 2019. **15**(12): p. 150.
224. Giskeodegard, G.F., et al., *Multivariate Modeling and Prediction of Breast Cancer Prognostic Factors Using MR Metabolomics.* Journal of Proteome Research, 2010. **9**(2): p. 972-979.
225. Kopper, A., et al., *Model Selection and Evaluation for Machine Learning: Deep Learning in Materials Processing.* Integrating Materials and Manufacturing Innovation, 2020. **9**(3): p. 287-300.
226. Krawczyk, B., *Learning from imbalanced data: open challenges and future directions.* Progress in Artificial Intelligence, 2016. **5**(4): p. 221-232.
227. Bekkar, M., H.K. Djemaa, and T.A. Alitouche, *Evaluation measures for models assessment over imbalanced data sets.* J Inf Eng Appl, 2013. **3**(10).
228. Huang, T., et al., *Promises and Challenges of Big Data Computing in Health Sciences.* Big Data Research, 2015. **2**(1): p. 2-11.
229. Mendez, K.M., D.I. Broadhurst, and S.N. Reinke, *Migrating from partial least squares discriminant analysis to artificial neural networks: a comparison of functionally equivalent visualisation and feature contribution tools using jupyter notebooks.* Metabolomics, 2020. **16**(2): p. 17.
230. Asakura, T., Y. Date, and J. Kikuchi, *Application of ensemble deep neural network to metabolomics studies.* Anal Chim Acta, 2018. **1037**: p. 230-236.
231. Trainor, P.J., A.P. DeFilippis, and S.N. Rai, *Evaluation of Classifier Performance for Multiclass Phenotype Discrimination in Untargeted Metabolomics.* Metabolites, 2017. **7**(2).
232. Knight, K. and W.J. Fu, *Asymptotics for Lasso-type estimators.* Annals of Statistics, 2000. **28**(5): p. 1356-1378.
233. Zou, H., *The adaptive lasso and its oracle properties.* Journal of the American Statistical Association, 2006. **101**(476): p. 1418-1429.

Paper I

Journal of proteome research

Cite This: J. Proteome Res. XXXX, XXX, XXX−XXX

Article

pubs.acs.org/jpr

Assessing Treatment Response and Prognosis by Serum and Tissue Metabolomics in Breast Cancer Patients

Julia Debik,*[,†] Leslie R. Euceda,[†,‡] Steinar Lundgren,[§,∥] Hedda von der Lippe Gythfeldt,[⊥] Øystein Garred,[¶] Elin Borgen,[¶] Olav Engebraaten,[⊥,¶,#] Tone F. Bathen,[†] and Guro F. Giskeødegård*[,†]

[†]Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences and [∥]Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, NTNU—Norwegian University of Science and Technology, 7491 Trondheim, Norway

[‡]CAMO Analytics, 0349 Oslo, Norway

[§]Department of Oncology, St. Olav's University Hospital, 7006 Trondheim, Norway

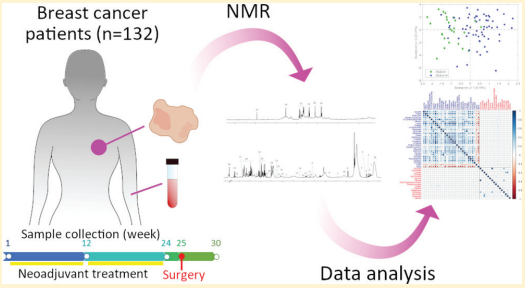
[⊥]Department of Oncology, Oslo University Hospital, 0424 Oslo, Norway

[¶]Department of Tumor Biology, Oslo University Hospital, 0424 Oslo, Norway

[#]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, 0318 Oslo, Norway

Supporting Information

ABSTRACT: Patients with locally advanced breast cancer have a worse prognosis compared to patients with localized tumors and require neoadjuvant treatment before surgery. The aim of this study was to characterize the systemic metabolic effect of neoadjuvant chemotherapy in patients with large primary breast cancers and to relate these changes to treatment response and long-term survival. This study included 132 patients with large primary breast tumors randomized to receive neoadjuvant chemotherapy with or without the addition of the antiangiogenic drug Bevacizumab. Tumor biopsies and serum were collected before and during treatment and, serum additionally 6 weeks after surgery. Samples were analyzed by nuclear magnetic resonance spectroscopy (NMR). Correlation analysis showed low correlations between metabolites measured in cancer tissue and serum. Multilevel partial least squares discriminant analysis (PLS-DA) showed clear changes in serum metabolite levels during treatment (p-values ≤ 0.001), including unfavorable changes in lipid levels. PLS-DA revealed metabolic differences between tissue samples from survivors and nonsurvivors collected 12 weeks into treatment with an accuracy of 72% (p-value = 0.005); however, this was not evident in serum samples. Our results demonstrate a potential clinical application for serum-metabolomics for patient monitoring during and after treatment, and indicate potential for tissue NMR spectroscopy for predicting patient survival.



KEYWORDS: *metabolomics, breast cancer, serum, tissue, NMR, response, survival*

INTRODUCTION

Breast cancer (BC) is the most frequent cancer type in women in Norway. Compared to cancer-free women of the same age, five-year survival of BC patients is 90% in Norway, but ranges from 28 to 100%, depending on the stage of the disease at the beginning of treatment.[1] It is however challenging to accurately predict outcome for individual patients, as there is high diversity in prognosis and response to treatment. This is due to the heterogeneous biology of the disease, resulting in patients with similar histology, clinical diagnosis, and stage of disease having a different prognosis.[2,3] BC is often divided into five genetic intrinsic subtypes, however, many studies have shown that there are many subgroups within these groups.[4−6] One type of treatment will thus not be beneficial for all patients and stratification of patients followed by application of targeted therapy may improve the overall long-term outcome of BC patients.

Locally advanced BC (LABC) patients, that is patients with large tumors or extension to lymph nodes, constitute 10−15% of diagnosed patients with a higher risk of future metastasis.[7] Neoadjuvant chemotherapy (NAC) is administered routinely in LABC patients. This treatment was initially developed to reduce the size of inoperable tumors prior to surgery and for eradication of potential micrometastasis, but is now also a tool to enable breast-conserving surgery.[8,9]

Received: May 15, 2019
Published: September 4, 2019

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

Angiogenesis, the formation of new blood vessels from existing vasculature, has an essential role for supplying nutrients and oxygen to rapidly growing tumors.[10] This process can be therapeutically targeted by antiangiogenetic treatment.[11] Bevacizumab has the ability to inhibit the proangiogenetic vascular endothelial growth factor.[12]

Because of improvements in treatment together with earlier diagnosis, mortality due to BC has decreased during the last years.[1] However, despite intensive treatment regimes, a great proportion of LABC patients will develop metastatic disease.[13,14] Additionally, treatment may induce unwanted long-term side effects, such as fatigue, increased risk of cardiovascular diseases (CVDs), and cardiotoxicity.[15−19] Characterizing the systemic effect of cancer treatment may further enhance our understanding of unwanted side-effects and potentially identify mechanisms to prevent late effects.

Metabolomics is a rapidly growing field in medical research and makes it possible to look at the contents of a biological matrix at the molecular level. Metabolites are downstream biochemical products in the omics cascade, and altered metabolism has been defined as a hallmark of cancer.[11] Following a minimal sample preparation, a wide range of metabolites can be detected within a short amount of time using nuclear magnetic resonance (NMR) spectroscopy.[20] NMR metabolomics has already shown potential in stratification of BC patients with respect to treatment response and long-term survival.[21,22] Most studies so far have focused on metabolomics of invasive tissue biopsies.[3,21−24] Metabolomics of biofluids is minimally invasive and repeated sampling is simple. A recent review concludes that many studies have shown impressive associations between biofluid metabolomics and cancer progression, suggesting that NMR metabolomics can be used to provide information with prognostic or predictive value.[25]

The NeoAva study is a phase II randomized clinical trial assessing the effect of antiangiogenesis treatment by bevacizumab in combination with standard NAC. We have previously shown that metabolic profiling of tumor tissue by magnetic resonance spectroscopy has a potential in predicting treatment response in this cohort.[21] Furthermore, both clinical and gene expression response was shown to differ between patients receiving combination therapy with bevacizumab and chemotherapy alone, and circulating cytokine profiles were found to correlate with different immune cell types at the tumor site.[26−28]

In this study, we performed metabolic profiling of serum samples from patients in the NeoAva study. The main aim was to characterize systemic metabolic effects of NAC in BC patients and to relate these changes to treatment response and long-term survival. Additionally, the metabolic information in serum and tissue samples from the same patients were compared, allowing for a better understanding of the difference in their metabolic information.

MATERIALS AND METHODS

Patient and Tumor Characteristics

Details of the inclusion criteria are fully described elsewhere.[21] In brief, 132 women of age ≥ 18 years with large (≥2.5 cm), non-metastatic, human epidermal growth factor receptor 2 (HER 2) negative tumors were recruited in the period November 2008−July 2012 in Norway. The study was approved for all centers involved by the Regional Ethics Committee (approval number S-08354a) and the Norwegian Medical Agency and an informed written consent was obtained from all patients. All patients included in this study received NAC in the form of FEC100 (5-fluorouracil 600 mg/m^2, epirubicine 100 mg/m^2, and cyclophosphamide 600 mg/m^2) followed by taxane-based therapy for 12 weeks, whereas they were randomized to receive bevacizumab or not. Tissue samples were obtained by ultrasound-guided needle biopsies prior to treatment (TP1) and 12 weeks into treatment (TP2), whereas surgical biopsies were obtained from the surgically removed tumor (TP3). Nonfasting serum was sampled at TP1, TP2, and TP3, in addition to 6 weeks after surgery (TP4). See Figure S1 for a graphical representation of the study design. The study cohort for further analyses has been restricted to contain subjects with full clinical data and available sample material from at least one sampling time point, giving N = 118 subjects. In total, 357 serum samples and 270 tissue samples were analyzed. Details on the patient and tumor characteristics are summarized in Table 1, whereas sample availability, including survival data, for each time point is illustrated in Figure S2.

Table 1. Patient Cohort and Tumor Characteristics[a]

	survivors ≥ 5 years	non-survivors
N	105	13
	Age (years)	
mean (range)	49.3 (25−70)	45.7 (31−55)
	Treatment	
Bev + chemo	53	7
chemo only	52	6
	RCB Class	
0	19	1
I	13	1
II	58	8
III	15	3
	ER Status	
positive	90	10
negative	15	3
	PgR Status	
positive	62	6
negative	43	7
	Histology	
ductal	84	11
lobular	19	1
other	2	1
	Metastasis during follow-up	
yes	5	13
no	100	0

[a]Sample availability varied for each time point, giving a slightly different number of survivors and nonsurvivors used in the prediction models. Details on sample availability are illustrated in Figure S2. Survivors are patients alive 5 years after treatment start; Bev + chemo: bevacizumab-treated in addition to NAC; chemo only: chemotherapy only, no bevacizumab; RCB: residual cancer burden; ER: estrogen receptor; PgR: progesterone receptor.

Prognostic Measures and Survival Evaluation

Residual cancer burden (RCB) is a measurement of patient response to NAC. It is a continuous index, which combines pathologic measurements of the primary tumor (size and cellularity) and nodal metastases (number and size).[29] RCB can be divided into four classes, where class 0 is equivalent to

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

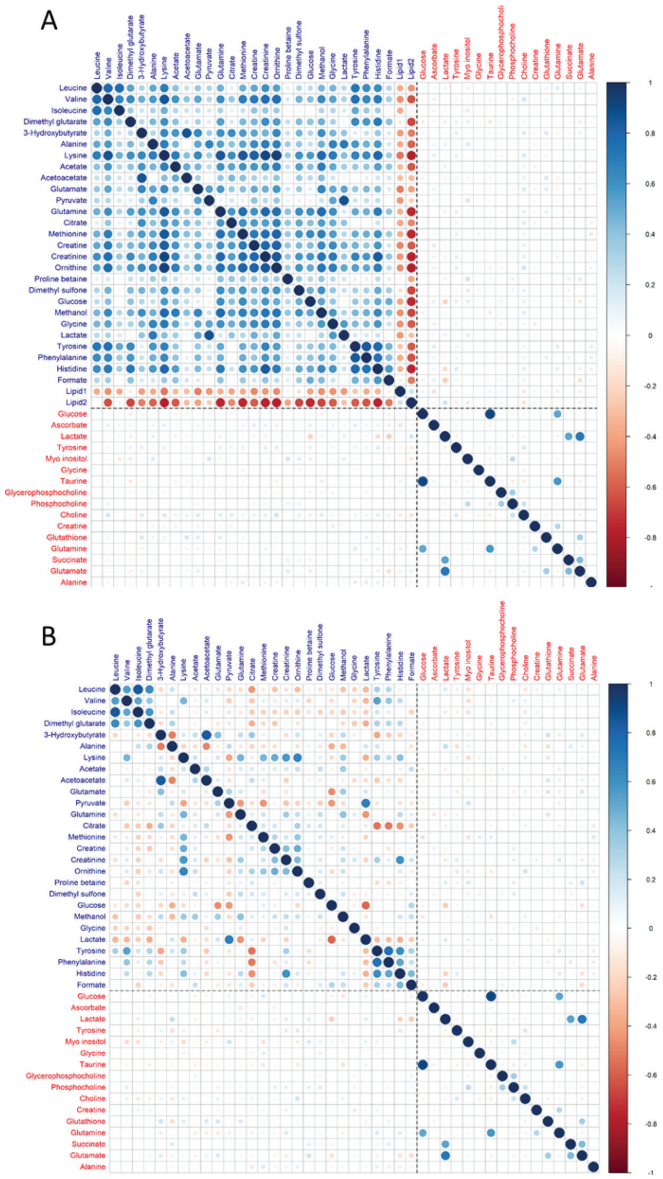


Figure 1. Significant Pearson-correlations between metabolites in serum (blue) and tissue (red) samples. (A) Whole data set; (B) after removal of lipid peaks from serum data and a second normalization. Color intensity and circle sizes are proportional to the correlation coefficients. Red and blue circles indicate negative and positive correlations, respectively. Only patients with both serum and tissue samples available (TP1, TP2, and TP3) have been included in this analysis.

pathologic complete response (pCR), meaning that no cancer cells are present after treatment.

Patients deceased within 5 years after diagnosis were classified as nonsurvivors, whereas patients surviving ≥5 years were classified as survivors.

NMR Experiments and Data Preprocessing

Analysis and Preprocessing of Serum Samples. NMR spectra were obtained on a Bruker AVANCE III Ultrashield Plus spectrometer operating at 600 MHz (Bruker BioSpin GmbH, Rheinstetten, Germany) equipped with a 5 mm QCI CryoProbe. The serum samples were thawed at 4 °C prior to the analysis. The serum (150 μL) was gently mixed with 150 μL of buffer (D_2O with 0.075 mM Na_2HPO_4, 5 mM NaN_3, 3.5 mM TSP, pH 7.4). The samples were analyzed in 3 mm NMR tubes. Data acquisition and sample handling was fully automated using a SampleJet with Icon-NMR on TopSpin

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

3.1 (Bruker BioSpin). Carr−Purcell−Meiboom−Gill (CPMG) spectra with water presaturation were acquired at a temperature of 37 °C. The spectra were Fourier transformed to 128 K after 0.3 Hz exponential line broadening.

The spectral data were transferred to MATLAB R2017b for preprocessing.[30] The left peak of the alanine doublet at 1.47 ppm was used as a chemical shift reference. Three spectra were removed from the analysis due to poor water suppression after visual inspection. Spectral peaks were aligned to the peaks of the spectrum with the highest correlation to the other spectra using the function icoshift.[31] The water region (4.3−5.0 ppm) was removed, and the spectral area between 0.2 and 9.2 ppm was used for further analysis. The NMR signals were assigned to metabolites both using the human metabolome database, published literature, and an in-house overview over previously assigned spectral peaks in serum based on 2D HSQC acquisitions, and the STOCSY algorithm.[32] The spectra were mean-normalized prior to quantification. Quantification was performed by integrating the region under each peak, giving the relative amounts of metabolites in each sample. If a metabolite had more than one identifiable peak, the mean value of the multiple peaks were calculated and used for further analysis. Signals from ethanol at 1.17 ppm were removed, resulting in 29 distinct peaks (27 metabolites and two lipid signals, see Table S1). The lipid signals arise from the methyl ($-CH_3$) groups at 0.85 ppm (lipid1) and methylene ($-CH_2-$) groups at 1.57 ppm (lipid2), mainly from triglycerides and esterified cholesterol within the lipoprotein particles.[33] A representative spectrum with annotated metabolite peaks is shown in Figure S3.

As evidenced by very high negative correlations (see Figure 1A) between the serum metabolites and lipid peaks, including the lipids in the analyses overshadowed changes in the low-molecular weight serum metabolites. We therefore removed the lipid peaks and normalized the metabolites a second time prior to statistical analyses.

Analysis and Preprocessing of Tissue Samples. A total of 270 tissue samples were analyzed by high resolution magic angle spinning (MAS) NMR. Details of NMR experiments, preprocessing, and quantification of the tissue samples have been described previously.[21] In brief, tissue samples (mean weight: 4.1 mg) were analyzed at 5 °C on a Bruker AVANCE DRX600 spectrometer equipped with a $^1H/^{13}C$ MAS rotor. A spin-echo one dimensional experiment with presaturation (cpmgpr1d, Bruker BioSpin, Germany) was recorded for all samples, with effective echo time of 77 ms, a spectral width of 20 ppm (−5 to 15 ppm), and 256 scans. Spectra were baseline corrected, peak aligned using the icoshift algorithm,[31] and normalized by probabilistic quotient normalization (PQN)[34] after removal of lipid residuals. Quantified metabolites were normalized by PQN.

Statistical Analysis

Multivariate Analyses. All variables were auto-scaled prior to multivariate analyses. Principal component analysis[35] (PCA) was performed on the quantified serum metabolites as a first step in the exploratory analysis.

Partial least squares discriminant analyses (PLS-DA) were employed to fit classification models for different clinical variables.[36] PLS-DA models were fitted and validated using 10-fold cross-validation, which was repeated 20 times. The optimal number of latent variables (LVs) was chosen to be the number of LVs corresponding to the first minima in the cross-validated classification error. Averaged sensitivities and specificities of the 20 iterations are reported. To verify the statistical significance of the models, permutation testing was employed, where the original class labels were shuffled among the individuals.[37] New models were fit to these permuted data sets and the classification error was calculated. The proportion of classifications equal to or better than the original classification was used to calculate the p-values. The permutations were repeated 1000 times and p-values ≤ 0.05 were considered significant. For the PLS-DA plots, the y-variance was condensed into the first LV through orthogonal projection to latent structures in cases where the optimal model had more than one LV. This orthogonalization does not improve the model accuracy, but rather the model interpretation, as the predictive from nonpredictive variation is separated.[36]

Metabolomics data is complex and many factors (such as age, disease state, and genetics) influence the metabolic profile of a biological sample, thus the variations between samples of different individuals are often higher than the variations within the samples of one individual. Variations, as a result of the treatment effect, can be overshadowed by the between-subject variations. The total effect is thus undetectable if the main focus is the average effect. Multilevel PLS-DA is an extension of PLS-DA and consists of two steps.[38] First, the variation between individuals is separated from the variation within the samples. Second, PLS-DA analysis is performed on the within-subject variation. This is an effective tool for longitudinal data, where there are two or more multivariate measurements per subject. As the multilevel PLS-DA models contain multiple measurements for each patient, 10% of the patients were left out during each iteration, which was repeated 20 times.

PCA, PLS-DA, and multilevel PLS-DA analyses were carried out in MATLAB R2017b using the PLS Toolbox 8.6.2.[39] The loading plots of the orthogonalized PLS-DA and multilevel PLS-DA analyses were colored according to the variable (here metabolite) importance in projection score (VIP score). The VIP score is a measure of how important each variable was for creating the discrimination model. It is calculated as a weighted sum of squares of the PLS loadings, where the weights are based on the amount of y-variance explained in each dimension.[40] A metabolite with a VIP score larger than or equal to 1 was considered to be important in the discrimination model.

Univariate Data Analysis. Because of non-normality of the serum metabolites, the nonparametric Wilcoxon-signed-rank test was used to test the significance of the changes in serum metabolite levels between time points.[41] p-values were adjusted using the Benjamini−Hochberg procedure and significance was considered for q-values ≤ 0.05.[42]

In this study, both serum and tumor samples from the same BC patients were analyzed, enabling us to investigate how much of the tissue-metabolic profile is reflected in the serum metabolome. To investigate this, Pearson-correlations between all quantified metabolites in the serum and tissue samples were calculated. p-values for significance were adjusted for multiple comparisons using the Benjamini−Hochberg procedure, and significance was considered for q-values ≤ 0.05. The calculations and graphical representations of the correlation were performed in the R software environment using the corrplot package.[43,44]

Statistical analyses of serum metabolites were performed on quantified metabolites. For tissue samples, multivariate

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

analyses were performed on the whole NMR spectra as in Euceda et al.,[21] whereas correlation analysis was performed using quantified metabolite levels.

RESULTS

Correlation Analysis of Serum and Tissue Metabolic Profiles

Availability of both tissue biopsies and serum samples from the same BC patients enabled to investigate how much of the tumor metabolism is reflected in the serum. The majority of the correlations between the serum metabolites were high (Figure 1A). The low-molecular weight serum metabolites had a highly negative correlation with the lipid peaks, whereas they were positively correlated with each other. There were fewer high correlations between tissue metabolites. However, tissue levels of taurine and glucose, and glutamate and lactate were highly correlated (ϱ = 0.903 and 0.714, respectively; q-values < 0.001). This figure also shows that correlations between serum and tissue metabolites, although some were significant, were low ($0.005 \leq |\varrho| \leq 0.269$; q-values ≤ 0.05). Serum lactate was not correlated with tissue lactate (ϱ = 0.061, q-value = 0.835). In addition, choline stands out from the other tissue metabolites, with low but significant correlations with the majority of the serum metabolites ($0.074 \leq |\varrho| \leq 0.269$). To emphasize correlations between the low-molecular weight metabolites, the analyses were repeated with the lipid peaks removed. Figure 1B shows correlation analyses of serum and tissue metabolic profiles after the removal of lipid peaks in serum data and a second normalization. The correlations in serum metabolites are then highly affected in both magnitude and direction.

Effect of NAC on Serum Metabolic Profiles

PCA analyses of serum metabolites did not show any clear trend or grouping of the patients with respect to the time point at which the samples were obtained (Figure S4). However, by employing multilevel PLS-DA and thus removing the between-subject variation in the data, significant changes in the serum metabolic profiles between each time point during the treatment were revealed. Table 2 summarizes the fit of the multilevel PLS-DA models on serum data without lipid peaks included. PLS-DA results for separating TP1 and TP2 with and without including the lipids are shown in Figure 2. First, when the lipid peaks are included in the multilevel analyses, it is clear that the amount of lipids in serum increase during treatment (Figure 2A). The same is evident throughout the treatment period as seen in Figure S5, which shows multilevel analysis comparing TP1 with TP4, when lipids are included. Removal of lipid peaks to emphasize changes within the metabolic profile did not have a significant influence on the prediction accuracy of the models. Further results are derived from the serum metabolic data without including the lipid peaks.

Table 2. Summary of Multilevel PLS-DA Applied on Serum Metabolites after the Removal of Lipid Peaks and a Second Normalization[a]

	no. of LV's	class accuracy (%)	sensitivity/specificity (%)	p-value
TP1 vs TP2	4	90	90/90	<0.001
TP2 vs TP3	2	77	77/77	<0.001
TP3 vs TP4	4	87	87/87	<0.001

[a]Sensitivities and specificities are averaged on 20 repetitions of 10-fold cross validation. The reported p-values are based on permutation testing, with 1000 random permutations of the original class labels. LV: latent variable.

Scores and loading plots of the multilevel PLS-DA models separating different time points are displayed in Figure S6, where the loadings are colored according to the VIP scores. The most important metabolites in discriminating between serum metabolic profiles at TP1 and TP2 are creatinine (↓), creatine (↓), isoleucine (↑), ornithine (↓), and histidine (↑) (Figure 2B), where the arrow shows the direction of the change with the treatment course. For discriminating TP2 from TP3, creatine (↑), valine (↑), dimethyl glutarate (↓), and pyruvate (↓) are of highest importance. Finally, for discriminating between serum metabolic profiles at TP4 and TP3, valine (↑), glycine (↓), dimethyl glutarate (↑), and methionine (↑) are the most important metabolites.

The median percentage change of each metabolite level between the different time points is displayed in Table 3, with corresponding q-values to assess statistical significance. Most significant changes occur between TP1 and TP2; however, the metabolic profiles change significantly throughout the treatment period. Only two metabolites exhibited significant changes across all sampling time points during the treatment course: dimethyl glutarate (↑↓↑) and acetate (↓↑↑).

Effect of Bevacizumab on Serum Metabolic Profiles

We further examined if the serum metabolites are affected by treatment with the drug bevacizumab in addition to chemotherapy. A significant discrimination model for separating patients receiving and not receiving bevacizumab was obtained at TP2, but not at later time points (accuracy = 64%; p-value = 0.014, Figure 3A and Table 4), even though the administration of bevacizumab was continued until TP3. The most important metabolites in the discrimination model for TP2 are higher levels of leucine, acetoacetate, and tri-hydroxybutyrate and lower levels of formate (VIP scores 1.76, 1.59, 1.56, and 1.47, respectively) for the group of patients treated with Bevacizumab compared to patients treated with chemotherapy only.

Serum Metabolic Differences between Responders and Nonresponders of Neoadjuvant Treatment

PLS-DA classification models were fitted to the serum metabolites for each time point separately to examine if there were metabolic differences between patients with a good or poor response to treatment. The model results are summarized in Table 4. Patients with a good response (RCB 0 or I) could be significantly discriminated from patients with a poor response (RCB II or III) at TP4 with an accuracy of 69% (p-value = 0.001, Figure 3B). The most important metabolites in the discrimination were citrate, phenylalanine, and histidine (VIP scores 2.25, 1.75, and 1.53, respectively), with higher levels of citrate and lower of the latter in RCB II or III compared to RCB 0 or I patients.

Predicting Survival from Serum and Tissue Metabolic Profiles

Discrimination models were fitted to assess if there is predictive power in the serum and tissue metabolites to predict long-term outcome. Results of the analyses show that the serum metabolites have no predictive power for 5 year

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

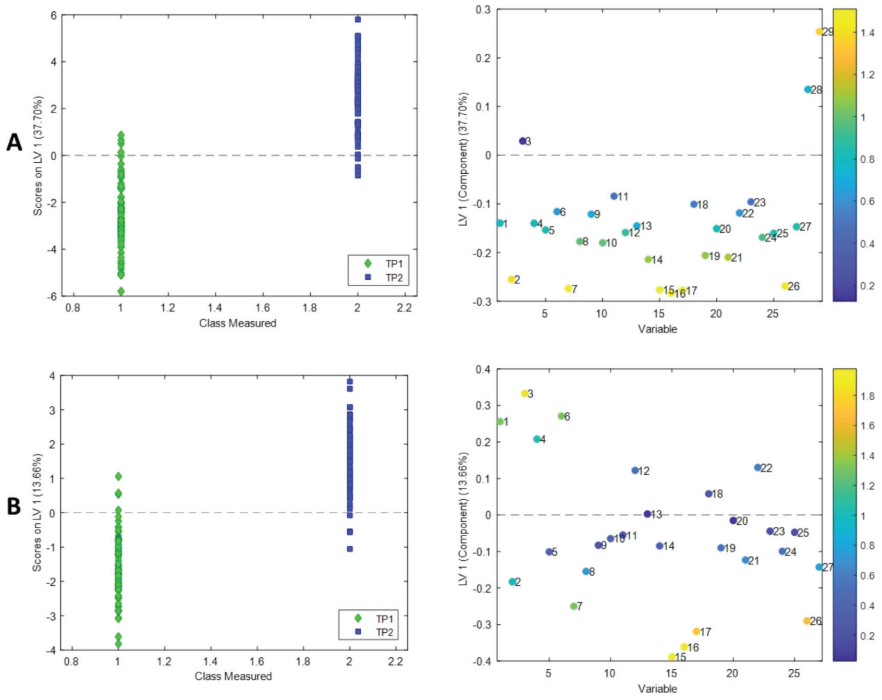


Figure 2. Scores and loadings plots from the multilevel PLS-DA analyses for discriminating between the serum metabolic profiles at TP2 from TP1. (A) Analysis including lipid peaks. (B) Analysis after excluding lipid peaks and a second normalization. Orthogonalized loadings colored according to VIP scores. LV: latent variable. 1: leucine; 2: valine; 3: isoleucine; 4: dimethyl glutarate; 5: tri-hydroxybutyrate; 6: alanine; 7: lysine; 8: acetate; 9: acetoacetate; 10: glutamate; 11: pyruvate; 12: glutamine; 13: citrate; 14: methionine; 15: creatine; 16: creatinine; 17: ornithine; 18: proline-betaine; 19: dimethyl sulfone; 20: glucose; 21: methanol; 22: glycine; 23: lactate; 24: tyrosine; 25: phenylalanine; 26: histidine; 27: formate; 28: lipid1; 29: lipid2.

survival (Table 4). Similar models were employed on the tissue metabolic profiles at the different time points and showed that tissue metabolic profiles at TP2 have a predictive potential for discriminating survivors from nonsurvivors, with a prediction accuracy of 72% (p-value = 0.005). Scores and loading plots for the corresponding PLS-DA model at TP2 are displayed in Figure 4.

DISCUSSION

In this study, we show that the NMR-based metabolic profile of serum from BC patients undergoing NAC changes significantly throughout treatment. Furthermore, we show that 5-year survival can be predicted from metabolic profiles in tissue, but not serum. Significant associations between serum metabolic profiles and response to treatment, in addition to changes in the serum metabolic profiles in patients receiving bevacizumab, were detected.

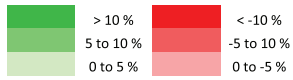
Several factors affect the serum metabolome, such as diet, age, body mass index (BMI), drug use, and diurnal variations.[45−47] The serum metabolome will contain metabolic signals from both the tumor itself and the host organism, both affected by treatment. Some studies have investigated the difference in the serum metabolic profiles of women with BC compared to healthy controls, showing that presence of the tumor has an evident effect on the serum metabolome,[48−52] whereas only few have looked into treatment-induced changes.[53,54] A previous study revealed baseline levels of formate and acetate as potential predictive biomarkers of treatment response in metastatic BC patients, linking these changes to the accelerated proliferation of aggressive BC cells.[55] In this study, we describe significant serum metabolic changes in response to treatment at all time points, showing that BC treatment has an effect on the overall metabolism. Particularly lipid levels in serum increased throughout the treatment course (Figures 2A and S5). These results are in agreement with a previous study where we describe serum metabolic changes from adjuvant BC treatment, where unfavorable changes in the lipoprotein profiles were observed during treatment.[56] Altered lipid metabolism may predispose for weight gain, increased risk of CVD, and a worse overall health and quality of life. Increased lipid levels in serum post treatment have additionally been observed and associated with an increased risk of disease recurrence.[57]

The most evident effect of BC treatment on the serum metabolome occurred during the first weeks of treatment (TP1 to 2) and from surgery to 6 weeks follow-up (TP3 to 4). When comparing samples acquired before treatment onset and 12 weeks into treatment, 11 of the 27 metabolites changed significantly, mainly to decreased levels. Comparing the first weeks of treatment revealed decreased histidine, creatine, creatinine, and ornithine levels and increased isoleucine, to be of highest importance (Figure 2B). Serum levels of isoleucine were previously shown to be upregulated in metastatic

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 3. Median Percentage Changes in the Serum Metabolite Levels During Treatment[a]

	Metabolite name	TP1 to TP2 (%)	q-value	TP2 to TP3 (%)	q-value	TP3 to TP4 (%)	q-value
1	Leucine	4.88	**0.001**	-2.52	0.613	4.64	**0.029**
2	Valine	-1.76	0.082	5.55	**0.004**	7.19	**<0.001**
3	Isoleucine	12.46	**<0.001**	-0.95	0.706	-0.73	0.512
4	Dimethylglutarate	3.28	**0.036**	-8.27	**0.001**	9.36	**0.001**
5	Alanine	-3.26	0.294	2.06	0.386	2.48	0.271
6	Lysine	5.06	**0.001**	-4.24	**0.010**	1.95	0.319
7	Acetate	-3.43	**<0.001**	1.81	**0.010**	3.57	**0.002**
8	Acetoacetate	-2.16	0.156	5.82	**0.030**	-0.20	0.589
9	3-Hydroxybutyrate	0.84	0.562	3.65	0.212	5.82	0.178
10	Glutamate	1.63	0.974	-0.06	0.613	1.22	0.280
11	Pyruvate	-3.83	0.808	-6.28	**0.010**	-0.98	0.722
12	Glutamine	1.81	0.244	0.16	0.955	-2.37	0.062
13	Citrate	-2.38	0.974	-1.82	0.953	-7.37	**0.039**
14	Methionine	-1.83	0.294	3.40	0.187	4.80	**0.002**
15	Creatine	-13.30	**<0.001**	9.01	**<0.001**	5.34	0.089
16	Creatinine	-7.80	**<0.001**	2.30	0.185	4.19	**0.040**
17	Ornithine	-6.33	**<0.001**	1.38	0.355	4.17	**0.002**
18	Proline-betaine	-3.10	0.974	4.22	0.585	1.58	0.604
19	Dimethyl-sulfone	-2.14	0.294	4.66	0.207	3.31	0.163
20	Methanol	-2.74	0.294	-0.54	0.706	-1.48	0.452
21	Glucose	-3.83	0.095	1.61	0.491	-0.51	0.798
22	Glycine	3.08	0.156	0.81	0.603	-6.08	**0.001**
23	Lactate	-1.19	0.974	-3.01	0.813	-14.40	**0.002**
24	Tyrosine	-3.83	0.303	-2.38	0.603	4.43	0.452
25	Phenylalanine	-1.50	0.887	0.83	0.799	8.24	**0.025**
26	Histidine	-9.81	**<0.001**	-0.33	0.706	9.84	**0.010**
27	Formate	-10.60	**0.036**	-6.57	0.603	-5.17	0.936



[a]Only patients with samples available at each of the two time points were included when calculating the percentage changes. *q*-values show *p*-values obtained from Wilcoxon-signed-rank test, adjusted for multiple comparisons. Significant changes are marked in bold.

compared to early BC[50,58] and higher isoleucine has also been associated with pCR.[54] Thus, the predictive value of changes in isoleucine levels should be further investigated. Creatinine is a breakdown product of phosphocreatine in muscles and is usually produced at a constant rate by the body; it is thus plausible that the observed increase is induced by treatment. Creatine, creatinine, and ornithine are amino acids closely linked together through the arginine and proline metabolism pathway, through which glutamate is synthetized from arginine and proline.

Twelve weeks into treatment, increased levels of valine and creatine, and decreased levels of dimethyl glutarate, lysine, and pyruvate were observed, compared to 6 weeks into treatment. Similarly, increased levels of valine and creatine during BC treatment, compared to baseline levels, were observed in a longitudinal study with HER-2 positive BC patients in the trastuzumab and everolimus treatment arm.[53] Increased valine levels have also been shown to be important in discriminating BC patients from healthy controls (post-treatment).[49] Pyruvate is a key intermediate in several metabolic pathways throughout the cell, including gluconeogenesis and the Krebs cycle; lower pyruvate levels therefore possibly reflect an increased energy metabolism due to the treatment.

Patients switched from FEC treatment to taxane-based therapy 12 weeks into the treatment (TP2), followed by no further treatment, other than surgery, between the last two sampling points (TP3 to 4). It appears that the serum metabolism tends to return to its pretreatment state in this period; valine, acetate, creatine, ornithine, and histidine, all experienced a decrease at the beginning of treatment, followed by an increase after surgery. Glycine levels remained relatively constant throughout treatment, but decreased significantly after treatment. Low levels of circulating glycine have previously been associated with metabolic syndrome; this decrease may thus indicate a negative side-effect of the treatment.[59]

Five year survival was predicted with an accuracy of 72% at TP2. Nonsurvivors had higher lactate and glycine levels compared to survivors at TP2, which is in accordance with previous studies in similar patient cohorts.[22,23] Elevated lactate and glycine levels have also been associated with lower survival rates in ER-positive BC patients receiving surgery as primary treatment.[60] Furthermore, lactate has been associated with poor prognosis in other cancers and is a generally accepted marker for tumor aggressiveness, as high levels of lactate have been correlated to low survival rates, high incidence of distant metastasis, and recurrence.[61,62] Increased lactate production and rapid glucose consumption are known characteristics of the Warburg effect, which can be observed in most cancer cells.[63] Glycine has been linked to cancer-induced metabolic reprogramming, and glycine consumption and expression of the mitochondrial glycine biosynthetic pathway have been

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

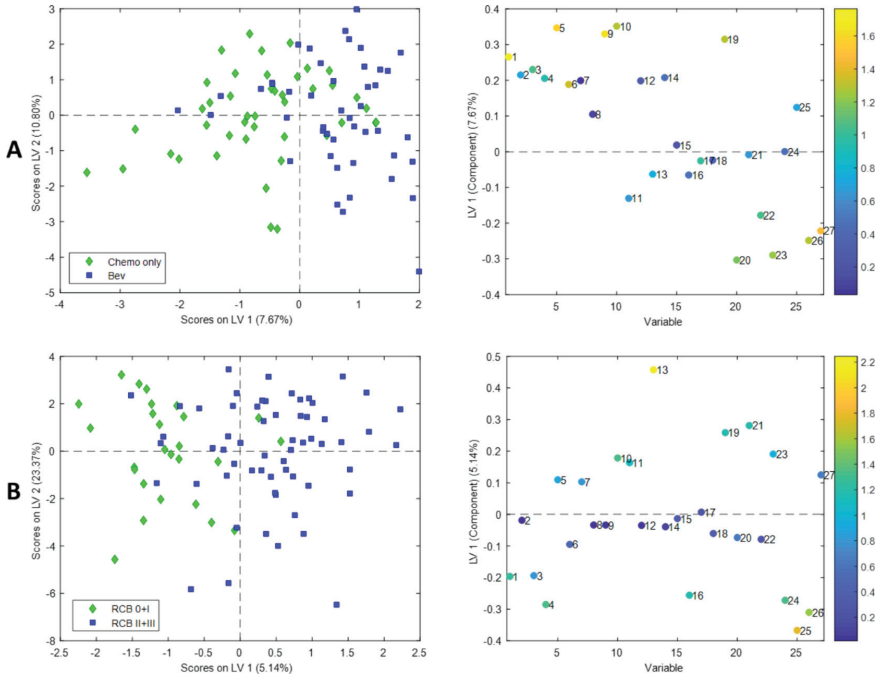


Figure 3. Scores and loading plots of the PLS-DA models for serum metabolic profiles. (A) Bevacizumab-treated vs Chemotherapy only at TP2. (B) RCB 0 or I vs RCB II or III at TP4. Orthogonalized loadings colored according to the VIP scores. 1: leucine; 2: valine; 3: isoleucine; 4: dimethyl glutarate; 5: tri-hydroxybutyrate; 6: alanine; 7: lysine; 8: acetate; 9: acetoacetate; 10: glutamate; 11: pyruvate; 12: glutamine; 13: citrate; 14: methionine; 15: creatine; 16: creatinine; 17: ornithine; 18: proline-betaine; 19: dimethyl sulfone; 20: glucose; 21: methanol; 22: glycine; 23: lactate; 24: tyrosine; 25: phenylalanine; 26: histidine; 27: formate.

Table 4. Summary of PLS-DA Classification Models Fitted to the Serum and Tissue Metabolic Profiles at Different Time Points[a]

	discriminated classes	time point	*n*	class accuracy (%)	sensitivity/specificity (%)	permutation *p*-value
	Bev-treat./chemo treat. only	**TP2**	**89**	**64**	**58/70**	**0.0140**
		TP3	93	59	60/57	0.0870
		TP4	86	57	67/47	0.0960
	RCB class 0 + I/RCB class II + III	TP1	89	36	27/44	0.9580
		TP2	89	48	33/63	0.6500
serum		TP3	93	58	58/57	0.1700
		TP4	**86**	**69**	**65/73**	**0.0010**
	5 year survival	TP1	89	37	5/70	0.7700
		TP2	89	64	48/81	0.2570
		TP3	93	61	43/79	0.1780
		TP4	86	48	23/73	0.5620
	5 year survival	TP1	105	58	30/86	0.2190
tissue		**TP2**	**78**	**72**	**55/90**	**0.0050**
		TP3	87	57	26/88	0.2210

[a]Sensitivities and specificities are averaged on 20 repetitions of 10-fold cross validation. The reported *p*-values are based on permutation testing, with 1000 random permutations of the original class labels. Significant classification models are marked in bold. *n*: number of samples included in model.

identified to be strongly correlated with the rates of proliferation across cancer cells.[64]

The RCB response measure represents an independent prognostic factor of distant relapse-free survival (DRFS) in multivariate Cox regression analyses of cancer patients.[29] RCB 0 and I are associated with good prognosis, whereas RCB II and III are associated with poor prognosis. Based on serum metabolic profiles, we could not predict patient response to treatment before or during treatment. However, patients with a good prognosis could be discriminated from patients with a poor prognosis 6 weeks after treatment completion (TP4) with an accuracy of 69% (*p*-value = 0.001). RCB II or III patients

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

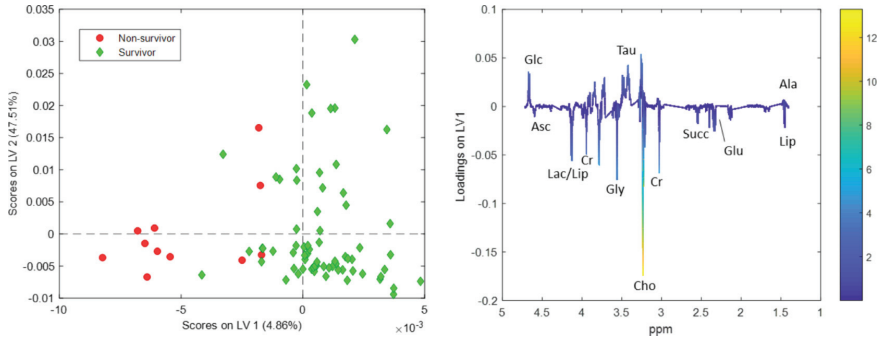


Figure 4. Scores and loadings plots for predicting survival from tissue metabolic profiles at TP2. Orthogonalized loadings colored according to their VIP score. LV: latent variable; Glc: glucose; Asc: ascorbate; Lac: lactate; Lip: lipid; Gly: glycine; Tau: taurine; Cho: cholines (glycerophosphocholine, phosphoscholine and choline); Cr: creatine; Succ: succinate; Glu: glutamate; Ala: alanine.

had higher serum levels of citrate and lower levels of phenylalanine and histidine. Significantly higher serum levels of citrate and lower levels of phenylalanine and histidine have been observed in metabolic profiles of metastatic compared to early BC implying that they play a role in the formation of metastasis.[58]

Patients receiving bevacizumab were significantly discriminated from those treated only with chemotherapy 12 weeks into treatment (TP2). Discriminating metabolites were lower levels of leucine, acetoacetate, and tri-hydroxybutyrate and higher levels of formate in patients receiving bevacizumab. A previous study has linked the rate of β-hydroxybutyrate and acetoacetate in blood to mitochondrial activity.[65] The effect of bevacizumab on the serum metabolome of BC patients has, to our knowledge, not been described previously. A study on metastatic renal cell carcinoma identified changes in glucose, *N*-acetyl glycoproteins, lipids, and lipoproteins as an effect of treatment, relating these to known side effects of the drugs bevacizumab and temsirolimus.[66] Our previous study on tissue metabolites from the same patient cohort[21] showed weak associations between bevacizumab and tissue metabolic profiles.

An advantage of this study cohort is that both tissue biopsies and serum samples were available from the same patients, allowing for a comparison of metabolic information. Importantly, the metabolic information from these two types of biological samples is different, with some significant, but low correlations (Figure 1B). This explains why we could predict patient survival from tissue, but not serum metabolites. Although tumors are often characterized by high lactate production, there was no correlation between tissue and serum lactate levels. A study linking tumor information in early BC patients with plasma metabolites showed an inverse correlation between plasma lactate levels and the tumor size.[67] In general, despite possible leakage of metabolites from the cancer tissue into the bloodstream of the host organism, the overall serum metabolism has larger variation that may mask these tumor-derived metabolites; thus, metabolites which have been associated with treatment response when analyzing tumor tissue are not necessarily relevant in the context of serum metabolomics.

Multivariate analysis, taking advantage of the multilevel structure of the data focusing on the within-subject variations resulted in models with high classification accuracy for characterizing the serum metabolic changes from treatment. Our study also pinpoints that awareness regarding the effect of normalization procedures is necessary, given the different results observed with the exclusion of lipid signals prior to a second normalization of the serum metabolic profiles. Although different normalization strategies did not affect the quality of the multivariate models per se, making their robustness evident, variables important for the classifications were affected, making comparisons of potential biomarkers across studies challenging.

CONCLUSIONS

By metabolic profiling of serum sampled before, during, and after neoadjuvant treatment in BC patients, we have revealed significant metabolic changes in serum as a response to treatment. This gives an insight into how the body is affected by treatment, and provides a possible tool for understanding negative side-effects of treatment. Serum metabolomics therefore has a potential for longitudinal patient-monitoring during and after BC treatment.

Tissue metabolic profiles during treatment were significantly correlated to five-year survival, whereas no such information was apparent in the serum metabolic profiles. Importantly, we demonstrate low correlations between serum and tissue metabolites, emphasizing the complementary nature of the metabolic information in these biological matrices.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.9b00316.

Flow diagram showing the experimental setup of the study; sample availability at each sampling time point, including survival data; details on quantification of serum metabolites; representative spectrum with annotated metabolite peaks; PCA scores plot of the serum metabolites, colored according to the time point at which they were obtained; scores and loading plot of multilevel PLS-DA analyses on serum metabolites with lipid peaks, comparing TP1 with TP4; and scores and loading plots of multilevel PLS-DA analyses on serum metabolites (PDF)

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

AUTHOR INFORMATION

Corresponding Authors

*E-mail: julia.b.debik@ntnu.no (J.D.).
*E-mail: guro.giskeødegård@ntnu.no (G.F.G.).

ORCID

Julia Debik: 0000-0002-2765-8043

Funding

The NeoAva study was co-sponsored by Roche Norway and Sanofi-Aventis Norway. Oslo University Hospital is the main sponsor for the NeoAva study.

Notes

The authors declare the following competing financial interest(s): The NeoAva study was co-sponsored by Roche Norway and Sanofi-Aventis Norway. Oslo University Hospital is the main sponsor for the NeoAva study.

ACKNOWLEDGMENTS

The NMR analysis was performed at the MR Core Facility, Norwegian University of Science and Technology (NTNU), which is funded by the Faculty of Medicine and Health Sciences at NTNU and the Central Norway Regional Health Authority. Part of this work was supported by the Norwegian Cancer Society (grant 163243).

ABBREVIATIONS

BC, breast cancer; BMI, body mass index; CPMG, Carr−Purcell−Meiboom−Gill; CVD, cardio vascular disease; DRFS, distant relapse-free survival; ER, estrogen receptor; HER, human epidermal growth factor receptor; LABC, locally advanced BC; LV, latent variable; MR, magnetic resonance; NAC, neoadjuvant chemotherapy; NMR, nuclear magnetic resonance; NOESY, nuclear overhauser effect spectroscopy; PC, principal component; PCA, principal component analysis; pCR, pathologic complete response; PgR, progesterone receptor; PLS-DA, partial least squares discriminant analysis; RCB, residual cancer burden; TP1, TP2, TP3, TP4, time points for sampling, before treatment, 12 weeks into treatment, 25 weeks into treatment, and 6 weeks after treatment, respectively; VIP, variable importance in projection

REFERENCES

(1) Cancer Registry of Norway. *Cancer in Norway 2017—Cancer Incidence, Mortality, Survival and Prevalence in Norway*, 2018.

(2) Sørlie, T.; Perou, C. M.; Tibshirani, R.; Turid Aas, S. G.; Johansen, H.; Hastie, T.; Eisen, M. B.; Rijn, M. v. d.; Jefferey, S. S.; Thorsen, T.; Quist, H.; Matese, J. C.; Brown, P. O.; Botstein, D.; Lønning, P. E.; Børresen-Dale, A.-L. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *98*, 10869.

(3) Sitter, B.; Bathen, T. F.; Singstad, T. E.; Fjøsne, H. E.; Lundgren, S.; Halgunset, J.; Gribbestad, I. S. Quantification of metabolites in breast cancer patients with different clinical prognosis using HR MAS MR spectroscopy. *NMR Biomed.* **2010**, *23*, 424.

(4) Haukaas, T. H.; Euceda, L. R.; Giskeodegard, G. F.; Lamichhane, S.; Krohn, M.; Jernstrom, S.; Aure, M. R.; Lingjaerde, O. C.; Schlichting, E.; Garred, O.; Due, E. U.; Mills, G. B.; Sahlberg, K. K.; Borresen-Dale, A. L.; Bathen, T. F.; Oslo Breast Cancer Consortium. Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. *Cancer Metab* **2016**, *4*, 12.

(5) Aure, M. R.; Vitelli, V.; Jernstrom, S.; Kumar, S.; Krohn, M.; Due, E. U.; Haukaas, T. H.; Leivonen, S. K.; Vollan, H. K.; Luders, T.; Rodland, E.; Vaske, C. J.; Zhao, W.; Moller, E. K.; Nord, S.; Giskeodegard, G. F.; Bathen, T. F.; Caldas, C.; Tramm, T.; Alsner, J.; Overgaard, J.; Geisler, J.; Bukholm, I. R.; Naume, B.; Schlichting, E.; Sauer, T.; Mills, G. B.; Karesen, R.; Maelandsmo, G. M.; Lingjaerde, O. C.; Frigessi, A.; Kristensen, V. N.; Borresen-Dale, A. L.; Sahlberg, K. K.; Osbreac. Integrative clustering reveals a novel split in the luminal A subtype of breast cancer with impact on outcome. *Breast Cancer Res.* **2017**, *19*, 44.

(6) Zhao, X.; Rodland, E. A.; Tibshirani, R.; Plevritis, S. Molecular subtyping for clinically defined breast cancer subgroups. *Breast Cancer Res.* **2015**, *17*, 29.

(7) *The Norwegian Cancer Registry Data and Statistics*; Institute of Population-based Cancer Research: Oslo, 2019.

(8) Makhoul, I.; Kiwan, E. Neoadjuvant systemic treatment of breast cancer. *J. Surg. Oncol.* **2011**, *103*, 348−357.

(9) Miller, E.; Lee, H. J.; Lulla, A.; Hernandez, L.; Gokare, P.; Lim, B. Current treatment of early breast cancer: adjuvant and neoadjuvant therapy. *F1000Research* **2014**, *3*, 198.

(10) Sherwood, L. M.; Parris, E. E.; Folkman, J. Tumor angiogenesis: therapeutic implications. *N. Engl. J. Med.* **1971**, *285*, 1182−1186.

(11) Hanahan, D.; Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **2011**, *144*, 646−674.

(12) Ranieri, G.; Patruno, R.; Ruggieri, E.; Montemurro, S.; Valerio, P.; Ribatti, D. Vascular endothelial growth factor (VEGF) as a target of bevacizumab in cancer: from the biology to the clinic. *Curr. Med. Chem.* **2006**, *13*, 1845−1857.

(13) Monneur, A.; Goncalves, A.; Gilabert, M.; Finetti, P.; Tarpin, C.; Zemmour, C.; Extra, J. M.; Tallet, A.; Lambaudie, E.; Jacquemier, J.; Houvenaeghel, G.; Boher, J. M.; Viens, P.; Bertucci, F. Similar response profile to neoadjuvant chemotherapy, but different survival, in inflammatory versus locally advanced breast cancers. *Oncotarget* **2017**, *8*, 66019−66032.

(14) Wang, M.; Hou, L.; Chen, M.; Zhou, Y.; Liang, Y.; Wang, S.; Jiang, J.; Zhang, Y. Neoadjuvant Chemotherapy Creates Surgery Opportunities For Inoperable Locally Advanced Breast Cancer. *Sci. Rep.* **2017**, *7*, 44673.

(15) Berger, A. M.; Mooney, K.; Alvarez-Perez, A.; Breitbart, W. S.; Carpenter, K. M.; Cella, D.; Cleeland, C.; Dotan, E.; Eisenberger, M. A.; Escalante, C. P.; Jacobsen, P. B.; Jankowski, C.; LeBlanc, T.; Ligibel, J. A.; Loggers, E. T.; Mandrell, B.; Murphy, B. A.; Palesh, O.; Pirl, W. F.; Plaxe, S. C.; Riba, M. B.; Rugo, H. S.; Salvador, C.; Wagner, L. I.; Wagner-Johnston, N. D.; Zachariah, F. J.; Bergman, M. A.; Smith, C. Cancer-Related Fatigue, Version 2.2015. *J. Natl. Compr. Cancer Network* **2015**, *13*, 1012.

(16) Bower, J. E.; Ganz, P. A.; Desmond, K. A.; Rowland, J. H.; Meyerowitz, B. E.; Belin, T. R. Fatigue in breast cancer survivors: occurrence, correlates, and impact on quality of life. *J. Clin. Oncol.* **2000**, *18*, 743.

(17) Bower, J. E.; Wiley, J.; Petersen, L.; Irwin, M. R.; Cole, S. W.; Ganz, P. A. Fatigue after breast cancer treatment: Biobehavioral predictors of fatigue trajectories. *Health Psychol* **2018**, *37*, 1025−1034.

(18) Oh, P.-J.; Cho, J.-R. Changes in Fatigue, Psychological Distress, and Quality of Life After Chemotherapy in Women with Breast Cancer: A Prospective Study. *Cancer Nursing* **2018**, 1.

(19) Aleman, B. M. P.; Moser, E. C.; Nuver, J.; Suter, T. M.; Maraldo, M. V.; Specht, L.; Vrieling, C.; Darby, S. C. Cardiovascular disease after cancer therapy. *Eur. J. Cancer Suppl.* **2014**, *12*, 18−28.

(20) Markley, J. L.; Brüschweiler, R.; Edison, A. S.; Eghbalnia, H. R.; Powers, R.; Raftery, D.; Wishart, D. S. The future of NMR-based metabolomics. *Curr. Opin. Biotechnol.* **2017**, *43*, 34−40.

(21) Euceda, L. R.; Haukaas, T. H.; Giskeødegård, G. F.; Vettukattil, R.; Engel, J.; Silwal-Pandit, L.; Lundgren, S.; Borgen, E.; Garred, Ø.; Postma, G.; Buydens, L. M. C.; Børresen-Dale, A.-L.; Engebraaten, O.; Bathen, T. F. Evaluation of metabolomic changes during neoadjuvant chemotherapy combined with bevacizumab in breast cancer using MR spectroscopy. *Metabolomics* **2017**, *13*, 37.

(22) Cao, M. D.; Giskeodegard, G. F.; Bathen, T. F.; Sitter, B.; Bofin, A.; Lonning, P. E.; Lundgren, S.; Gribbestad, I. S. Prognostic

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

value of metabolic response in breast cancer patients receiving neoadjuvant chemotherapy. *BMC Cancer* **2012**, *12*, 39.
(23) Cao, M. D.; Sitter, B.; Bathen, T. F.; Bofin, A.; Lønning, P. E.; Lundgren, S.; Gribbestad, I. S. Predicting long-term survival and treatment response in breast cancer patients receiving neoadjuvant chemotherapy by MR metabolic profiling. *NMR Biomed.* **2012**, *25*, 369−378.
(24) Giskeødegård, G. F. Identification and Characterization of Prognostic Factors in Breast Cancer Using MR metabolomics. Doctoral Thesis; Norwegian University of Science and Technology: Trondheim, 2011.
(25) Giskeødegård, G. F.; Madssen, T. S.; Euceda, L. R.; Tessem, M.-B.; Moestue, S. A.; Bathen, T. F. NMR-based metabolomics of biofluids in cancer. *NMR Biomed.* **2018**, No. e3927, . (Special Issue Review Article) .
(26) Jabeen, S.; Zucknick, M.; Nome, M.; Dannenfelser, R.; Fleischer, T.; Kumar, S.; Lüders, T.; von der Lippe Gythfeldt, H.; Troyanskaya, O.; Kyte, J. A.; Børresen-Dale, A.-L.; Naume, B.; Tekpli, X.; Engebraaten, O.; Kristensen, V. Serum cytokine levels in breast cancer patients during neoadjuvant treatment with bevacizumab. *Oncoimmunology* **2018**, *7*, No. e1457598.
(27) Höglander, E. K.; Nord, S.; Wedge, D. C.; Lingjærde, O. C.; Silwal-Pandit, L.; Gythfeldt, H. V.; Vollan, H. K. M.; Fleischer, T.; Krohn, M.; Schlitchting, E.; Borgen, E.; Garred, Ø.; Holmen, M. M.; Wist, E.; Naume, B.; Van Loo, P.; Borresen-Dale, A. L.; Engebraaten, O.; Kristensen, V. Time series analysis of neoadjuvant chemotherapy and bevacizumab-treated breast carcinomas reveals a systemic shift in genomic aberrations. *Genome Med.* **2018**, *10*, 92.
(28) Silwal-Pandit, L.; Nord, S.; von der Lippe Gythfeldt, H.; Møller, E. K.; Fleischer, T.; Rødland, E.; Krohn, M.; Borgen, E.; Garred, Ø.; Olsen, T.; Vu, P.; Skjerven, H.; Fangberget, A.; Holmen, M. M.; Schlitchting, E.; Wille, E.; Nordberg Stokke, M.; Moen Vollan, H. K.; Kristensen, V.; Langerød, A.; Lundgren, S.; Wist, E.; Naume, B.; Lingjærde, O. C.; Børresen-Dale, A.-L.; Engebraaten, O. The Longitudinal Transcriptional Response to Neoadjuvant Chemotherapy with and without Bevacizumab in Breast Cancer. *Clin. Cancer Res.* **2017**, *23*, 4662−4670.
(29) Symmans, W. F.; Peintinger, F.; Hatzis, C.; Rajan, R.; Kuerer, H.; Valero, V.; Assad, L.; Poniecka, A.; Hennessy, B.; Green, M.; Buzdar, A. U.; Singletary, S. E.; Hortobagyi, G. N.; Pusztai, L. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J. Clin. Oncol.* **2007**, *25*, 4414−4422.
(30) *MATLAB, R2017b*; The MathWorks Inc.: Natick, Massachusetts, 2017.
(31) Tomasi, G.; Savorani, F.; Engelsen, S. B. icoshift: An effective tool for the alignment of chromatographic data. *J. Chromatogr. A* **2011**, *1218*, 7832−7840.
(32) Cloarec, O.; Dumas, M.-E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J. C.; Holmes, E.; Nicholson, J. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Anal. Chem.* **2005**, *77*, 1282−1289.
(33) Aru, V.; Lam, C.; Khakimov, B.; Hoefsloot, H. C. J.; Zwanenburg, G.; Lind, M. V.; Schäfer, H.; van Duynhoven, J.; Jacobs, D. M.; Smilde, A. K.; Engelsen, S. B. Quantification of lipoprotein profiles by nuclear magnetic resonance spectroscopy and multivariate data analysis. *TrAC, Trends Anal. Chem.* **2017**, *94*, 210−219.
(34) Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal. Chem.* **2006**, *78*, 4281−4290.
(35) Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37−52.
(36) Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109−130.
(37) Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; van Velzen, E. J. J.; van Duijnhoven, J. P. M.; van Dorsten, F. A. Assessment of PLSDA cross validation. *Metabolomics* **2008**, *4*, 81−89.
(38) Westerhuis, J. A.; van Velzen, E. J. J.; Hoefsloot, H. C. J.; Smilde, A. K. Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. *Metabolomics* **2010**, *6*, 119−128.
(39) Eigenvector Research, Incorporated. *PLS Toolbox, 8.6.2*, Manson, WA USA 98831, 2018.
(40) Mehmood, T.; Liland, K. H.; Snipen, L.; Sæbø, S. A review of variable selection methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* **2012**, *118*, 62−69.
(41) Rosner, B. *Fundamentals of Biostatistics*, 8th ed.; Cengage Learning: Boston, 2015.
(42) Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289−300.
(43) Wei, T.; Simko, V. *R Package "corrplot": Visualization of a Correlation Matrix*, 2017.
(44) R Development Core Team. *R: A Language and Environment for Statistical Computing*, Vienna, Austria, 2009.
(45) Yu, Z.; Zhai, G.; Singmann, P.; He, Y.; Xu, T.; Prehn, C.; Römisch-Margl, W.; Lattka, E.; Gieger, C.; Soranzo, N.; Heinrich, J.; Standl, M.; Thiering, E.; Mittelstrass, K.; Wichmann, H.-E.; Peters, A.; Suhre, K.; Li, Y.; Adamski, J.; Spector, T. D.; Illig, T.; Wang-Sattler, R. Human serum metabolic profiles are age dependent. *Aging Cell* **2012**, *11*, 960−967.
(46) Esko, T.; Hirschhorn, J. N.; Feldman, H. A.; Hsu, Y.-H. H.; Deik, A. A.; Clish, C. B.; Ebbeling, C. B.; Ludwig, D. S. Metabolomic profiles as reliable biomarkers of dietary composition. *Am. J. Clin. Nutr.* **2017**, *105*, 547−554.
(47) Sato, S.; Parr, E. B.; Devlin, B. L.; Hawley, J. A.; Sassone-Corsi, P. Human metabolomics reveal daily variations under nutritional challenges specific to serum and skeletal muscle. *Mol. Metab.* **2018**, *16*, 1−11.
(48) Jové, M.; Collado, R.; Quiles, J. L.; Ramírez-Tortosa, M. C.; Sol, J.; Ruiz-Sanjuan, M.; Fernandez, M.; de la Torre Cabrera, C.; Ramírez-Tortosa, C.; Granados-Principal, S.; Sánchez-Rovira, P.; Pamplona, R. A plasma metabolomic signature discloses human breast cancer. *Oncotarget* **2017**, *8*, 19522−19533.
(49) Cala, M. P.; Aldana, J.; Medina, J.; Sánchez, J.; Guio, J.; Wist, J.; Meesters, R. J. W. Multiplatform plasma metabolic and lipid fingerprinting of breast cancer: A pilot control-case study in Colombian Hispanic women. *PLoS One* **2018**, *13*, No. e0190958.
(50) Xie, G.; Zhou, B.; Zhao, A.; Qiu, Y.; Zhao, X.; Garmire, L.; Shvetsov, Y. B.; Yu, H.; Yen, Y.; Jia, W. Lowered circulating aspartate is a metabolic feature of human breast cancer. *Oncotarget* **2015**, *6*, 33369−33381.
(51) Qiu, Y.; Zhou, B.; Su, M.; Baxter, S.; Zheng, X.; Zhao, X.; Yen, Y.; Jia, W. Mass spectrometry-based quantitative metabolomics revealed a distinct lipid profile in breast cancer patients. *Int. J. Mol. Sci.* **2013**, *14*, 8047−8061.
(52) Gu, H.; Pan, Z.; Xi, B.; Asiago, V.; Musselman, B.; Raftery, D. Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: application to the detection of breast cancer. *Anal. Chim. Acta* **2011**, *686*, 57−63.
(53) Jobard, E.; Tredan, O.; Bachelot, T.; Vigneron, A. M.; Ait-Oukhatar, C. M.; Arnedos, M.; Rios, M.; Bonneterre, J.; Dieras, V.; Jimenez, M.; Merlin, J. L.; Campone, M.; Elena-Herrmann, B. Longitudinal serum metabolomics evaluation of trastuzumab and everolimus combination as pre-operative treatment for HER-2 positive breast cancer patients. *Oncotarget* **2017**, *8*, 83570−83584.
(54) Wei, S.; Liu, L.; Zhang, J.; Bowers, J.; Gowda, G. A. N.; Seeger, H.; Fehm, T.; Neubauer, H. J.; Vogel, U.; Clare, S. E.; Raftery, D. Metabolomics approach for predicting response to neoadjuvant chemotherapy for breast cancer. *Mol. Oncol.* **2013**, *7*, 297−307.
(55) Jiang, L.; Lee, S. C.; Ng, T. C. Pharmacometabonomics Analysis Reveals Serum Formate and Acetate Potentially Associated with Varying Response to Gemcitabine-Carboplatin Chemotherapy in

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

Metastatic Breast Cancer Patients. *J. Proteome Res.* **2018**, *17*, 1248−1257.
(56) Madssen, T. S.; Thune, I.; Flote, V. G.; Lundgren, S.; Bertheussen, G. F.; Frydenberg, H.; Wist, E.; Schlichting, E.; Schäfer, H.; Fjøsne, H. E.; Vettukattil, R.; Lømo, J.; Bathen, T. F.; Giskeødegård, G. F. Metabolite and lipoprotein responses and prediction of weight gain during breast cancer treatment. *Br. J. Cancer* **2018**, *119*, 1144−1154.
(57) Tenori, L.; Oakman, C.; Morris, P. G.; Gralka, E.; Turner, N.; Cappadona, S.; Fornier, M.; Hudis, C.; Norton, L.; Luchinat, C.; Di Leo, A. Serum metabolomic profiles evaluated after surgery may identify patients with oestrogen receptor negative early breast cancer at increased risk of disease recurrence. Results from a retrospective study. *Mol. Oncol.* **2015**, *9*, 128−139.
(58) Hart, C. D.; Vignoli, A.; Tenori, L.; Uy, G. L.; Van To, T.; Adebamowo, C.; Hossain, S. M.; Biganzoli, L.; Risi, E.; Love, R. R.; Luchinat, C.; Di Leo, A. Serum Metabolomic Profiles Identify ER-Positive Early Breast Cancer Patients at Increased Risk of Disease Recurrence in a Multicenter Population. *Clin. Cancer Res.* **2017**, *23*, 1422−1431.
(59) Li, X.; Sun, L.; Zhang, W.; Li, H.; Wang, S.; Mu, H.; Zhou, Q.; Zhang, Y.; Tang, Y.; Wang, Y.; Chen, W.; Yang, R.; Dong, J. Association of serum glycine levels with metabolic syndrome in an elderly Chinese population. *Nutr. Metab.* **2018**, *15*, 89.
(60) Giskeødegård, G. F.; Lundgren, S.; Sitter, B.; Fjøsne, H. E.; Postma, G.; Buydens, L. M.; Gribbestad, I. S.; Bathen, T. F. Lactate and glycine-potential MR biomarkers of prognosis in estrogen receptor-positive breast cancers. *NMR Biomed.* **2012**, *25*, 1271−1279.
(61) Walenta, S.; Mueller-Klieser, W. F. Lactate: mirror and motor of tumor malignancy. *Semin Radiat Oncol* **2004**, *14*, 267−274.
(62) Walenta, S.; Wetterling, M.; Lehrke, M.; Schwickert, G.; Sundfør, K.; Rofstad, E. K.; Mueller-Klieser, W. High lactate levels predict likelihood of metastases, tumor recurrence, and restricted patient survival in human cervical cancers. *Cancer Res.* **2000**, *60*, 916−921.
(63) Vander Heiden, M. G.; Cantley, L. C.; Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **2009**, *324*, 1029−1033.
(64) Jain, M.; Nilsson, R.; Sharma, S.; Madhusudhan, N.; Kitami, T.; Souza, A. L.; Kafri, R.; Kirschner, M. W.; Clish, C. B.; Mootha, V. K. Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science* **2012**, *336*, 1040−1044.
(65) Galán, A.; Hernández, J.; Jimenez, O. Measurement of blood acetoacetate and β-hydroxybutyrate in an automatic analyser. *J. Autom. Methods Manage. Chem.* **2001**, *23*, 69−76.
(66) Jobard, E.; Blanc, E.; Négrier, S.; Escudier, B.; Gravis, G.; Chevreau, C.; Elena-Herrmann, B.; Trédan, O. A serum metabolomic fingerprint of bevacizumab and temsirolimus combination as first-line treatment of metastatic renal cell carcinoma. *Br. J. Cancer* **2015**, *113*, 1148−1157.
(67) Richard, V.; Conotte, R.; Mayne, D.; Colet, J. M. Does the 1H-NMR plasma metabolome reflect the host-tumor interactions in human breast cancer? *Oncotarget* **2017**, *8*, 49915−49930.

DOI: 10.1021/acs.jproteome.9b00316
J. Proteome Res. XXXX, XXX, XXX−XXX

Supporting Information for Publication

Assessing treatment response and prognosis by serum and tissue metabolomics in breast cancer patients.

Julia Debik[1*], Leslie R. Euceda[1,2], Steinar Lundgren[3,4], Hedda von der Lippe Gythfeldt[5], Øystein Garred[7], Elin Borgen[7], Olav Engebraaten[5,6,8], Tone F. Bathen[1], Guro F. Giskeødegård[1*]

[1]Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, NTNU - Norwegian University of Science and Technology, 7491 Trondheim, Norway.

[2]CAMO Analytics, 0349 Oslo, Norway.

[3]Department of Oncology, St. Olav's University Hospital, 7006 Trondheim, Norway.

[4]Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, NTNU - Norwegian University of Science and Technology, 7491 Trondheim, Norway.

[5]Department of Oncology, Oslo University Hospital, 0424 Oslo, Norway.

[6]Department of Tumor Biology, Oslo University Hospital, 0424 Oslo, Norway.

[7]Department of Pathology, Oslo University Hospital, 0372 Oslo, Norway.

[8]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, 0318 Oslo, Norway.

Corresponding Authors

*Guro F. Giskeødegård, Email: guro.giskeødegård@ntnu.no

*Julia Debik, Email: julia.b.debik@ntnu.no

Table of contents

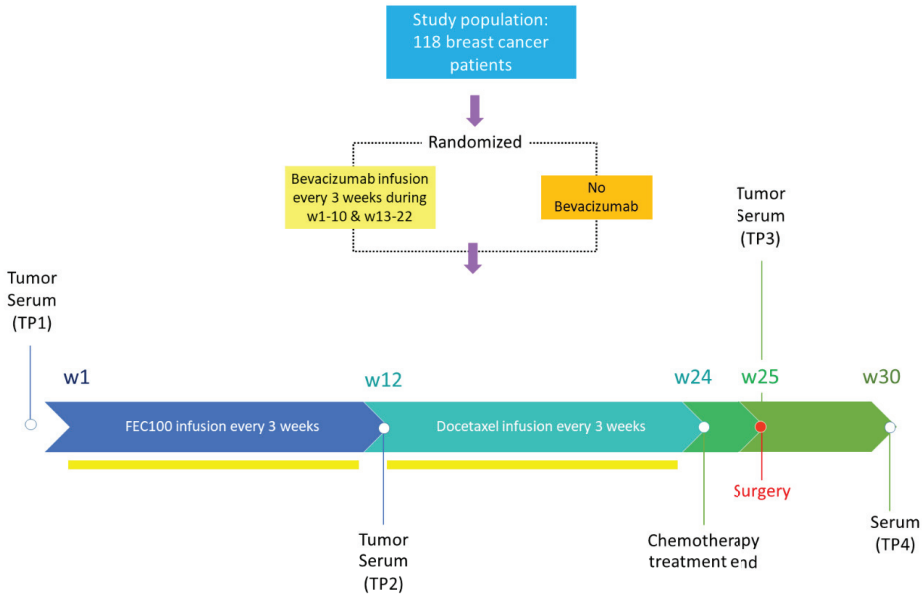


Figure S1. Diagram showing the treatment regime and experimental set up of the study.

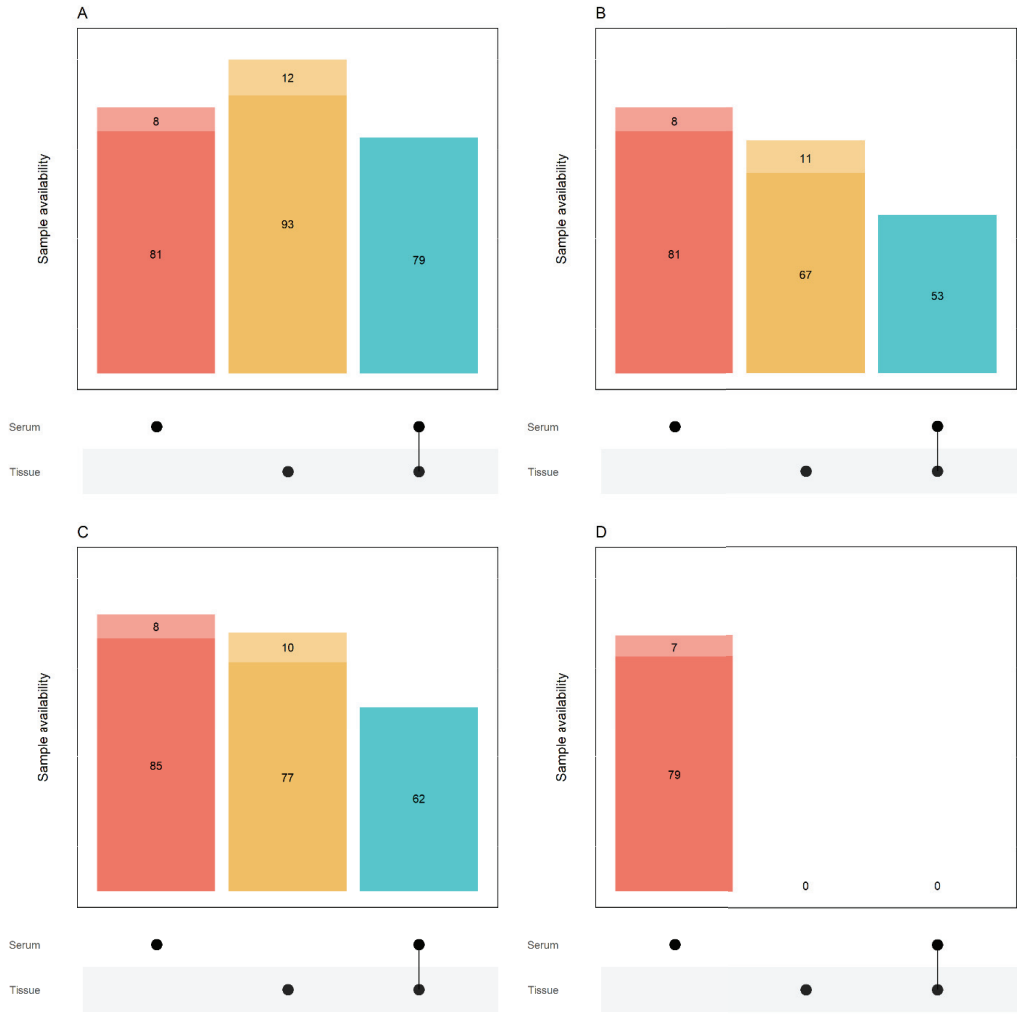


Figure S2. Sample availability of serum and tissue samples at the distinct sampling time points (TP's) in this study. Figures A, B, C, and D correspond to time points TP1, TP2, TP3, and TP4, respectively. For serum and tissue samples, a darker color denotes the number of survivors, while a lighter color denotes the number of non-survivors.

Table S1. Details on quantification of serum metabolites

	Metabolite	Chemical shift (peak multiplicity)
1	Leucine	0.95 (tr)
2	Valine	0.98 (d) 1.02 (d) 3.60 (d)
3	Isoleucine	1.00 (d)
4	2-methylglutarate	1.06 (d)
5	3-hydroxybutyrate	1.19 (d) 2.30 (q) 2.40 (q)
6	Alanine	1.47 (d)
7	Lysine	1.77 (m) 3.03 (right d)
8	Acetate	1.91 (s)
9	Acetoacetate	2.27 (s)
10	Glutamate	2.35 (d)
11	Pyruvate	2.36 (s)
12	Glutamine	2.45 (m)
13	Citrate	2.68 (right d)
14	Methionine	2.63 (right d)
15	Creatine	3.03 (s) 3.92 (s)
16	Creatinine	3.04 (s) 4.05 (s)
17	Ornithine	3.06 (d)
18	Proline betaine	3.1(s) 3.29 (s)
19	Dimethylsulfone	3.15 (s)
20	Glucose	3.25 (t) 3.42 (m) 3.36 (m) 3.49(m) 3.34 (q) 3.72 (m) 3.84 (m) 3.90 (q) 5.25 (d)
21	Methanol	3.36 (s)
22	Glycine	3.55 (s)
23	Lactate	4.12 (q)
24	Tyrosine	6.90 (d) 7.20 (d)
25	Phenylalanine	7.34 (d) 7.38 (t) 7.43 (t)
26	Histidine	7.05 (s) 7.78 (d)
27	Formate	8.46 (s)
28	Lipid 1	0.85 (br)
29	Lipid 2	1.57 (br)
*	Ethanol	1.17 (d)

Quantified metabolites and their placement on the ppm scale. ppm: parts per million; br: broad peak; s: singlet; d: doublet; dd: double doublet; t: triplet; q: quartet; m: multiplet; *: removed from analyses.

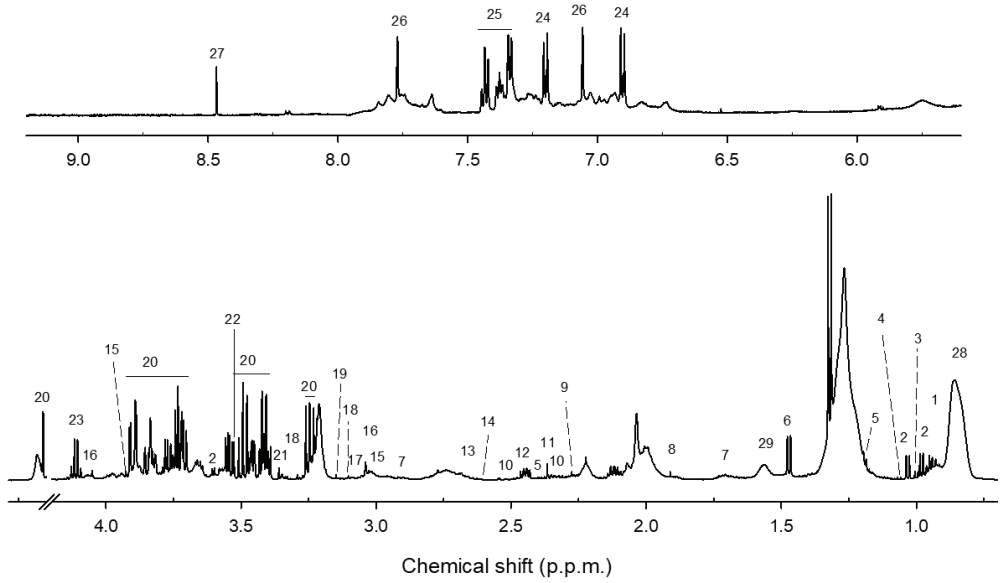


Figure S3: A representative spectrum with annotated metabolite peaks. 1: leucine; 2: valine; 3: isoleucine; 4: dimethylglutarate; 5: tri-hydroxybutyrate; 6: alanine; 7: lysine; 8: acetate; 9: acetoacetate; 10: glutamate; 11: pyruvate; 12: glutamine; 13: citrate; 14: methionine; 15: creatine; 16: creatinine; 17: ornithine; 18: proline-betaine; 19: dimethylsulfone; 20: glucose; 21: methanol; 22: glycine; 23: lactate; 24: tyrosine; 25: phenylalanine; 26: histidine; 27: formate; 28: lipid1; 29: lipid2.

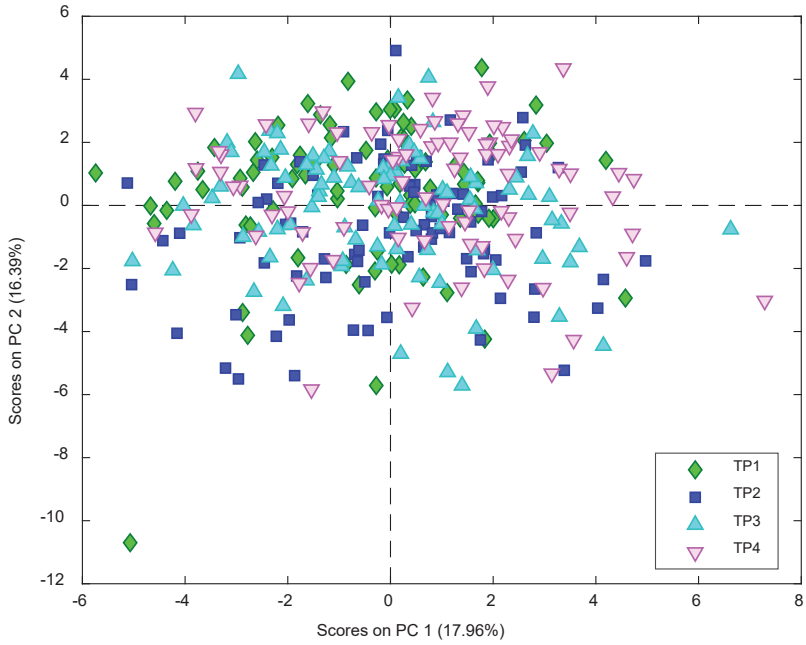


Figure S4. Scores plot for the PCA analysis of the serum metabolites. Each point corresponds to a sample colored according to the time point at which it was collected.

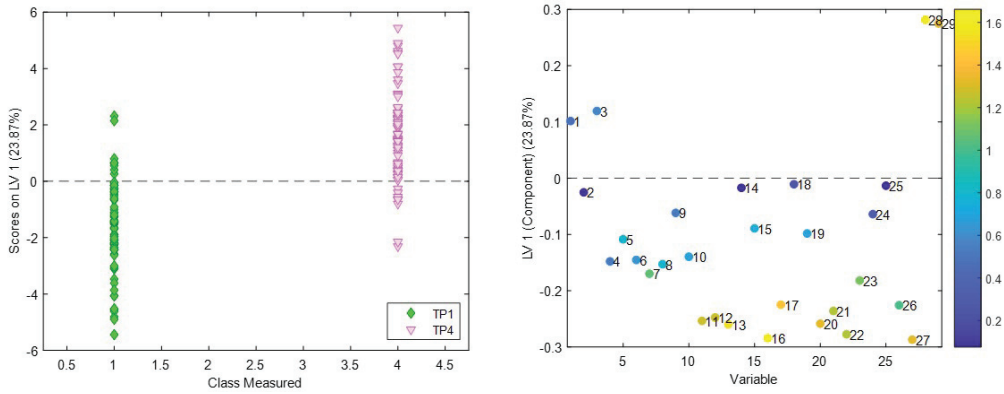


Figure S5. Scores (left) and loading plots (right) of multilevel PLS-DA analyses on serum metabolites comparing TP1 and TP4, with lipid peaks. Accuracy = 77%; p-value < 0.001. LV: latent variable. Loadings colored according to the VIP scores. 1: leucine; 2: valine; 3: isoleucine; 4: dimethylglutarate; 5: trihydroxybutyrate; 6: alanine; 7: lysine; 8: acetate; 9: acetoacetate; 10: glutamate; 11: pyruvate; 12: glutamine; 13: citrate; 14: methionine; 15: creatine; 16: creatinine; 17: ornithine; 18: proline-betaine; 19: dimethylsulfone; 20: glucose; 21: methanol; 22: glycine; 23: lactate; 24: tyrosine; 25: phenylalanine; 26: histidine; 27: formate; 28: lipid1; 29: lipid2.

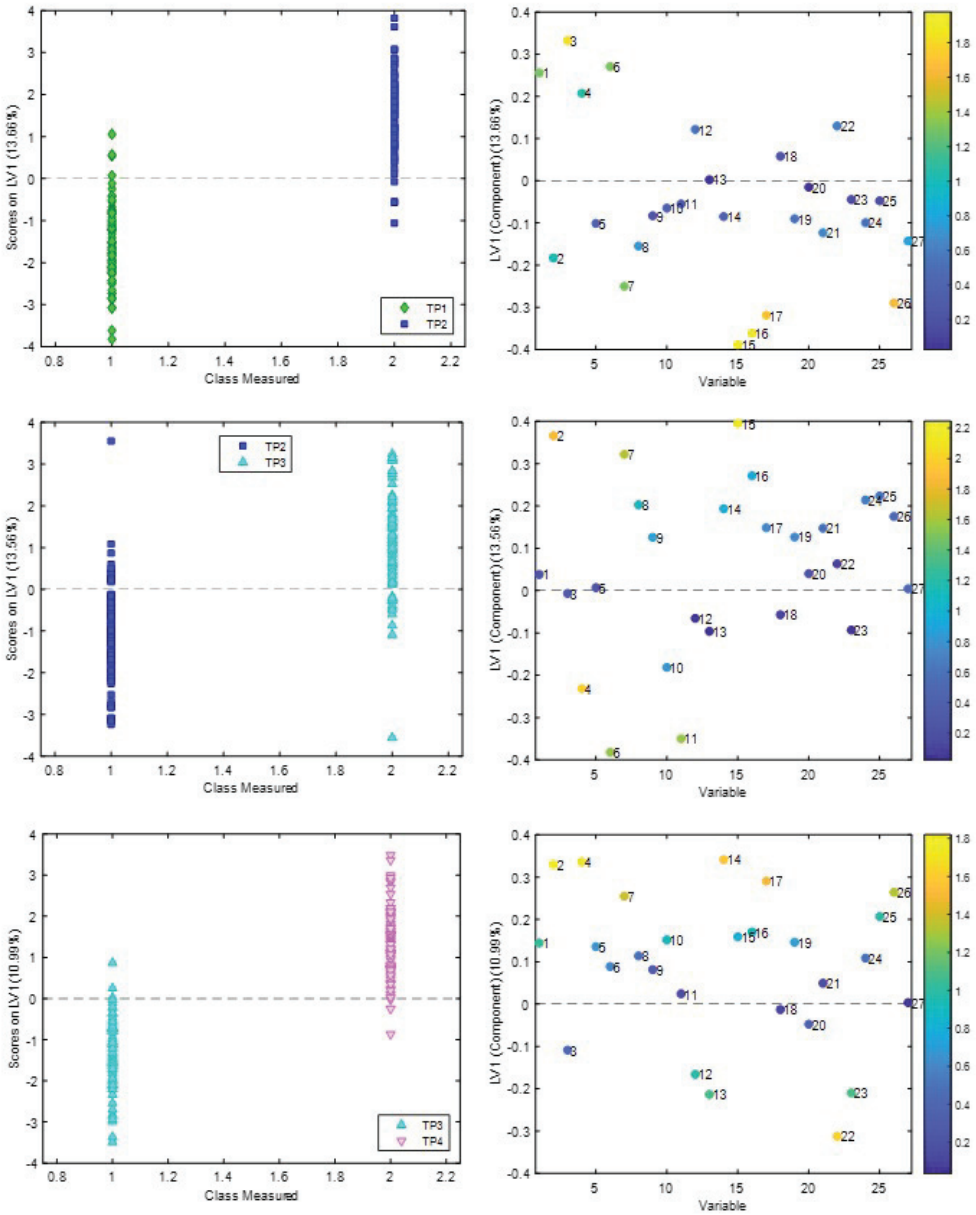


Figure S6. Scores (left) and loading plots (right) of multilevel PLS-DA analyses on serum metabolites. LV: latent variable. Top: TP2 vs TP1; Middle: TP3 vs TP2; Bottom: TP4 vs TP3. Loadings colored according to the VIP scores. 1: leucine; 2: valine; 3: isoleucine; 4: dimethylglutarate; 5: trihydroxybutyrate; 6: alanine; 7: lysine; 8: acetate; 9: acetoacetate; 10: glutamate; 11: pyruvate; 12: glutamine; 13: citrate; 14: methionine; 15: creatine; 16: creatinine; 17: ornithine; 18: proline-betaine; 19: dimethylsulfone; 20: glucose; 21: methanol; 22: glycine; 23: lactate; 24: tyrosine; 25: phenylalanine; 26: histidine; 27: formate

Paper II

Journal of proteome research

Cite This: J. Proteome Res. XXXX, XXX, XXX−XXX

Article

pubs.acs.org/jpr

Effect of Repeated Freeze−Thaw Cycles on NMR-Measured Lipoproteins and Metabolites in Biofluids

Feng Wang,[†] Julia Debik,[†] Trygve Andreassen,[‡] Leslie R. Euceda,[†,§] Tonje H. Haukaas,[†,∥] Claire Cannet,[⊥] Hartmut Schäfer,[⊥] Tone F. Bathen,[*,†] and Guro F. Giskeødegård[*,†]

[†]Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences and [‡]MR Core Facility, Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway

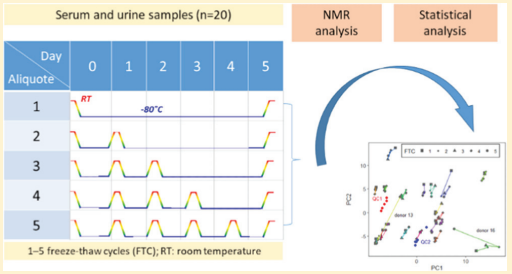
[§]Camo Analytics, Oslo Science Park, Gaustadalléen 21, 0349 Oslo, Norway

[∥]SINTEF Industry, Richard Birkelands vei 3, 7034 Trondheim, Norway

[⊥]Bruker Biospin GmbH, Silberstreifen, 76287 Rheinstetten, Germany

Supporting Information

ABSTRACT: Metabolic profiling of biofluids by nuclear magnetic resonance (NMR) spectroscopy serves as an important tool in disease characterization, and its accuracy largely depends on the quality of samples. We aimed to explore possible effects of repeated freeze−thaw cycles (FTCs) on concentrations of lipoprotein parameters in serum and metabolite concentrations in serum and urine samples. After one to five FTCs, serum and urine samples (n= 20) were analyzed by NMR spectroscopy, and 112 lipoprotein parameters, 20 serum metabolites, and 35 urine metabolites were quantified by a commercial analytical platform. Principal component analysis showed no systematic changes related to FTCs, and samples from the same donor were closely clustered, showing a higher between-subject variation than within-subject variation. The coefficients of variation were small (medians of 4.3%, 11.0%, and 4.9% for lipoprotein parameters and serum and urine metabolites, respectively). Minor, but significant accumulated freeze−thaw effects were observed for 32 lipoprotein parameters and one serum metabolite (acetic acid) when comparing FTC1 to further FTCs. Remaining lipoprotein and metabolite concentrations showed no significant change. In conclusion, five FTCs did not significantly alter the concentrations of urine metabolites and introduced only minor changes to serum lipoprotein parameters and metabolites evaluated by the NMR-based platform.



KEYWORDS: *NMR analysis, freeze−thaw cycle, quantification, metabolite, lipoprotein parameter*

INTRODUCTION

Metabolomics has become an important tool in medical research and involves analytical approaches with the capability to detect a wide range of metabolites in biofluids and tissues.[1] One of the main analytical approaches for metabolic characterization is proton nuclear magnetic resonance (NMR) spectroscopy. NMR spectroscopy is high-throughput, has high reproducibility, and requires minor sample preparation, thus having a great potential in human population studies.[2,3]

The quality and reproducibility of NMR analysis can be affected by the sample quality, which is influenced by factors resulting from preanalytical processes, including sample collection, storage, and preparation.[4−6] For example, chemical degradation processes, such as oxidation and decomposition of chemically unstable compounds, can severely influence sample composition. Biobanks are organized to collect, store, and distribute samples of human tissues and biofluids for a variety of clinical research purposes.[4] Blood plasma, serum, and urine are commonly available biofluids in biobanks and are usually stored in −80 °C freezers or in liquid nitrogen tanks (−196 °C) prior to analysis. Standard operating procedures for preanalytical handling of blood and urine samples for biobank metabolomic studies have been published.[6,7] Samples are recommended to be analyzed right after collection or to be stored at −80 °C (blood samples) or in liquid nitrogen/liquid nitrogen vapor (urine samples) until further analysis.[7] In practice, however, avoiding repeated thawing and refreezing of aliquots of biobank samples can be challenging.

Based on multivariate analysis of serum NMR spectra, a previous study showed that one freeze−thaw cycle (FTC) caused relatively small differences in the contents of lipids, alanine, glucose, and lactate.[8] Other studies have reported that

Received: May 27, 2019
Published: September 2, 2019

DOI: 10.1021/acs.jproteome.9b00343
J. Proteome Res. XXXX, XXX, XXX−XXX

serum or plasma composition was altered by multiple FTCs.[9,10] However, these studies included a small number of subjects and did not report quantitative concentrations. While high reproducibility of NMR-measured urine samples from rats was observed for up to eight FTCs,[11] no corresponding data are available from human urine samples.

Lipoproteins consisting of lipids and apolipoproteins are important constituents of lipid fractions, and their functions are to transfer water-insoluble lipids within the bloodstream. They have been identified as the primary drivers of atherosclerotic processes for several decades. However, recent data have suggested that, compared to lipoprotein particle numbers, measuring the total amount of cholesterol (CH) in each major lipoprotein fraction has less predictive power for cardiovascular risk in patients with metabolic disorders such as obesity and diabetes mellitus.[12,13] Improved characterization of lipoprotein subclass composition may help not only to understand the pathophysiology of atherosclerosis and diabetic dyslipidemia, but also to develop and monitor novel diet and drug therapies.[14] In addition, several epidemiological studies have indicated possible correlations between cholesterol and lipoprotein levels and risks for several cancers (such as breast cancer and prostate cancer).[15−17]

Ultracentrifugation-based lipoprotein quantification is time-consuming and labor-intensive, and NMR spectroscopy serves as an alternative rapid method for quantifying lipoproteins from plasma and serum samples. NMR spectroscopy allows for the measurement of lipoprotein subclasses, giving their lipid and apolipoprotein concentrations, as well as particle numbers and sizes.[18−20] A recently released commercial lipoprotein subclass analysis, based on an NMR-based metabolomics platform, can simultaneously quantify 112 lipoprotein parameters and 26 low-molecular-weight metabolites in blood.[21] In addition, an NMR-based quantification method for 50 urine metabolites has been established. While a previous study explored the effect of multiple FTC on lipoprotein particle numbers,[22] the effects of FTCs on NMR-measured concentrations of lipids and apolipoproteins in lipoprotein subclasses have not yet been reported.

In this study, we investigated the effect of repeated FTCs prior to NMR analysis on concentrations of serum lipoprotein parameters and metabolite concentrations in serum and urine samples. The aim was to gain insight into the extent that FTCs can affect the composition of these biological samples so as to avoid misinterpretation of findings arising from sample handling.

EXPERIMENTAL SECTION

Sample Collection and Experimental Design

Nonfasting serum and spot-urine samples were obtained from two sets of 20 anonymized healthy female and male adult donors. Blood samples were collected into serum tubes with no additives from the Blood bank, St. Olavs University Hospital, and left to coagulate for approximately 1 h before centrifugation (3000 rpm, 10 min). Urine samples were refrigerated (4 °C) after collection and were transferred to aliquots within 3 h. Each serum and urine sample was divided into five aliquots on ice and stored at −80 °C. The five aliquots of each sample were subjected to one to five FTCs before NMR analyses. For each FTC, serum and urine samples were thawed at room temperature for approximately 1 and 2 h, respectively, and refrozen at −80 °C for approximately 24 h. After the FTCs, the samples were stored in −80 °C until NMR analysis. The study was approved by the data protection officer at The Norwegian University of Science and Technology (NTNU). According to communication with the Regional Committee for Medical and Health Research Ethics in Central Norway, this quality control study using completely anonymized samples from healthy volunteers could be performed without formal ethical approval.

NMR Analysis and Data Preprocessing

After thawing at room temperature, 150 μL of serum was mixed with 150 μL of buffer [D_2O with 0.075 mM Na_2HPO_4, 5 mM NaN_3, 3.5 mM trimethylsilylpropanoic acid (TSP), pH 7.4]. Thawed urine samples were centrifuged at 12 121g at 4 °C for 5 min, and 540 μL of supernatant was mixed with 60 μL of buffer (pH 7.4, 1.5 mM KH_2PO_4 in D_2O, 0.1% TSP). Serum and urine mixtures were transferred to 3 and 5 mm NMR tubes, respectively. To assess the reproducibility of the NMR method and to compare with the variability of the FTC samples, four sets of quality control (QC) samples were prepared from pooled serum or urine samples: serum QC set 1 and set 2 (QC1 and QC2) and urine QC1 and QC2. Each set of QC consisted of five samples. NMR analysis was carried out on a Bruker Avance III Ultrashield Plus 600 MHz spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany) equipped with a 5 mm QCI Cryoprobe. Sample handling and data acquisition were automatically performed using SampleJet sample changer and Icon-NMR on Topspin 3.5 (Bruker BioSpin). NMR spectra were recorded using a one-dimensional nuclear Overhauser effect spectroscopy pulse sequence (noesygppr1d), with irradiation (25 Hz) on the water resonance during relaxation delay (4 s) and mixing time (10 ms). The urine spectra were recorded at 300 K, with 64k data points and 20 ppm spectral width. The serum spectra were recorded at 310 K using 96k data points and 30 ppm spectral width. For both biofluids, 32 scans were recorded and the free induction decays were Fourier-transformed after 0.3 Hz line broadening to 128k real data points. For urine, two-dimensional JRES spectra (jresgpprqf) were also recorded, using two scans, water presaturation (25 Hz) during relaxation delay (2 s), 8k direct and 40 indirect data points with 10 026.7 and 78 Hz spectral width, respectively, and 12 820.51 μs delay incrementation.

Lipoprotein Parameter Analysis and Metabolite Quantification

Lipoprotein parameter analysis was automatically performed using the commercial Bruker IVDr Lipoprotein Subclass Analysis (B.I.LISA™) method from Bruker BioSpin.[21] This method provides the concentrations of lipids [cholesterol (CH), free cholesterol (FC), triglycerides (TG), and phospholipids (PL)] in serum, and in four main lipoprotein classes: very low-, intermediate-, low-, and high-density lipoproteins (VLDL, IDL, LDL, and HDL, respectively) as well as 15 subclasses (VLDL 1−5, LDL 1−6, and HDL 1−4). Simultaneously, it quantifies the concentrations of apolipoproteins (Apo-A1, Apo-A2, and Apo-B) in serum, two main classes (HDL and LDL) and 10 subclasses (HDL 1−4 and LDL 1−6). In addition, the model gives 12 calculated parameters, including ratios of LDL-CH/HDL-CH and Apo-B/Apo-A1, and 10 particle numbers (particle numbers of total serum, VLDL, IDL, LDL, and LDL 1−6). In total, this yields 112 quantitative lipoprotein parameters (Table S1). The

DOI: 10.1021/acs.jproteome.9b00343
J. Proteome Res. XXXX, XXX, XXX−XXX

density ranges of different lipoprotein subclasses are listed in Table S1.

A total of 26 serum and 50 urine metabolite concentrations were automatically quantified using Bruker B.I. Quant-PS™ and Bruker B.I.Quant-UR™ methods, respectively, based on algorithms developed for fitting predefined proton signals.[21] By automatic quantification applying limits of detection (LODs) as threshold for quantification, concentrations were not reported for several of the metabolites. We thus disregarded LODs in the quantification of these samples. Metabolites with concentrations >0 for more than 30% of serum or urine samples were classified as quantifiable and were included in the analysis.

Multivariate Modeling and Statistical Analysis

The effect of repeated FTCs on the serum and urine samples was evaluated using both multivariate and univariate statistical analyses. Zero values were replaced by half of the lowest detected value of the corresponding metabolite or lipoprotein. Principal component analysis (PCA)[23] was carried out to visually assess variation in the metabolic profiles within and between donors. PCA is an unsupervised dimension reduction technique, which makes it possible to visualize the majority of the variance in the data, projected onto a lower-dimensional space. Each point on the PCA scores plot represents one sample. PCA was performed in Matlab R2017b[24] using the PLS-toolbox version 8.6.2.[25]

Coefficients of variation (CV) and intraclass correlations (ICCs) were calculated to evaluate the reproducibility of quantified lipoprotein parameters and metabolites, given repeated FTCs.[26] CVs were calculated for each lipoprotein/metabolite parameter, across all FTCs of a donor separately, to determine their relative extent of variation (within-donor variation). ICCs were calculated for each lipoprotein/metabolite parameter, across all FTCs and donors to evaluate the reproducibility. For each lipoprotein parameter and metabolite separately, Wilcoxon signed-rank tests were used to test if the lipoprotein or metabolite levels between two consecutive FTCs were significantly different. To test for accumulated effects of FTCs, Wilcoxon signed-rank tests were used to compare lipoprotein and metabolite levels of FTC2−5 samples with FTC1. p-values were adjusted using the Benjamini−Hochberg procedure[27] with significance considered for corrected p-values ≤ 0.05. Percentage changes were calculated for individual lipoprotein parameters and metabolites to obtain a visual representation of the changes in their levels between the FTCs. All statistical analyses were carried out in R.[28] The nmle package[29] was used for calculating the ICCs.

RESULTS

Concentrations of Lipoprotein Parameters and Metabolites in Serum and Urine Samples

Figure 1 shows the five FTC serum and urine spectra obtained from two representative donors. Variations within the same donor among FTCs are clearly much lower compared to the variations across different donors. We successfully quantified 112 lipoprotein parameters and 20 serum metabolites from the serum NMR spectra (Tables S1 and S2), and 35 urine metabolites from the urine NMR spectra (Table S3).

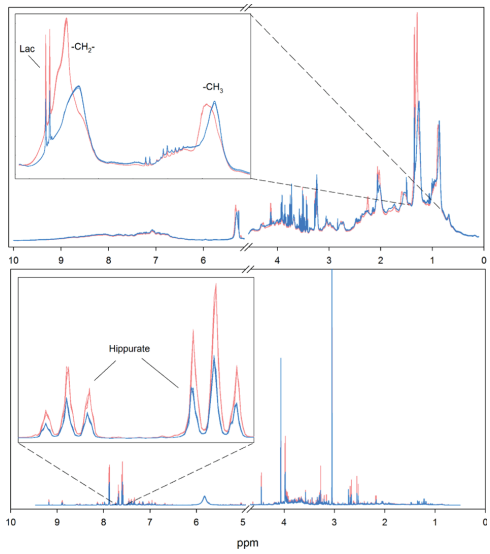


Figure 1. NMR spectra of serum and urine samples. Serum (top) and urine (bottom) spectra from two representative donors, colored in orange and blue, respectively. All five spectra, one from each FTC are plotted. For the serum spectra, the area in focus shows part of the spectral region where the lipoprotein signals appear. The spectral position of the $-CH_2-$ and $-CH_3$ signals reflect the lipoprotein particle size.[43] The area in focus for the urine spectra shows signals from hippurate as an example.

Effect of Freeze−Thaw Cycles on Lipoprotein Parameter Concentrations

The impact of repeated FTCs on concentrations of lipoprotein parameters in serum samples was visualized by PCA (Figure 2B,C). No systematic variation due to repeated FTCs was apparent. The score plot (Figure 2B) shows that most of the samples with different FTCs were well grouped according to donors, and the variations within five FTCs were lower than those between different donors. From FTC4 to FTC5, samples from two donors (numbers 13 and 16) had larger spread in the PCA scores than the QCs. However, these changes were not in the same direction, indicating that these variations are not caused by FTCs, but rather represent variations from sample preparation.

As shown in Figure 2A, CVs of lipoprotein parameters in samples thawed and refrozen one to five times were comparable to those of the QCs. Median CV for the lipoprotein parameters were between 1.1-16.8%,ith a median of 4.3% across all parameters. Only 4/112 lipoprotein parameters had median CVs > 15% (IDCH, IDFC, V2FC, V4FC). The ranges of CVs for quality control sets QC1 and QC2 were 0.0-82.5% (median: 5.0%) and 0.6-27.9% (median: 6.0%), respectively. ICCs of FTC samples are presented in Figure S1A. ICC was larger than 0.8 for most of the parameters (94 of 105), giving a mean ICC of 0.91 across all lipoprotein parameters.

Wilcoxon signed-rank tests were used to test whether the differences in concentrations between each pair of subsequent FTCs were significant. Most of the lipoprotein parameters did not differ significantly between any of the consecutive FTCs (Table S4). Five VLDL-related lipoprotein parameters had

DOI: 10.1021/acs.jproteome.9b00343
J. Proteome Res. XXXX, XXX, XXX−XXX

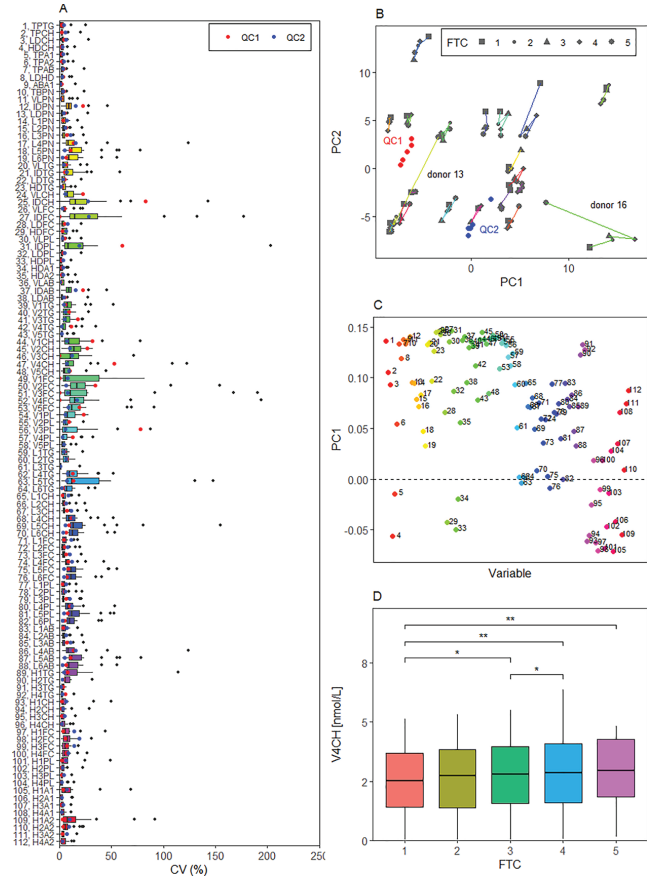


Figure 2. Effects of five FTCs on concentrations of 112 lipoprotein parameters in serum samples. (A) Box plots of coefficients of variation for all lipoprotein parameters. Median values, interquartile range, and outliers are presented. Values from quality control samples [QC1 (red) and QC2 (blue)] are shown for comparison. • represents outliers. (B) PCA scores plot. Samples from the same donor are connected by lines of different colors. (C) Loadings of PC1 from PCA. Numbers represent each lipoprotein parameter, as shown in (A). Colors for loadings are similar to those for corresponding coefficients of variation in (A). (D) V4CH concentrations of samples with different numbers of FTCs. Significance tested by Wilcoxon signed-rank tests. *: adjusted p-value $\leq$ 0.05; **: adjusted p-value $\leq$ 0.01.

significantly different concentrations between FTC3 and FTC4: VLCH, VLFC, V3FC, V4CH, and V1CH (adjusted p-values = 0.015, 0.015, 0.038, 0.031, and 0.003, respectively) (exemplified by V4CH in Figure 2D). However, relative changes in median concentrations for these five lipoprotein parameters were low (range: 2.93−6.88% increase in concentration).

Although few significant changes in the lipoprotein levels were found between consecutive FTCs, some significant accumulated effects of FTCs were observed (Table S4). This is exemplified by V4FC in Figure 2D. A total of 32 lipoprotein parameters showed accumulated effects when comparing FTC1 to FTC5, of which significant accumulated effects were detectable from FTC3 for 22 parameters, from FTC4 for 8 parameters, and from FTC5 for 2 parameters. This resulted in decreased concentrations for 10 parameters and increased concentrations for 22 parameters compared to the levels of FTC1 samples.

Effect of Freeze−Thaw Cycles on Metabolite Concentrations in Serum Samples

No systematic effects of one to five FTCs on serum metabolite concentrations were observed as visualized by PCA (Figure 3A). With the exception of samples from donors 13 and 16, all samples were clustered according to the corresponding donors and the variances of samples in each donor were lower than or of the same order of magnitude as those across different donors. As shown in Figure 3B, median CVs of all of the 20 metabolites were between 1.7-67.2%. with a median of 11.0% across all metabolites. 13/20 metabolites had median CV < 15%, and 18/20 metabolites had median CV < 25%. . Among all of the 20 quantified metabolites, trimethylamine N-oxide had the highest CV (median: 67.2%). The ranges of CVs for QC1 and QC2 samples were 3.2-51.9% (median: 10.4%) and 4.2-79.5% (median: 11.1%), respectively. ICCs of all serum metabolites were between 0.3 and 0.9, with a mean of 0.7 (Figure S1B). Results of PCA and ICC together show a small

DOI: 10.1021/acs.jproteome.9b00343
J. Proteome Res. XXXX, XXX, XXX−XXX

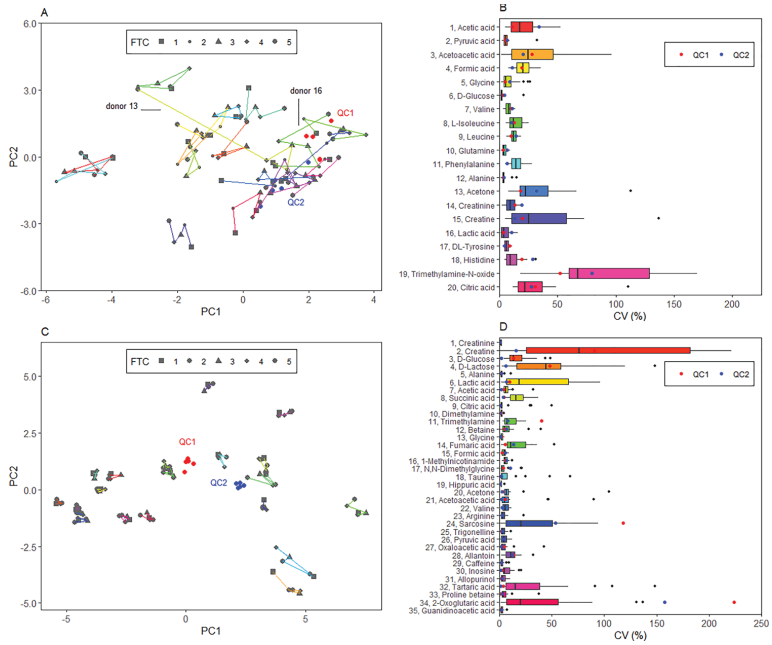


Figure 3. Effects of five FTCs on serum and urine metabolite concentrations. (A, C) PCA scores plots for serum (A) and urine (C) metabolites. Samples from the same donor are connected by lines of different colors. Values from quality control samples [QC1 (red) and QC2 (blue)] are shown for comparison. (B, D) Box plots of the coefficients of variation for serum (B) and urine (D) metabolites. Median values, interquartile ranges, and outliers are presented.

overall variation in the data, indicating that this cohort of healthy donors represents a homogeneous group.

Wilcoxon signed-rank tests showed that the concentrations of the 20 serum metabolites did not significantly differ between two consecutive FTCs (Table S5). The concentration of serum acetic acid was significantly increased between FTC1 and FTC5 (adjusted p-value = 0.01), while no significant accumulated effects of FTCs were found for other serum metabolites (Table S5).

Effect of Freeze–Thaw Cycles on Metabolite Concentrations in Urine Samples

We found that one to five FTCs had no systematic effect on the metabolite concentrations in urine samples, as shown by PCA (Figure 3C). Samples from the same donors were clustered and within-donor variations were lower compared to between-donor variations. Figure 3D demonstrates that the median CVs were between 1.3-76.1%, with a median of 4.9% across all metabolites. 28/35 metabolites had median CV < 15%. The CVs of creatine (median: 76.1%) were higher compared to other metabolites. CVs of the two sets of quality control samples varied from 0.7-223.6% (median: 3.5%)- and from 0.4-157.6% (median: 2.4%), respectively. ICCs of all urine metabolites were high (ICC > 0.9 for 33 of 35 metabolites) (Figure S1C).

Based on Wilcoxon signed-rank tests, no significant differences between two consecutive FTCs on the metabolite concentrations were found (Table S6). No accumulated effects of FTCs on the metabolite levels were found (Table S6).

DISCUSSION

In this study, we examined the impact of repeated FTCs on the absolute concentrations of lipoprotein parameters and metabolites determined by NMR spectroscopy of human biofluids. We found that up to five repeated FTCs induced minimal changes in the measured concentrations of serum and urine metabolites and lipoproteins.

NMR-based quantitative analysis has been applied widely to evaluate the association of lipoproteins with cardiovascular diseases, diabetic dyslipidemia, and cancer.[20,30,31] To the best of our knowledge, we are the first to report the effects of repeated FTCs on NMR-measured concentrations of lipids and apolipoproteins in lipoprotein subclasses. Using a commercial analytical platform for automatic quantification of lipoproteins and their subclasses, we found that systematic changes of lipoprotein parameter concentrations did not appear for up to five repeated FTCs. While the concentrations of five VLDL-derived parameters were significantly changed between FTC3 and FTC4, the relative changes in concentration values were low. The remaining 107 lipoprotein parameters did not change between consecutive FTCs. However, minor, significant accumulated effects of FTCs were observed for 32 lipoprotein parameters, occurring after FTC3, FTC4, or FTC5. The median CV across all of the parameters was 4.3%, showing a high reproducibility of the experiments and little variation of samples with multiple FTCs. Median CV for total cholesterol (TPCH), LDL cholesterol (LDCH) and HDL cholesterol (HDCH) was 2.0%, 3.0%, and 2.1%, respectively. This is in the range of the recommendations

DOI: 10.1021/acs.jproteome.9b00343
J. Proteome Res. XXXX, XXX, XXX−XXX

of the National Cholesterol Education Program (NCEP) Laboratory Standardization Panel (CV < 3.0% for cholesterol, CV < 4.0% for LDL cholesterol, and CV < 5.0% for HDL cholesterol).[32−35]

Our findings are consistent with previous studies on clinically measured cholesterol levels, showing that FTCs had either no significant effects or only minor effects (less than day-to-day variations) on serum or plasma LDL and HDL concentrations.[22,36,37] Conversely, compared to fresh samples, freezing prior to lipoprotein fractionation via density gradient ultracentrifugation was shown to cause large variations (up to 37%) in concentrations of HDL and LDL cholesterol, and VLDL-free fatty acids in serum.[38] This is in line with the certification protocols for determining total cholesterol, HDL and LDL cholesterol, which recommend to use fresh samples for the reference method and to use frozen samples with caution.[32−34] However, in large cohorts such as biobank studies, it is not feasible to analyze fresh serum; thus, our study is limited to the comparison of lipoprotein concentrations in frozen samples with different FTCs. Moreover, in our study, lipoprotein particle numbers did not significantly change (LDL, and subclasses of LDL-2-6) or slightly increased (VLDL, IDL, and LDL-1) after multiple FTCs. Similarly, a recent study demonstrated that three or more FTCs, when thawing plasma samples in a cold room, caused significant, but relatively small, changes in lipoprotein sizing (a shift of HDL particle size from large to small, decrease in large LDL and increase in IDL), and no changes in the proteome.[22]

No significant changes in urine metabolite concentrations were observed with up to five FTCs, while only acetic acid changed significantly within small-molecule serum metabolites. Compared to urine metabolites, the spread of donors in the PCA scores plot was lower for serum metabolites, possibly reflecting high homogeneity in this cohort of healthy donors. The coefficients of variation for some of the serum metabolites were high. These high CVs could result from mean concentration values close to zero, in which case CV is not an appropriate measure of reproducibility. However, the range of ICCs of serum metabolites was also high (from 0.3 to 0.9), eight of which were below 0.7. This indicates a lower reproducibility for some of the serum metabolites, such as trimethylamine-N-oxide and acetoacetic acid.

Unlike our findings, previous NMR-based studies have reported changes in plasma composition or relative quantitative levels of metabolites by either two or multiple FTCs.[8−10] Based on three subjects, Pinto et al.[9] reported that the effects of repeated freezing and thawing on plasma metabolites appeared after four and five cycles, but not within the first three cycles. Upon five consecutive FTCs, alterations consisted of decreased lipids and acetone and increased choline phospholipid, alanine, glucose, and pyruvate. Similarly, compared to one FTC, 5 or 10 FTCs were shown to have a visible impact on the metabolic profile of serum samples, and the levels of several small molecular metabolites were decreased, including choline, glycerol, methanol, ethanol, and proline.[10] Four of these mentioned metabolites were analyzed in our study. Among them, acetone and alanine displayed nonsignificant changes across five FTCs, and pyruvate and ethanol were undetectable. Moreover, Teahan et al.[8] detected minor spectral differences in samples with only one additional FTC. However, their freeze−thaw conditions are different compared to ours and the others: they thawed and refroze sample mixtures consisting of fresh sera and saline in NMR tubes at −40 °C prior to NMR analysis, while we subjected serum samples more in accordance with a biobank setting. It cannot be ruled out that dilution with saline might change the stability of serum. To assess possible differences in the stability of serum and plasma metabolites, studies should be performed to compare the effects of FTCs in serum and plasma samples under the same experimental conditions.

The effect of FTCs has also been evaluated by mass spectrometry-based analysis, reporting small changes in metabolic profiles in blood samples. Up to four FTCs only affected the plasma metabolome slightly but increased the individual variability.[39] By use of a targeted liquid chromatography−mass spectrometry (LC−MS) approach, Breier et al. discovered that, compared to those in fresh-frozen serum samples, all 159 investigated metabolites, except methionine sulfoxide, maintained stable concentrations in samples with two FTCs.[40] After three FTCs, 11 metabolites had significantly decreased concentrations, among which are the amino acids isoleucine, tryptophan, and valine. Similarly, Anton et al. found quite stable metabolite concentrations with up to four FTCs compared to fresh-frozen samples.[41] Among a total of 163 measured metabolites, only five amino acids, including glycine, methionine, phenylalanine, tryptophan, and tyrosine, had slightly, but not significantly increased concentrations with increased numbers of FTCs. Most metabolite concentrations were stable in serum samples subjected to two FTCs.

We found that urine metabolite concentrations did not significantly differ between consecutive FTCs and that no systematic changes appeared after five FTCs. Our findings are consistent with data reported from NMR analysis of rat urine samples, which showed that up to five FTCs had no influence on single metabolite concentrations or metabolomics classification approach.[11] In contrast, data from metabolic profiling of human urine samples measured by an LC−MS/MS method showed that three FTCs, but not one or two cycles, significantly increased concentrations of two metabolites (propionylcarnitine and hexose), compared to samples immediately frozen at −80 °C.[42] However, these two metabolites were not quantified in our study.

CONCLUSIONS

In conclusion, evaluated by using a commercially available NMR-based platform, quantification of lipoprotein parameters and metabolites was reproducible even with five repeated FTCs. No significant effects on concentrations of small-molecule metabolites in urine samples were observed. Minor accumulated changes were observed in concentrations of 32/112 lipoprotein parameters and 1/20 metabolites in serum samples. Although significant, the variation was still within the recommended NCEP guidelines for total cholesterol, LDL cholesterol and HDL cholesterol. Taken together, our results show that using this platform clearly allows for utilizing samples with different FTCs for quantification of metabolites and lipoprotein parameters in the same study.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.9b00343.

Intraclass correlation coefficients for lipoproteins, serum, and urine metabolites (Figure S1); measured lipoprotein

DOI: 10.1021/acs.jproteome.9b00343
J. Proteome Res. XXXX, XXX, XXX−XXX

parameters (Table S1); measured serum metabolites (Table S2); measured urine metabolites (Table S3); effects of FTCs on lipoprotein parameter concentrations (Table S4); effects of FTCs on serum metabolite concentrations (Table S5); and effects of FTCs on urine metabolite concentrations (Table S6) (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: tone.f.bathen@ntnu.no (T.F.B.).
*E-mail: guro.giskeodegard@ntnu.no (G.F.G.).

ORCID

Feng Wang: 0000-0002-8773-6963
Julia Debik: 0000-0002-2765-8043

Author Contributions

F.W. and J.D. contributed equally as first authors, and T.F.B. and G.F.G. contributed equally as last authors to this work. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

Part of this work was supported by the Norwegian Cancer Society (grant 163243).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors sincerely thank all of the volunteers who donated their blood and urine samples to this study. The NMR analyses were performed at the MR Core Facility, Norwegian University of Science and Technology (NTNU). MR Core facility is funded by the Faculty of Medicine at NTNU and Central Norway Regional Health Authority. H.S. and C.C. are employees of Bruker Biospin who offer the methods described here (B.I.LISA™, B.I.Quant-PS™, and B.I.Quant-UR™) as a commercial product.

ABBREVIATIONS

NMR, nuclear magnetic resonance; FTC, freeze−thaw cycle; QC, quality control; PCA, principal component analysis; CV, coefficient of variation; ICC, intraclass correlation

REFERENCES

(1) Markley, J. L.; Bruschweiler, R.; Edison, A. S.; Eghbalnia, H. R.; Powers, R.; Raftery, D.; Wishart, D. S. The future of NMR-based metabolomics. *Curr. Opin. Biotechnol.* **2017**, *43*, 34−40.

(2) Holmes, E.; Loo, R. L.; Stamler, J.; Bictash, M.; Yap, I. K.; Chan, Q.; Ebbels, T.; De Iorio, M.; Brown, I. J.; Veselkov, K. A.; Daviglus, M. L.; Kesteloot, H.; Ueshima, H.; Zhao, L.; Nicholson, J. K.; Elliott, P. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* **2008**, *453*, 396−400.

(3) Barton, R. H.; Nicholson, J. K.; Elliott, P.; Holmes, E. High-throughput 1H NMR-based metabolic analysis of human serum and urine for large-scale epidemiological studies: validation study. *Int. J. Epidemiol.* **2008**, *37*, i31−i40.

(4) Yin, P.; Lehmann, R.; Xu, G. Effects of pre-analytical processes on blood samples used in metabolomics studies. *Anal. Bioanal. Chem.* **2015**, *407*, 4879−4892.

(5) Beckonert, O.; Keun, H. C.; Ebbels, T. M.; Bundy, J.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* **2007**, *2*, 2692−2703.

(6) Emwas, A. H.; Roy, R.; McKay, R. T.; Ryan, D.; Brennan, L.; Tenori, L.; Luchinat, C.; Gao, X.; Zeri, A. C.; Gowda, G. A.; Raftery, D.; Steinbeck, C.; Salek, R. M.; Wishart, D. S. Recommendations and Standardization of Biomarker Quantification Using NMR-Based Metabolomics with Particular Focus on Urinary Analysis. *J. Proteome Res.* **2016**, *15*, 360−373.

(7) Bernini, P.; Bertini, I.; Luchinat, C.; Nincheri, P.; Staderini, S.; Turano, P. Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. *J. Biomol. NMR* **2011**, *49*, 231−243.

(8) Teahan, O.; Gamble, S.; Holmes, E.; Waxman, J.; Nicholson, J. K.; Bevan, C.; Keun, H. C. Impact of analytical bias in metabonomic studies of human blood serum and plasma. *Anal. Chem.* **2006**, *78*, 4307−4318.

(9) Pinto, J.; Domingues, M. R.; Galhano, E.; Pita, C.; Almeida Mdo, C.; Carreira, I. M.; Gil, A. M. Human plasma stability during handling and storage: impact on NMR metabolomics. *Analyst* **2014**, *139*, 1168−1177.

(10) Fliniaux, O.; Gaillard, G.; Lion, A.; Cailleu, D.; Mesnard, F.; Betsou, F. Influence of common preanalytical variations on the metabolic profile of serum samples in biobanks. *J. Biomol. NMR* **2011**, *51*, 457−465.

(11) Schreier, C.; Kremer, W.; Huber, F.; Neumann, S.; Pagel, P.; Lienemann, K.; Pestel, S. Reproducibility of NMR analysis of urine samples: impact of sample preparation, storage conditions, and animal health status. *BioMed Res. Int.* **2013**, *2013*, No. 878374.

(12) Rosenson, R. S.; Davidson, M. H.; Pourfarzib, R. Under-appreciated opportunities for low-density lipoprotein management in patients with cardiometabolic residual risk. *Atherosclerosis* **2010**, *213*, 1−7.

(13) Davidson, M. H.; Ballantyne, C. M.; Jacobson, T. A.; Bittner, V. A.; Braun, L. T.; Brown, A. S.; Brown, W. V.; Cromwell, W. C.; Goldberg, R. B.; McKenney, J. M.; Remaley, A. T.; Sniderman, A. D.; Toth, P. P.; Tsimikas, S.; Ziajka, P. E.; Maki, K. C.; Dicklin, M. R. Clinical utility of inflammatory markers and advanced lipoprotein testing: advice from an expert panel of lipid specialists. *J. Clin. Lipidol.* **2011**, *5*, 338−367.

(14) Mallol, R.; Rodriguez, M. A.; Brezmes, J.; Masana, L.; Correig, X. Human serum/plasma lipoprotein analysis by NMR: application to the study of diabetic dyslipidemia. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *70*, 1−24.

(15) Furberg, A. S.; Veierod, M. B.; Wilsgaard, T.; Bernstein, L.; Thune, I. Serum high-density lipoprotein cholesterol, metabolic profile, and breast cancer risk. *J. Natl. Cancer Inst.* **2004**, *96*, 1152−1160.

(16) Chang, S. J.; Hou, M. F.; Tsai, S. M.; Wu, S. H.; Hou, L. A.; Ma, H.; Shann, T. Y.; Wu, S. H.; Tsai, L. Y. The association between lipid profiles and breast cancer among Taiwanese women. *Clin. Chem. Lab. Med.* **2007**, *45*, 1219−1223.

(17) Katzke, V. A.; Sookthai, D.; Johnson, T.; Kuhn, T.; Kaaks, R. Blood lipids and lipoproteins in relation to incidence and mortality risks for CVD and cancer in the prospective EPIC-Heidelberg cohort. *BMC Med.* **2017**, *15*, No. 218.

(18) Bathen, T. F.; Krane, J.; Engan, T.; Bjerve, K. S.; Axelson, D. Quantification of plasma lipids and apolipoproteins by use of proton NMR spectroscopy, multivariate and neural network analysis. *NMR Biomed.* **2000**, *13*, 271−288.

(19) Jeyarajah, E. J.; Cromwell, W. C.; Otvos, J. D. Lipoprotein particle analysis by nuclear magnetic resonance spectroscopy. *Clin. Lab. Med.* **2006**, *26*, 847−870.

(20) Soininen, P.; Kangas, A. J.; Wurtz, P.; Tukiainen, T.; Tynkkynen, T.; Laatikainen, R.; Jarvelin, M. R.; Kahonen, M.; Lehtimaki, T.; Viikari, J.; Raitakari, O. T.; Savolainen, M. J.; Ala-Korpela, M. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. *Analyst* **2009**, *134*, 1781−1785.

DOI: 10.1021/acs.jproteome.9b00343
J. Proteome Res. XXXX, XXX, XXX−XXX

(21) Jiménez, B.; Holmes, E.; Heude, C.; Tolson, R. F.; Harvey, N.; Lodge, S. L.; Chetwynd, A. J.; Cannet, C.; Fang, F.; Pearce, J. T. M.; Lewis, M. R.; Viant, M. R.; Lindon, J. C.; Spraul, M.; Schafer, H.; Nicholson, J. K. Quantitative Lipoprotein Subclass and Low Molecular Weight Metabolite Analysis in Human Serum and Plasma by (1)H NMR Spectroscopy in a Multilaboratory Trial. *Anal. Chem.* **2018**, *90*, 11962−11971.
(22) Rebholz, S. L.; Melchior, J. T.; Welge, J. A.; Remaley, A. T.; Davidson, W. S.; Woollett, L. A., Effects of Multiple Freeze/Thaw Cycles on Measurements Novel Biomarkers Associated With Adverse Pregnancy Outcomes. *J. Clin. Lab. Med.* **2017**, *2* DOI: 10.16966/2572-9578.107.
(23) Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37−52.
(24) *MATLAB*, R2017b; The MathWorks Inc.: Natick, MA, 2017.
(25) *PLS Toolbox*, version 8.6.2; Eigenvector Research Inc.: Manson, WA, 2018.
(26) Rosner, B. *Fundamentals of Biostatistics*, 8th ed.; CENGAGE Learning: Boston, 2016.
(27) Benjamini, Y.; Hockberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc., Ser. B* **1995**, *57*, 289−300.
(28) *R: A Language and Environment for Statistical Computing*; R Development Core Team: Vienna, Austria, 2008.
(29) Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D. R Core Team (2019), *nmle: Linear and Nonlinear Mixed Effects Models*, R package version 3.1-141, https://CRAN.R-project.org/package=nlme.
(30) Madssen, T. S.; Thune, I.; Flote, V. G.; Lundgren, S.; Bertheussen, G. F.; Frydenberg, H.; Wist, E.; Schlichting, E.; Schafer, H.; Fjosne, H. E.; Vettukattil, R.; Lomo, J.; Bathen, T. F.; Giskeodegard, G. F. Metabolite and lipoprotein responses and prediction of weight gain during breast cancer treatment. *Br. J. Cancer* **2018**, *119*, 1144−1154.
(31) Flote, V. G.; Vettukattil, R.; Bathen, T. F.; Egeland, T.; McTiernan, A.; Frydenberg, H.; Husoy, A.; Finstad, S. E.; Lomo, J.; Garred, O.; Schlichting, E.; Wist, E. A.; Thune, I. Lipoprotein subfractions by nuclear magnetic resonance are associated with tumor characteristics in breast cancer. *Lipids Health Dis.* **2016**, *15*, No. 56.
(32) National Reference System for Cholesterol. *Total Cholesterol Certification Protocol for Manufactures (Revised)*; Cholesterol Reference Method Laboratory Network (CRMLN), 2004.
(33) Centers for Disease Control and Prevention. *HDL Cholesterol Certification Protocol for Manufactures*; Cholesterol Reference Method Laboratory Network (CRMLN), 2018.
(34) Centers for Disease Control and Prevention. *LDL Cholesterol Certification Protocol for Manufacturers*; Cholesterol Reference Method Laboratory Network (CRMLN), 2018.
(35) *Recommendations for Improving Cholesterol Measurement*; A Report from the Laboratory Standardization Panel of the National Cholesterol Education Program; NIH Publication No. 90-2964; U.S. Department of Health and Human Services, Public Health Service, National Institute of Health: Bethesda, MD, 1990.
(36) Paltiel, L.; Ronningen, K. S.; Meltzer, H. M.; Baker, S. V.; Hoppin, J. A. Evaluation of Freeze Thaw Cycles on stored plasma in the Biobank of the Norwegian Mother and Child Cohort Study. *Cell Preserv. Technol.* **2008**, *6*, 223−230.
(37) Cuhadar, S.; Koseoglu, M.; Atay, A.; Dirican, A. The effect of storage time and freeze-thaw cycles on the stability of serum samples. *Biochem. Med.* **2013**, *23*, 70−77.
(38) Zivkovic, A. M.; Wiest, M. M.; Nguyen, U. T.; Davis, R.; Watkins, S. M.; German, J. B. Effects of sample handling and storage on quantitative lipid analysis in human serum. *Metabolomics* **2009**, *5*, 507−516.
(39) Yin, P.; Peter, A.; Franken, H.; Zhao, X.; Neukamm, S. S.; Rosenbaum, L.; Lucio, M.; Zell, A.; Haring, H. U.; Xu, G.; Lehmann, R. Preanalytical aspects and sample quality assessment in metabolomics studies of human blood. *Clin. Chem.* **2013**, *59*, 833−845.
(40) Breier, M.; Wahl, S.; Prehn, C.; Fugmann, M.; Ferrari, U.; Weise, M.; Banning, F.; Seissler, J.; Grallert, H.; Adamski, J.; Lechner, A. Targeted metabolomics identifies reliable and stable metabolites in human serum and plasma samples. *PLoS One* **2014**, *9*, No. e89728.
(41) Anton, G.; Wilson, R.; Yu, Z. H.; Prehn, C.; Zukunft, S.; Adamski, J.; Heier, M.; Meisinger, C.; Romisch-Margl, W.; Wang-Sattler, R.; Hveem, K.; Wolfenbuttel, B.; Peters, A.; Kastenmuller, G.; Waldenberger, M. Pre-analytical sample quality: metabolite ratios as an intrinsic marker for prolonged room temperature exposure of serum samples. *PLoS One* **2015**, *10*, No. e0121495.
(42) Rotter, M.; Brandmaier, S.; Prehn, C.; Adam, J.; Rabstein, S.; Gawrych, K.; Bruning, T.; Illig, T.; Lickert, H.; Adamski, J.; Wang-Sattler, R. Stability of targeted metabolite profiles of urine samples under different storage conditions. *Metabolomics* **2017**, *13*, No. 4.
(43) Lounila, J.; Ala-Korpela, M.; Jokisaari, J.; Savolainen, M. J.; Kesaniemi, Y. A. Effects of orientational order and particle size on the NMR line positions of lipoproteins. *Phys. Rev. Lett.* **1994**, *72*, 4049−4052.

DOI: 10.1021/acs.jproteome.9b00343
J. Proteome Res. XXXX, XXX, XXX−XXX

Effect of repeated freeze-thaw cycles on NMR measured lipoproteins and metabolites in biofluids

Feng Wang[1, ‡], Julia Debik[1, ‡], Trygve Andreassen[2], Leslie R. Euceda[1, 3], Tonje H. Haukaas[1, 4], Claire Cannet[5], Hartmut Schäfer[5], Tone F. Bathen[1, ‡, *], Guro F. Giskeødegård[1, ‡, *]

[1] Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7491, Trondheim, Norway.

[2] MR Core Facility, Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7491, Trondheim, Norway.

[3] Camo Analytics, Oslo Science Park, Gaustadalléen 21, 0349, Oslo, Norway.

[4] SINTEF Industry, Richard Birkelands vei 3, 7034, Trondheim, Norway.

[5] Bruker Biospin GmbH, Silberstreifen, 76287 Rheinstetten, Germany.

[‡] Equal contributions

[*] Authors for correspondence

Tone F. Bathen, tone.f.bathen@ntnu.no;

Guro F. Giskeødegård, guro.giskeodegard@ntnu.no

SUPPORTING INFORMATION

Table of contents

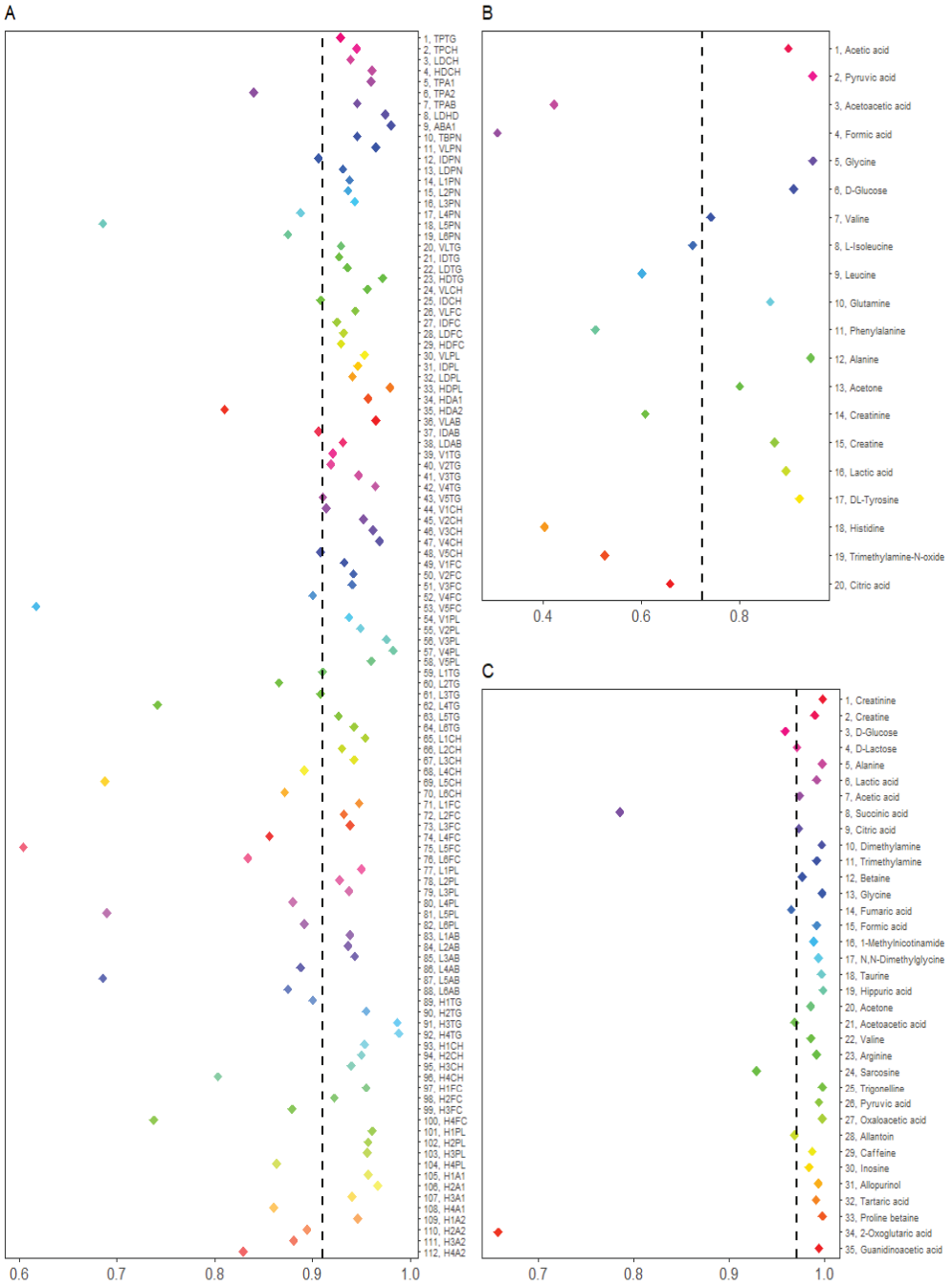


Figure S1. Intraclass correlation coefficients for lipoproteins (A), serum (B) and urine (C) metabolites. The dotted line represents the mean of ICCs across all lipoprotein parameters, all serum metabolites and all urine metabolites, respectively.

Table S1. List of measured lipoprotein parameters

Number	Lipoprotein parameter	Matrix	Measured (or cacluated) Analyte	Unit
1	TPTG	Total Serum	Triglycerides	mg/dL
2	TPCH	Total Serum	Cholesterol	mg/dL
3	LDCH	LDL	Cholesterol	mg/dL
4	HDCH	HDL	Cholesterol	mg/dL
5	TPA1	Total Serum	Apo-A1	mg/dL
6	TPA2	Total Serum	Apo-A2	mg/dL
7	TPAB	Total Serum	Apo-B	mg/dL
8	LDHD*	LDL/HDL	LDL-Chol/HDL-Chol	-/-
9	ABA1*	Apo-B/Apo-A1	Apo-B/Apo-A1	-/-
10	TBPN*	Total Serum	Particle Number	nmol/L
11	VLPN*	VLDL	Particle Number	nmol/L
12	IDPN*	IDL	Particle Number	nmol/L
13	LDPN*	LDL	Particle Number	nmol/L
14	L1PN*	LDL-1	Particle Number	nmol/L
15	L2PN*	LDL-2	Particle Number	nmol/L
16	L3PN*	LDL-3	Particle Number	nmol/L
17	L4PN*	LDL-4	Particle Number	nmol/L
18	L5PN*	LDL-5	Particle Number	nmol/L
19	L6PN*	LDL-6	Particle Number	nmol/L
20	VLTG	VLDL	Triglycerides	mg/dL
21	IDTG	IDL	Triglycerides	mg/dL
22	LDTG	LDL	Triglycerides	mg/dL
23	HDTG	HDL	Triglycerides	mg/dL
24	VLCH	VLDL	Cholesterol	mg/dL
25	IDCH	IDL	Cholesterol	mg/dL
26	VLFC	VLDL	Free Cholesterol	mg/dL
27	IDFC	IDL	Free Cholesterol	mg/dL
28	LDFC	LDL	Free Cholesterol	mg/dL
29	HDFC	HDL	Free Cholesterol	mg/dL
30	VLPL	VLDL	Phospholipids	mg/dL
31	IDPL	IDL	Phospholipids	mg/dL
32	LDPL	LDL	Phospholipids	mg/dL
33	HDPL	HDL	Phospholipids	mg/dL
34	HDA1	HDL	Apo-A1	mg/dL
35	HDA2	HDL	Apo-A2	mg/dL
36	VLAB	VLDL	Apo-B	mg/dL
37	IDAB	IDL	Apo-B	mg/dL
38	LDAB	LDL	Apo-B	mg/dL
39	V1TG	VLDL-1	Triglycerides	mg/dL
40	V2TG	VLDL-2	Triglycerides	mg/dL
41	V3TG	VLDL-3	Triglycerides	mg/dL
42	V4TG	VLDL-4	Triglycerides	mg/dL
43	V5TG	VLDL-5	Triglycerides	mg/dL

44	V1CH	VLDL-1	Cholesterol	mg/dL
45	V2CH	VLDL-2	Cholesterol	mg/dL
46	V3CH	VLDL-3	Cholesterol	mg/dL
47	V4CH	VLDL-4	Cholesterol	mg/dL
48	V5CH	VLDL-5	Cholesterol	mg/dL
49	V1FC	VLDL-1	Free Cholesterol	mg/dL
50	V2FC	VLDL-2	Free Cholesterol	mg/dL
51	V3FC	VLDL-3	Free Cholesterol	mg/dL
52	V4FC	VLDL-4	Free Cholesterol	mg/dL
53	V5FC	VLDL-5	Free Cholesterol	mg/dL
54	V1PL	VLDL-1	Phospholipids	mg/dL
55	V2PL	VLDL-2	Phospholipids	mg/dL
56	V3PL	VLDL-3	Phospholipids	mg/dL
57	V4PL	VLDL-4	Phospholipids	mg/dL
58	V5PL	VLDL-5	Phospholipids	mg/dL
59	L1TG	LDL-1	Triglycerides	mg/dL
60	L2TG	LDL-2	Triglycerides	mg/dL
61	L3TG	LDL-3	Triglycerides	mg/dL
62	L4TG	LDL-4	Triglycerides	mg/dL
63	L5TG	LDL-5	Triglycerides	mg/dL
64	L6TG	LDL-6	Triglycerides	mg/dL
65	L1CH	LDL-1	Cholesterol	mg/dL
66	L2CH	LDL-2	Cholesterol	mg/dL
67	L3CH	LDL-3	Cholesterol	mg/dL
68	L4CH	LDL-4	Cholesterol	mg/dL
69	L5CH	LDL-5	Cholesterol	mg/dL
70	L6CH	LDL-6	Cholesterol	mg/dL
71	L1FC	LDL-1	Free Cholesterol	mg/dL
72	L2FC	LDL-2	Free Cholesterol	mg/dL
73	L3FC	LDL-3	Free Cholesterol	mg/dL
74	L4FC	LDL-4	Free Cholesterol	mg/dL
75	L5FC	LDL-5	Free Cholesterol	mg/dL
76	L6FC	LDL-6	Free Cholesterol	mg/dL
77	L1PL	LDL-1	Phospholipids	mg/dL
78	L2PL	LDL-2	Phospholipids	mg/dL
79	L3PL	LDL-3	Phospholipids	mg/dL
80	L4PL	LDL-4	Phospholipids	mg/dL
81	L5PL	LDL-5	Phospholipids	mg/dL
82	L6PL	LDL-6	Phospholipids	mg/dL
83	L1AB	LDL-1	Apo-B	mg/dL
84	L2AB	LDL-2	Apo-B	mg/dL
85	L3AB	LDL-3	Apo-B	mg/dL
86	L4AB	LDL-4	Apo-B	mg/dL
87	L5AB	LDL-5	Apo-B	mg/dL
88	L6AB	LDL-6	Apo-B	mg/dL

89	H1TG	HDL-1	Triglycerides	mg/dL
90	H2TG	HDL-2	Triglycerides	mg/dL
91	H3TG	HDL-3	Triglycerides	mg/dL
92	H4TG	HDL-4	Triglycerides	mg/dL
93	H1CH	HDL-1	Cholesterol	mg/dL
94	H2CH	HDL-2	Cholesterol	mg/dL
95	H3CH	HDL-3	Cholesterol	mg/dL
96	H4CH	HDL-4	Cholesterol	mg/dL
97	H1FC	HDL-1	Free Cholesterol	mg/dL
98	H2FC	HDL-2	Free Cholesterol	mg/dL
99	H3FC	HDL-3	Free Cholesterol	mg/dL
100	H4FC	HDL-4	Free Cholesterol	mg/dL
101	H1PL	HDL-1	Phospholipids	mg/dL
102	H2PL	HDL-2	Phospholipids	mg/dL
103	H3PL	HDL-3	Phospholipids	mg/dL
104	H4PL	HDL-4	Phospholipids	mg/dL
105	H1A1	HDL-1	Apo-A1	mg/dL
106	H2A1	HDL-2	Apo-A1	mg/dL
107	H3A1	HDL-3	Apo-A1	mg/dL
108	H4A1	HDL-4	Apo-A1	mg/dL
109	H1A2	HDL-1	Apo-A2	mg/dL
110	H2A2	HDL-2	Apo-A2	mg/dL
111	H3A2	HDL-3	Apo-A2	mg/dL
112	H4A2	HDL-4	Apo-A2	mg/dL

Density ranges for lipoprotein main fractions: VLDL: 0.950-1.006 kg/L, IDL: 1.006-1.019 kg/L, LDL: 1.019-1.063 kg/L and HDL: 1.063-1.210 kg/L.

Density ranges for lipoprotein subfractions: LDL1: 1.019-1.031 kg/L, LDL2: 1.031-1.034 kg/L, LDL3: 1.034-1.037 kg/L, LDL4: 1.037-1.040 kg/L, LDL5: 1.040-1.044 kg/L, and LDL6: 1.044-1.063 kg/L. HDL1: 1.063-1.100 kg/L, HDL2: 1.100-1.112 kg/L, HDL3: 1.112-1.125 kg/L, and HDL4: 1.125-1.210 kg/L. Properties of VLDL subfractions are specified in the following reference: Lindgren FT, Jensen LL, Hatch FT (1972) The isolation and quantitative analysis of serum lipoproteins. In Nelson GJ (ed.) Blood lipids and lipoproteins: Quantitation, composition and metabolism. Wiley-Interscience, New York, p 181-274.

Numbers for lipoprotein parameters are corresponding to those used in **Figure 2**.

* : Calculated from the original ones.

Table S2. List of measured serum metabolites

Quantified	Not quantifiable
Acetic acid	Ethanol
Pyruvic acid	3-Hydroxybutyric acid
Acetoacetic acid	Glycerol
Formic acid	Glutamic acid
Glycine	Ca-EDTA*
D-Glucose	
Valine	
L-Isoleucine	
Leucine	
Glutamine	
Phenylalanine	
Alanine	
Acetone	
Creatinine	
Creatine	
Lactic acid	
DL-Tyrosine	
Histidine	
Trimethylamine-N-oxide	
Citric acid	
K-EDTA*	

*: Ca-EDTA and K-EDTA are not endogenous serum metabolites, and thus K-EDTA is not included as a quantified serum metabolite for further analysis.

Table S3. List of measured urine metabolites

Quantified	Not quantifiable
Creatinine	D-Galactose
D-Lactose	Myo-Inositol
Alanine	3-Hydroxybutyric acid
Acetic acid	Methionine
Succinic acid	1-Methylhistidine
Citric acid	Benzoic acid
Dimethylamine	4-Aminobutyric acid
Trimethylamine	D-Mannitol
Betaine	D-Mannose
Glycine	Adenosine
Fumaric acid	Imidazole
Formic acid	D-Mandelic acid
1-Methylnicotinamide	2-Furoylglycine
N,N-Dimethylglycine	2-Methylsuccinic acid
Hippuric acid	1-Methyladenosine
Valine	
Arginine	
Trigonelline	
Pyruvic acid	
Oxaloacetic acid	
Allantoin	
Caffeine	
Inosine	
Allopurinol	
Tartaric acid	
Proline betaine	
Guanidinoacetic acid	
D-Glucose	
Lactic acid	
Taurine	
Acetone	
Acetoacetic acid	
Creatine	
Sarcosine	
2-Oxoglutaric acid	

Table S4. Effects of FTCs on lipoprotein parameter concentrations

Lipoprotein parameters	Consecutive FTCs								Compared to FTC1							
	Adjusted p-values of Wilcox tests				Median percentage change*				Adjusted p-values of Wilcox tests				Median percentage change*			
	FTC1-2	FTC2-3	FTC3-4	FTC4-5	FTC1-2	FTC2-3	FTC3-4	FTC4-5	FTC1-2	FTC1-3	FTC1-4	FTC1-5	FTC1-2	FTC1-3	FTC1-4	FTC1-5
TPTG [mg/dL]	0,731	0,324	0,440	0,955	0,42	-0,73	0,57	-0,15	0,731	0,364	0,583	0,792	0,42	-0,28	0,06	0,04
TPCH [mg/dL]	0,977	0,311	0,291	0,955	0,83	-0,41	1,13	-0,28	0,977	0,435	0,871	0,878	0,83	-0,56	-0,33	0,41
LDCH [mg/dL]	0,093	0,230	0,617	0,990	-0,84	-1,12	-0,24	-0,77	0,093	0,014	0,003	0,029	-0,84	-3,31	-5,48	-3,38
HDCH [mg/dL]	0,896	0,311	0,357	0,990	-0,42	-0,64	1,51	-0,28	0,896	0,132	0,485	0,633	-0,42	-1,34	-0,51	-0,22
TPA1 [mg/dL]	0,896	0,311	0,303	0,955	-0,44	-0,72	0,78	-0,14	0,896	0,382	0,721	0,832	-0,44	-0,87	-0,54	0,36
TPA2 [mg/dL]	0,541	0,820	0,202	0,955	-1,40	0,21	2,28	-1,00	0,541	0,364	0,796	0,863	-1,40	-1,16	-0,86	0,84
TPAB [mg/dL]	0,533	0,747	0,634	0,955	-0,29	-0,67	0,42	-0,50	0,533	0,143	0,174	0,189	-0,29	-1,57	-1,68	-1,76
LDHD [-/-]	0,134	0,311	0,065	0,955	-1,35	-0,86	-1,69	-0,17	0,134	0,081	0,004	0,086	-1,35	-1,97	-3,35	-3,34
ABA1 [-/-]	0,167	0,747	0,185	0,955	0,00	0,00	-0,78	0,00	0,167	0,276	0,016	0,088	0,00	0,00	-1,58	-1,44
TBPN [nmol/L]	0,533	0,747	0,634	0,955	-0,29	-0,67	0,42	-0,51	0,533	0,143	0,174	0,176	-0,29	-1,56	-1,68	-1,76
VLPN [nmol/L]	0,237	0,132	0,071	0,662	1,44	1,59	1,62	1,19	0,237	0,015	0,004	0,022	1,44	2,35	4,32	4,43
IDPN [nmol/L]	0,177	0,983	0,071	0,662	5,53	1,21	8,20	5,96	0,177	0,112	0,004	0,010	5,53	6,11	11,44	15,69
LDPN [nmol/L]	0,345	0,740	0,905	0,955	-0,64	-0,62	0,45	-1,20	0,345	0,072	0,066	0,176	-0,64	-2,20	-2,69	-2,50
L1PN [nmol/L]	0,093	1,000	0,308	0,955	4,17	0,01	1,11	0,83	0,093	0,013	0,001	0,009	4,17	4,32	5,47	7,84
L2PN [nmol/L]	0,533	0,827	0,431	0,955	1,75	-1,34	-0,96	3,81	0,533	0,933	0,779	0,832	1,75	-0,75	-1,01	0,84
L3PN [nmol/L]	0,942	0,132	0,585	0,955	-0,15	-2,18	-2,44	0,68	0,942	0,686	0,382	0,751	-0,15	-0,06	-0,72	0,72
L4PN [nmol/L]	0,093	0,132	0,774	0,955	-8,57	-4,82	-1,76	-2,62	0,093	0,013	0,004	0,129	-8,57	-8,68	-12,27	-6,62
L5PN [nmol/L]	0,093	0,820	0,717	0,955	-10,30	-1,59	-0,29	-4,31	0,093	0,014	0,003	0,053	-10,30	-8,91	-14,58	-16,12
L6PN [nmol/L]	0,345	0,747	0,569	0,955	-4,07	1,13	4,77	0,84	0,345	0,288	0,445	0,194	-4,07	-4,70	-4,93	-9,82
VLTG [mg/dL]	0,499	0,311	0,717	0,955	-0,95	-1,73	1,16	-1,51	0,499	0,012	0,006	0,043	-0,95	-3,75	-3,36	-5,36
IDTG [mg/dL]	0,134	0,196	0,617	0,955	-1,50	-7,63	-0,61	-3,22	0,134	0,014	0,004	0,016	-1,50	-7,64	-8,54	-12,76
LDTG [mg/dL]	0,535	0,700	0,963	0,990	0,60	-0,52	-0,07	-0,11	0,535	0,967	0,792	0,747	0,60	-0,41	0,50	2,98
HDTG [mg/dL]	0,533	0,867	0,869	0,955	0,47	0,61	0,61	-0,21	0,533	0,364	0,583	0,378	0,47	0,52	2,64	2,03

VLCH [mg/dL]	0,093	0,700	0,015	0,662	6,93	0,66	3,91	3,29	0,093	0,017	0,003	0,020	6,93	8,79	11,96	17,94
IDCH [mg/dL]	0,285	0,912	0,119	0,922	6,02	0,00	12,29	4,16	0,285	0,214	0,015	0,009	6,02	7,67	18,83	20,49
VLFC [mg/dL]	0,438	0,955	0,015	0,662	1,78	-0,25	2,93	1,28	0,438	0,522	0,090	0,129	1,78	2,65	5,56	6,78
IDFC [mg/dL]	0,134	0,834	0,185	0,943	7,94	0,00	8,69	5,38	0,134	0,107	0,009	0,011	7,94	4,80	14,52	21,77
LDFC [mg/dL]	0,345	0,176	0,951	0,955	-1,16	-1,92	-0,76	-0,31	0,345	0,045	0,014	0,230	-1,16	-2,17	-2,82	-2,41
HDFC [mg/dL]	0,787	0,311	0,303	0,990	1,67	-2,50	2,23	-0,53	0,787	0,329	0,505	0,832	1,67	-0,53	-0,81	1,07
VLPL [mg/dL]	1,000	0,584	0,273	0,955	0,21	-0,81	1,94	1,08	1,000	0,364	0,894	0,863	0,21	-1,26	1,75	0,46
IDPL [mg/dL]	0,859	0,132	0,440	0,955	0,05	-3,61	2,09	-1,83	0,859	0,045	0,164	0,218	0,05	-5,72	-3,22	-2,98
LDPL [mg/dL]	0,093	0,132	0,590	0,990	-0,86	-0,93	-0,62	-0,95	0,093	0,012	0,004	0,021	-0,86	-3,23	-3,89	-2,49
HDPL [mg/dL]	0,448	0,311	0,313	0,955	0,82	-0,96	0,90	0,03	0,448	0,655	0,295	0,194	0,82	-0,32	1,05	0,75
HDA1 [mg/dL]	0,723	0,311	0,440	0,955	-0,29	-0,81	0,62	-0,32	0,723	0,153	0,404	0,856	-0,29	-1,42	-1,12	-0,04
HDA2 [mg/dL]	0,535	0,747	0,185	0,955	-1,17	0,35	2,20	-1,21	0,535	0,435	0,908	0,832	-1,17	-1,07	-1,15	0,58
VLAB [mg/dL]	0,237	0,132	0,071	0,662	1,41	1,64	1,61	1,22	0,237	0,018	0,005	0,025	1,41	2,38	4,30	4,32
IDAB [mg/dL]	0,170	0,983	0,082	0,662	5,92	0,95	8,29	5,89	0,170	0,112	0,006	0,011	5,92	5,98	11,36	15,58
LDAB [mg/dL]	0,345	0,740	0,905	0,955	-0,64	-0,63	0,45	-1,20	0,345	0,072	0,066	0,176	-0,64	-2,21	-2,69	-2,50
V1TG [mg/dL]	0,137	0,311	0,905	0,776	-5,27	-3,59	-2,37	-3,92	0,137	0,001	0,015	0,000	-5,27	-7,30	-9,52	-11,39
V2TG [mg/dL]	0,535	0,324	0,590	0,933	-0,82	-4,13	-1,60	-2,79	0,535	0,252	0,145	0,176	-0,82	-1,80	-4,13	-4,50
V3TG [mg/dL]	0,631	0,834	0,964	0,955	2,03	-1,13	0,59	-0,75	0,631	0,999	1,000	0,863	2,03	2,41	1,40	4,57
V4TG [mg/dL]	0,476	0,761	0,617	0,955	1,96	-0,96	0,15	0,81	0,476	0,859	0,340	0,529	1,96	1,48	2,13	3,18
V5TG [mg/dL]	0,531	0,985	0,452	0,955	2,56	-0,12	0,58	-0,40	0,531	0,872	0,617	0,318	2,56	0,60	0,77	2,56
V1CH [mg/dL]	0,161	0,615	0,003	0,922	2,96	1,09	4,08	3,09	0,161	0,069	0,010	0,075	2,96	5,55	13,56	12,77
V2CH [mg/dL]	0,170	0,747	0,065	0,955	12,92	2,20	8,62	3,59	0,170	0,097	0,014	0,061	12,92	12,22	23,47	23,53
V3CH [mg/dL]	0,093	0,747	0,071	0,955	9,70	0,00	3,20	0,91	0,093	0,015	0,006	0,055	9,70	7,84	20,73	24,93
V4CH [mg/dL]	0,093	0,840	0,031	0,662	4,18	0,18	6,88	4,96	0,093	0,013	0,003	0,009	4,18	6,70	14,01	21,52
V5CH [mg/dL]	0,610	0,740	0,523	0,955	1,88	0,84	-2,03	-0,04	0,610	0,132	0,560	0,387	1,88	2,16	2,08	3,48
V1FC [mg/dL]	0,631	0,179	0,653	0,955	-0,22	-4,53	0,00	-0,55	0,631	0,053	0,066	0,025	-0,22	-4,26	-1,64	-2,91
V2FC [mg/dL]	0,134	0,311	0,062	0,662	7,00	5,23	5,64	5,32	0,134	0,030	0,006	0,033	7,00	17,81	30,83	32,02
V3FC [mg/dL]	0,120	0,311	0,038	0,662	8,11	0,58	3,55	1,79	0,120	0,013	0,004	0,025	8,11	10,10	19,73	25,24

V4FC [mg/dL]	0,093	0,700	0,065	0,706	9,00	0,00	8,90	4,01	0,093	0,028	0,004	0,009	9,00	12,09	24,93	39,05
V5FC [mg/dL]	0,304	0,176	0,105	0,706	2,85	7,45	10,04	7,06	0,304	0,090	0,015	0,011	2,85	6,50	22,28	24,74
V1PL [mg/dL]	0,237	0,345	0,634	0,662	-3,35	-4,27	-2,44	-4,43	0,237	0,042	0,014	0,009	-3,35	-4,02	-6,80	-13,71
V2PL [mg/dL]	1,000	0,420	0,894	0,662	2,19	-2,64	-0,44	-2,75	1,000	0,591	0,779	0,523	2,19	-0,77	-0,22	1,57
V3PL [mg/dL]	0,112	1,000	0,273	0,955	3,84	0,00	2,54	-0,15	0,112	0,110	0,008	0,043	3,84	3,30	6,80	6,04
V4PL [mg/dL]	0,162	0,820	0,071	0,922	3,45	-1,25	1,93	3,29	0,162	0,112	0,013	0,056	3,45	3,77	5,51	7,16
V5PL [mg/dL]	0,295	0,596	0,734	0,969	1,76	0,79	-0,86	0,00	0,295	0,070	0,066	0,102	1,76	1,67	3,70	1,64
L1TG [mg/dL]	0,093	0,615	0,617	0,955	3,62	1,16	-0,32	1,31	0,093	0,013	0,004	0,017	3,62	3,92	5,60	7,69
L2TG [mg/dL]	0,120	0,910	0,979	0,955	7,14	0,65	-1,46	0,00	0,120	0,108	0,019	0,078	7,14	3,00	4,76	9,38
L3TG [mg/dL]	0,170	0,311	0,717	0,955	-0,52	-0,56	0,00	-0,78	0,170	0,030	0,009	0,041	-0,52	-1,41	-1,45	-1,91
L4TG [mg/dL]	0,457	0,492	0,685	0,990	-5,08	-3,44	-1,25	-4,50	0,457	0,112	0,066	0,482	-5,08	-7,15	-12,58	-4,66
L5TG [mg/dL]	0,735	0,747	0,774	0,955	5,33	-3,38	0,00	-0,39	0,735	0,706	0,478	0,789	5,33	0,00	0,00	1,21
L6TG [mg/dL]	0,345	0,834	0,774	0,955	-2,23	0,28	0,72	-1,91	0,345	0,252	0,340	0,088	-2,23	-3,73	-3,09	-7,14
L1CH [mg/dL]	0,093	0,910	0,291	0,955	4,86	-0,82	1,78	1,06	0,093	0,018	0,001	0,011	4,86	4,97	5,77	7,15
L2CH [mg/dL]	0,535	0,634	0,357	0,955	1,21	-1,70	-0,90	2,92	0,535	0,933	0,634	0,878	1,21	-0,60	-1,33	0,18
L3CH [mg/dL]	0,956	0,132	0,542	0,955	-0,69	-2,13	-1,84	0,45	0,956	0,476	0,164	0,218	-0,69	-1,59	-1,95	-2,14
L4CH [mg/dL]	0,093	0,132	0,532	0,955	-10,15	-4,97	-0,52	-2,60	0,093	0,011	0,004	0,053	-10,15	-9,88	-12,88	-10,87
L5CH [mg/dL]	0,093	0,747	0,634	0,955	-12,07	-0,84	-4,81	-3,13	0,093	0,012	0,001	0,053	-12,07	-11,32	-15,10	-17,90
L6CH [mg/dL]	0,295	0,906	0,590	0,955	-5,89	1,17	6,38	0,56	0,295	0,131	0,346	0,124	-5,89	-5,40	-7,18	-11,30
L1FC [mg/dL]	0,093	0,834	0,403	0,955	4,79	-0,70	2,03	1,25	0,093	0,017	0,004	0,011	4,79	3,79	6,52	8,47
L2FC [mg/dL]	0,631	0,584	0,928	0,955	0,42	-1,69	0,15	0,45	0,631	0,469	0,511	0,751	0,42	-1,78	-0,71	1,69
L3FC [mg/dL]	0,731	0,311	0,979	0,955	-0,60	-2,80	0,28	0,90	0,731	0,159	0,064	0,474	-0,60	-2,77	-3,33	-1,41
L4FC [mg/dL]	0,112	0,311	0,590	0,955	-5,89	-3,75	-2,29	-2,83	0,112	0,013	0,004	0,093	-5,89	-6,19	-10,49	-9,32
L5FC [mg/dL]	0,126	0,752	0,738	0,955	-6,69	-1,73	2,21	-3,33	0,126	0,017	0,009	0,090	-6,69	-7,66	-10,82	-14,83
L6FC [mg/dL]	0,285	0,521	0,590	0,955	-7,60	-2,93	-2,72	-2,84	0,285	0,014	0,009	0,053	-7,60	-6,14	-9,09	-10,72
L1PL [mg/dL]	0,093	0,958	0,303	0,955	4,01	-0,24	1,06	0,75	0,093	0,036	0,001	0,020	4,01	4,52	5,34	6,33
L2PL [mg/dL]	0,535	0,615	0,373	0,955	1,90	-1,34	-0,90	2,84	0,535	1,000	0,721	1,000	1,90	-0,32	-0,77	0,88
L3PL [mg/dL]	0,956	0,132	0,494	0,955	-0,52	-1,96	-1,98	0,39	0,956	0,602	0,346	0,308	-0,52	-0,72	-1,32	-0,18

L4PL [mg/dL]	0,093	0,132	0,483	0,955	-9,57	-3,53	-1,75	-2,45	0,093	0,003	0,002	0,049	-9,57	-10,72	-12,35	-14,88
L5PL [mg/dL]	0,093	0,747	0,590	0,955	-8,96	-1,32	-0,23	-2,78	0,093	0,011	0,001	0,053	-8,96	-8,98	-15,28	-14,67
L6PL [mg/dL]	0,304	0,955	0,634	0,955	-4,45	0,50	3,76	-0,04	0,304	0,143	0,320	0,088	-4,45	-3,71	-5,86	-9,73
L1AB [mg/dL]	0,093	1,000	0,308	0,955	4,17	-0,04	1,14	0,84	0,093	0,013	0,003	0,009	4,17	4,34	5,49	7,84
L2AB [mg/dL]	0,533	0,823	0,431	0,955	1,74	-1,34	-0,95	3,80	0,533	0,933	0,761	0,832	1,74	-0,76	-1,03	0,85
L3AB [mg/dL]	0,940	0,149	0,585	0,955	-0,17	-2,13	-2,45	0,71	0,940	0,665	0,404	0,751	-0,17	-0,08	-0,71	0,73
L4AB [mg/dL]	0,093	0,132	0,774	0,955	-8,58	-4,79	-1,71	-2,54	0,093	0,014	0,004	0,129	-8,58	-8,66	-12,27	-6,61
L5AB [mg/dL]	0,093	0,820	0,717	0,955	-10,27	-1,63	-0,16	-4,32	0,093	0,013	0,003	0,053	-10,27	-8,96	-14,56	-16,22
L6AB [mg/dL]	0,345	0,747	0,569	0,955	-4,06	1,15	4,79	0,84	0,345	0,276	0,445	0,194	-4,06	-4,69	-4,96	-9,83
H1TG [mg/dL]	0,254	0,747	0,590	0,955	3,65	5,41	2,02	1,55	0,254	0,131	0,066	0,176	3,65	4,38	6,96	6,81
H2TG [mg/dL]	0,373	0,955	0,440	0,955	3,20	-0,97	-1,75	-0,15	0,373	0,706	0,758	0,509	3,20	0,00	1,29	2,73
H3TG [mg/dL]	0,428	0,521	0,759	0,955	0,20	1,66	0,00	-0,65	0,428	0,339	0,242	0,328	0,20	0,28	1,17	-0,38
H4TG [mg/dL]	0,170	0,311	0,202	0,955	-1,01	1,26	-0,92	-0,14	0,170	0,686	0,404	0,175	-1,01	-0,39	-0,83	-1,36
H1CH [mg/dL]	0,170	0,311	0,569	0,955	2,42	-0,79	0,49	0,74	0,170	0,602	0,164	0,194	2,42	1,77	1,69	2,54
H2CH [mg/dL]	0,535	0,634	0,303	0,990	1,08	-1,01	1,42	0,65	0,535	0,758	0,518	0,516	1,08	-0,90	0,06	1,53
H3CH [mg/dL]	0,856	1,000	0,253	0,955	0,47	0,29	1,60	0,08	0,856	0,933	0,589	0,489	0,47	0,27	0,56	-0,04
H4CH [mg/dL]	0,093	0,834	0,905	0,955	-2,98	0,10	-0,01	-0,75	0,093	0,013	0,008	0,043	-2,98	-3,00	-3,82	-4,81
H1FC [mg/dL]	0,631	0,230	0,313	0,955	0,97	-3,34	2,21	-1,64	0,631	0,418	0,896	0,832	0,97	-0,69	-0,28	1,61
H2FC [mg/dL]	0,523	0,747	0,331	0,955	2,76	-1,71	3,90	0,85	0,523	0,956	0,295	0,093	2,76	-0,47	2,65	4,36
H3FC [mg/dL]	0,762	0,560	0,743	0,955	-0,07	-2,85	0,25	-0,19	0,762	0,382	0,404	0,751	-0,07	-1,62	-2,72	1,10
H4FC [mg/dL]	0,428	0,747	0,979	0,969	-1,77	-0,31	-0,16	-2,34	0,428	0,128	0,135	0,269	-1,77	-1,59	-2,84	-2,00
H1PL [mg/dL]	0,093	0,311	0,403	0,955	4,15	-1,35	0,80	1,19	0,093	0,105	0,007	0,057	4,15	3,42	2,75	4,98
H2PL [mg/dL]	0,123	0,761	0,308	0,990	2,66	-0,53	0,72	0,76	0,123	0,118	0,059	0,043	2,66	1,59	2,53	4,86
H3PL [mg/dL]	0,177	0,984	0,303	0,955	1,34	-0,82	1,08	0,68	0,177	0,252	0,054	0,053	1,34	1,04	2,30	2,07
H4PL [mg/dL]	0,093	0,607	0,990	0,955	-1,37	-0,48	-0,21	-0,69	0,093	0,017	0,033	0,088	-1,37	-1,64	-1,55	-2,61
H1A1 [mg/dL]	0,237	0,311	0,313	0,955	2,40	-2,02	2,02	1,39	0,237	0,597	0,230	0,124	2,40	-1,21	1,92	4,02
H2A1 [mg/dL]	0,533	0,747	0,273	0,955	0,99	0,16	0,74	0,03	0,533	0,602	0,174	0,176	0,99	0,57	1,64	1,33
H3A1 [mg/dL]	0,648	0,955	0,237	0,955	1,00	-0,31	1,59	0,32	0,648	0,721	0,174	0,172	1,00	0,36	1,38	0,83

H4A1 [mg/dL]	0,123	0,747	0,905	0,955	-1,70	0,10	0,11	-0,54	0,123	0,012	0,010	0,053	-1,70	-2,32	-2,77	-3,52
H1A2 [mg/dL]	0,112	0,958	0,136	0,955	6,16	-0,47	6,43	1,44	0,112	0,090	0,010	0,011	6,16	4,48	7,90	13,95
H2A2 [mg/dL]	0,145	0,837	0,131	0,955	5,05	-0,52	3,58	0,27	0,145	0,107	0,011	0,020	5,05	4,99	9,60	10,77
H3A2 [mg/dL]	0,467	0,955	0,185	0,955	0,92	-0,09	2,39	0,09	0,467	0,660	0,066	0,104	0,92	0,12	2,64	1,98
H4A2 [mg/dL]	0,156	0,834	0,348	0,955	-1,88	0,70	1,55	-0,71	0,156	0,129	0,202	0,150	-1,88	-2,88	-2,05	-4,52

*: An example for calculating median percentage change: median percentage change of FTC1-2 = median ((FTC2-FTC1)/FTC1*100)

Significant accumulated increase at FTC5 compared to FTC1

Significant accumulated decrease at FTC5 compared to FTC1

Table S5. Effects of FTCs on serum metabolite concentrations

Serum metabolites	Consecutive FTCs								Compared to FTC1							
	Adjusted p-values of Wilcox tests				Median percentage change*				Adjusted p-values of Wilcox tests				Median percentage change*			
	FTC1-2	FTC2-3	FTC3-4	FTC4-5	FTC1-2	FTC2-3	FTC3-4	FTC4-5	FTC1-2	FTC1-3	FTC1-4	FTC1-5	FTC1-2	FTC1-3	FTC1-4	FTC1-5
Acetic acid	0,990	0,875	0,783	0,814	11,88	0,75	2,29	2,76	0,990	0,315	0,118	0,010	11,88	6,65	15,66	29,06
Pyruvic acid	0,990	0,875	0,783	0,753	-0,57	-1,62	1,13	2,49	0,990	0,945	0,840	0,753	-0,57	-3,15	1,23	3,99
Acetoacetic acid	0,990	0,875	0,990	0,753	8,52	-2,27	0,80	6,22	0,990	0,945	0,668	0,604	8,52	7,92	1,74	5,33
Formic acid	0,990	0,875	0,807	0,977	-6,21	6,23	0,76	-1,99	0,990	0,990	0,728	0,945	-6,21	-5,80	3,21	4,74
Glycine	0,990	0,921	0,807	0,814	4,78	-1,96	-1,98	2,03	0,990	0,945	0,705	0,604	4,78	1,02	1,31	1,82
D-Glucose	0,990	0,977	0,783	0,852	-0,48	0,18	0,91	-0,16	0,990	0,945	0,840	0,945	-0,48	-0,64	0,14	0,00
Valine	0,990	0,962	0,693	0,977	1,42	-5,43	4,58	-0,88	0,990	0,990	0,668	0,693	1,42	2,47	2,90	1,01
L-Isoleucine	0,990	0,875	0,807	0,753	7,16	3,25	-2,34	-3,63	0,990	0,315	0,668	0,945	7,16	11,90	5,69	-3,83
Leucine	0,990	0,875	0,807	0,735	-1,10	5,43	-2,88	-13,30	0,990	0,958	0,990	0,532	-1,10	2,10	-0,74	-7,70
Glutamine	0,990	0,875	0,807	0,753	0,14	-1,61	-0,08	-1,78	0,990	0,958	0,668	0,945	0,14	1,11	2,06	-1,32
Phenylalanine	0,990	0,875	0,807	0,753	-2,54	3,78	3,87	8,39	0,990	0,958	0,840	0,604	-2,54	-2,90	4,34	5,03
Alanine	0,990	0,962	0,783	0,814	0,07	0,45	1,71	1,57	0,990	0,958	0,668	0,693	0,07	-0,90	1,83	1,21
Acetone	0,990	1,000	0,783	0,103	-7,32	-3,20	-9,26	24,27	0,990	0,958	0,668	0,945	-7,32	-10,08	-18,76	-7,90
Creatinine	0,990	0,875	0,788	0,753	4,91	-3,15	2,43	-1,41	0,990	0,980	0,705	0,945	4,91	1,83	3,95	-4,04
Creatine	0,990	0,875	0,783	0,735	-1,13	5,52	-6,32	10,07	0,990	0,945	0,705	0,604	-1,13	6,86	-6,61	17,10
Lactic acid	0,990	0,962	0,990	0,735	0,20	-0,66	0,43	1,63	0,990	0,945	0,668	0,604	0,20	1,71	1,90	2,07
DL-Tyrosine	0,990	0,962	0,783	0,990	2,69	-1,47	1,13	3,17	0,990	0,945	0,557	0,315	2,69	3,90	3,11	5,87
Histidine	0,990	0,875	0,783	0,753	2,76	-2,85	4,01	-2,29	0,990	0,990	0,840	0,945	2,76	2,69	3,61	2,62
Trimethylamine-N-oxide	0,990	0,875	0,783	0,753	1,60	65,74	-12,14	33,32	0,990	0,945	0,977	0,921	1,60	40,40	-16,81	58,65
Citric acid	0,990	0,962	0,807	0,814	-7,65	0,36	-7,13	5,34	0,990	0,980	0,668	0,990	-7,65	-11,72	-11,17	3,72

*: An example for calculating median percentage change: median percentage change of FTC1-2 = median ((FTC2-FTC1)/FTC1*100)

Table S6. Effects of FTCs on urine metabolite concentrations

Urine Metabolites	Consecutive FTCs								Compared to FTC1							
	Adjusted p-values of Wilcox tests				Median percentage change*				Adjusted p-values of Wilcox tests				Median percentage change*			
	FTC1-2	FTC2-3	FTC3-4	FTC4-5	FTC1-2	FTC2-3	FTC3-4	FTC4-5	FTC1-2	FTC1-3	FTC1-4	FTC1-5	FTC1-2	FTC1-3	FTC1-4	FTC1-5
Creatinine	0,919	0,506	1,000	0,844	0,20	0,41	0,32	-0,08	0,919	0,792	0,986	0,758	0,20	0,42	0,27	-0,32
Creatine	0,919	0,902	1,000	1,000	-0,09	1,16	-2,45	-29,24	0,919	0,834	0,986	0,862	-0,09	-9,01	-1,67	0,49
D-Glucose	0,919	0,902	1,000	0,844	3,85	1,53	3,49	-2,90	0,919	0,792	0,464	0,754	3,85	9,53	6,74	4,77
D-Lactose	0,829	0,957	1,000	1,000	-1,42	0,49	-2,60	0,38	0,829	0,804	0,986	0,758	-1,42	1,15	0,13	-4,10
Alanine	0,919	0,506	1,000	1,000	0,44	0,74	-0,06	-0,03	0,919	0,420	0,986	0,525	0,44	1,76	0,73	0,78
Lactic acid	0,541	0,902	1,000	1,000	-2,94	3,28	-1,28	-0,70	0,541	0,834	0,986	0,758	-2,94	0,18	-1,00	-1,03
Acetic acid	0,541	0,525	1,000	1,000	3,54	3,95	-0,32	-1,41	0,541	0,245	0,464	0,263	3,54	4,60	4,74	4,78
Succinic acid	0,541	0,902	1,000	1,000	6,20	1,07	0,77	0,02	0,541	0,792	0,986	0,758	6,20	5,85	1,90	2,71
Citric acid	0,583	0,420	1,000	0,550	-0,17	1,52	0,84	-1,36	0,583	0,804	0,986	0,525	-0,17	0,02	0,55	-0,54
Dimethylamine	0,919	0,420	1,000	1,000	0,53	1,02	-0,94	-0,73	0,919	0,792	0,986	0,758	0,53	0,80	0,08	-0,34
Trimethylamine	0,829	0,957	1,000	1,000	1,10	-0,18	-1,66	0,06	0,829	0,804	0,986	0,758	1,10	2,48	1,24	1,16
Betaine	0,829	0,957	1,000	0,805	0,42	0,65	0,81	-1,65	0,829	0,792	0,986	0,926	0,42	2,65	1,33	0,11
Glycine	0,919	0,506	1,000	1,000	0,56	0,96	-1,09	0,53	0,919	0,804	0,986	0,926	0,56	-0,02	0,10	-0,50
Fumaric acid	0,541	0,817	1,000	1,000	-4,69	2,95	-1,58	-1,18	0,541	0,834	0,679	0,525	-4,69	0,08	-5,52	-3,34
Formic acid	0,919	0,420	1,000	0,844	-0,18	2,41	1,28	-2,73	0,919	0,245	0,464	0,754	-0,18	3,53	4,24	2,58
1-Methylnicotinamide	0,829	0,817	1,000	1,000	-1,50	2,17	-0,01	-0,06	0,829	0,804	0,986	0,869	-1,50	2,34	2,71	-0,40
N,N-Dimethylglycine	0,829	0,902	1,000	1,000	-0,21	1,61	0,15	0,40	0,829	0,834	0,986	0,881	-0,21	1,06	2,53	0,45
Taurine	0,919	0,913	1,000	0,778	0,62	-0,20	0,60	-1,61	0,919	0,804	0,986	0,525	0,62	-0,53	-0,10	-1,75
Hippuric acid	0,541	0,917	1,000	1,000	1,08	-0,78	-0,41	0,41	0,541	0,792	0,986	0,758	1,08	0,59	0,09	-0,68

Acetone	0,923	0,902	1,000	0,550	0,55	-0,85	-2,07	2,55	0,923	0,834	0,986	1,000	0,55	-0,99	-1,79	-0,01
Acetoacetic acid	0,541	0,902	1,000	1,000	-1,58	0,59	-0,46	0,37	0,541	0,792	0,986	0,881	-1,58	2,96	-0,48	-0,34
Valine	0,541	0,902	1,000	0,817	1,70	-1,61	-1,00	2,51	0,541	0,792	0,988	0,840	1,70	1,17	-1,09	-0,22
Arginine	0,829	0,506	1,000	1,000	0,34	1,24	0,29	-0,45	0,829	0,792	0,464	0,754	0,34	1,49	2,21	2,25
Sarcosine	0,919	0,957	1,000	0,550	0,22	0,55	7,27	-5,32	0,919	0,792	0,986	0,525	0,22	0,00	4,74	-2,98
Trigonelline	0,541	0,902	1,000	1,000	2,49	-1,39	-0,17	0,78	0,541	0,792	0,986	0,754	2,49	0,44	0,38	2,35
Pyruvic acid	0,919	0,420	1,000	0,550	-0,48	2,26	-2,91	1,40	0,919	0,804	0,986	0,758	-0,48	1,65	0,14	0,20
Oxaloacetic acid	0,541	0,817	1,000	1,000	1,65	-0,40	0,26	0,02	0,541	0,930	0,986	0,881	1,65	0,21	0,12	0,86
Allantoin	0,541	0,902	1,000	0,835	7,21	1,10	-1,78	3,57	0,541	0,245	0,986	0,525	7,21	7,72	-0,98	6,18
Caffeine	0,919	0,957	1,000	0,817	-0,03	0,15	0,29	-0,75	0,919	0,834	0,986	0,758	-0,03	0,31	0,75	-0,64
Inosine	0,541	0,817	1,000	0,550	-3,79	2,76	0,32	-0,74	0,541	0,792	0,986	0,754	-3,79	-1,61	-2,73	-1,03
Allopurinol	0,829	0,902	1,000	1,000	-0,81	0,47	0,98	-0,13	0,829	0,804	0,986	0,758	-0,81	-0,22	0,44	1,90
Tartaric acid	1,000	0,907	1,000	0,844	-3,27	-1,97	0,30	-3,07	1,000	0,792	0,986	0,263	-3,27	-1,51	-2,47	-5,14
Proline betaine	0,919	0,506	1,000	0,778	0,26	0,99	-0,25	-1,49	0,919	0,792	0,986	0,886	0,26	3,66	0,25	-0,06
2-Oxoglutaric acid	0,919	0,990	1,000	0,550	-1,12	-1,37	-2,22	6,25	0,919	0,834	0,986	0,881	-1,12	0,03	-9,98	-0,73
Guanidinoacetic acid	1,000	0,817	0,665	0,550	0,38	0,80	-1,24	0,94	1,000	0,792	1,000	0,758	0,38	0,88	0,07	0,24

*: An example for calculating median percentage change: median percentage change of FTC1-2 = median((FTC2-FTC1)/FTC1*100)

Paper III

Paper 3: Debik, Julia Barbara; Schäfer, Hartmut; Andreassen, Trygve; Wang, Feng; Fang, Fang; Cannet, Claire; Spraul, Manfred; Bathen, Tone Frost; Giskeødegård, Guro F.. Serum metabolic profiling for assessment of breast cancer risk in women participating in the HUNT2 study.

This article is awaiting publication and is therefore not included.

ISBN 978-82-326-6712-3 (printed ver.)
ISBN 978-82-326-6662-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology