Liquid chromatography – Orbitrap mass spectrometry is a useful tool in untargeted metabolomics analysis of dried blood spots in clinical chemistry

Elise Mørk Sandås

Thesis for the Master's Degree in Chemistry

60 credits

Department of Chemistry
Faculty of Mathematics and Natural Sciences
UNIVERSITY OF OSLO

August 2018

Liquid chromatography – Orbitrap mass spectrometry is a useful tool in untargeted metabolomics analysis of dried blood spots in clinical chemistry

Elise Mørk Sandås

Thesis for the Master's Degree in Chemistry

60 credits

Department of Chemistry

Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

August 29th 2018

2018

Liquid chromatography – Orbitrap mass spectrometry is a useful tool in untargeted metabolomics analysis of dried blood spots in clinical chemistry

Elise Mørk Sandås

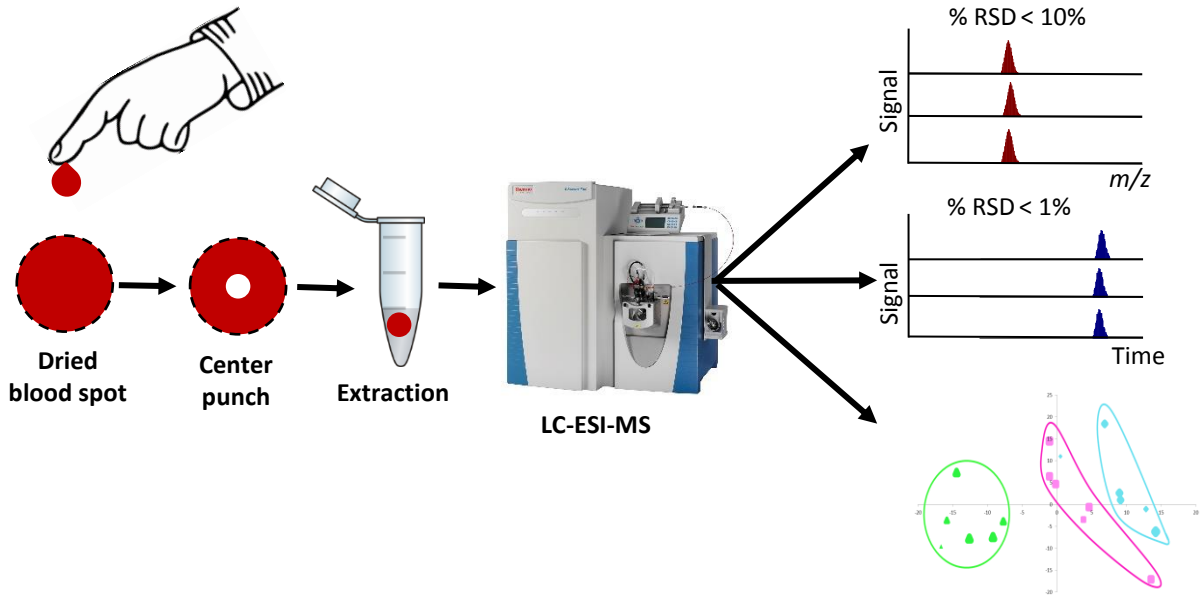
http://www.duo.uio.no/

Printed at Reprosentralen, Universitetet i Oslo

IV

Abstract

There is a need for comprehensive analysis of the dried blood spot (DBS) metabolome to study both new and known inborn errors of metabolism (IEM). The purpose of this study was to complete and evaluate an untargeted metabolomics method using liquid chromatography – electrospray ionization – Q Exactive Orbitrap mass spectrometry (MS), for analysis of one punch (3.2 mm diameter corresponding to about 3 µL whole blood) of a DBS. The criteria for evaluation and measurements were inspired by stringency applied by the World Anti-Doping Agency. In this regard, the instrument repeatability and assay reproducibility were satisfactory with relative standard deviation (% RSD) of peak areas <10 % and retention times <1 %, using a pooled control sample, injected three times each day during sample analysis for 11 days. Compared to perimeter punches, analyzing center punches improved the repeatability. An increased organic solvent amount in the reconstitution solution, and reduced *m/z* range of the MS (focusing more on each *m/z* increases sensitivity) increased the number of hydrophobic and low-abundant compounds detected, respectively. The quantification of most of the method evaluation compounds (acylcarnitines and amino acids) was satisfactory, with linear correlation of R^2>0.99 in signal vs. concentration plots, using 1-4 punches (about 3, 6, 9 and 12 µL whole blood) of the DBS, but some polar compounds were affected by matrix effects. Tandem MS data was acquired to secure better identification. Combining the method with bioinformatics, one could identify changes between samples taken after free diet, overnight fasting (12 hours) and prolonged fasting (36 hours), using only 3 µL blood on paper. The method will be used to detect differences in the metabolome of patient and healthy samples in research and in future diagnostics.



Preface

The present work was performed at the National Unit of Screening and Diagnosis of Congenital Pediatric Metabolic Disorders at Oslo University Hospital, Rikshospitalet from January 2016 to August 2018. My supervisors have been Katja B. P. Elgstøen, Anja Østeby and Hanne B. Skogvold at Rikshospitalet and Steven R. H. Wilson at the Department of Chemistry, University of Oslo. I would like to thank them for giving me work that has been interesting and challenging, but most of all fun.

I would like to give a huge thanks to Katja for always helping me. I admire her knowledge and ability to make people around her feel well. She always made me feel good about my work and decisions. A meeting with her resulted in hundreds of new thoughts to be used in my work. A common phrase from people in the corridors has been "Where is Katja?". A lot of people compete for her time, and in some way, she always managed to make time for everyone.

Thank you to Helge Rootwelt for the medical view and for showing a genuine interest for my work, always asking questions and making sure that I was fine.

Thank you to Anja and Hanne for making my days perfect and for taking care of me throughout this period. I love to see their smiling faces and I am forever grateful for all our time together; the lunches, the troubleshooting (a lot) and the discussions (work and non-work subjects). A special thanks to Hanne for helping me with the practical work and in the writing process and for all the dog pictures she shared with me. Thanks to Camilla Løkken for introducing me to the method and everyone at IKB for all the lunches and conversations.

Steven deserves a huge thank you for the excellent guidance and constructive feedback throughout my master's degree. I would also like to thank the Bioanalytics group at the Department of Chemistry for the generosity and sharing of knowledge.

I would like to thank Norsk Kjemisk Selskap – Faggruppe for Analytisk Kjemi for the economic support so that I got the opportunity to attend "Det 18. Norske seminar i væskekromatografi" in Sandefjord. My work was presented with a poster shown in **Appendix** section 6.7.

Thank you to Per Ola Rønning at OsloMet for getting me in contact with Katja. He is the reason that I ended up with this work and I am forever grateful for that.

Finally, I want to thank my family and friends for their endless love and support throughout my degree: the members of Laila  for believing in me and for cheering me up every day, Martin for taking care of my man, Camilla for the workouts and talks, Nils for letting me "passe døra hans", mor and far for their interest and prayers, Haakon and Gerd for taking care of Nala, Jakob for letting me play with him and hug him, my sisters, that I love to the moon and back, for always being there for me, mamma and pappa for believing in me and for always supporting me no matter what. A special thanks to my beloved at home, Espen, for taking care of me when I was at home (and the apartment when I was gone) and for the endless support, and Nala for all the hours beside me and the computer and for always being happy to see me.

Oslo, Norway, august 2018

Elise Mørk Sandås

Table of content

1 Abbreviations

Abbreviations	Term
AH^+	Protonated acid
AIF	All ion fragmentation
Ala	Alanine
APCI	Atmospheric pressure chemical ionization
APPI	Atmospheric pressure photoionization
Arg	Arginine
BH^-	Deprotonated base
C18	Octadecyl
C8	Octyl
Cit	Citrulline
Da	Dalton
DBS	Dried blood spot
DC	Direct current
DDA	Data-dependent acquisition
$ddMS^2$	Data-dependent tandem mass spectrometry
DIA	Data-independent acquisition
DNA	Deoxyribonucleic acid
DPS	Dried plasma spot
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted ion chromatogram
ESI	Electrospray ionization
FA	Formic acid
FDA	Food and Drug Administration
FT-ICR	Fourier transform ion cyclotron resonance
FWHM	Full width at half maximum
GC	Gas chromatography
GC - MS	Gas chromatography – mass spectrometry

Gly	Glycine
HCD	Higher energy collision-induced dissociation
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High-performance liquid chromatography
I	Intensity
I_{th}	Intensity threshold
IEM	Inborn error of metabolism
IS	Internal standard
LC	Liquid chromatography
LC - MS	Liquid chromatography – mass spectrometry
Leu	Leucine
LLOD	Low limit of detection
M	Molecule
MeOH	Methanol
Met	Methionine
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MS^2	Tandem mass spectrometry
m/z	Mass-to-charge ratio
N	Number of replicates
NMR	Nuclear magnetic resonance
Or	Ornithine
PA	Peak area
PCA	Principal component analysis
Phe	Phenylalanine
PLS	Partial least square
ppm	Parts per million
QC	Quality control
Q-TOF	Quadrupole - time-of-flight
R_s	Resolution
REC	Regional Committee for Medical and Health Research Ethics
RF	Radio frequency
RNA	Ribonucleic acid

RSD	Relative standard deviation
Rt	Retention time
SIM	Single ion monitoring
STD	Absolute standard deviation
SWATH	Sequential window acquisition of all theoretical fragment-ion spectra
TIC	Total ion chromatogram
TOF	Time-of-flight
Tyr	Tyrosine
Val	Valine
WADA	World Anti-Doping Agency

2 Introduction

2.1 Inborn errors of metabolism

The metabolism involves all biochemical reactions that take place within each cell of a living organism. Biochemical reactions are linked in series as biochemical pathways that start with one metabolite which is converted by specific proteins (enzymes) to other metabolites [1]. The recipes of specific proteins are in the genes, stored as deoxyribonucleic acid (DNA) information which transcribes into messenger ribonucleic acid (mRNA), and mRNA information translates into proteins. Mutations in genes give the wrong recipes for protein synthesis and may result in dysfunctional proteins [2]. Inborn errors of metabolism (IEMs) are a group of genetic disorders causing deficiency or abnormality of specific proteins, most often enzymes, their cofactors or transporters, leading to abnormal levels of metabolites in biochemical pathways. **Figure 1** illustrates an example of the result of an IEM when a deficiency or dysfunction of a necessary enzyme protein occurs. The result may be toxic substrate accumulation or deficiency of downstream products. The altered concentrations of metabolites (upstream substrates and downstream products) may affect other biochemical pathways creating further biochemical imbalances and pathological metabolites [3].

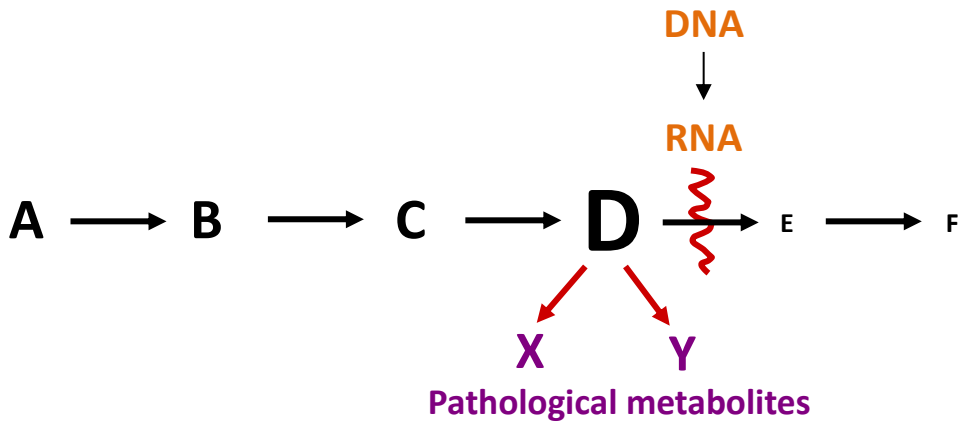


Figure 1: **General pathophysiology in inborn errors of metabolism**. A, B, C, D, E and F represent metabolites in a biochemical pathway. The arrows represent enzymes. Mutation in DNA creates dysfunctional or abnormal enzymes leading to a metabolic block in the biochemical pathway after metabolite D. The results may be deficiency of products (E and F), toxic accumulation of substrate (D) and creation of pathological metabolites (X and Y) when the accumulated substrate is involved in other biochemical pathways.

An example of an IEM is glycogen storage disease type I. This is a result of a defect in the gene for the enzyme glucose-6-phosphatase which is essential to the cell's ability to metabolize glycogen. The reduced ability to metabolize glycogen results in a low concentration of glucose, which is life threatening. In addition, the intermediate metabolites can interact in other pathways leading to increased lipid, uric acid and lactic acidosis levels [4].

An IEM can cause irreversible mental retardation, physical disability, neurological damage and even mortality if undiagnosed and untreated [5]. Early diagnostics is crucial to start early treatment to avoid severe irreversible damage or, in the worst cases, mortality. The study of IEMs to identify new diseases, to better characterize known diseases and to identify relevant biomarkers is of high importance to improve both diagnosis and monitoring of the diseases. Tools and methodologies to study and characterize the different "-omics" are important to increase knowledge of, and improve treatment of IEMs.

2.2 Omics to study inborn errors of metabolism

Biological information in an organism starts with genes and continues with transcripts through proteins and finally to metabolites. Genomics, epigenomics, transcriptomics, proteomics and metabolomics are the comprehensive study of DNA (genome), chromatin structure and gene regulation (epigenome), mRNA (transcriptome), proteins (proteome) and metabolites (metabolome), respectively, in a biological system. The study of the omics cascades improves understanding of biological systems [6].

Genomics, transcriptomics and proteomics reveal information about potential and actual biochemical processes in an organism and metabolomics can tell us which processes that are really happening; that is, the actual biochemical status or results at the time of sampling. Genomics is used to investigate what *can* happen in an organism. Transcriptomics makes it possible to see what *appears to be happening*. Proteomics tells what *makes this happen*. Metabolomics shows what *has happened and is happening*. All this results in the phenotype of an organism [7].

Metabolomics is the endpoint of the omics cascade and the metabolome reflects changes in the genome, transcriptome and proteome. In addition, the metabolome reflects the ingestion

or effects of exogenous factors, such as environmental and dietary factors, contrary to the other omics. The metabolome is therefore dynamic and provides a "snapshot" of the physiology of the organism compared to the genome, transcriptome and proteome [8]. Metabolites are tightly related to the pathophysiology of IEMs and the study of all metabolites are therefore relevant for diagnostics of IEMs [9]. This makes metabolomics a promising and useful tool in laboratory research, diagnostics and monitoring of IEMs [6].

2.2.1 Metabolomics

Metabolomics is the comprehensive study of low-molecular weight metabolites <1500 Da [10, 11], like sugars, nucleotides, amino acids, fatty acids and lipids, in a biological sample [12]. Metabolomics approaches are divided into targeted and untargeted metabolomics. Targeted metabolomics is specific assays which identifiy and quantify specific metabolites coupled to specific biochemical pathways. Untargeted metabolomics is the comprehensive approach to identify and quantify as many metabolites in a biological sample as possible. The latter provides a fingerprint of the pathophysiological state of the organism [13].

Targeted metabolomics

Targeted metabolomics is used in standard clinical chemistry to measure the levels of known metabolites related to biochemical pathways of interest. The chemical properties of the metabolites are known so that specific sample preparation and analytical methods can be optimized. This gives the potential to obtain specific and sensitive measurements, but it results in multiple sample preparation and analytical methods when measuring several metabolites [14]. Targeted metabolomics is a good approach to measure known metabolites related to known biochemical pathways, but there is a high chance of missing information about other interesting metabolites [15].

Untargeted metabolomics

Untargeted metabolomics is an approach to identify and quantify as many metabolites as possible in a biological sample. Non-specific sample preparation and analytical assays are used which determine the number of detected metabolites. Data obtained from the analysis are processed and metabolites of interest may be identified [16]. The development of bioinformatics tools over the years, such as databases and software, facilitates the data

processing and identification of metabolites [17]. The use shows an increase in clinical research of IEMs due to more sensitive and accurate analytical techniques developed over the last years [18, 19]. New technologies allow to screen for many different metabolites in one biological sample, in one single analysis [20].

Targeted approaches provide better quantification, but untargeted approaches provide broader coverage [21]. Targeted metabolomics approach has been used in laboratory diagnostics of IEMs since the discovery of the first IEM [22, 23], while untargeted metabolomics has the potential to improve future diagnostics and monitoring of IEMs by rapid identification of a broad range of biochemical disturbances reflecting the specific biochemical aberration and the resulting dysfunctional pathways and biochemical status of the patient [20, 24, 25].

2.3 Diagnostic practice for inborn errors of metabolism

Each individual IEM is rare, but IEMs as a group are collectively common, with over 1500 described IEMs [26]. Many symptoms can be avoided or relieved by treatments such as dietary restriction and supplements. An example is the treatment of phenylketonuria with a low-phenylalanine formula. Maintaining a strict diet prevents an otherwise inevitable progression of mental retardation [27]. It is important to diagnose these patients early to start treatment, as some of these diseases can lead to rapid death or severe conditions if not properly treated.

Targeted metabolomics is used in diagnostics of IEMs by measuring the amount of substrate, products and pathological metabolites in known biochemical pathways. The metabolites are biomarkers for IEMs to see if the measured amount is within a reference range for healthy individuals or in levels compatible with an IEM. When a specific diagnosis is suspected, targeted metabolomics is a useful tool. Multiple targeted approaches may be required if the result doesn’t support the hypothesis or if additional findings are necessary to conclude, resulting in multiple sample types and analytical assays used. Targeted approaches may be time and resource demanding when there is no clear hypothesis. This can hamper the process when an early diagnosis is needed. In untargeted approaches, in contrast, one sample and one assay may be used to screen for virtually all possible metabolites, resulting in a hypothesis

followed up by confirmation. This can be done by using one single targeted follow-up method [24], and/or DNA sequencing confirming disease causing mutation(s) in the relevant gene. This is illustrated in **Figure 2.**

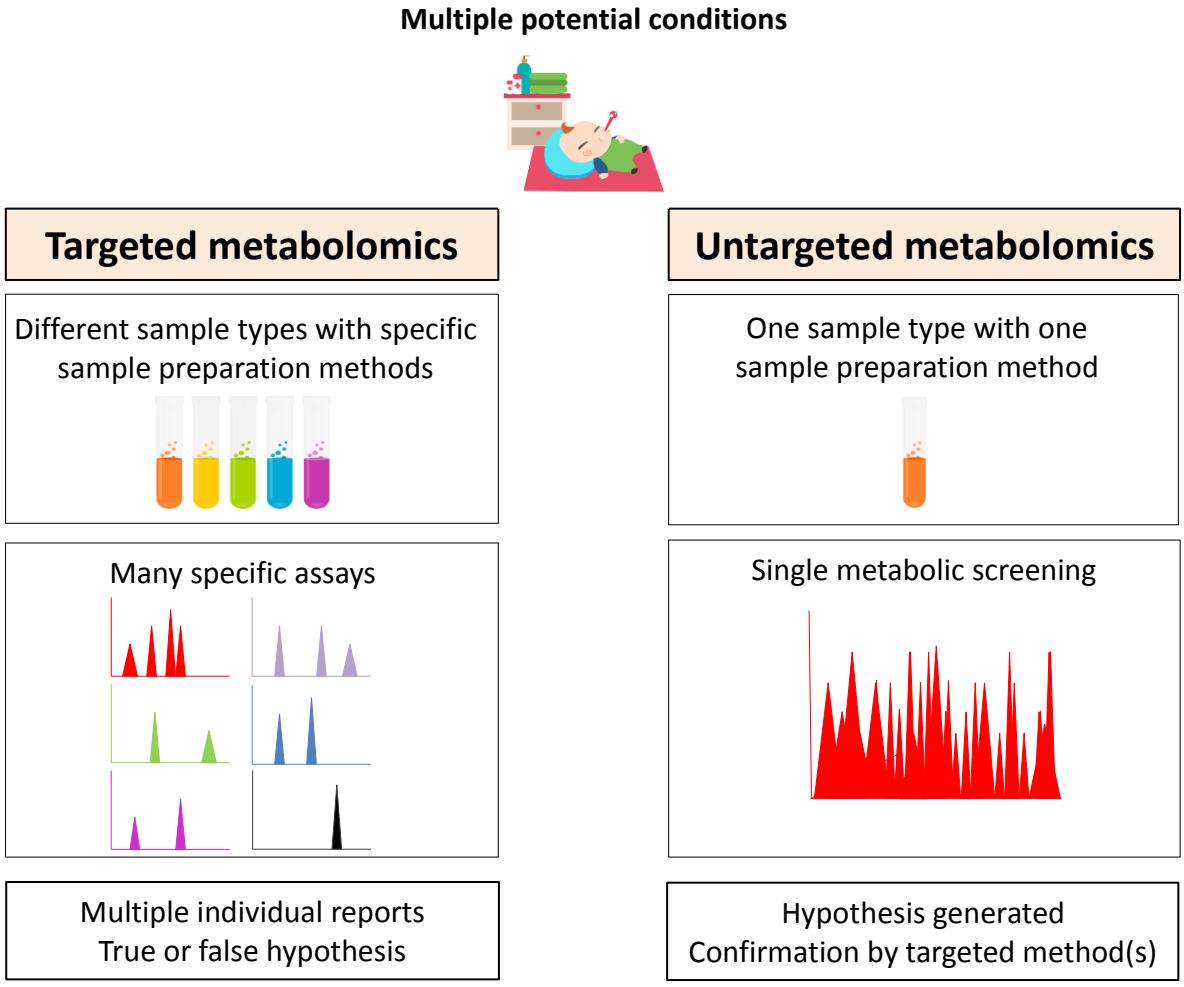


Figure 2: Metabolomics has the potential to efficiently screen metabolites in biological samples for diagnostics of inborn errors of metabolism. Targeted metabolomics may require multiple sample types and assays, while untargeted metabolomics has the potential to use one sample type and one comprehensive analysis to screen for all possible metabolites. The figure was adapted from [24].

Diagnostics of IEMs is based on the specific presenting clinical phenotype in patients, family history of IEMs and results of newborn screening [28]. Experienced health personnel consider the symptoms and signs of the patient and recommend targeted metabolomics analyses. The most used method is to search for specific metabolites in various body fluids. Urine is for example used to analyze organic acids and plasma is used to analyze acylcarnitines [29, 30]. A lot of time and resources can therefore be used. The future goal of diagnostics of IEMs is therefore to start the diagnostic examination with untargeted metabolomics analyses of biological samples, generate an hypothesis after studying the metabolic pathways and then do experimental validation with targeted analyses or DNA sequencing [9].

All newborns in Norway are since 2012 screened for 21 IEMs. This screening is performed by the Department of Newborn Screening at Oslo University Hospital. Patients with other suspected IEMs are diagnosed at the National Unit of Screening and Diagnosis of Congenital Pediatric Metabolic Disorders.

2.3.1 Department of Newborn screening

Even though a newborn seems healthy at birth, it may be born with an IEM. Early diagnostics is important to be able to start treatment as soon as possible. The Norwegian Newborn Screening Programme offers parents in Norway screen of their newborns for 25 rare disorders where early treatment is of vital importance. Out of these rare disorders, 21 are IEMs; the others are two endocrinological conditions, severe combined immunodeficiency and cystic fibrosis. Dried blood spot (DBS) samples are collected from a puncture of the newborn's heel 48-72 hours after birth. This is a common sample collection method for newborn screening in many countries [31]. The samples are shipped for analysis at the Department of Newborn Screening at Oslo University Hospital along with contact information of the mother, and information such as time of birth, gestational length, weight and sex of the baby, as these factors affect the metabolite concentrations. Information about the newborn is important to consider all the factors that may affect the test results. The Norwegian Newborn Screening Programme is a good service, but is only used to detect some of the IEMs. There are strict rules for newborn screening laboratory diagnostics, for example that the disease must be severe, the treatment must be efficient and there must be a good diagnostic marker. This explains the low number of diseases that are screened for compared to the huge number of individual diseases described. When it is suspected that a patient is suffering from an IEM

that is not included in the newborn screening, samples are collected and analyzed at the National Unit of Screening and Diagnosis of Congenital Pediatric Metabolic Disorders [32].

2.3.2 National Unit of Screening and Diagnosis of Congenital Pediatric Metabolic Disorders

The National Unit of Screening and Diagnosis of Congenital Pediatric Metabolic Disorders performs diagnostic analyses when there is a suspicion of an IEM. Information about clinical symptoms, medical history, family members with IEMs and relevant investigations already performed on the patient is used to generate a diagnostic hypothesis and determine the diagnostic workflow.

The unit wants to establish an untargeted metabolomics method and use it in diagnostics of IEMs as a door opener to new targeted metabolomics methods. The goal is to identify pathological metabolites that stand out from a healthy biochemical profile and study the biochemical processes of the identified metabolites to find the biochemical fault and to identify new or better biomarkers of the disease and its severity. **Figure 3** illustrates a healthy vs. a pathological biochemical process that result in healthy and biochemical profiles.

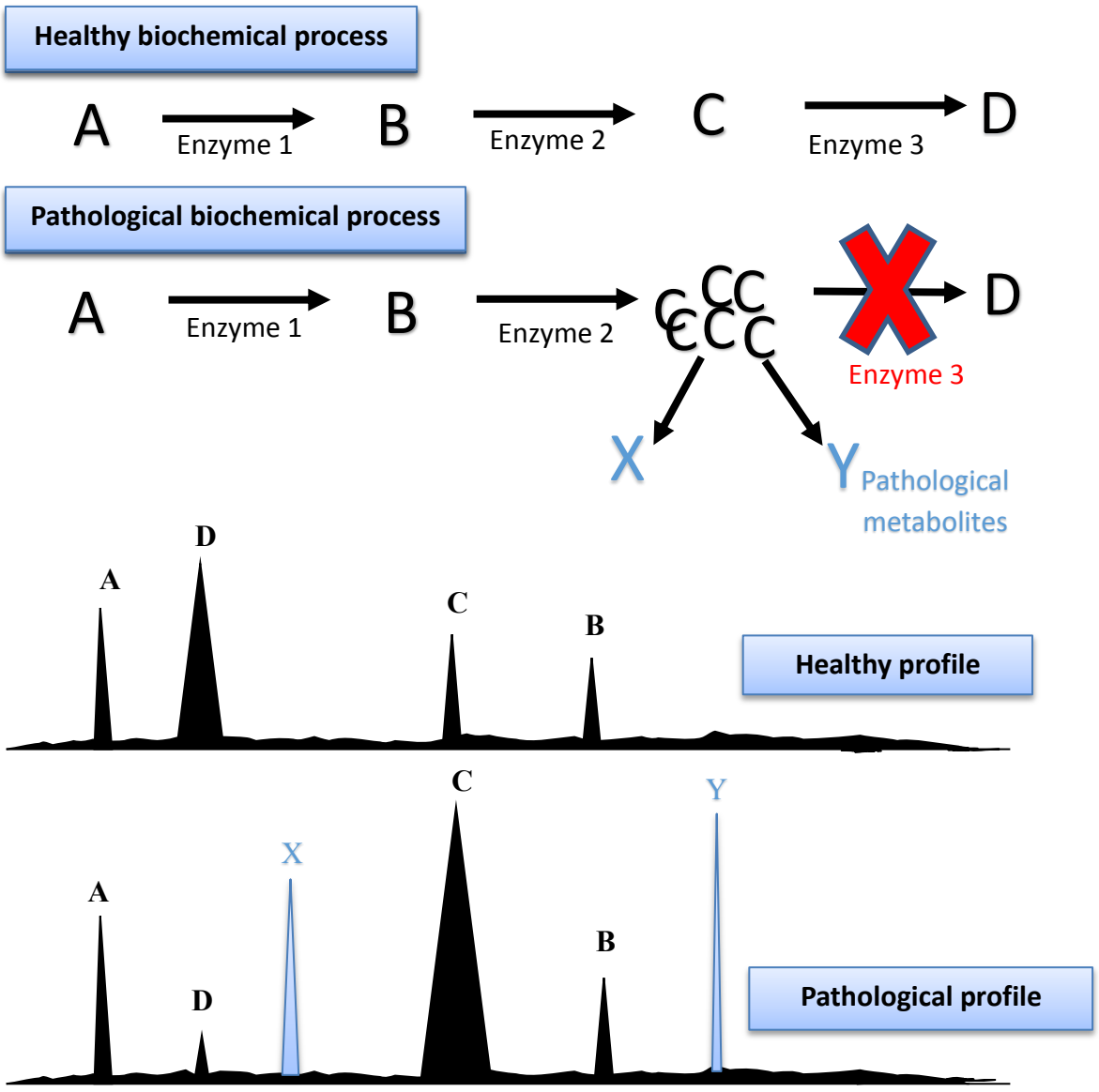


Figure 3: An inborn error of metabolism causing a defect enzyme results in a pathological biochemical process. The comparison of a normal vs. a pathological profile may show the accumulation of substrate, the deficiency of product and/or the pathological metabolites from other biochemical processes where the substrate is involved. The area of the peaks represents the amount.

Most of the IEMs are named by the laboratory findings, such as tyrosinemia (high levels of tyrosine in blood) and phenylketonuria (high levels of phenylalanine in urine).This IEM would be called C-emia if the diagnostic test was in blood and C-uria if the diagnostic test was in urine. Targeted approaches, that are used today, can for example measure metabolite "C", but untargeted metabolomics has the potential to measure all the metabolites in the biochemical pathway and pathological metabolites created as a consequence of an IEM. In

general, normal time for a complete diagnostic service of an IEM is one month [33]. The use of an untargeted metabolomics approach may reduce the response time. Likewise, the use of next generation DNA sequencing of all relevant genes might be more cost effective and time consuming than analyzing one or a few suspected genes.

Metabolic screening is performed in urine, serum, plasma, spinal fluid, whole blood and DBS dependent on what metabolites that are measured. There can be specific pre-analytical procedures and sampling, like overnight fasting, immediate freezing and shipping on dry ice. The use of DBS facilitates the storage and transport.

2.4 Dried blood spots

Whole blood contains red blood cells, white blood cells and platelets (about 45 %) which are suspended in plasma (about 55 %). Plasma contains water, proteins, nutrients and vitamins, hormones, electrolytes, trace elements, blood gases and a whole range of metabolites. The cells gravitate towards the bottom, thus, plasma is obtained by separating the red blood cells from the liquid portion of whole blood. Centrifugation accelerates the process of separating plasma and cells. The difference between plasma and serum is that plasma contains all plasma proteins and is obtained by adding anticoagulant in the whole blood sample, whereas serum contains no anticoagulant and has lost the proteins and factors used in generating the blood clot in the sample vial. Venipuncture collection (collection of blood directly from a vein in a tube) is the most common way to collect blood samples. Blood may also be collected from the capillaries (tiny blood vessels near the surface of the skin) by puncturing the skin and using capillary blood collection tubes or collection on filter paper card as DBS [34].

Plasma, serum and urine are the most used sample materials for diagnostics of metabolic disorders [20, 24, 30, 35]. DBS collection is of high interest in diagnostics of metabolic disorders due to its many benefits. DBS sampling requires only a small amount of sample material, which is beneficial for collection of blood samples from infants, and for frequent sampling, such as in follow-up treatment. The benefits of using DBS samples compared to plasma and urine are the easy collection, less storage space and easy transport [36, 37]. The risk of infections for the laboratory personnel using DBS compared to venous sampling is reduced [38].

Generally, a puncture in the finger or the heel is performed with a lancet. The droplet of blood is applied in a dashed ring of 1.2 cm (used in this project) in diameter that holds about 50 µL of blood, see **Figure 4**. Sometimes, the whole spot is used for analysis, but commonly, only a portion of the spot is analyzed. When only a portion of the spot is used, a punch is taken from the spot with either a manual or an automatic puncher (often 0.32 cm in diameter, which is equivalent to approximately 3 µL whole blood [39]). Compounds can be extracted from the spot into a solution.

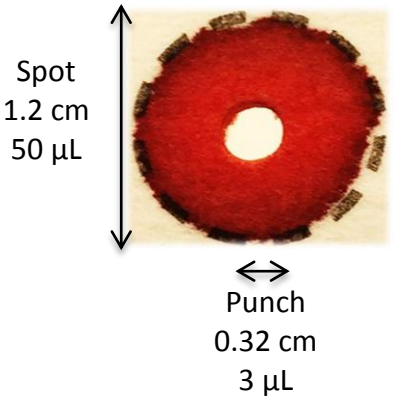


Figure 4: Dried blood spot with a center punch. The diameter of the spot is 1.2 cm (holds approximately 50 µL blood) and the diameter of the punch is 0.32 cm (equivalent to approximately 3 µL whole blood).

A DBS is a droplet of blood that is applied on a filter paper card and dried. There are many filter paper cards available for collection of whole blood. The properties of different types of filter paper cards, for example material, thickness and pore size, influence the rate of adsorption and distribution of whole blood. Whatman 903 filter paper cards consist of 100 % cellulose and were used in this project. These are one of two card types approved by the Food and Drug Administration (FDA) [40].

The location of the spot where the punch is taken (punch location), hematocrit value (the percentage of red blood cells in a unit volume of whole blood) and volume of blood spotted are factors that potentially affect DBS quantification and may give false positive or negative measurements. These variables must be controlled or adjusted for, if possible, in quantitative measurements [41]. Studies that have investigated the punch location have found different amounts of metabolites in punches taken from the center of the DBS compared to the perimeter dependent on the targeted metabolite [36, 41]. The hematocrit affects the sample volume at a fixed punch size of a DBS due to the greater viscosity of blood at high hematocrit values compared to low. For the same volume applied, higher hematocrit results in smaller blood spots and therefore higher observed concentrations of analytes [38, 41]. Studies show that the correction of measurements using hematocrit values increases the accuracy of measurements by measuring, for example, light transmission [41], potassium measurements [42] and reflectance [38], as surrogates of hematocrit. Volume spotted may be controlled

using the dashed line and only DBS of approximately equal area should be compared. The effect of hematocrit values and volume spotted was not investigated in this project as untargeted metabolomics methods analyze the metabolic profile with the potential to observe increased or decreased concentrations of all detected compounds.

The use of DBS in targeted metabolomics approaches requires that the effect of storage, matrix effects and extraction efficiency on the quantification should be evaluated [43]. Newborn screening all over the world have been using DBS samples for about 60 years, resulting in an accumulation of considerable experience and thorough validation of several diagnostic markers, amino acids and acylcarnitines, specifically. There are, for example, many studies of the storage of metabolites using DBS samples. A study that investigated the effect of storage of DBS in different temperatures found that storage over two years at -20 ˚C and -80 ˚C reduces changes of metabolite concentrations significantly over time compared to storage at room temperature, but some compounds are more affected than others and this need to be considered when analyzing stored samples [44]. Another study found that acylcarnitines were stable for over 330 days at -18 ˚C [45]. The DBS cards in this project were stored at -80 ˚C based on these findings and DBS samples that will be compared using untargeted metabolomics approaches should be stored and handled in the same way. The investigation of the matrix effect, extraction efficiency and the absolute quantification is impossible to do for all compounds when performing untargeted metabolomics, but these factors can be investigated for some chosen metabolites, which were done in this project.

DBS samples require analytical techniques that are able to detect low concentrations and separate many metabolites.

2.5 Analytical techniques used in untargeted metabolomics

The goal of an untargeted metabolomics method is to pick up differences in the metabolome of test and control groups and to identify them. Thus, the analytical techniques used must be able to perform quantitative and qualitative measurements of as many metabolites as possible in a biological sample. There is no single analytical technique that covers the entire spectrum of the metabolome, due to the wide differences in chemical properties and concentrations of

the metabolites. This means that the analytical techniques used determine the coverage of the metabolome in the analyzed sample [46, 47]. Nuclear magnetic resonance (NMR) and mass spectrometry (MS) are the two main analytical techniques used in untargeted metabolomics today [19, 47-49]. The analyses may be qualitative and/or quantitative [50].

NMR is reproducible and non-destructive, and requires little sample preparation [51]. The disadvantage of NMR compared to MS is the sensitivity [46]. The sensitivity of NMR has improved due to technological improvements, but is still about 1000 times less sensitive than MS [52]. The low limit of detection (LLOD) of NMR is normally in µM, but reaches down to nM (10^{-9}) with the latest improvements. The LLOD of MS is less than pM (10^{-12}) [53]. The concentration of many metabolites in DBS samples is low, thus high sensitivity is important to detect as many metabolites as possible. MS was used in this project, as the LLOD of MS results in a higher number of detected metabolites, see **Figure 5**. This gives a wider coverage of the metabolome by increasing the coverage of the low abundant metabolites in a given biological sample compared to the use of NMR [54].

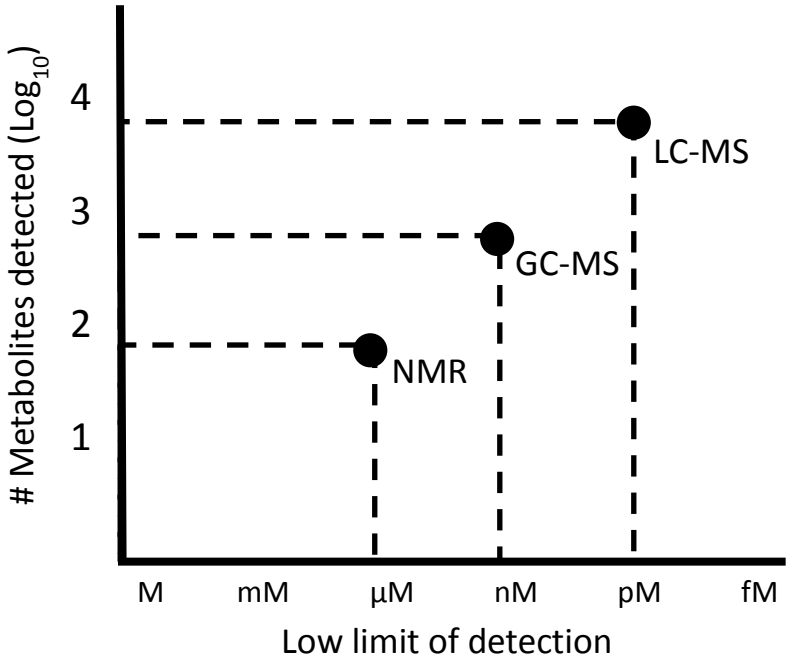


Figure 5: Relative sensitivity and number of metabolites detected by analytical techniques used in untargeted metabolomics approaches. The detection limit when using liquid chromatography-mass spectrometry (LC-MS) is lower than when using gas chromatography-mass spectrometry (GC-MS) and nuclear magnetic resonance (NMR). The figure was adapted from [53].

MS measures ions under high vacuum, thus ions must be transferred into gas phase and neutral gas molecules must be removed prior to ion separation to avoid molecular collision [55]. Different ion sources can be used to solve these challenges. Commonly used ion sources in metabolomics are atmospheric pressure chemical ionization (APCI), atmospheric pressure photoionization (APPI) and electrospray ionization (ESI) [56]. The ionization techniques have different abilities to ionize. The use of different ionization techniques increases the number and types of analytes observed [57, 58]. A comparison of the compounds that are ionized with the different ionization techniques based on polarity and molecular weight of compounds is shown in **Figure 6**.

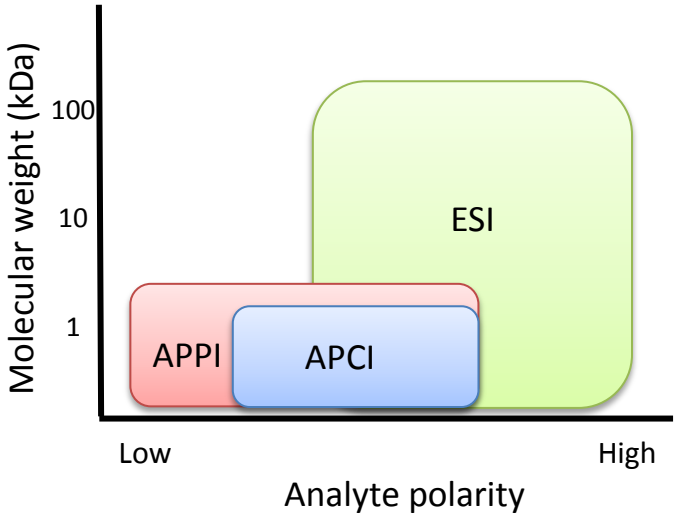


Figure 6: A comparison of ionization techniques used in LC-MS metabolomics based on analyte polarity and molecular weight. Atmospheric pressure photoionization (APPI) and atmospheric pressure chemical ionization (APCI) are more suitable for non-polar and low-molecular weighted analytes than electrospray ionization (ESI). The figure was adapted from [59].

ESI is one of the most used ion sources in the chemical and biochemical analyses [60, 61], and was used in this project as the majority of metabolites in the DBS metabolome are polar [47]. It is a favored ionization technique in untargeted metabolomics due to the ability to produce a large number of ions, as compounds of moderate to high polarity and low to high masses [62]. The technique is soft, with little or no uncontrolled fragmentation of analytes due to little residual energy retained by the analytes [63]. The result is mass spectra mostly containing peaks from intact molecular ions [11]. This is an advantage in untargeted metabolomics as many of the metabolites measured are unknown. A consequence of direct injection of a sample into the ESI source is that molecules of unknown amounts and

properties are introduced to the source simultaneously, resulting in poor ionization efficiency for multiple analytes. Ion signal may be suppressed or enhanced in the ion source, thus a separation step of a complex sample prior to the introduction into the ion source is necessary [47, 55].

The compounds can be separated in time by separation techniques such as liquid chromatography (LC) and gas chromatography (GC). This allows compounds to enter the ion source at different times, which improves the sensitivity and facilitates the identification of compounds. GC uses temperature to separate compounds and is used to separate volatile, or derivatized into volatile and thermally stable compounds [56]. GC-MS is a robust and reproducible technique with a lot of databases for identification, but it is less sensitive than LC-MS [53], see **Figure 5**. The majority of the metabolites are polar and involatile, thus the separation with GC requires derivatization [47]. LC is a technique that is used to separate compounds in a solution based on different equilibriums between a solid stationary phase and a liquid mobile phase. It separates very polar to non-polar compounds dependent on the column used, and is the most used separation method in untargeted metabolomics based MS due to its versatility [56].

LC-ESI-MS is a much used analytical tool in untargeted metabolomics [47, 64-66]. Further details on each technique are described below.

2.5.1 Mass spectrometry

Mass spectrometry (MS) is an analytical technique that separates charged compounds in gas phase based on their mass-to-charge ratio (*m/z*) using an electric and/or magnetic field. The result is a mass spectrum where the signal intensity of a charged compound is plotted against its *m/z*. A sample is introduced in to the instrument through a sample inlet, and the analytes are transformed into gas phase ions by an ion source. Further, the ions are focused on to a mass analyzer which separates the ions according to their *m/z*, and focuses the ions into the detector which determines the signal intensities of each separated *m/z*. A computer records and processes the data generated in addition to control the operation of the instrument. The mass analyzer, detector and sometimes the ion source are kept under vacuum to prevent loss of ions due to collision with neutral molecules. The components of a mass spectrometer are shown in **Figure 7** [67].

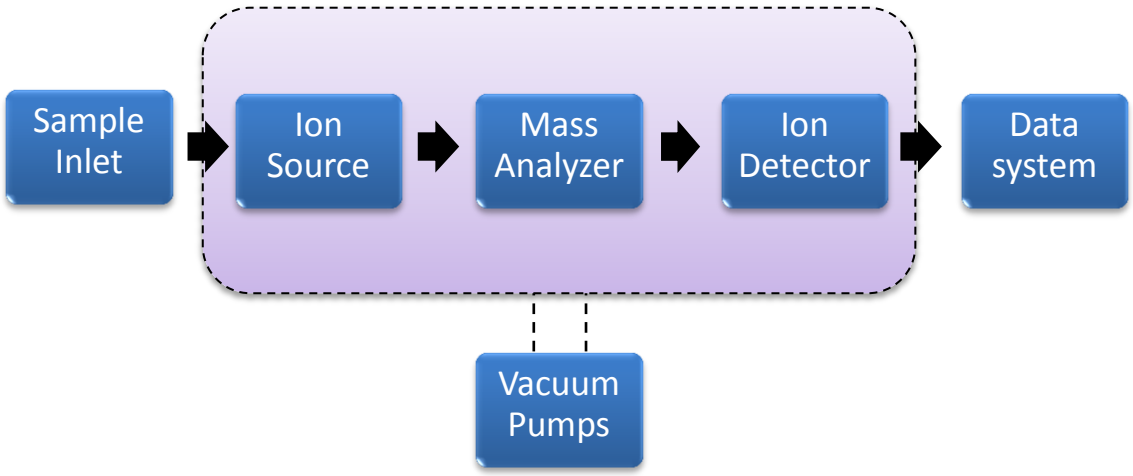


Figure 7: General components of a mass spectrometer. The sample inlet introduces the sample. **The ion source** transforms analytes to gas phase ions. **The mass analyzer** separates ions according to their *m/z*. **The ion detector** detects ions where the signal intensities of each m/z value are determined**. The vacuum system** prevents loss of ions through collisions with neutral gas molecules and the walls of the ion source, mass analyzer and ion detector**. The data system** controls the instrument operation, and records and processes generated data. The figure was adapted from [68].

Tandem MS (MS^2) can be performed to investigate structural information in addition to the mass of a compound. In MS^2 analyses, one mass analyzer is commonly used to select an ion of interest (precursor ion), induce fragmentation and the fragments (product ions) are analyzed in a second mass analyzer [67].

High resolution and mass accuracy are important performance parameters of the MS in untargeted metabolomics as this increase the confidence in metabolite identification [47, 69]. Mass resolution is a measure of the ability of a mass analyzer to separate ions with different *m/z*. The most commonly used definition of mass resolution is to use the full width of the peak at half its maximum height (FWHM), see **Equation 1.**

$$R_S = \frac{m/z}{w_{0.5}} \qquad \textbf{Eq. 1}$$

Where R_s is the mass resolution, *m/z* is the mass-to-charge ratio and $w_{0.5}$ is the peak width at half height (FWHM) [67].

Mass accuracy is the difference between the measured accurate mass and the calculated exact mass and can be expressed as the mass error in parts per million (ppm) [67], which is defined in **Equation 2**.

$$Mass\ accuracy\ (ppm) = \frac{Exact\ mass - Accurate\ mass}{Exact\ mass} \times 10^6 \qquad \textbf{Eq. 2}$$

There are many different mass analyzers available, with different performance abilities. Triple quadrupoles are often used in targeted metabolomics, due to their ability to scan for masses in a certain mass range and select specific masses to be analyzed and/or fragmented. The disadvantages are low resolution and mass accuracy. Time-of-flight (TOF) is often used in untargeted metabolomics analyses due to the ability to obtain many spectra in a short time, but TOF instruments have lower mass accuracy compared to Orbitraps and Fourier transform ion cyclotron resonance (FT-ICR) instruments. The latter has better mass resolution and accuracy than the Orbitrap, but needs a superconducting magnet and is relatively expensive [67, 70]. **Table 1** shows the comparison of frequently used mass analyzers.

Table 1: Comparison of frequently used mass analyzers based on mass resolution (full width at half maximum (FWHM)), mass accuracy in parts per million (ppm) and mass range in Da (Dalton). Table adapted from [52, 67]. SIM: Single ion monitoring. FT-ICR: Fourier transform ion cyclotron resonance.

	Triple Quadrupole	**Time-of-flight**	**Orbitrap**	**FT-ICR**
Mass resolution (FWHM)	1,000-5,000	10,000-50,000	50,000-200,000	1,000,000-2,000,000
Mass accuracy (ppm)	5-500	2-5	0.5-1	0.5-1
Upper mass limit (Da)	4,000	>350,000	4,000	4,000
Spectra collection rate (Hz)	20	10,000	10	<1
Advantages	Scan SIM	Fast Spectra collection Mass range	Resolution Mass accuracy	Resolution Mass accuracy
Disadvantages	Nominal mass only	No SIM	Resolution decreases with data collection and increased mass	Resolution decreases with data collection and increased mass Superconducting magnet Expensive

The Orbitrap MS is widely used in untargeted metabolomics methods due to the high resolution and mass accuracy and was used in this project [47].

Q-Exactive Orbitrap mass spectrometer

The Q-Exactive mass spectrometer consists of a bent flatapole, a quadrupole mass filter, a C-trap, a higher collision-induced dissociation (HCD) cell with the opportunity to perform MS and MS^2 analysis and an Orbitrap mass analyzer [37]. A schematic view is shown in **Figure 8**. Gas phase ions from the ion source are introduced into a bent flatapole that prevents neutral

ions to pass since they are not able to follow the curved shape [71]. The ions then enter a quadrupole which uses an electrical field to filter the *m/z* of interest. The quadrupole consists of four electrodes (rods) where the opposite rods are in pair. Radio frequency (RF) and direct current (DC) voltages are applied to the rods, one pair with positive DC voltage applied combined with RF voltage and the other pair with negative DC voltage applied combined with RF voltage 180º out of phase. The applied RF voltages results in the rods constantly changing between positive and negative polarities. Ions start to oscillate in the electrical field and the combination of RF and DC voltages determines if the ions with a specific *m/z* oscillate in a stable trajectory along the axis of the quadrupole to the detector. All the other ions are lost due to collision with the rods. When filtering masses in an *m/z* range, the DC and RF are changed progressively, but with constant ratio [67]. The ions that are filtered enter the C-trap, which is a curved RF-quadrupole filled with gas that can store and inject ions either into the Orbitrap in packages, or to the HCD cell for fragmentation, back to the C-Trap and then to the Orbitrap. In the Orbitrap, ions are trapped in axial oscillations in a homogenous electrostatic field where each *m/z* achieve a specific axial oscillation and is used to determine the *m/z* of the ions [70].

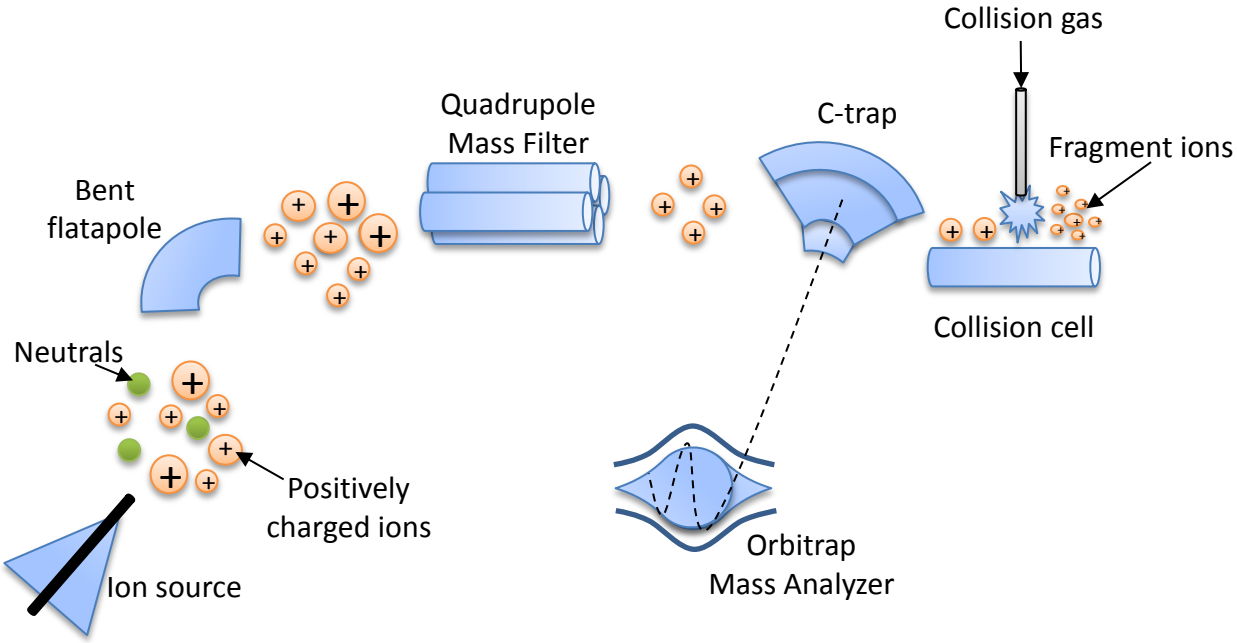


Figure 8: A schematic view of the most important parts of the Q Exactive Orbitrap mass spectrometer. Ions and neutrals enter the mass spectrometer from the ion source. The bent flatapole prevents neutrals to enter the quadrupole where ions in the preselected *m/z* range pass. The ions enter the C-trap that shoots the ions into the Orbitrap mass analyzer or send them to the collision cell for fragmentation. The fragments are sent back to the C-trap and further to the Orbitrap mass analyzer for detection. The figure was adapted from [55].

The mass analyzer consists of two types of electrodes that are electrically isolated from each other: a spindle-like central electrode and a pair of barrel-like outer electrodes with a DC voltage applied. Ions enter between the electrodes from the C-trap. The ions oscillate around the central electrode and oscillate axially in trapped stable orbits simultaneously. These orbits cause image currents in the outer electrodes, which can be detected using a differential amplifier [55]. The axial frequency of the movements is related to the *m/z* of ions. The signal of all axial frequencies of the ions in the Orbitrap mass analyzer is detected, decomposed into individual frequencies and converted to an *m/z* spectrum by Fourier Transformation [72].

The Q Exactive mass spectrometers can be operated in full MS mode, where precursor ions within a selected *m/z* range are detected, and MS^2 mode, where the precursor ions are fragmented and detected for identification. The latter requires a HCD cell [11].

2.5.2 Electrospray ionization

ESI uses electrical energy to transform ions in solution into gas phase ions by the formation of charged droplets [61]. Positively and negatively charged compounds in solution may be transferred to a gaseous phase dependent on the voltage applied, called positive and negative ionization mode, respectively. **Figure 9** illustrates ESI in positive mode. Easily ionized compounds, e.g., protonated bases (BH^+) and deprotonated acids (AH^-) may already be ionized in the solvent by regulating the pH in the mobile phase. Eluent from the LC column passes through a steel capillary with a coaxial flow of nebulizing gas (often N_2) [55]. A potential is applied between the MS inlet and the tip of the capillary, usually 2-6 kV [60, 63], which creates a strong electric field in the spray chamber. In positive ionization mode, a negative voltage is applied to the MS inlet and positive voltage is applied to the capillary tip. The negatively charged MS inlet attracts positive ions and the capillary wall attracts negative ions, leading to positive ions at the end of the open capillary. Repulsive forces between the positively charged ions overcome the surface tension and create a Taylor cone, which turns into a plume of droplets with positive ions moving towards the MS inlet [55]. The volatile solvent in the droplets is evaporated by heated inert gas (often N_2) or heated air [60], which increases the repulsive forces further, resulting in the creation of smaller droplets. This process continues until the ions are in the gas phase ready to enter the MS which is held under high vacuum [63].

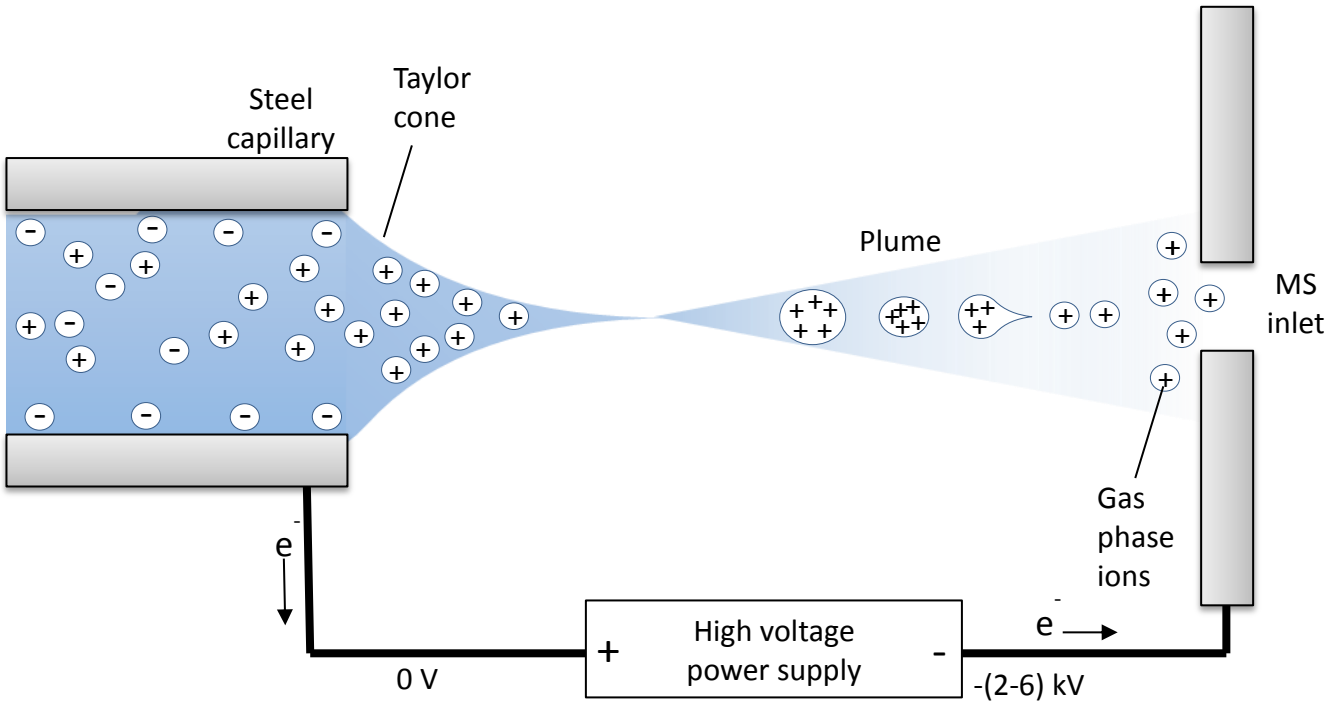


Figure 9: Illustration of electrospray ionization in positive mode. The liquid forms a Taylor cone due to the electric potential and spread into a plume of droplets. The droplets form smaller droplets due to evaporation of the solvent and repulsive forces between the ions until the ions are the in gas phase. The gas phase ions are charged and are attracted to the MS inlet. The figure was adapted from [55].

All voltages applied are opposite in negative polarity mode, resulting in the formation of negative ions in the gas phase entering the MS [55]. The ions will be detected and are called molecular ions: $[M+nH]^{n+}$ for positive ions and $[M-nH]^{n-}$ for negative ions, where M is the monoisotopic mass if the compounds, H is hydrogen and the mass of a proton and n is the number of protons donated or accepted. The ions formed are most often with -1 or +1 charge [55]. Compounds may form adducts in solution by cationization or anionization. Cations such as Na^+ are always present in ESI samples that may come from glass storage bottles, impurities or solvents. The measured molecule (M) with Na^+ adducts are represented as $[M+Na]^+$ [73].

A common problem that occurs when using electrospray mass spectrometry is ion suppression or ion enhancement. This can be caused by high concentration sample compounds of other than the analyte of interest. There are a limited number of charge sites in the droplets, making ESI a competitive process where sample compounds eluting at the same time as the analyte can be ionized at the expense of the analyte. If another compound with the same *m/z* co-elutes with the analyte of interest, signal enhancement may occur. To reduce ion suppression and enhancement, separation of the sample compounds is important to reduce the degree of ions entering simultaneously. This is often done by chromatographic techniques [60].

2.5.3 High-performance liquid chromatography

High-performance liquid chromatography (HPLC) is an analytical technique that separates compounds in solution. A mobile phase transports sample compounds through a closed column containing a stationary phase using high pressure generated by a pump. Each compound in the sample will interact differently with the stationary phase and the mobile phase based on their molecular structure and composition [55]. The compounds are separated in time, and compounds with high affinity to the stationary phase will be strongly retained, thereby eluting late from the column. The time a compound takes from the sample introduction to the elution from the column is called the compound's retention time [74].

The basic HPLC instrumentation consists of solvent reservoirs, pumps, an injector, one or several columns, a detector and a data system, see **Figure 10.** The reservoirs consist of mobile phase and the pumps deliver the mobile phase at a given flow. The sample is introduced into the mobile phase flow by the injector, most often automatically, but can be manual, and the sample compounds are separated in time by the analytical column filled with stationary phase. The compounds are detected by the detector and the signals are visualized as a function of time and presented as a chromatogram. The data system controls the instrumentation and handles the data [74].

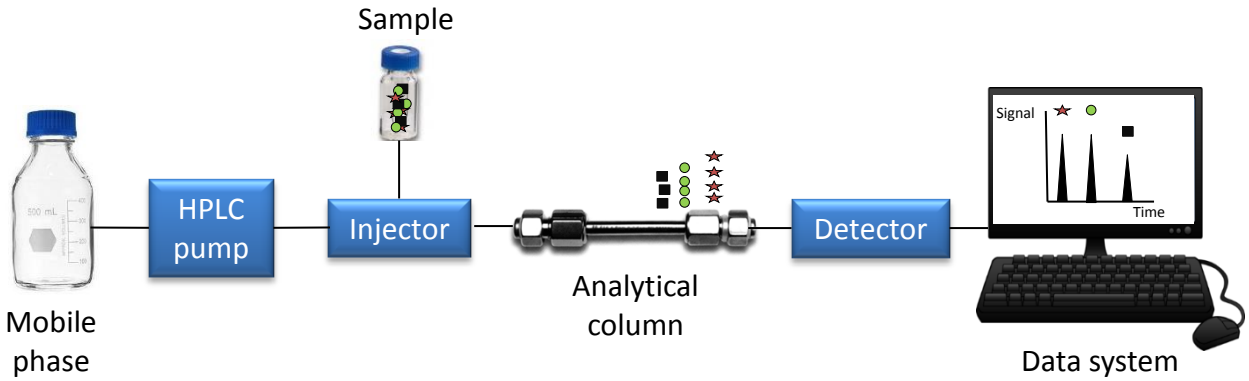


Figure 10: General high-performance liquid chromatography instrumentation consists of reservoirs with mobile phase, pumps that deliver the mobile phase, a manual or automatic injector where the sample is introduced, analytical column(s) that is filled with stationary phase and separates the compounds, a detector that detects the signal of the compounds and a data system that controls the instrument and handles the data. The figure was adapted from [55].

There are several different separation principles in HPLC. The stationary phase used determines how the compounds are separated. Silica is the most used material to prepare

stationary phases in HPLC, forming spherical porous beads of uniform particle size. The silica itself can be used as a stationary phase, or functional groups can be covalently bonded to the surface. Normal phase chromatography uses polar stationary phase (e.g., silica and silica with attached cyano or amino groups) and non-polar mobile phase (e.g., hexane), thus the polar compounds are the most retained. Hydrophilic interaction liquid chromatography (HILIC) uses polar stationary phase, as in normal phase, but with polar mobile phase (e.g., acetonitrile and methanol). In HILIC, the polar compounds will be the most retained. Reversed phase liquid chromatography is the most used separation principle in HPLC and is the reverse principle of normal phase chromatography. Non-polar stationary phases (e.g., octyl (C8) and octadecyl (C18)) and polar mobile phases are used. Here, the non-polar compounds are the most retained. The most common stationary phase is silica-based C18 materials [74].

A constant composition of the organic mobile phase and water may be used in the separation, called isocratic elution, but this is not beneficial when separating compounds with highly different hydrophobicity (e.g., biological samples). Gradient elution is when the mobile phase composition changes throughout the separation. The solvent gradient usually starts with a low organic amount (2-5 %) in water and gradually increases the amount of organic solvent [74]. This can reduce the retention of the highly retained compounds and give more narrow peaks, resulting in faster and better separation. Peak capacity is a measure of the efficiency of the column, how well it separates close peaks and is defined by the number of peaks separated within the gradient time [75]. Reversed phase liquid chromatography using gradient elution is the most commonly used separation technique in metabolome analysis and was used in this project [47].

HPLC – ESI – Q Exactive Orbitrap MS was used as the analytical platform in this project. Data acquisition by the instrument is an important factor for data processing and metabolite identification [76].

2.6 Untargeted metabolomics workflow

There are multiple ways to perform untargeted metabolomics using LC-MS [77]. A general workflow includes the preparation of a sample, data acquisition, data processing and statistical analyses of the data generated to reduce the complexity and reveal trends, use the information to identify metabolites and map them to their biochemical pathway. A hypothesis is hopefully generated and an experimental validation with quantitative measurements may be done to validate the identification [47, 76]. See **Figure 11** for an illustration of a general untargeted metabolomics workflow.

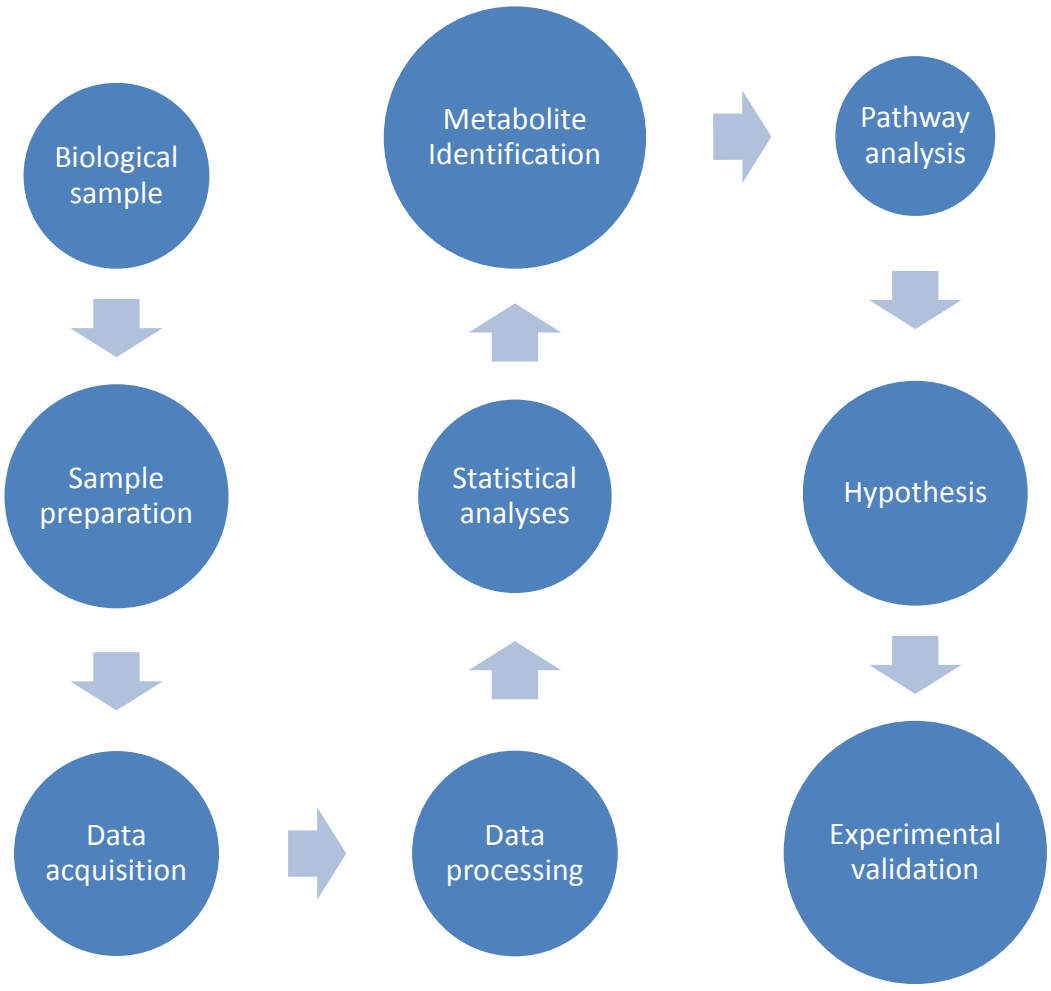


Figure 11: A general untargeted metabolomics workflow is hypothesis generating. A biological sample is prepared before data acquisition. The data are processed and statistically analyzed. Metabolites of interest are identified and their biochemical pathways are studied resulting in a hypothesis that may be validated with established methods.

Untargeted metabolomics LC-MS analysis compares peaks from a single ion (e.g., $[M+H]^+$, $[M-H]^-$ and $[M+Na]^+$), called features. The features detected must be grouped to their compound prior to metabolite identification [69]. Data acquisition and data processing methods are important factors of the workflow for metabolite identification.

2.6.1 Data acquisition

The analytical aim of untargeted metabolomics is to maximize the number and diversity of detected metabolites with different concentrations and in addition achieve high precision and good repeatability and reproducibility [78]. The data acquisition by the instrument is managed using software. A method is set up and a sequence file is generated which determines the samples that will be analyzed. The performance of the analytical method may change over time, especially for long runs, due to the need for periodical cleaning and calibration. The use of control samples or internal standards facilitates observation of possible analytical variations during a run [79].

Quality control samples

A quality control (QC) sample is a sample that is injected multiple times throughout an experiment to control the analytical drifts. The relative standard deviation of peak areas and retention times are commonly observed to evaluate the analytical performance [47]. All samples in a run pooled together are commonly used as a QC sample, but one sample or standards may also be used. The relative standard deviation peak areas in QC samples under 20 % is considered to be good reproducibility [79].

Internal standards

The use of quality control samples makes it possible to discover changes in instrument performance over time. A possibility to control the instrument signal variation in each sample analysis is to add internal standards (IS) of predetermined concentrations to each sample. The *m/z*, retention time, peak shape and peak area of each IS compound in each sample are observed and compared with expected values to determine parameter drift over time [78].

ISs are commonly used in targeted LC-MS metabolomics analyses to quantify compounds [47, 65]. ISs are added to the samples and, ideally, they have the same physical and chemical

properties as the analyte, meaning that the analyte of interest and the ISs are affected by instrument performance in the same way. Instead of using an absolute response, the ratio of response between the analyte and the IS is used. The analyte/IS response ratio is used to compensate for losses during sample preparation, volume injection variations and matrix effects as ion suppression and enhancement. The use of ISs improves the accuracy and precision of measurements [80]. This would be ideal for untargeted metabolomics as well, but, unfortunately, this is not so easily done. First, there are too many metabolites, and second, there are many unknown metabolites detected. It is impossible to use internal standards and specific calibration curves for all analytes. It is therefore impossible to achieve absolute quantification, accuracy and precision and detection limits [78]. ISs for each compound group (e.g., amino acids, sugars, nucleotides) and ISs grouped by retention time have been used in untargeted metabolomics approaches without complete success [81].

Full scan and tandem mass spectrometry

Data acquisition by LC-MS in untargeted metabolomics typically relies on full scan analyses where metabolites of interest are identified and confirmed with targeted MS^2 analysis. This procedure may be time- and sample-consuming, thus, multi-event acquisition which detects, quantifies and performs structure analysis in one single run is of high interest [76].

There are two commonly used multi-event acquisition methods used to acquire data in untargeted LC-MS based metabolomics: data-independent acquisition and data-dependent acquisition.

All ion fragmentation is a data-independent acquisition method which fragments all precursor ions in a selected *m/z* range simultaneously. This acquisition mode produces complicated fragmentation spectra which make it difficult to link the precursor and fragment ions in untargeted analyses. Sequential window acquisition of all theoretical fragment-ion spectra systematically fragments all precursor ions in predefined retention time and *m/z* ranges and facilitates the coupling of precursor and fragment ions [76]. Fragment ions are matched with precursor ions based on retention time and *m/z* [77].

The MS switches automatically from full MS to MS^2 in data dependent acquisition when precursor ions fulfill the predefined criteria such as a minimum intensity. All ions above the intensity threshold are fragmented [76]. Low abundant metabolites will not be fragmented in

data dependent acquisition mode. The choice of the intensity threshold is of high importance since relevant compounds are not selected for fragmentation if they do not exceed the intensity threshold and MS^2 spectra of irrelevant ions with high intensity might be generated, resulting in a complex fragmentation – precursor puzzle. Collection of many MS^2 spectra may result in an insufficient number of data points of a chromatographic peak [76]. Most, if not all, metabolomics fragmentation uses low energy collision (0-100 eV) to not produce many fragments resulting in a big puzzle of fragments [11]. **Figure 12** illustrates the acquisition modes.

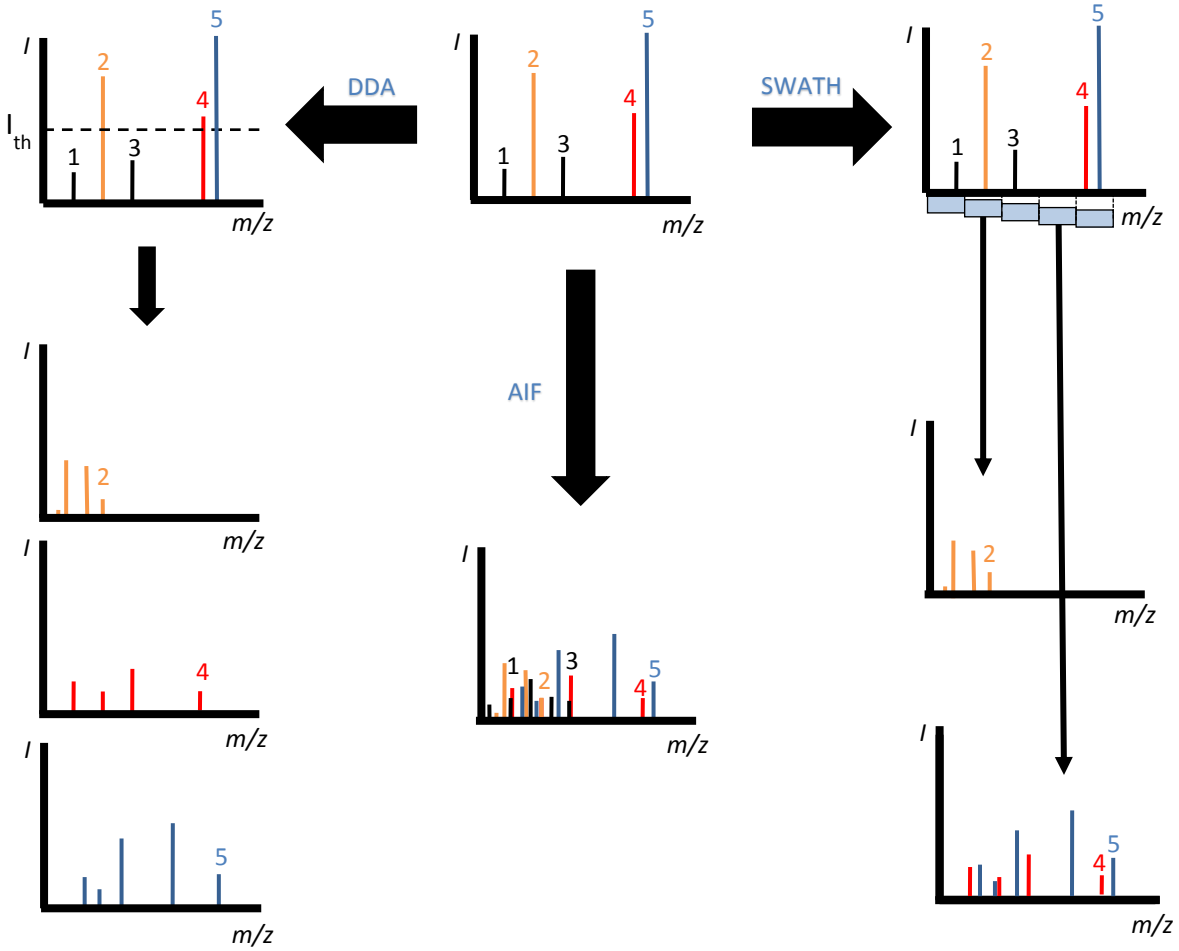


Figure 12: Schematic illustration of some acquisitions modes of Q Exactive Orbitrap MS: data-dependent acquisition (DDA), sequential window acquisition of all theoretical fragment-ion spectra (SWATH) and all ion fragmentation (AIF). Mass spectra with intensity (I) over mass-to-charge ratio (*m/z*). I_{th}: Intensity threshold. The figure was adapted from [76].

2.6.2 Data processing

The first goal of data processing is to correct for mass and retention time shifts using the observed signal from the internal standards or QC samples [79]. The ionization process may

generate several different ions from the same neutral molecule, with the result that the variables measured are often features and not metabolites. Retention times are used to group the features together as the same metabolite, including isotope ratios and common adducts. This is often done using algorithms with comprehensive software [79]. There are several bioinformatics software programs that generally provide the steps of spectral alignment and statistical analysis in a relatively automated process [49].

Several univariate and multivariate statistical methods can be used to the desired study analysis of the metabolite features. Univariate methods (e.g., t-test and ANOVA) analyze metabolomics features independently, features from the same metabolite will, therefore, be considered to be different metabolites. Multivariate methods take all the metabolomics features into account simultaneously, thus relationship patterns of features may be seen. The multivariate analyses that are unsupervised summarize the complex metabolomics data to detect data patterns, while the supervised identify metabolic patterns correlated with a specific phenotype. Principal component analysis (PCA) is the most used unsupervised method and partial least square (PLS) is the most used supervised method in untargeted metabolomics [19].

PCA analysis is a technique used to emphasize variation and bring out strong patterns in a dataset. PCA analysis results in a PCA plot which is used for easier exploration and visualization of the data. The distance between the points indicates the difference between them. The closer they are together, the more similar. In untargeted metabolomics, one point represents all detected metabolites in a specific sample [19].

The identification of new biomarkers is usually determined using a supervised analysis, such as PLS. Components in a PLS plot represent the measure of how much a feature contributes to the differentiation [19]. Differential analysis using volcano plot calculate the ratio of metabolites in two compared groups and the P-value to determine the metabolites that contribute to the variation and the significance of the change.

2.6.3 Metabolite identification

The identification of metabolites is the bottleneck of untargeted metabolomics [19, 46, 47, 69]. The identification method determines the level of metabolite identification confidence. The lowest level of confidence in the identification of metabolites is to use the measured mass

± ppm to determine a particular molecular formula. A potential composition may be determined and searched for in databases, which may result in a large number of potential metabolites. The number of hits is reduced with additional information about the retention time, isotope pattern, charge states, common adduct ions and fragmentation patterns from a MS^2 analysis. The highest level of metabolite identification confidence is when the metabolite is confirmed by a targeted MS^2 analysis of a reference compound and/or orthogonal information obtained by the analysis of the sample with different separation techniques and/or analytical techniques, such as NMR [49]. Confidence of the identification increases when the identified metabolite can be related to metabolic pathways [82]. Many metabolite features do not match a specific compound, thus searches in more databases or more data from other analytical techniques may help [49].

2.7 Untargeted metabolomics of dried blood spots

Detecting as many metabolites as possible in a biological sample is not a straightforward process. Several untargeted metabolomics approaches have been successfully developed to analyze biological samples using different analytical platforms. Miller et. al identified about 900 compounds where about 500 were known endogenous metabolites and successfully screened for 20 out of 21 IEMs tested using GC-MS and LC-MS with positive and negative ionization mode in one single plasma sample [20]. Kennedy et. al were able to identify about 1200 metabolites from about 100 urine samples, including biomarkers for 16 out of 18 IEMs tested using four different analytical platforms; single quadrupole MS, reversed phase LC-MS in positive ionization mode, and HILIC-MS in negative ionization mode [24]. Coene et. al diagnosed 42 out of 46 IEMs correctly in a plasma sample from each patient using Quadrupole-TOF (Q-TOF) MS in positive and negative ionization mode [25]. Petrick et. al found more than 1000 features using LC-Q-TOF MS in positive and negative ionization mode in punches of DBSs containing about 8 µL blood with additional measurement of potassium to correct for the hematocrit values [83].

At the National Unit of Screening and Diagnosis of Congenital Pediatric Metabolic Disorders at Oslo University Hospital, an untargeted metabolomics method for urine with HPLC-Q-TOF-MS was used as a starting point for the development of an untargeted metabolomics method for DBS analysis with HPLC-Orbitrap-MS [84, 85]. The sample preparation method of DBSs was developed based on the sample preparation method used at the Department of

Newborn Screening at Oslo University Hospital [86]. This method was developed with the purposes of achieving the best possible separation of as many compounds as possible and maximizing sensitivity and coverage of the metabolome [87, 88]. The developed untargeted metabolomics method of DBS samples was evaluated in this project.

An untargeted metabolomics method of DBS samples is wanted as the Department of Newborn Screening has archived DBS samples from many newborns in Norway with known and unknown IEMs. Samples from infants with a known IEM can be used to perform untargeted metabolomics methods to get a wider biochemical picture of an IEM. The untargeted metabolomics method is wanted to see the wider biochemical consequences of IEMs to narrow the gap between genotype and phenotype.

2.7.1 Lack of thorough evaluation of untargeted metabolomics methods

Targeted bioanalytical methods are commonly validated using, for example, criteria from FDA, European Medicines Agency or EuroChem. The validation of an untargeted metabolomics method is a difficult task. Many untargeted metabolomics studies determine the proof of concept of their methods by comparing sample groups to identify metabolites that stand out [20, 24, 25, 83]. Generally, there is a lack of thorough evaluation in untargeted metabolomics methods. It is necessary to know the limitations of the method as well as the strengths. Each step of the untargeted metabolomics workflow affects the number of metabolites detected and the quality of the data [16]. The evaluation of the coverage of the metabolome by the method is important to explain the metabolites not detectable. The evaluation of the peak area and retention time repeatability is important to achieve accurate measurements and confidently identify compounds [47]. Untargeted metabolomics methods are not absolute quantitative [78], but the quantitativity should be evaluated as the amounts of metabolites in groups are compared. Data workflow including acquisition and processing are important factors to evaluate as it is the foundation for metabolite identification [82].

The ability of the method to detect changes in metabolic profiles is obviously important, but the result may be easier to explain when all these factors are evaluated. In untargeted metabolomics, the evaluation of all these factors can be performed using a carefully chosen selection of metabolites representing several classes of compounds with the same physical and chemical properties, rather than evaluating each and every single metabolite.

2.8 Aim of study

Early diagnosis of IEMs can be crucial to avoid irreversible damage and death. Targeted metabolomics approaches are most often used in diagnostics. Untargeted metabolomics are hypothesis generating and may result in earlier diagnosis and, most importantly, saved lives. An ideal untargeted metabolomics method identifies and quantifies all metabolites in a biological sample, but the metabolites measured are restricted by the analytical techniques and the biological material used.

The aim of this study was to complete and evaluate the ability of a previously developed HPLC-Orbitrap MS method, using one punch from a DBS, to cover the DBS metabolome and to give repeatable and quantitative measurements, to establish an untargeted metabolomics data workflow, and to show the method's suitability for detection of changes in the metabolome after normal physiological changes. The evaluation was done by investigating the method's ability to detect a set of chosen compounds, to perform repeatable measurements and to perform quantitative measurements by analyzing 1, 2, 3 and 4 punches of a dried blood spot (equivalent to about 3, 6, 9 and 12 µL whole blood). To evaluate the method's ability to detect changes in the metabolome after normal physiological changes, samples taken from healthy volunteers after free diet intake, overnight fasting (12 hours) and prolonged fasting (36 hours) were compared. **Figure 13** shows the untargeted metabolomics workflow used in this project with factors that were evaluated in this thesis.

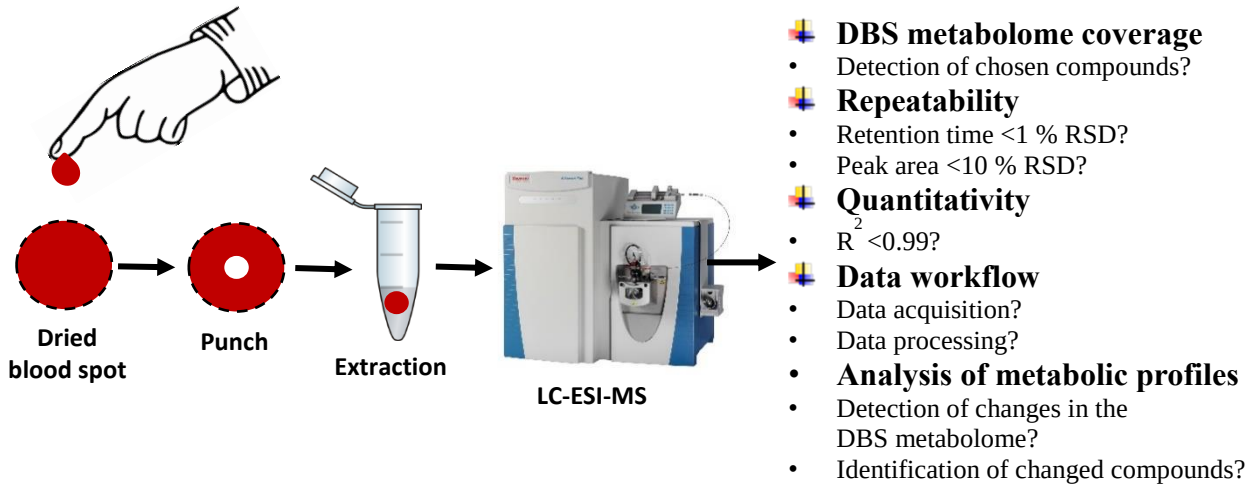


Figure 13: The untargeted metabolomics workflow: sampling, sample preparation, HPLC-MS analysis, and factors that were evaluated in this project. RSD: Relative standard deviation. R^2=1 when there is a linear correlation.

3 Experimental

3.1 Small equipment

All pipettes used were PIPETMAN purchased from Gilson, Inc. (Middleton, WI, USA). The analytical balance weight used was AG 245 from Mettler-Toledo (Columbus, OH, USA).

3.2 Chemicals

3.2.1 Solvents

Methanol (MeOH) (≥ 99.9 % purity) and formic acid (FA) (98-100 % purity) were obtained from Merck (Dramstadt, Germany). All water used was type 1 water (resistivity of 18.2 MΩ•cm at 25 ˚C) taken from a Millipore Milli-Q purification system with a Q-guard cartridge, a Quantum cartridge and a filter membrane with 0.22 µm pores purchased from Merck.

3.2.2 Reagents

Creatinine was obtained from Merck. Uric acid was purchased from Sigma Aldrich (Dramstadt, Germany). Creatine was obtained from Nutritional Biochemical Corporation (Twinsburg, OH, USA). C2-, C12- and C16 acylcarnitine were purchased from Larodan (Malmö, Sweden).

Isotopically labeled internal standards were obtained from Cambridge Isotope Laboratories, Inc (London, England). Three sets of a dry mixture in a vial, of 25 isotopically labeled internal standards were purchased. One set of 12 isotopically labeled amino acids (^{15}N; 2-^{13}C-Glycine, $^{2}H_4$-Alanine, $^{2}H_8$-Valine, $^{2}H_3$-Leucine, $^{2}H_3$-Methionine, $^{13}C_6$-Phenylalanine, $^{13}C_6$-Tyrosine, $^{2}H_3$-Aspartate, $^{2}H_3$-Glutamate, $^{2}H_2$-Ornithine, $^{2}H_2$-Citrulline and $^{2}H_4$; 5-^{13}C-Arginine) (NSK-A), another set of 8 isotopically labeled carnitine and acylcarnitines ($^{2}H_9$-C0, $^{2}H_3$-C2, $^{2}H_3$-C3, $^{2}H_3$-C4, $^{2}H_9$-C5, $^{2}H_3$-C8, $^{2}H_9$-C14 and $^{2}H_3$-C16) (NSK-B) and a third set of 5

isotopically labeled acylcarnitines ($^{2}H_3$-C18OH, $^{2}H_3$-C5OH, $^{2}H_9$-C12, $^{2}H_3$-C18 and $^{2}H_3$-C16OH) (NSK-B-G1).

3.2.3 Solutions

Standard solution

A mix of creatinine, creatine, uric acid and acylcarnitines C2-, C12- and C16 was made. 13.940 mg creatinine (1.233 mM), 5.931 mg uric acid (0.353 mM), 1.831 mg creatine (0.1397 mM), 1 mL 0.303 mM (30.3 µM) acylcarnitine C2, 1 mL 0.053 mM (5.3 µM) acylcarnitine C12 and 1 mL 0.045 mM (4.5 µM) acylcarnitine C16 were dissolved in water to a final volume of 100 mL. The concentrations in the parenthesis are the final concentrations in the standard solution (stored at 4 ˚C). The solution is further referred to as "standard solution" in this thesis.

Extraction solution

The extraction solution consisted of 80 % methanol (MeOH), 20 % water with 0.1 % formic acid (FA) and was stored at 4 ˚C.

Reconstitution solution

The reconstitution solution consisted of 98 % water, 2 % MeOH with 0.1 % FA and was stored at 4 ˚C.

Mobile phases

Mobile phase A consisted of water with 0.1 % FA. Mobile phase B consisted of MeOH with 0.1 % FA. Mobile phases were stored at room temperature (about 30 ˚C) after preparation. Mobile phases were prepared prior to each experiment.

Calibration solutions

Pierce LTQ Velos ESI Positive Ion Calibration solution and Pierce ESI Negative Ion Calibration solution were obtained from Thermo Fisher Scientific (Waltham, MA, USA) (stored at -18 ˚C).

Acylcarnitine standard solution

All compounds referred to as CX are acylcarnitines with X carbon atoms attached. Acylcarnitines (C0-C18) were purchased from European Research Network for evaluation and improvement of screening, Diagnosis and treatment of Inherited disorders of Metabolism (Winterswijk, The Netherlands), and stored at 4 ˚C. The standard concentration in µmol/L and the reference range in blood of the acylcarnitines are shown in **Table 2.**

Table 2: Concentrations of acylcarnitines in the acylcarnitine standard solution.

Acylcarnitine	Standard concentration (µmol/L)	Ref. range in blood (µmol/L)		Acylcarnitine	Standard concentration (µmol/L)	Ref. range in blood (µmol/L)
C3	0.41	0.08-1.2		C10	0.21	0.02-0.33
C4	0.32	0.06-0.36		C12	0.063	0.3-0.016
C5	0.21	0.05-0.54		C14	0.03	0.02-0.12
C6	0.041	0.011-0.094		C16	0.094	0.03-0.51
C8	0.14	0.02-0.22		C18	0.033	0.02-0.07

Internal standard solution

NSK-A, NSK-B and NSK-B-G1 were reconstituted in the vial with 1 mL of 80 % MeOH in water and mixed well to make concentrated standards. The three standards were mixed and diluted in 9 mL of 80 % MeOH in water and this diluted standard is referred to as the extraction solution with internal standards. The extraction solution with IS was used to prepare DBS samples. Concentrations of the isotopically labeled compounds in the extraction solution with IS are shown in **Tables 3-5**. The extraction solution with IS was stored in a tightly sealed vial at 4 ˚C.

Table 3: Concentration in standard solution and in extraction solution of the isotopically labeled compounds in NSK-A.

Standard solution NSK-A		**Extraction solution with IS**
Reference Standard	**Concentration (nmol/mL)**	**Concentration (nmol/mL)**
^{15}N; 2-^{13}C-Glycine	2500	208.3
$^{2}H_4$-Alanine	500	41.7
$^{2}H_8$-Valine	500	41.7
$^{2}H_3$-Leucine	500	41.7
$^{2}H_3$-Methionine	500	41.7
$^{13}C_6$-Phenylalanine	500	41.7
$^{13}C_6$-Tyrosine	500	41.7
$^{2}H_3$-Aspartate	500	41.7
$^{2}H_3$-Glutamate	500	41.7
$^{2}H_2$-Ornithine	500	41.7
$^{2}H_2$-Citrulline	500	41.7
$^{2}H_4$; 5-^{13}C-Arginine	500	41.7

Table 4: Concentration in standard solution and in extraction solution of the isotopically labeled compounds in NSK-B.

Standard solution NSK-B		**Extraction solution with IS**
Reference Standard	**Concentration (nmol/mL)**	**Concentration (nmol/mL)**
$^{2}H_{9}$-Carnitine (free carnitine, CN)	152.0	12.7
$^{2}H_{3}$-Acetylcarnitine (C2)	38.0	3.2
$^{2}H_{3}$-Propionylcarnitine (C3)	7.6	0.6
$^{2}H_{3}$-Butyrylcarnitine (C4)	7.6	0.6
$^{2}H_{9}$-Isovalerylcarnitine (C5)	7.6	0.6
$^{2}H_{3}$-Octanoylcarnitine (C8)	7.6	0.6
$^{2}H_{9}$-Myristoylcarnitine (C14)	7.6	0.6
$^{2}H_{3}$-Palmitoylcarnitine (C16)	15.2	1.3

Table 5: Concentration in standard solution and in extraction solution of the isotopically labeled compounds in NSK-B-G1.

Standard solution NSK-B-G1		**Extraction solution with IS**
Reference Standard	**Concentration (nmol/mL)**	**Concentration (nmol/mL)**
$^{2}H_{3}$-Glutarylcarnitine (C18OH)	15.20	1.3
$^{2}H_{3}$-Hydroxyisovalerylcarnitine (C5OH)	7.6	0.6
$^{2}H_{9}$-Dodecanoylcarnitine (C12)	7.6	0.6
$^{2}H_{3}$-Octadecanoylcarnitine (C18)	15.20	1.3
$^{2}H_{3}$-Hydroxypalmitoylcarnitine (C16OH)	15.20	1.3

3.3 Sample preparation of dried blood spots

The filter paper cards used in all experiments were Whatman 903 Protein Saver cards from GE Healthcare Life Science (Chicago, IL, USA). **Figure 14** shows droplets of blood spotted on a filter paper card with a punch of 3.2 mm (about 3 µL full blood) diameter from the center of the spot. The diameter of the circle on the card is 1.2 cm and holds about 75-80 µL of sample.

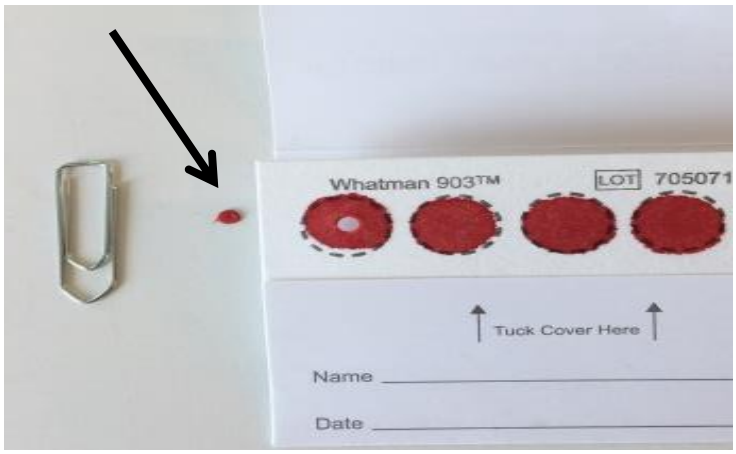


Figure 14: Dried blood on a filter paper card with a punch taken from the center of the spot.

The following procedure was used for sample preparation of all DBS samples at the beginning of the project: DBS samples were taken from a healthy volunteer. Either directly from the finger or collected in an EDTA (ethylenediaminetetraacetic acid, prevents the whole blood to coagulate) tube prior to spotting. A manual puncher (from McGill (Jacksonville, FL, USA)) was used to punch out 3.2 mm, containing approximately 3 µL whole blood [39]. A punch of clean filter paper was done in between spot punches to prevent contamination. 100 µL 80 % aqueous MeOH with 0.1 % FA was added and extraction was performed by mixing (700 rpm) for 45 minutes at 45 ˚C with a Thermomixer Comfort from Eppendorf (Hamburg, Germany). This was performed in a micro tube from Sarstedt (Nümbrecht, Germany). The sample was transferred to a glass tube from VWR (Radnor, PA, USA), and evaporated to dryness with a TurboVap LV, Caliper Life Sciences (Waltham, MA, USA) under a stream of dry nitrogen at 40 ˚C. After evaporation to dryness, approximately 10 minutes, the sample was reconstituted in 100 µL 2 % aqueous MeOH with 0.1 % FA. The sample solution was mixed on a table vortex (Genie2, Scientific Industries, Bohemia, NY, USA) and transferred to an HPLC vial (from Matriks AS, Oslo, Norway) with insert glass and caps from Matriks. The sample was placed in the autosampler and 2 µL was injected. The samples were analyzed in

positive and negative mode. However, only results from positive ionization are presented due to temporary technical problems with negative ionization.

3.4 Liquid chromatography – mass spectrometry instrumentation and settings

The liquid chromatography – mass spectrometer instrumentation used was Dionex Ultimate 3000 UHPLC system coupled to a Q Exactive Orbitrap MS, both from Thermo Scientific. The analytical column used in all experiments was Pursuit XRs C18-diphenyl (250 x 2.0 mm, particle size 3 µm) purchased from Agilent Technologies (Santa Clara, CA, USA).

The liquid chromatography settings have been optimized to achieve separation of as many compounds as possible by investigating the analytical column, mobile phase composition, gradient elution profile, injection volume, column temperature and mobile phase flow rate [87]. The mass spectrometric and electrospray settings have been optimized with the purpose of maximizing the number of compounds detected and peak intensity by investigating electrospray voltage, electrospray needle position, resolution, automatic gain control value and mobile phase flow rate [88]. The liquid chromatography, mass spectrometric and electrospray settings used are shown in **Tables 6-9** and **Figure 15**.

Table 6: Liquid chromatography settings used in all experiments.

Parameter	**Setting**
Mobile phase A	Water + 0.1 % formic acid
Mobile phase B	Methanol + 0.1 % formic acid
Gradient	See **Table 7** and **Figure 15**
Injection volume	2 µL
Column temperature	30 ˚C
Flow rate	300 µL/min
Analysis time	27.5 minutes
Re-equilibration time	10 minutes

Table 7: Flow gradient.
Mobile phase:
A: Water + 0.1 % formic acid.
B: Methanol + 0.1 % formic acid.

Time	% B
0	2
6	10
8.5	75
25	100
27.5	100

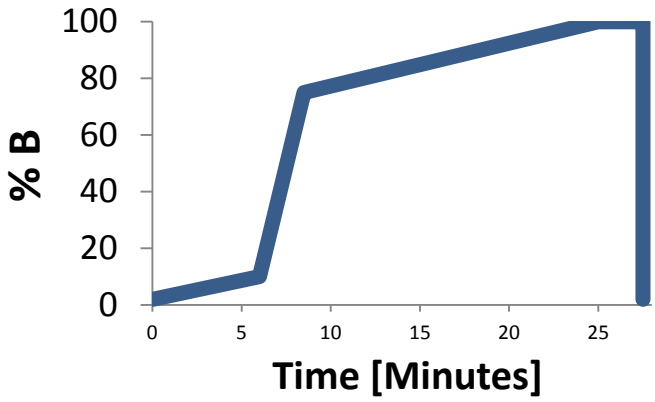


Figure 15: Flow gradient.
Mobile phase: A: Water + 0.1 % formic acid.
B: Methanol + 0.1% formic acid.

Table 8: Initial mass spectrometric settings used in this project.

Parameter	Setting
Scan type	Full MS
Scan ranges (*m/z*)	50-750 and 750-1700
Fragmentation	None
Resolution at *m/z* 200	70,000
Polarity	Positive and negative
Micro scans	1
Lock masses	Off
Automatic gain control target value	1.00E+06
Maximum injection time	250 ms
Analysis time	27.5 minutes
Re-equilibration time	10 minutes

Table 9: Electrospray settings used in all experiments.

Electrospray parameter	Setting
Sheath gas (N_2) flow rate	40 a.u.
Auxiliary gas (N_2) flow rate	10 a.u.
Sweep gas (N_2) flow rate	2 a.u.
Capillary temperature	250 ˚C
S-lens RF level	50.0
Auxiliary gas heater temperature	300 ˚C
Electrospray voltage	3.5 kV
Electrospray needle position	C
Capillary temperature	250 ˚C

3.5 Computer software

Xcalibur (Version 4.0) was used for data acquisition and processing. An automatic chromatographic peak detection strategy was used by creating layout files with exact *m/z*, using Qual Browser in Xcalibur to observe the chromatographic peaks, retention times and *m/z*. Processing SetUp in Xcalibur was used with *m/z* and expected retention time to automatically integrate the chromatographic peaks to obtain peak area.

Tune (version 2.5 Build 2042) was used to calibrate and observe the mass spectrometric parameters, and Mass Calculator in Tune was used to calculate the exact masses of the compounds. Chromeleon Xpress (version 6.80) was used to control the instruments. Compound discoverer (version 2.1 SP1) was used to process the data, perform statistics and search in online database (chemspider, mzCloud and KEGG). All the software programs are from Thermo Scientific.

4 Results and discussion

In the present study, further development and evaluation of an untargeted HPLC-ESI-MS metabolomics method of dried blood spot was done, referred to as the "initial method" [87, 88]. The workflow of the initial method is shown in **Figure 16.**

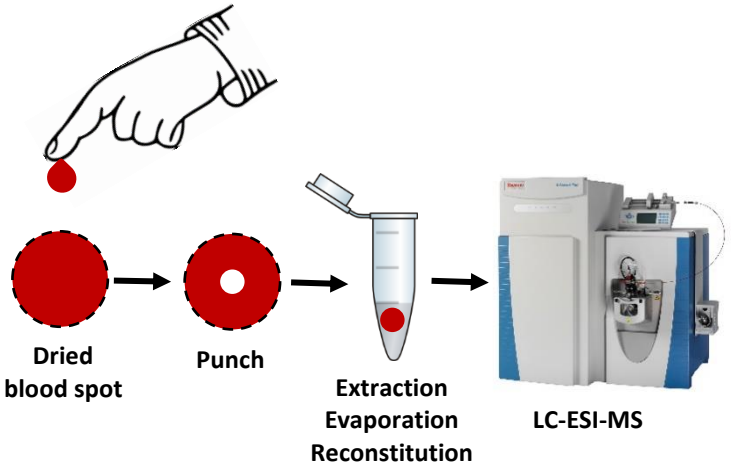


Figure 16: Workflow of the initial untargeted LC-ESI-MS metabolomics method.

The results of the experiments in this project are presented in each section below (**Section 4.2-4.8**) with a schematic conclusion at the end of each section. The schematics show the workflow of the method where the parts in black and white were held constant, while the parts in color were changed.

4.1 Framework of study

4.1.1 Initial compounds that represent the dried blood spot metabolome

The initial method was developed using DBS samples spiked with tobramycin, bilirubin, creatinine, creatine, glucose and acylcarnitines C2, C12 and C16. These compounds were chosen to obtain diversity in molecular weight and hydrophobicity, and they can all be measured by established methods at the laboratory at the National Unit of Screening and

Diagnosis of Congenital Pediatric Metabolic Disorders, Oslo University Hospital. Tobramycin is a medicament, bilirubin was not detected by the method and glucose was only detected in negative ionization mode. All the compounds except from tobramycin are present in whole blood (endogenous metabolites) [87]. In this project, endogenous metabolites in whole blood from a healthy volunteer were chosen to be observed for method evaluation. The endogenous metabolites chosen were creatinine, creatine, uric acid and acylcarnitines C2, C12 and C16, referred to as the "initial compounds", as the compounds observed in the method evaluation was changed during the project. **Table 10** shows their exact mass, molecular formula and logP value (obtained from PubChem database (https://pubchem.ncbi.nlm.nih.gov/)). The logP value indicates the degree of the hydrophobicity of a compound: the higher the value, the more hydrophobic the compound. P is the partition coefficient, that is the ratio between the concentration of a compound in a non-polar phase (e.g, octanol) and its concentration in a polar phase (e.g., water) [89].

Table 10: Exact mass, molecular formula and logP value of initial endogenous compounds used for method evaluation.

Compound	Exact mass	Molecular formula	logP
Creatinine	113.05836	$C_4H_7N_3O$	-1.8
Uric acid (UA)	168.02779	$C_5H_4N_4O_3$	-1.9
Creatine	131.06893	$C_4H_9N_3O_2$	-1.2
Acylcarnitine C2	203.11521	$C_9H_{17}NO_4$	0.4
Acylcarnitine C12	343.27171	$C_{19}H_{37}NO_4$	5.5
Acylcarnitine C16	399.33431	$C_{23}H_{45}NO_4$	7.7

4.1.2 Defined criteria of detected compounds

The minimum HPLC-MS criteria for identification of a compound were chosen based on the criteria used by the World Anti-Doping Agency (WADA) to confirm the identity of analytes for doping control purposes [90]. The chromatographic criteria for identification of known metabolites used, was that the sample compounds should not differ more than ± 0.1 minute. The mass spectrometric identification criteria used was mass accuracy set to the exact mass of the molecular ion ±5 ppm. This mass accuracy is commonly used in untargeted LC-MS based

metabolomics [91, 92]. WADA uses ±0.5 Da from the mass of standard compounds, but WADA laboratories use quadrupole mass analyzers with lower mass accuracy (±5-500 ppm) than the Orbitrap (± 0.5-1 ppm) [52]. Another mass spectrometric criteria for detection used were 10 or more data points across the chromatographic peak and peak area $>10^4$ a.u. These criteria were chosen to be able to consider the peak shape and obtain reproducible instrument signals. Relative standard deviation in % (% RSD) of the peak area and retention time was calculated after analysis of samples. Accepted % RSD of peak area and retention time were <10 % and <1 %, respectively.

% RSD for the measured peak areas and retention times of all compounds observed in all experiments are shown in **Appendix**, Tables 13-24, section 6.1.

4.2 The hydrophobic and low-abundant compounds were not detected by the initial method

4.2.1 Detection of compounds in spotted standard solution

The retention time and the *m/z* of the initial compounds were used in this project to identify the compounds. To find the retention times, the standard solution was spotted on to a filter paper card and analyzed five times by the initial method. The result was a chromatogram containing signals from all measured ions, called a total ion chromatogram (TIC). The exact masses of the molecule (M) with an added hydrogen atom $[M+H]^+$ were found in the Tune software using the Mass calculator. These masses ±5 ppm were used to extract the peaks of the compounds from the TIC, resulting in extracted ion chromatograms (EIC). The EICs with peaks, extracted *m/z* ranges and retention times observed of the charged compounds are shown in **Figure 17.**

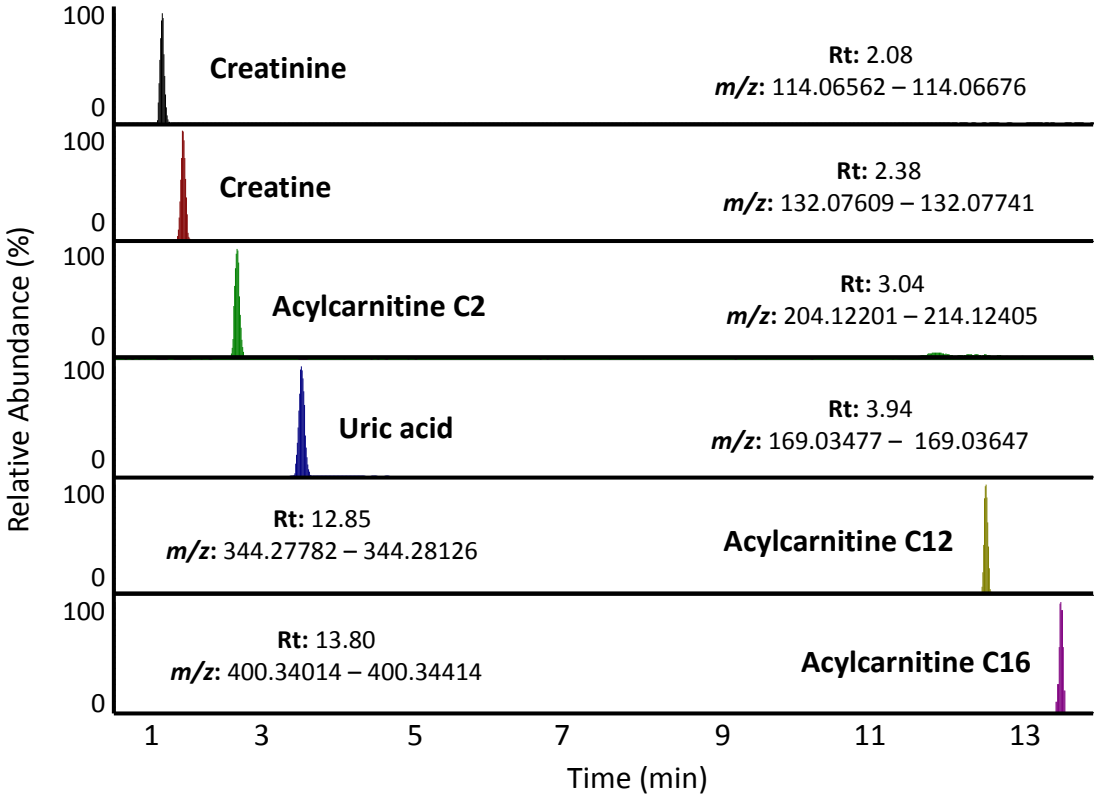


Figure 17: Extracted ion chromatograms of compounds in the standard solution analyzed by the initial method with retention time (Rt) and extracted mass-to-charge ratio range (*m/z*).

All the peaks of the chosen compounds met the criteria to be detected except for acylcarnitine C16 with 49 % RSD of peak area and 7 points across the chromatographic peak. The % RSD of the retention times and peak areas of all the other compounds was satisfactory, <0.4 % and of <5 %, respectively.

4.2.2 Detection of compounds in dried blood spot samples

One DBS sample was prepared and analyzed 20 times with the initial method. The exact masses (±5 ppm) and observed retention times (±0.1 minutes) (see **Figure 17)** of the compounds in the spotted standard solution were used to extract peaks of endogenous compounds from the TIC. The EICs with peaks of the initial compounds detected by the initial method in the DBS sample are shown in **Figure 18.**

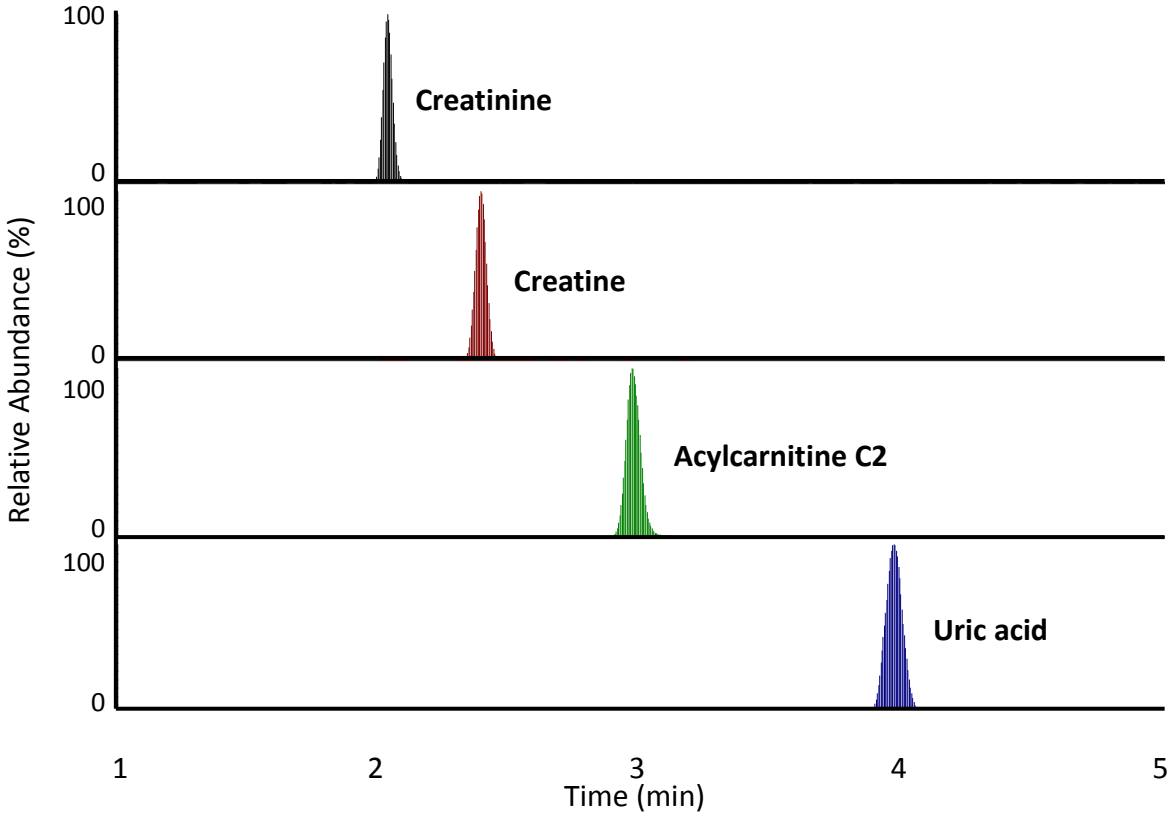


Figure 18: Extracted ion chromatograms (EIC) of the initial compounds detected by the initial method in a DBS sample.

Creatinine, creatine, acylcarnitine C2 and uric acid met the criteria of a detected peak with % RSD of peak areas <2 % and retention times <0.2 %.

The hydrophobic and low-abundant compounds were not detected by the initial method. Efforts to include these compounds were done in this project. The instrument repeatability was satisfying for the compounds detected.

4.3 Random punch locations contributed to lower assay reproducibility

The instrument repeatability of the detected compounds was satisfactory. Assay reproducibility was assessed by analyzing 20 DBS samples from one healthy volunteer with the initial method. The peak areas of each compound were observed and the % RSDs were calculated, shown in **Figure 19**.

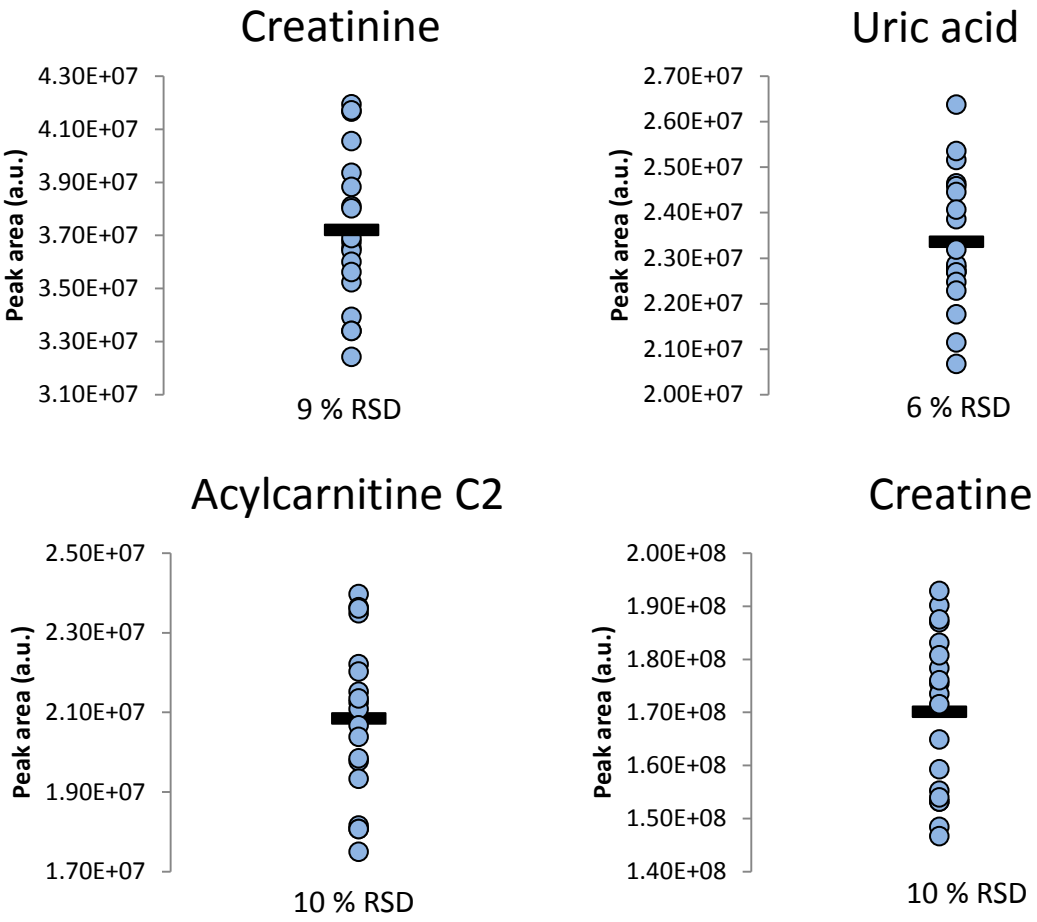


Figure 19: Assay reproducibility: Peak areas with calculated relative standard deviation (% RSD) of initial compounds detected by the initial method in 20 DBS samples.

The assay reproducibility of the initial compounds was satisfactory for all compounds detected. The 20 DBS samples analyzed were prepared with punches taken from random locations of the DBS. Analyses of DBS samples prepared from specific punch locations of the DBS were done to see if punching from one location reduced the % RSD of peak areas.

4.3.1 Center punches improved the assay reproducibility

To observe the effect of punch location, DBS samples from one healthy volunteer were prepared using 10 punches from each location; A, B, C and D, where A was in center and B and C had increased distance to D in the perimeter, see **Figure 20**.

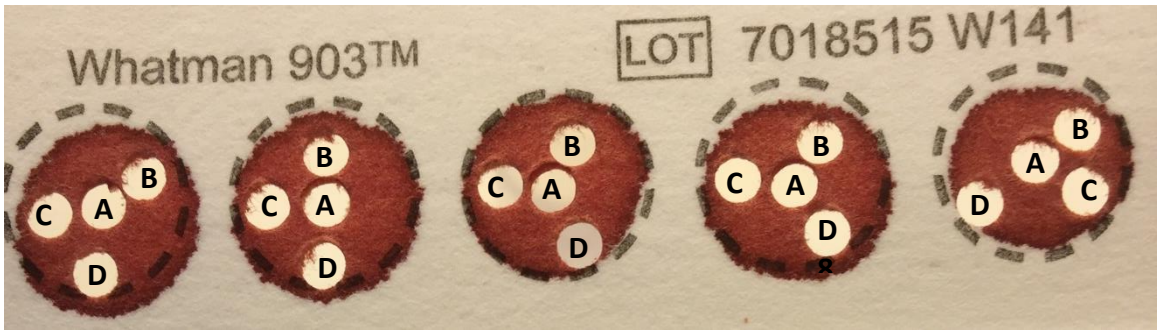


Figure 20: A filter paper card with locations of the punch used in the punch location experiment.

Figure 21 shows the peak areas of the initial compounds detected by the initial method, the calculated % RSD for each compound in the four locations and the % increase of peak area in perimeter to center punches (D/A). Creatinine and C2 had more than 10 % higher peak area in punches taken from the perimeter than from the center of the spot, while creatine and uric acid differed less than 10 %, implying that the difference may be due to instrumental variations. Uric acid was the only compound that was observed with higher peak area in the center than in the perimeter of the DBS, but the difference in peak area was only 2 %.

Larger % RSD was observed in punches taken from the perimeter than from the center of the spot. Accurate measurements were considered to be more important than high sensitivity, thus; punches from the center of the spot were taken in all the experiments that followed.

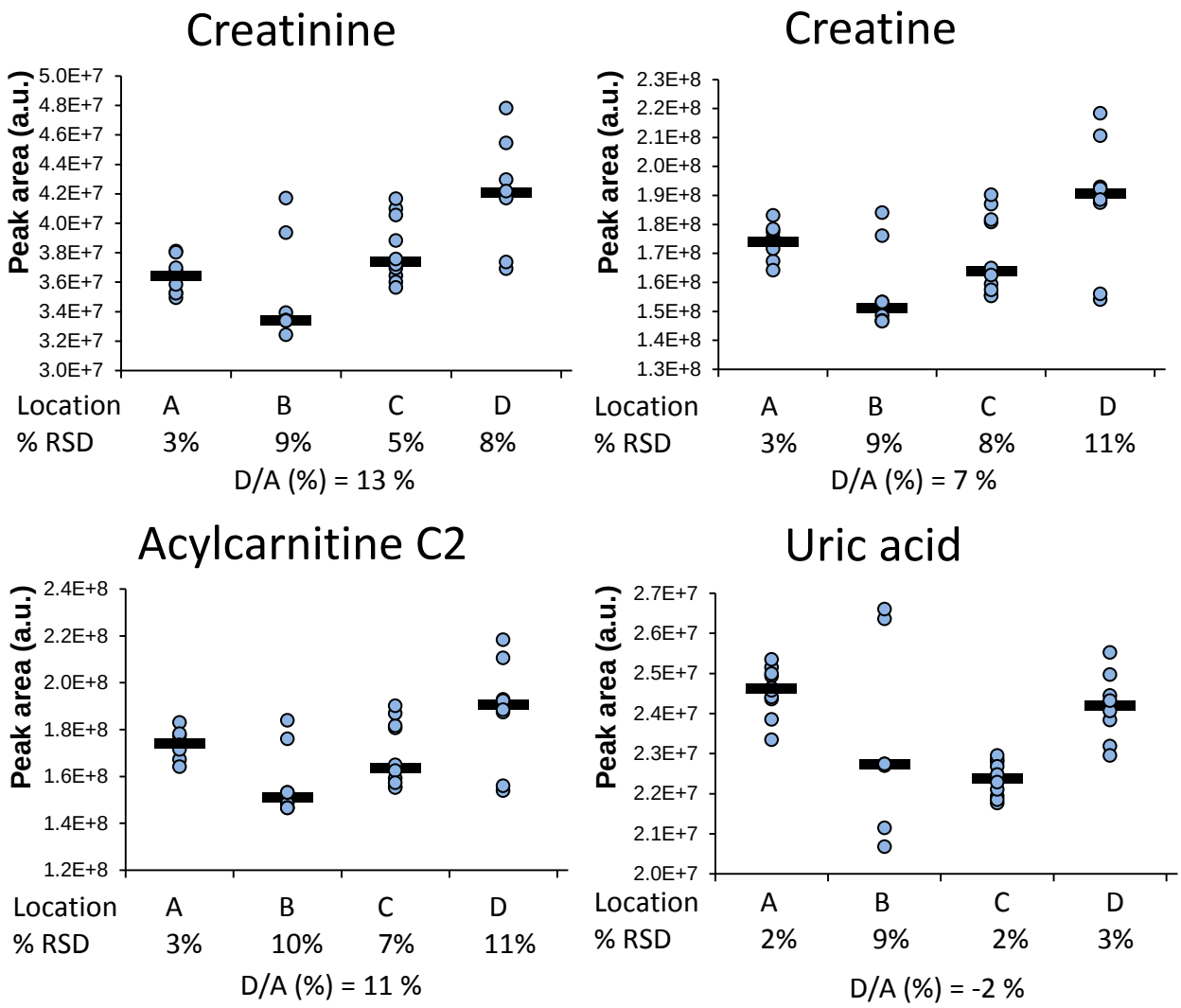


Figure 21: Peak areas with calculated % RSD of initial compounds detected by the initial method in DBS samples prepared with punches taken from location A, B, C and D. (A was in the center, and B and C had increased distance to the center, while D was in the perimeter). 10 samples from each location were prepared.

The different amount of compounds in perimeter punches compared to in center punches may be due to the inhomogeneous nature of blood, as whole blood may not distribute homogenously on the filter paper as it contains cells and plasma [93]. Samples with punches from the center and the perimeter in DBS were compared with dried plasma spots, see **Appendix**, section 6.2. The presence of red blood cells seemed to result in higher abundance of compounds in the perimeter of the DBS. This indicates that correcting for hematocrit values (amount red blood cells) may result in more accurate measurements.

To summarize, the instrument repeatability and assay reproducibility were satisfactory with % RSD of peak areas and retention times of the detected chosen initial compounds <10 % and <1 %, respectively, with the initial method. The assay reproducibility was improved by the analysis of punches taken from the center of the DBS, see ***Figure 22****.*

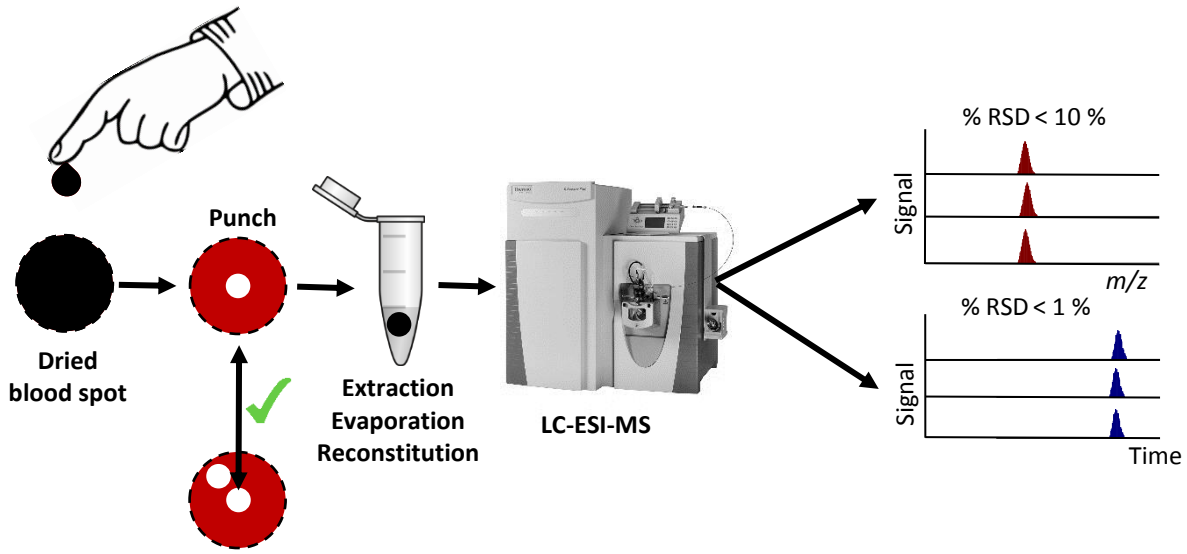


Figure 22: Center punches improved the assay reproducibility, resulting in % relative standard deviation of peak areas and retention times of the initial compounds detected by the initial method <10 % and <1 %, respectively.

4.4 Changing the initial method increased the detection of hydrophobic and low-abundant compounds

4.4.1 Increased amount of organic solvent in the reconstitution solution increased the coverage of the hydrophobic compounds

Acylcarnitine C12 and C16 were the most hydrophobic compounds of the initial compounds and were not detected by the initial method. Compounds from whole blood in a punch of a DBS were extracted by using 80 % aqueous MeOH and thermal mixing. The solution was evaporated to dryness and reconstituted in the mobile phase start gradient (2 % aqueous MeOH). Three DBS samples prepared with 40 % and 80 % aqueous methanol in the reconstitution solution, and with the exclusion of the evaporation and reconstitution steps were tested. Retention time and peak area of initial compounds were observed, see **Figure 23.**

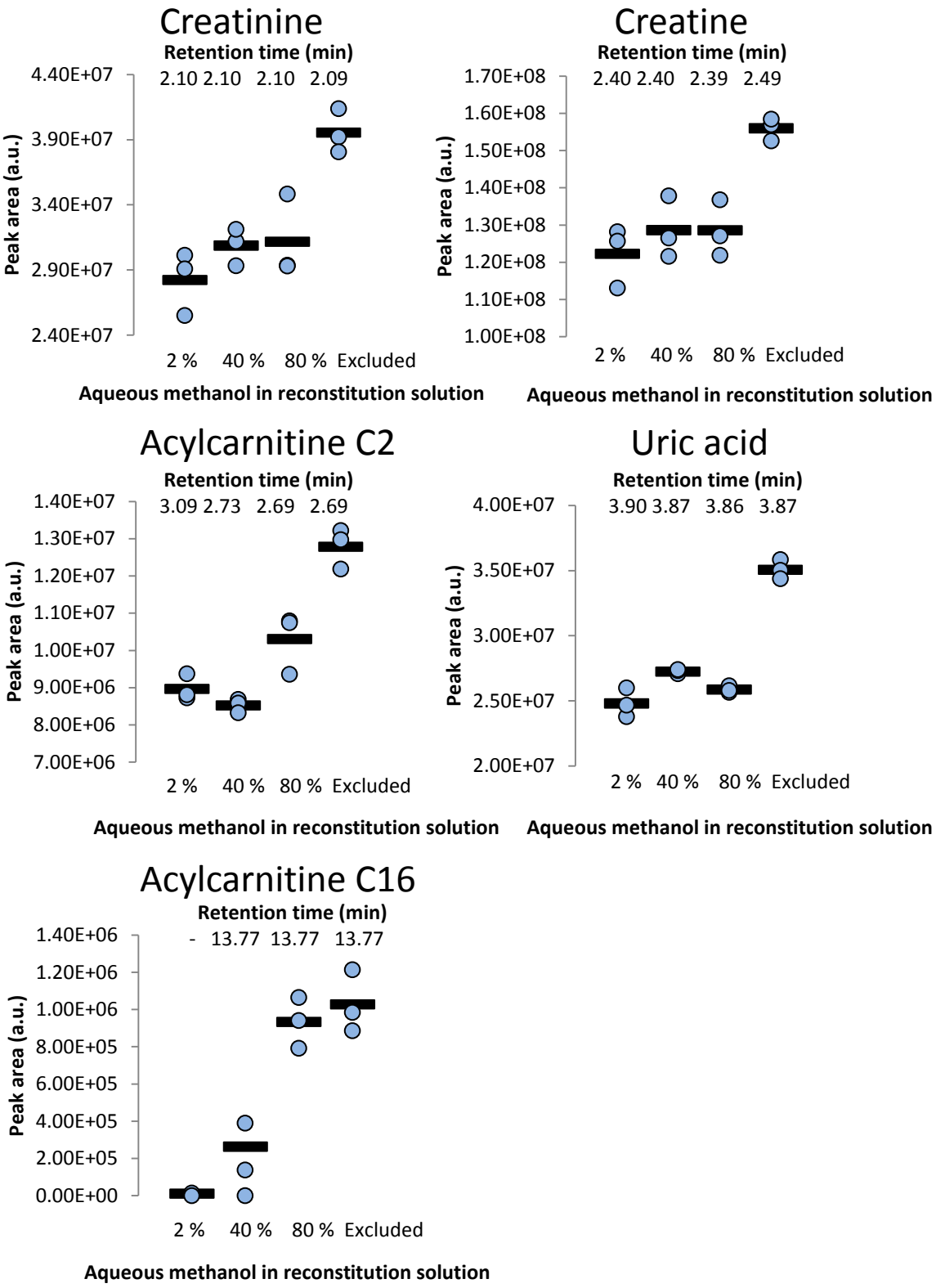


Figure 23: Peak area of compounds detected in DBS samples prepared with 2 %, 40 % and 80 % aqueous methanol in the reconstitution solution, and with the exclusion of the evaporation and reconstitution steps. With retention times of each organic solvent concentration tested.

All compounds detected increased the peak area (28-40 % increase for all compounds) when excluding the evaporation to dryness and reconstitution steps compared to the initial method. Comparison of reconstitution in 80 % organic solvent with exclusion of the evaporation and reconstitution steps showed that the amount of compounds lost was decreased. This was most likely due to evaporated compounds and/or compounds being "stuck" in the evaporation glass tube.

Acylcarnitine C12 was still not detected. C16 increased the peak area with increased amount of organic solvent amount in the reconstitution solution, but did only have 7 points across the chromatographic peak. The % RSD of peak area was about 13 %, probably due to few data points across the chromatographic peak. A standard solution containing acylcarnitines (C0-C18) was analyzed to observe the instrument method's ability to detect other hydrophobic compounds than the initial ones, see **Appendix**, Section 6.3.1. The mid-long acylcarnitines: C6, C8, C12 and C14 were detected with high peak area % RSD (15-35 %). Poorer extraction efficiency of these compounds compared to the other acylcarnitines was observed and affected the detection of low-abundant compounds, like acylcarnitines, see **Appendix**, section 6.3.2.

Decreased retention of compounds was expected, as the eluent strength increases as the percentage of organic solvent increases [55]. The retention time of acylcarnitine C2 was most affected, but the retention time of the polar compounds (creatinine and creatine) did not change. The changes in retention time were not considered to be large, most likely due to the low volume injected (2 µL). Taken this under consideration, it was decided to exclude the evaporation to dryness and reconstitution steps. This leads to less potential errors and reduction of the sample preparation time.

The poor detection of hydrophobic compounds seemed to be partly caused by solubility issues, but excluding the evaporation to dryness and reconstitution steps did not solve the problem completely.

4.4.2 Reduction of the mass range increased the coverage of the low-abundant compounds

The initial method was set to perform two scan ranges (*m/z* 50-750 and 750-1700) to cover a wide range of metabolites. However, more of the metabolites of interest are in the scan range

of *m/z* 50-750 rather than the scan range of *m/z* 750-1700. The results from one scan range of *m/z* 50-750 were compared with results from two scan ranges (*m/z* 50-750 and 750-1700) to observe if the low-abundant compounds were detected by increasing the sensitivity.

With two scan ranges, the quadrupole scans for *m/z* 50-750 and *m/z* 750-1700 every other time and passes ions within the scan range in to the C-trap which sends the ions into the Orbitrap. **Figure 24** shows the peak of carnitine (*m/z* 113.05836) with one and two scan ranges. This *m/z* was only detected when scan range *m/z* 50-750 was performed. When scan range *m/z* 750-1700 was also performed, this *m/z* gave a lower signal. The result was that the peak of this *m/z* appeared to go up and down when using two scan ranges, and looked smooth when using one.

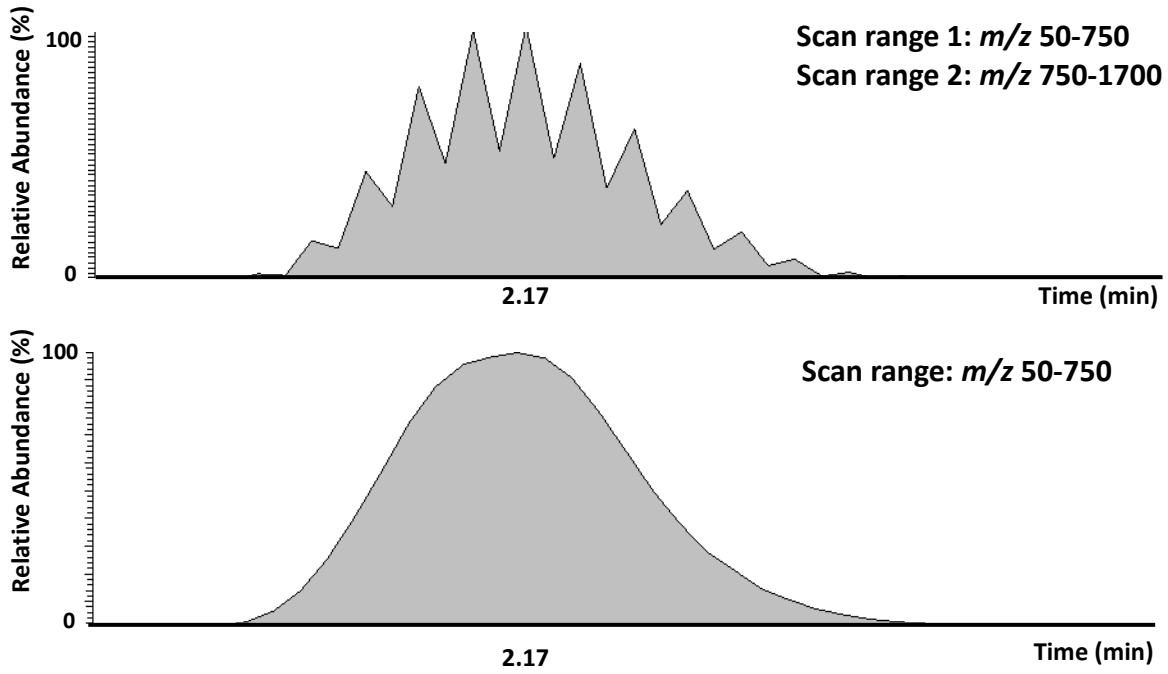


Figure 24: The peak of carnitine (*m/z* 113.05836) after analysis with two scan ranges (*m/z* 50-750 and 750-1700) and one scan range (*m/z* 50-750).

Using one scan range instead of two resulted in a doubling of peak area of all already detected compounds by the initial method and C12 was detected, see **Figure 25**. The increase in peak area from using two scan ranges to one scan range was 92-109 %**.**

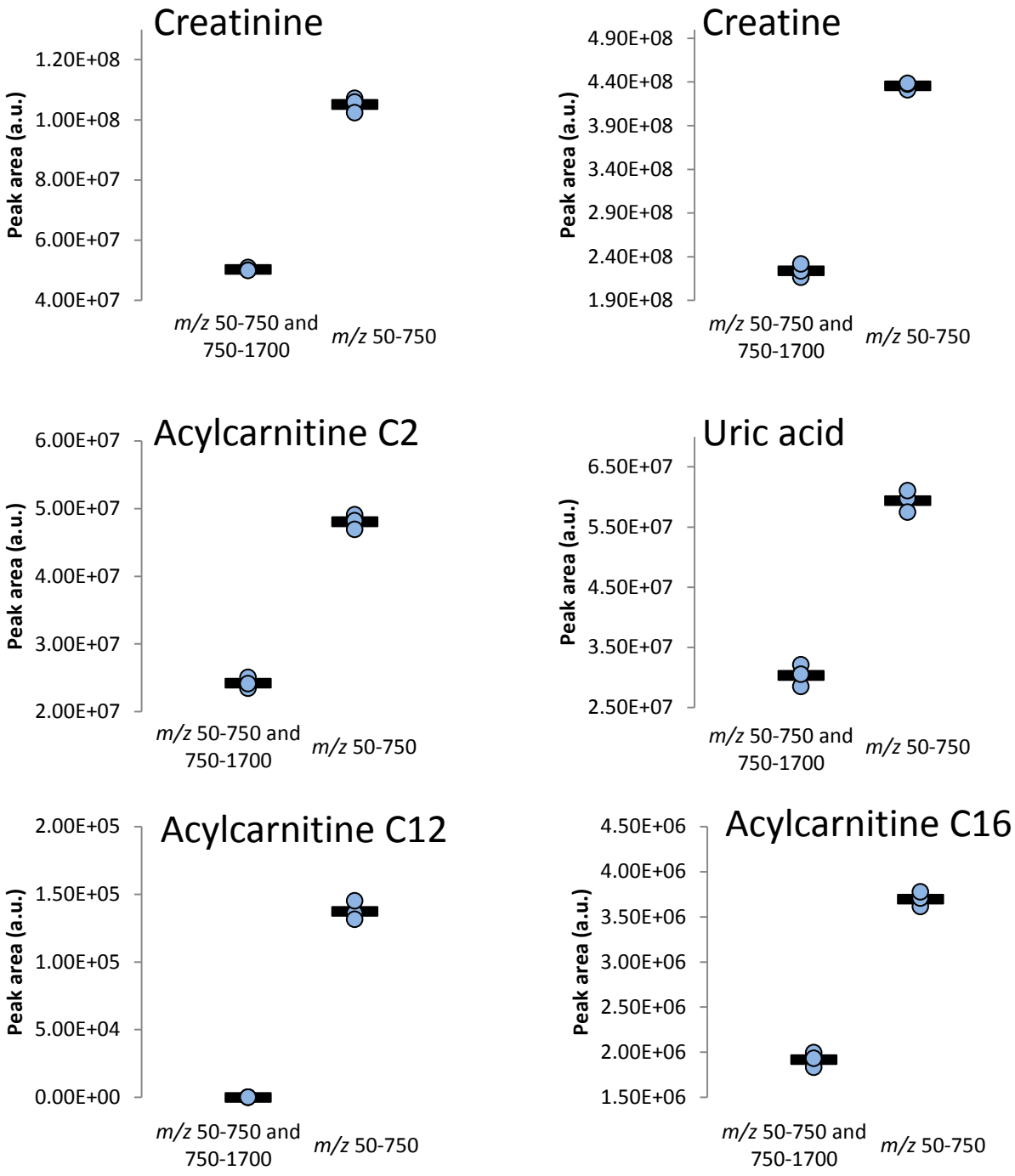


Figure 25: Peak area increase of all detected compounds when going from two scan ranges (*m/z* 50-750 and 750-1700) (initial method) and one scan range (*m/z* 50-750).

If greater sensitivity is wanted rather than a wide *m/z* range, it is preferred to use one scan range. If both ranges are preferred, two separate analyses of one sample, one with scan range *m/z* 50-750, and one with scan range *m/z* 750-1700 may be performed.

To summarize: An increased organic solvent amount in the reconstitution solution, and reduced m/z *range of the MS, increased the number of hydrophobic and low-abundant compounds detected, respectively, see* ***Figure 26****.*

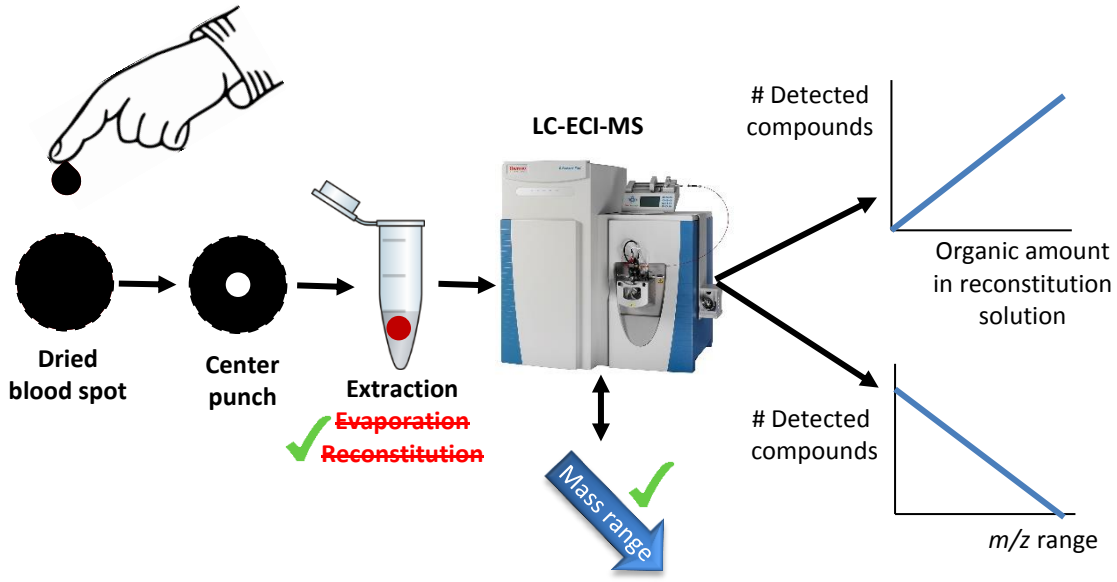


Figure 26: The exclusion of the evaporation to dryness and reconstitution steps and the decreased mass range of the mass analyzer resulted in more detected hydrophobic and low-abundant compounds.

4.5 The final method detected all the initial compounds

The initial method was able to detect all the chosen compounds except from the most hydrophobic and low-abundant, thus changes of the initial method were done to increase the coverage of the metabolome. The final sample preparation method in pictures is shown in **Appendix**, section 6.4, figure 42. **Figure 27** shows the initial and final method, where the evaporation to dryness and reconstitution steps were eliminated and the mass range was reduced, with the chromatographic peaks of acylcarnitines C6-C18 before and after the changes were made. This final method made it possible to detect all the chosen compounds and was used in all the experiments described below.

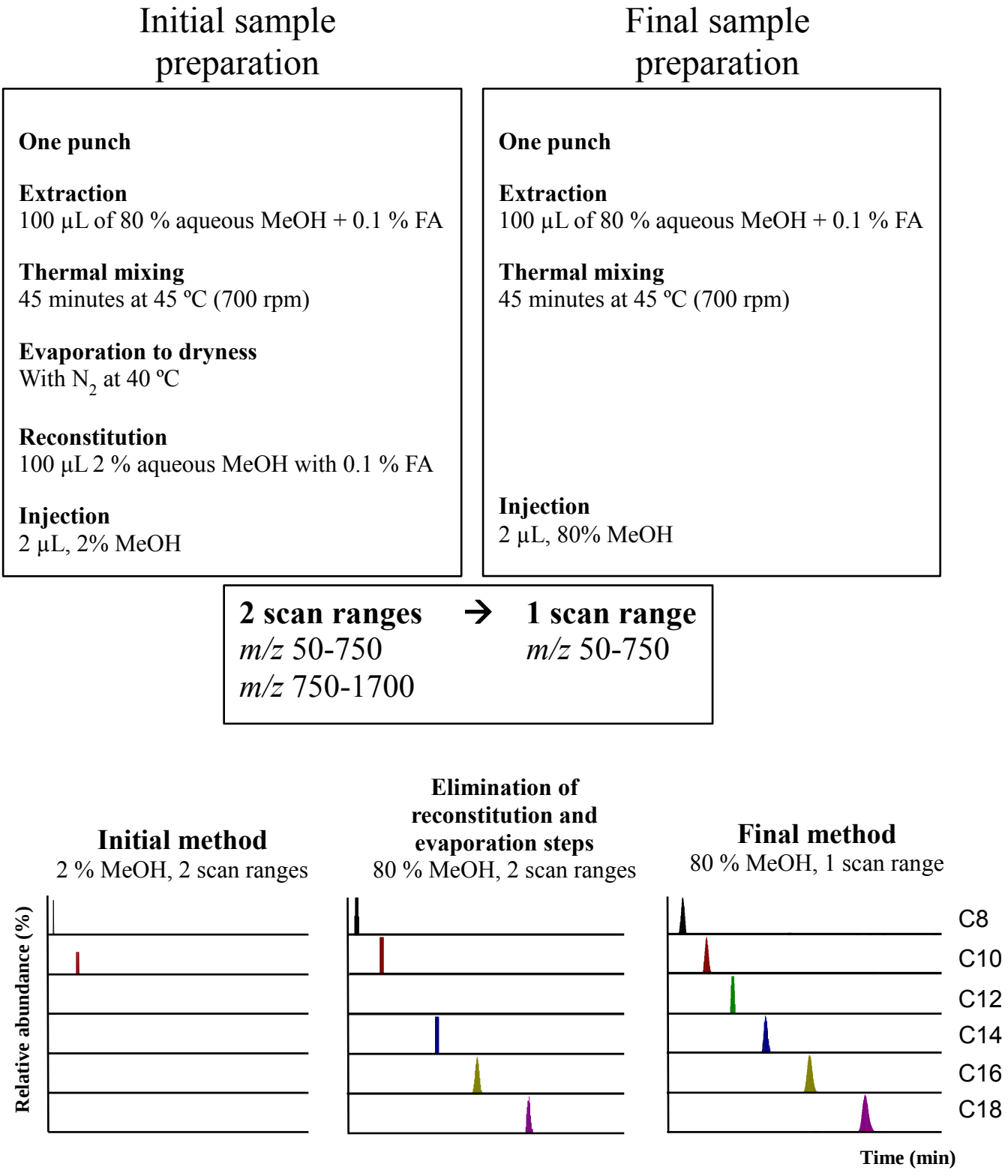


Figure 27: Initial and final method and the chromatographic peaks of acylcarnitines C6-C18 after analysis with the initial method, method with increased organic amount and the final method.

4.5.1 Final compounds that represented the dried blood spot metabolome

Isotopically labeled ISs were added to the extraction solution to control the instrumental drift. These isotopically labeled ISs have diversity in mass and polarity and have the same physical and chemical properties as the endogenous compounds in **Table 11.** It was chosen to observe the ISs' endogenous compounds, referred to as the "final compounds", when performing the evaluation of the final method. The IS compounds are shown in **Section 3.2.3,** Tables 3-5

Table 11: Final endogenous compounds chosen and their chemical formula, exact mass and logP value (obtained from PubChem database (https://pubchem.ncbi.nlm.nih.gov/)).

Amino acid	Chemical formula	Exact mass (Da)	logP		Acylcarnitine	Chemical formula	Exact mass (Da)	logP
Ornithine (Or)	$C_5H_{12}N_2O_2$	132.08987	-4.4		**C0**	$C_7H_{15}NO_3$	161.10519	-0.2
Citrulline (Cit)	$C_6H_{13}N_3O_3$	175.09569	-4.3		**C2**	$C_9H_{17}NO_4$	203.11575	0.4
Arginine (Arg)	$C_6H_{14}N_4O_2$	174.11167	-4.2		**C3**	$C_{10}H_{19}NO_4$	217.13140	0.9
Glycine (Gly)	$C_2H_5NO_2$	75.03202	-3.2		**C4**	$C_{11}H_{21}NO_4$	231.14705	1.2
Alanine (Ala)	$C_3H_7NO_2$	89.04767	-2.7		**C5**	$C_{12}H_{23}NO_4$	245.16270	1.6
Valine (Val)	$C_5H_{11}NO_2$	117.07897	-2.3		**C8**	$C_{15}H_{29}NO_4$	287.20965	3.4
Tyrosine (Tyr)	$C_9H_{11}NO_3$	181.07389	-2.3		**C12**	$C_{19}H_{37}NO_4$	343.27225	5.5
Methionine (Met)	$C_5H_{11}NO_2S$	149.05104	-1.9		**C14**	$C_{21}H_{41}NO_4$	371.3035	6.6
Leucine (Leu)	$C_6H_{13}NO_2$	131.09462	-1.5		**C16**	$C_{23}H_{45}NO_4$	399.33485	7.7
Phenylalanine (Phe)	$C_9H_{11}NO_2$	165.07897	-1.5		**C18**	$C_{25}H_{49}NO_4$	427.36615	8.8

4.5.2 All amino acids and acylcarnitines were detected by the final method

All the final compounds were detected by the final method. The EICs of the compounds are shown in **Figure 28**.

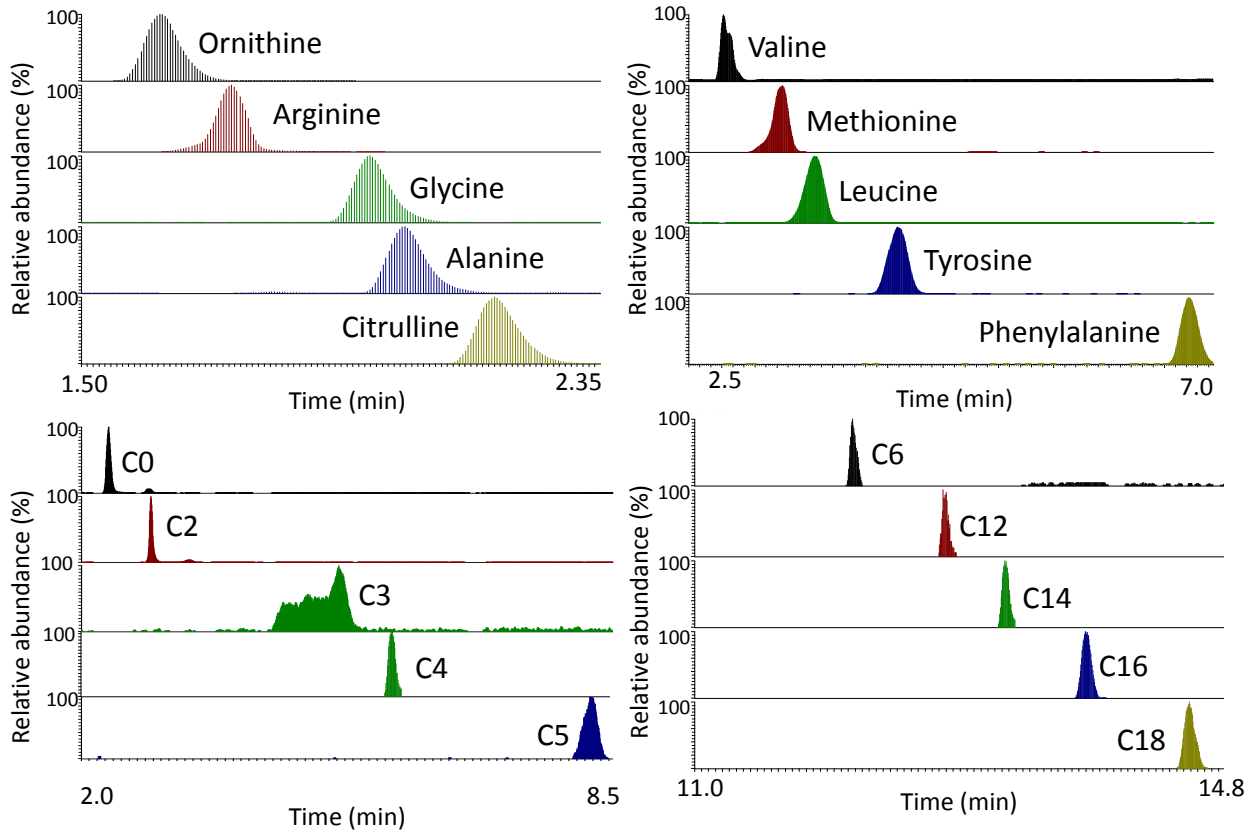


Figure 28: Extracted ion chromatograms of all the final compounds.

All the final compounds were detected by the final method.

4.6 The method was able to perform quantitative measurements

The ability of the final method to perform quantitative measurements of the final compounds was determined by the analysis of more than one punch of the DBS. DBS samples of 1, 2, 3 and 4 (3, 6, 9 and 12 µL) punches were prepared three times. Linear regression of the peak areas from 1, 2, 3 and 4 punches of each compound was performed. A perfect positive linear correlation is when R^2=1, meaning that all the points make up a straight line. In analytical

practice, quantitative graphs give numerical R^2-values greater than 0.99 [94]. **Figure 29** shows the peak area for the different number of punches tested with the R^2 value for glycine and alanine with R^2<0.99 and ornithine and arginine with R^2>0.99. All the rest of the final compounds had R^2<0.99 except from tyrosine and C3. The correlation of the rest of the compounds is shown in **Appendix,** Figure 43 in section 6.5.

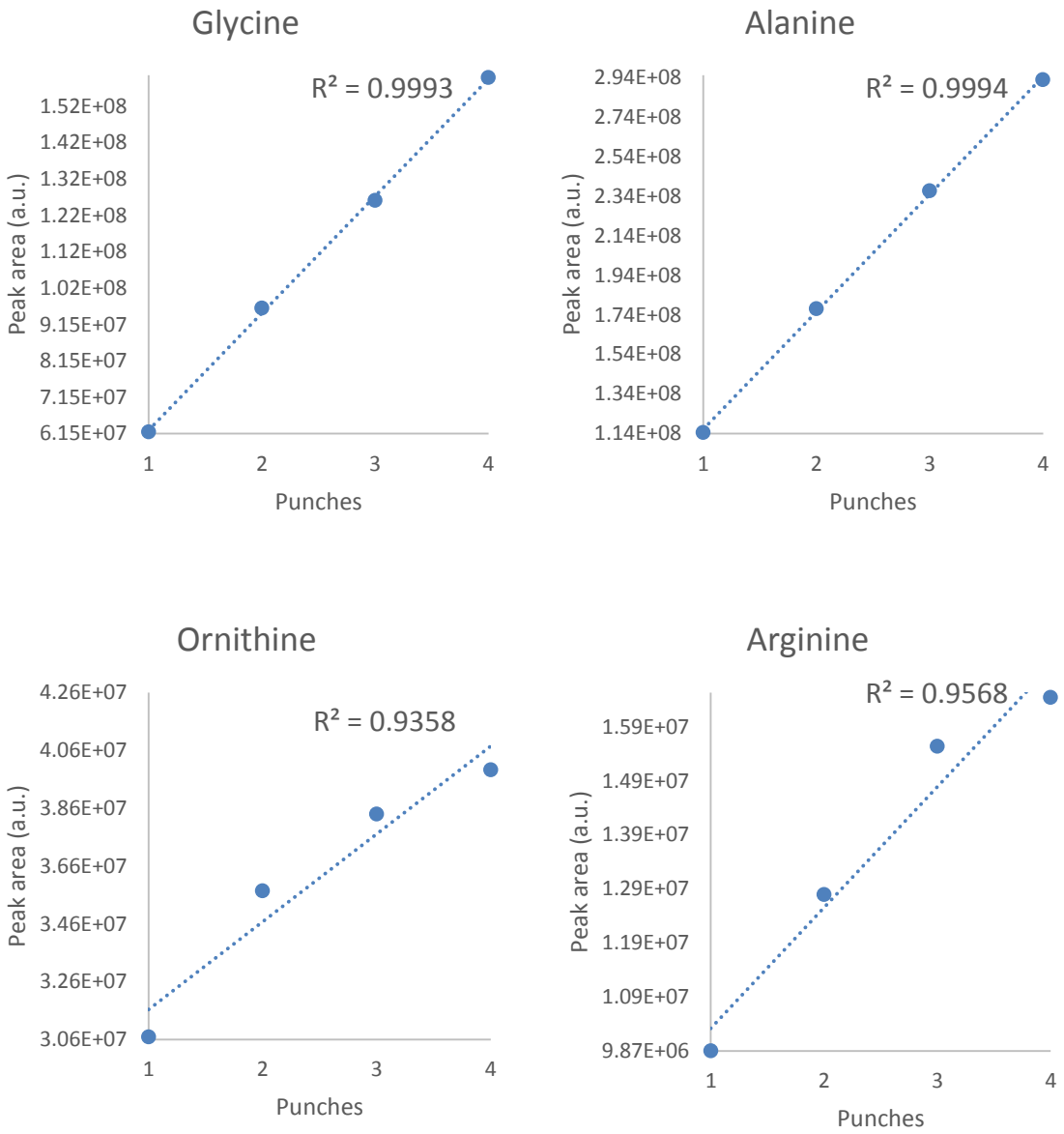


Figure 29: Peak area for the numbers of punches tested with the R^2 value for glycine, alanine, ornithine and arginine.

C12 had a high % RSD when using one punch, while the other acylcarnitines had <10 %. Using four punches compared to one improved the % RSD of C12, and therefore the detection.

Matrix effects of the final compounds were investigated in **Appendix** Section 6.5.1. The peak area of the polar compounds, ornithine, arginine and methionine were enhanced with ratios of 278, 166 and 140 %, respectively, and the peak area of C3 and C4 were slightly suppressed with ratios of 79 and 83 %, respectively. Ornithine, arginine and C3 showed poorer quantitativity than the other compounds and may be explained by the matrix effect.

To summarize: The quantification of most of the acylcarnitines and amino acids was satisfactory, with linear correlation of R^2>0.99 in signal vs. concentration plots, using 1-4 punches (about 3, 6, 9 and 12 µL whole blood) of the DBS, see ***Figure 30****, but some polar compounds were affected by matrix effects.*

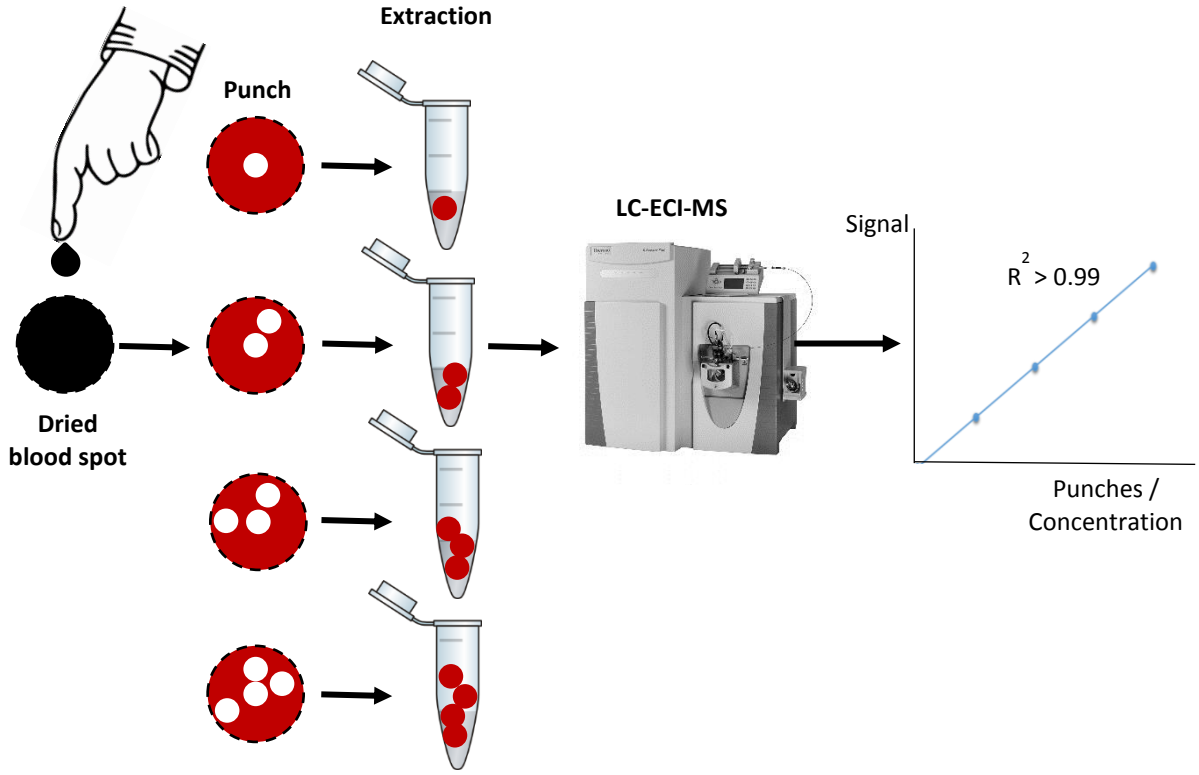


Figure 30: The analyses of 1, 2, 3 and 4 punches of the DBS using the final method showed that the method was able to perform quantitative measurements.

4.7 Data acquisition, data processing and metabolite identification

In the analytical evaluation, data acquisition of all possible metabolites (*m/z* 50-750) was performed by reverse phase LC – positive ESI mode – Orbitrap MS without fragmentation.

Fragmentation spectra are commonly used to secure better identification of metabolites [11, 82]. Thus, MS^2 data acquisition was wanted to be implemented in the final method. Data-dependent acquisition was tested and the identification of metabolites using MS and MS^2 was compared. Data-dependent acquisition with intensity threshold was tested, as observations throughout the experiment have been poor detection of peaks with intensity below 10^4 counts per second. Additionally, a study showed that higher quality spectra were obtained using data-dependent acquisition compared to data-independent acquisition [95].

Full bioinformatics software (Compound Discoverer 2.1 from Thermo Scientific) was used to process the untargeted metabolomics data after analysis of DBS using the final method. Untargeted metabolomics workflow of data processing was used to find and identify the differences between compared samples. The software aligned the retention times of all features to group features across all samples. Predicted elemental compositions of all metabolites in the samples and/or fragmentation spectra were used to search in databases for metabolite identification.

4.7.1 Tandem mass spectrometry facilitated the metabolite identification

The use of full scan MS^1 and data-dependent MS^2 acquisition in combination make it possible to quantify and identify target compounds and non-target compounds in the same workflow [77, 96]. This scan mode was tried with the MS settings shown in **Table 12**. The settings used were inspired by a study that successfully identified metabolites in food [97].

Table 12: Mass spectrometric settings in full scan and data-dependent MS^2 acquisition ($ddMS^2$): Resolution in full width at half maximum (FWHM), scan range, collision energy and intensity threshold.

Full scan	
Resolution setting	70,000 (FWHM) at *m/z* 200
Scan range (*m/z*)	50-750
$ddMS^2$	
Resolution setting	17,500 (FWHM) at *m/z* 200
HCD collision energy (stepped)	20, 80 eV
Intensity threshold	3.2e4

Data obtained after analysis of a DBS sample by the final method was processed in the software. All ions with intensity $>10^5$ counts per second acquired were used in the processing. 1226 features were detected and grouped to 536 compounds. Data acquisition with full scan MS resulted in the software proposing 229 names, while data acquisition with $ddMS^2$ proposed 334 names. A name was proposed for multiple compounds, as the names proposed using MS data are obtained by searching the elemental compositions in a database and the most referred compound is proposed.

One example was a compound with an elemental composition could that could be caffeine or enprophylline suggested by this database. Caffeine was proposed as it was the most referred compound. The additional MS^2 data made it possible to determine that the compound was caffeine, as the fragmentation spectrum matched 97.1 % with spectra in a database.

Another example was two compounds with molecular mass of 180.06439 Da with two different retention times. Three different structures were suggested by using MS data (theobromine, theophylline and paraxanthine) and both of them were named theophylline because of the highest number of references. MS^2 data was able to identify them. The first peak was theobromine (87.6 % match) and the second peak was paraxanthine (91.1 % match).

These two examples illustrates why MS^2 data is important to identify metabolites in untargeted LC-MS based metabolomics analyses.

Only the peaks with intensity over 10^5 counts per second were processed. The intensity threshold was decreased to 10^4 counts per second to observe the number of features, compounds and proposed names by the software. 30551 features were detected, grouped to 5333 compounds, and 1837 names were proposed. Decreasing the intensity threshold from 10^5 to 10^4 counts per second increased the number of compounds by a factor of 10.

Data processing using the software was a relatively automatic process. The proposed names of compounds must be thoroughly considered to be identified as a metabolite, as multiple compounds were named as the same compound. There were also many unidentified compounds. This made metabolite identification the most challenging step of the workflow.

Additional searches in other databases and the use of other analytical techniques could potentially secure better identification of metabolites and identify more metabolites.

Validation with established quantitative methods can be done to reach the highest level of quantification.

To summarize; acquisition of MS^2 data secured better metabolite identification, see ***Figure 31****, even though metabolite identification was the most challenging step of the untargeted metabolomics workflow, as the metabolite identification often requires additional steps.*

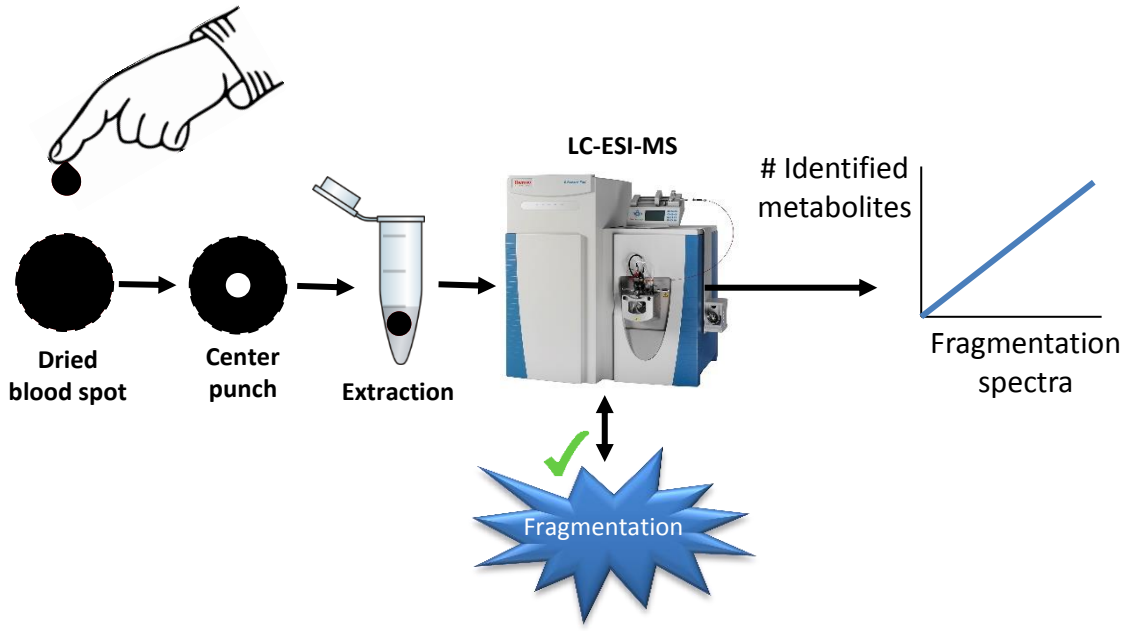


Figure 31: Acquiring MS^2 data increased the number of identified metabolites.

4.8 The method detected changes in the metabolome after a physiological change

4.8.1 Detected the intake of one metabolite

The ability of the method to detect one single metabolite among the other detected metabolites was investigated by comparing the measured caffeine amount in DBS samples collected from six healthy volunteers. Samples were collected during one day of free intake of caffeine (e.g., coffee and soda) and two days of no intake of caffeine. As shown in **Figure 32**, the measured amount in samples taken after free intake was more spread and higher than in samples taken during no intake and the amount decreased with time after the last intake, as expected.

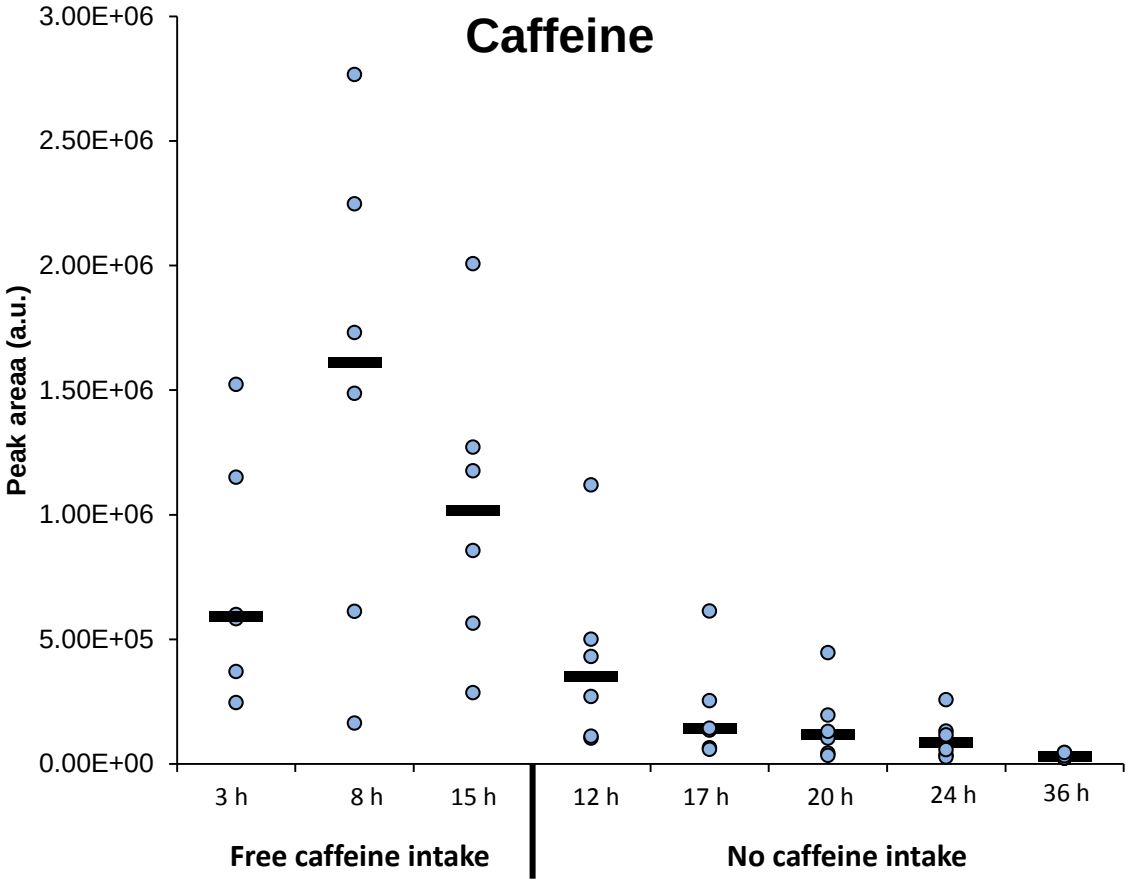


Figure 32: Measured amount of caffeine in samples from six individuals after free caffeine intake for 15 hours and after no caffeine intake for 36 hours.

There was a clear difference between the samples after free caffeine intake versus no intake in such a small amount of whole blood on a paper card using the final method.

4.8.2 Research project: Investigation of the dried blood spot metabolome after normal physiological changes

Medical and health research projects must be approved by the Regional Committee for Medical and Health Research Ethics (REC). The aim of REC is to promote good quality and ethically accepted medical and health research. The benefits versus risks and that the data is protected are considered when they receive an application.

An application for a research project to REC called NOMFASTE2018 was sent during this project. The research project involved the investigation of the DBS metabolome after normal physiological changes, such as fasting, workout etc., as these factors alter the metabolic

concentrations. The application included the purpose of the research project, an execution plan and the creation of a biobank. A biobank is the storage of biological material to be used in research projects. A coding list with letters given to each volunteer with their information was created to protect the identity of the healthy volunteers. The list was stored securely. The REC approved our research project and the informed consent form to all the healthy volunteers can be seen in **Appendix**, section 6.6.1. A pilot study was performed during this project to see if the method was able to detect changes in the metabolome after fasting.

4.8.3 Detected changes in the metabolome after fasting

Many metabolic analyses are taken only after overnight fasting to avoid positive false results that are caused by food intake. As a proof of concept, to see if the method was able to differentiate between the samples after food intake and after fasting, the investigation of the dried blood spot metabolome before, during and after fasting was done. Six healthy volunteers fasted for 36 hours. The volunteers were allowed to drink water during the fasting period. Urine samples and whole blood in tubes were also collected and stored in – 80℃, if targeted approaches to validate metabolite concentrations in the future are wanted. 16 samples were collected over five days during free diet, fasting for 36 hours and controlled diet. See **Appendix**, section 6.6.2, Figure 46 and Table 25 for the sampling programme.

The samples collected during free diet and overnight fasting were analyzed in one analysis run, using ISs to control the instrumental drift. Another analysis run analyzed all 16 DBS samples with control samples, see **Appendix**, section 6.6.3 for the measured peak areas and calculated % RSD from the IS and control samples. Peak area % RSD for most of the IS compounds were <10 % for all samples in the first analysis run. Compounds in the control samples showed excellent peak area % RSD <10 % in a DBS sample analyzed three times each day over 11 days. The software had troubles with the alignment of some of the samples (probably due to retention time drift), thus, two samples during free diet and a sample during fasting were not included in the results in the other analysis run. Thus, multivariate analyses were only performed with samples during free diet and during fasting from the first analysis run.

The software used was able to perform statistical analyses such as PCA and differential analysis using volcano plots. This was used to detect trends and specific altered compounds between groups of the data in the fasting project. **Figure 33** shows the PCA plot (from the

analysis run of samples taken during free diet and overnight fasting) used to easier explore and visualize the variation between samples. One point represents all the metabolites detected in that specific sample. The distance between points represents the variation in all metabolites compared to the other samples. The closer the points are together, the more similar. PC1 (principal component 1) on the *x*-axis represented the largest variation (18.2 %) over PC2 on the *y*-axis (13.4 %). Thus, the distance in the x-direction represented a larger difference than in the y-direction. PC3 represented an even lower variation (10.1 %) and was shown in **Figure 33** as larger and smaller points, where large points were positive variations and small points were negative variations. A 3D PCA plot may be seen at https://ggbm.at/ZTN8mGns. Right click, hold and drag the PCA plot to explore the data. Samples from the six individuals were spread, but there was minimal overlap after free diet, overnight fasting (12 hours) and prolonged fasting (36 hours). Samples taken after prolonged fasting clustered together. This was expected, as different food intake contributes to larger variation in the metabolome.

What separated samples taken during free diet and fasting overnight, were that samples taken during free diet were taken after breakfast and sample taken during overnight fasting were taken after no intake of breakfast. Surprisingly, one sample taken after breakfast clustered together with the samples taken after overnight fasting. The explanation was that individual F actually did not eat breakfast before collecting the DBS sample, as the other individuals did. These results showed that the method was able to pick up differences in the metabolome after a metabolic change in only 3 µL whole blood on a paper.

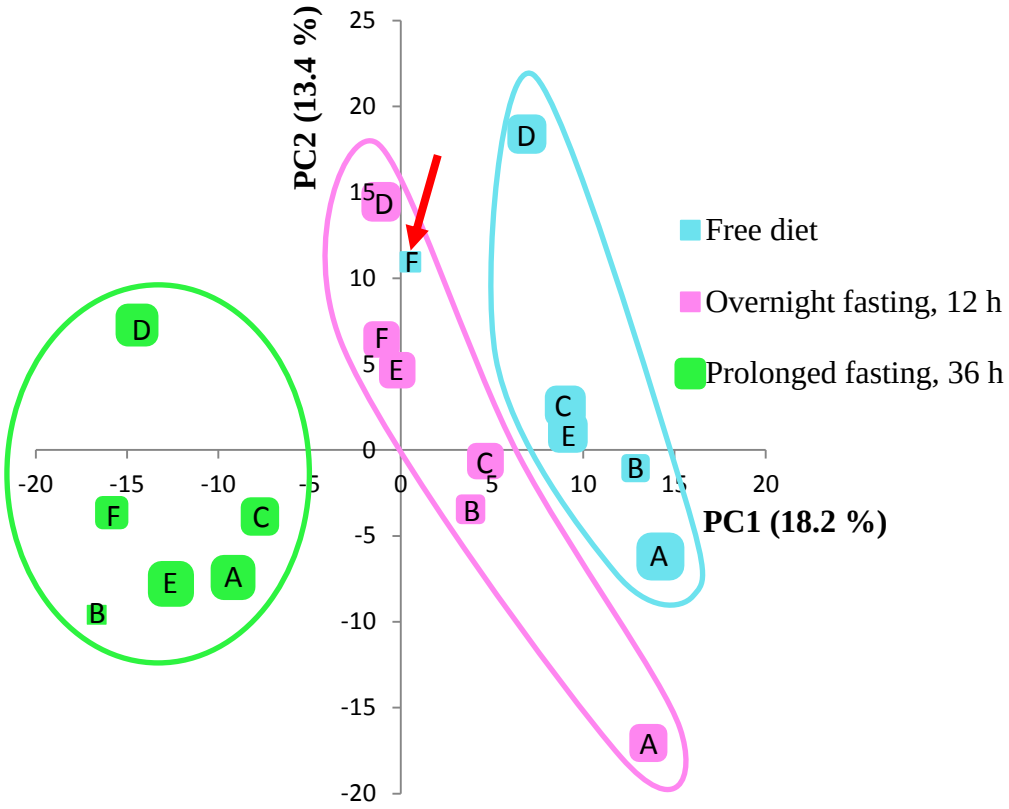


Figure 33: Principal component analysis (PCA) plot of samples from six individuals (A-F) after free diet, overnight fasting (12 hours) and prolonged fasting (36 hours). A point represents all the metabolites detected in that sample. Individual F did not eat before collecting the free diet sample, resulting in that it clustered together with the overnight fasting samples.

In this project, it was known that individual F did not eat breakfast prior to sampling, and individual C took paracetamol (500 mg) for headache during the fasting period. The method was able to determine that individual C had paracetamol in the whole blood. The intake of medications will affect metabolic concentrations, but the intake by a patient is not necessarily known prior to sampling. Many analyses in diagnostics of IEMs use targeted metabolomics approaches that require overnight fasting prior to sample collection. If a patient has been eating food containing proteins (amino acids) prior to sample collection, and the diagnostic analyses determine the amino acid levels to be normal, low or high, the potential of obtaining false positives will be high. That the untargeted metabolomics method was able to differentiate between samples after food intake and after fasting, and that it was able to detect the intake of medications, are advantages in metabolomics diagnostics, as there is a potential that a patient was non-fasting prior to sample collection, resulting in unexplained levels of measured metabolites.

Differential analysis using volcano plots, see **Figure 34**, was done to see the metabolites that changed the most in all samples taken after prolonged fasting and overnight fasting. The *x*-axis represents the log2 of the fold change between two sample groups (generated ratio), and the *y*-axis represents the negative log10 of the p-value (test of significance) of the fold change. The significantly changed metabolites (P<0.05) with Log2 Fold change >1 were presented in a higher level in prolonged fasting samples (red in the volcano plot) and <1 were presented in a lower level in prolonged fasting (green in the volcano plot).

20 compounds were detected in significantly higher levels in samples taken after prolonged fasting versus overnight fasting, without any successful metabolite identification. Seven compounds were in significantly lower levels in samples taken after prolonged fasting versus overnight fasting, with four successfully metabolite identifications: caffeine (from coffee), paraxanthine (caffeine metabolite), theobromine (from chocolate) and hippuric acid (from orange juice/tea). The peak areas of the identified metabolites in each sample are shown in **Appendix**, section 6.6.4, Figures 48-52.

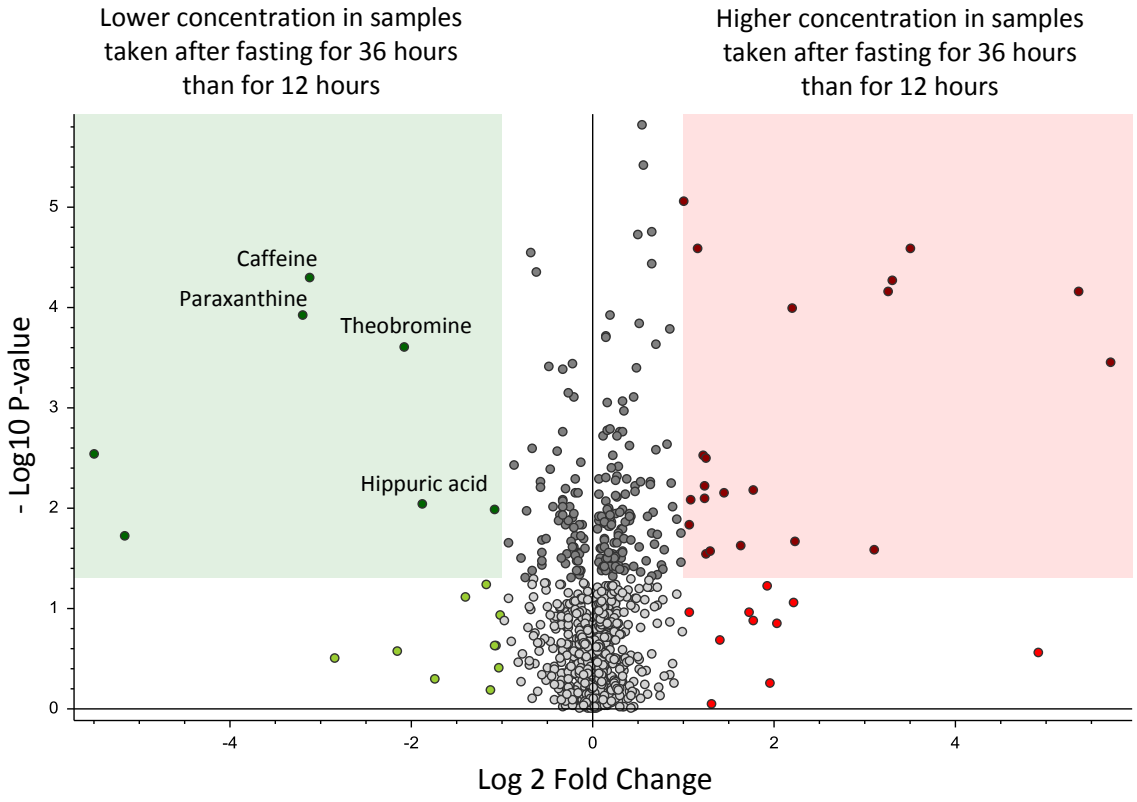


Figure 34: Differential analysis using volcano plot of data obtained from the analysis of samples taken during free diet and during fasting. Log2 Fold Change (prolonged fasting/overnight fasting) >1 (red) and <1 (green) (and P<0.05). Green: higher level in prolonged fasting samples than overnight fasting samples.

To summarize; Combining the method with bioinformatics, one could distinguish between samples taken after free diet, overnight fasting (12 hours) and prolonged fasting (36 hours), using only 3 µL blood on paper, see ***Figure 35****. Analyses of a DBS sample each day for 11 days showed excellent repeatability of peak areas and retention times, making the method able to perform long analysis batches. Taken together, the method has the potential to be used in diagnostics of IEMs and will be used to detect differences in the metabolome of patient samples and healthy samples.*

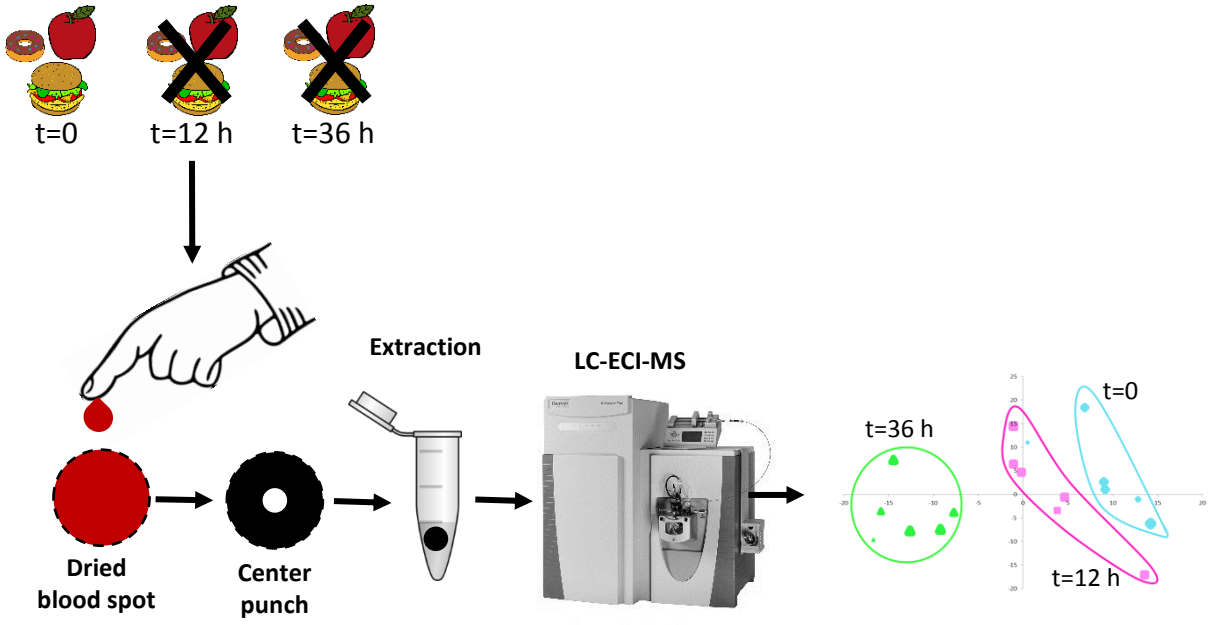


Figure 35: The final method was able to differentiate between metabolic profiles before and after fasting.

5 Conclusion and future work

DBS metabolome coverage: The initial method detected the hydrophobic and low-abundant compounds poorly. An increased organic solvent amount in the reconstitution solution and reduced *m/z* range of the MS increased the number of hydrophobic and low-abundant compounds detected, respectively. The poor detection of hydrophobic compounds was most likely due to poor extraction- and ionization efficiency. The final method detected all amino acids and acylcarnitines. **Repeatability:** The instrument repeatability and assay reproducibility were satisfactory with the % RSD of peak areas <10 % and retention times <1 %, using a pooled control sample, injected three times each day during sample analysis for 11 days. Compared to perimeter punches, analyzing center punches improved the assay reproducibility. **Quantitativity:** The quantification of most of the acylcarnitines and amino acids was satisfactory, with linear correlation of R^2>0.99 in signal vs. concentration plots, using 1-4 punches (about 3, 6, 9 and 12 µL whole blood) of the DBS, but some polar compounds were affected by matrix effects, most likely due to poor separation of polar compounds. **Data workflow:** Tandem MS data acquisition made it possible to secure better identification. The final method was combined with software for data processing. **Analysis of metabolic profiles:** The method was able to detect thousands of metabolites in 3 µL whole blood on a filter paper card and succeeded to detect and identify differences in the DBS metabolome after a physiological change, see **Figure 36**.

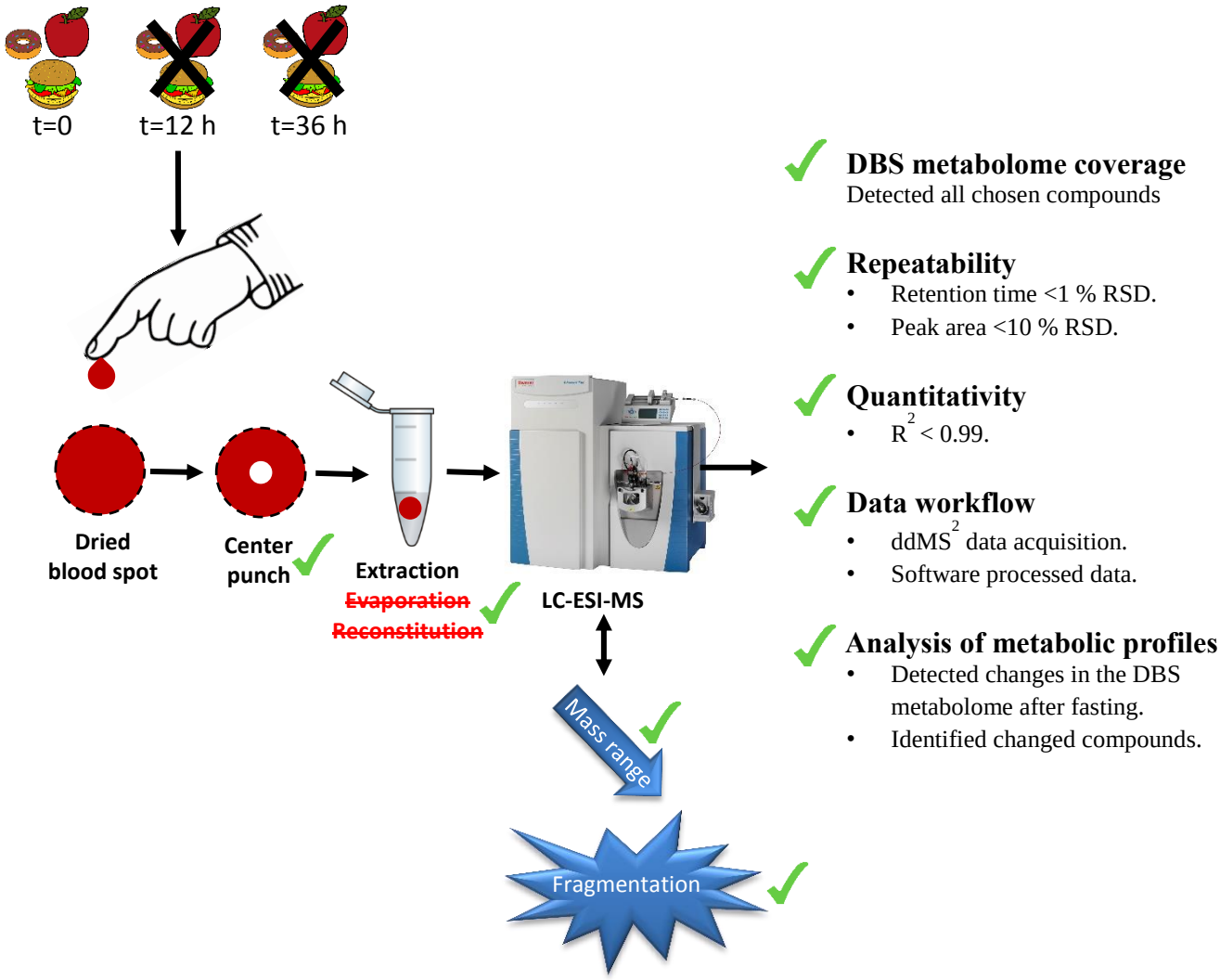


Figure 36: The final method with center punches, exclusion of the evaporation to dryness and reconstitution steps, reducing the mass range (compared to the initial method), MS^2 data acquisition and the data processing by software, compared to the initial method made it possible to differentiate between metabolic profiles after fasting and not fasting. RSD: Relative standard deviation. R^2=1 when there is a linear correlation.

The method will increase our knowledge of the biochemical and pathophysiological consequences of IEMs and will lead to faster laboratory diagnostics and better monitoring, thereby improving treatment and prognosis for patients with IEMs. The detailed biochemical phenotype provided will also narrow the gap between genotype and phenotype.

In the short term, the developed method will be implemented as a tool for diagnostics and monitoring of IEMs in the laboratory at The National Unit of Screening and Diagnosis of Congenital Pediatric Metabolic Disorders. In a wider perspective, the method can be used in research, diagnostics and monitoring of any disease where metabolites and biochemical patterns identified can be used as relevant biomarkers.

5.1 Future work

This method, as it is, will be implemented in the diagnostic laboratory at The National Unit of Screening and Diagnosis of Congenital Pediatric Metabolic Disorders. As any other analytical method, this method can also be further optimized.

The punch location was found to affect the concentration of metabolites in this project. If random punches are wanted, a method that measures the hematocrit value can be implemented to obtain more accurate measurements.

Long-term stability of analytes in stored DBS samples can be investigated, as DBS samples that have been stored for a long period of time will be analyzed by this method.

In this project, the detection of compounds was affected by the poor separation of polar compounds and poor ionization effect of hydrophobic compounds. The implementation of additional separation and ionization techniques may improve the detection of compounds.

Data was acquired in negative ionization mode during the optimization of LC-MS settings, but not during the evaluation due to temporary technical problems. Data in negative ionization mode will be acquired in the future to cover more of the DBS metabolome.

The metabolite identification was a difficult task. Other MS^2 data acquisition methods than $ddMS^2$, or other $ddMS^2$ settings may be tested to see if better metabolite identification is obtained.

Further investigation of statistics and metabolite identification of the data from the fasting project can be done to gain more knowledge of the fasting versus non-fasting metabolome. Investigation of the DBS metabolome after other normal physiological changes such as sleep, workout etc., can be done to determine the effect of these changes on metabolic concentrations.

The suitability for analyses in real clinical settings can be tested by analyzing samples from patients with known IEMs.

References

1. Berg, J.M., J.L. Tymoczko, and L. Stryer, *Section 14.1, Metabolism is Composed of Many Coupled, Interconnecting Reactions*, in *Biochemistry*. 5 ed. 2002, W H Freeman: New York.
2. Berg, J.M., J.L. Tymoczko, and L. Stryer, *Section 5.4, Gene Expression Is the Transformation of DNA Information Into Functional Molecules*, in *Biochemistry*. 5th ed. 2002, W H Freeman: New York.
3. Ezgu, F., *Inborn Errors of Metabolism.* Advances in Clinical Chemistry, 2016. **73**: p. 195-250.
4. Lanpher, B., N. Brunetti-Pierri, and B. Lee, *Inborn errors of metabolism: the flux from Mendelian to complex diseases.* Nature Reviews Genetics, 2006. **7**(6): p. 449-60.
5. Pandor, A., J. Eastham, C. Beverley, J. Chilcott, and S. Paisley, *Clinical effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review.* Health Technology Assessment, 2004. **8**(12): p. iii,1-121.
6. Bujak, R., W. Struck-Lewicka, M.J. Markuszewski, and R. Kaliszan, *Metabolomics for laboratory diagnostics.* Journal of Pharmaceutical and Biomedical Analysis, 2015. **10**(113): p. 108-20.
7. García-Sevillano, M.Á., T. García-Barrera, N. Abril, C. Pueyo, J. López-Barea, and J.L. Gómez-Ariza, *Omics technologies and their applications to evaluate metal toxicity in mice M. spretus as a bioindicator.* Journal of Proteomics, 2014. **2**(104): p. 4-23.
8. Nishiumi, S., M. Suzuki, T. Kobayashi, A. Matsubara, T. Azuma, and M. Yoshida, *Metabolomics for biomarker discovery in gastroenterological cancer.* Metabolites, 2014. **4**(3): p. 547-71.
9. Tebani, A., L. Abily-Donval, C. Afonso, S. Marret, and S. Bekri, *Clinical Metabolomics: The New Metabolic Window for Inborn Errors of Metabolism Investigations in the Post-Genomic Era.* International Journal of Molecular Sciences, 2016. **17**(7): p. 1167.
10. Wishart, D.S., *Current progress in computational metabolomics.* Briefings in Bioinformatics, 2007. **8**(5): p. 279-93.
11. Xiao, J.F., B. Zhou, and H.W. Ressom, *Metabolite identification and quantitation in LC-MS/MS-based metabolomics.* Trends in Analytical Chemistry, 2012. **1**(32): p. 1-14.
12. Milne, S.B., T.P. Mathews, D.S. Myers, P.T. Ivanova, and H.A. Brown, *Sum of the Parts: Mass spectrometry-Based Metabolomics.* Biochemistry, 2013. **52**(22): p. 3829-3840.
13. Imperlini, E., L. Santorelli, S. Orrù, E. Scolamiero, M. Ruoppolo, and M. Caterino, *Mass Spectrometry-Based Metabolomic and Proteomic Strategies in Organic Acidemias.* BioMed Research International, 2016. **2016**: p. 9210408.
14. Zhou, B., J.F. Xiao, L. Tuli, and H.W. Ressom, *LC-MS-based metabolomics.* Molecular BioSystems, 2012. **8**(2): p. 470-81.
15. Gertsman, I. and B.A. Barshop, *Promises and pitfalls of untargeted metabolomics.* Journal of Inherited Metabolic Disease, 2018. **41**(3): p. 355-366.
16. Roberts, L.D., A.L. Souza, R.E. Gerszten, and C.B. Clish, *Targeted metabolomics.* Current Protocols in Molecular Biology, 2012. **91**(1): p. 30.2.1-30.2.24.
17. Johnson, C.H., J. Ivanisevic, H.P. Benton, and G. Siuzdak, *Bioinformatics: The Next Frontier of Metabolomics.* Analytical Chemistry, 2015. **87**(1): p. 147-156.

18. Zhang, A.H., H. Sun, P. Wang, Y. Han, and X.-J. Wang, *Modern analytical techniques in metabolomics analysis.* The Analyst, 2012. **137**(2): p. 293-300.
19. Alonso, A., S. Marsal, and A. Julià, *Analytical Methods in Untargeted Metabolomics: State of the Art in 2015.* Frontiers in Bioengineering and Biotechnology, 2015. **3**(23).
20. Miller, M.J., A.D. Kennedy, A.D. Eckhart, L.C. Burrage, J.E. Wulff, L.A.D. Miller, M.V. Milburn, J.A. Ryals, A.L. Beaudet, Q. Sun, V.R. Sutton, and S.H. Elsea, *Untargeted metabolomic analysis for the clinical screening of inborn errors of metabolism.* Journal of Inherited Metabolic Disease, 2015. **38**(6): p. 1029-1039.
21. Nawaporn, V. and S. Alan, *Untargeted Metabolomics.* Current Protocols in Molecular Biology, 2010. **90**(1): p. 30.1.1-30.1.24.
22. Jacob, M., A. Malkawi, N. Albast, S. Al Bougha, A. Lopata, M. Dasouki, and A.M. Abdel Rahman, *A targeted metabolomics approach for clinical diagnosis of inborn errors of metabolism.* Analytica Chimica Acta, 2018. **1025**: p. 141-153.
23. Rashed, M.S., P.T. Ozand, M.P. Bucknall, and D. Little, *Diagnosis of Inborn Errors of Metabolism from Blood Spots by Acylcarnitined and Amino Acids Profiling Using Automated Electrospray Tandem Mass Spectrometry.* Pediatric Reasearch, 1995. **38**(3): p. 324-31.
24. Kennedy, A.D., M.J. Miller, K. Beebe, J.E. Wulff, A.M. Evans, L.A. Miller, V.R. Sutton, Q. Sun, and S.H. Elsea, *Metabolomic Profiling of Human Urine as a Screen for Multiple Inborn Errors of Metabolism.* Genetic Testing and Molecular Biomarkers, 2016. **20**(9): p. 485-95.
25. Coene, K.L.M., L.A.J. Kluijtmans, E. van der Heeft, U.F.H. Engelke, S. de Boer, B. Hoegen, H.J.T. Kwast, M. van der Vorst, M.C.D.G. Huigen, I.M.L.W. Keularts, M.F. Schreuder, C.D.M. van Karnebeek, S.B. Wortmann, M.C. de Vries, M.C.H. Janssen, C. Gilissen, J. Engel, and R.A. Wevers, *Next-generation metabolic screening: targeted and untargeted metabolomics for the diagnosis of inborn errors of metabolism in individual patients.* Journal of Inherited Metabolic Disease, 2018. **41**(3): p. 337-353.
26. *Seksjon for medfødte metabolske sykdommer*. 2018 [cited 2018 June 7th]; Available from: https://oslo-universitetssykehus.no/fag-og-forskning/laboratorietjenester/-medisinsk-biokjemi/seksjon-for-medfodte-metabolske-sykdommer.
27. Camp, K.M., M.A. Lloyd-Puryear, and K.L. Huntington, *Nutritional Treatment for Inborn Errors of Metabolism: Indications, Regulations, and Availability of Medical Foods and Dietary Supplements Using Phenylketonuria as an Example.* Molecular genetics and metabolism, 2012. **107**(1-2): p. 3-9.
28. Mohamed, S., *Recognition and diagnostic approach to acute metabolic disorders in the neonatal period.* Sudanese Journal of Paediatrics, 2011. **11**(1): p. 20-28.
29. Costa, C.G., E.A. Struys, A. Bootsma, H.J. ten Brink, L. Dorland, I. Tavares de Almeida, M. Duran, and C. Jakobs, *Quantitative analysis of plasma acylcarnitines using gas chromatography chemical ionization mass fragmentography.* Journal of Lipid Research, 1997. **38**(1): p. 173-82.
30. Jones, P. and M. Bennett, *Urine Organic Acid Analysis for Inherited Metabolic Disease Using Gas Chromatography-Mass Spectrometry.* Methods in Molecular Biology, 2010. **603**: p. 423-31.
31. Loeber, J.G., P. Burgard, M.C. Cornel, T. Rigter, S.S. Weinreich, K. Rupp, G.F. Hoffmann, and L. Vittozzi, *Newborn screening programmes in Europe; arguments and efforts regarding harmonization. Part 1 - From blood spot to screening result.* Journal of Inherited Metabolic Disease, 2012. **35**(4): p. 603-611.

32. *Expanded newborn screening programme*. 2018 [cited 2018 June 25th]; Available from: https://oslo-universitetssykehus.no/PublishingImages/avdelinger/barne-og-ungdomsklinikken/nyfodtscreeningen/nyfodtscreening/Engelsk.pdf.
33. *Medfødte metabolske sykdommer*. 2018 [cited 2018 June 25th]; Available from: http://ehandboken.oslo-universitetssykehus.no/api/File/GetFileDocument?entityId=135015.
34. Bishop, M.L., E.P. Fody, and L.E. Schoeff, *Clinical Chemistry: Techniques, Principles, Correlations*. 6 th ed. 2010, Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins. p. 28-29, 41, 64.
35. Gertsman, I., J.A. Gangoiti, and B.A. Barshop, *Validation of a dual LC–HRMS platform for clinical metabolic diagnosis in serum, bridging quantitative analysis and untargeted metabolomics*. Metabolomics, 2013. **10**(2): p. 312-323.
36. de Sain-van der Velden, M.G.M., M. van der Ham, J.J. Jans, G. Visser, P.M. van Hasselt, H.C.M.T. Prinsen, and N.M. Verhoeven-Duif, *Suitability of methylmalonic acid and total homocysteine analysis in dried bloodspots.* Analytica Chimica Acta, 2015. **853**: p. 435-441.
37. Wilson, I., *Global metabolomic profiling (metabonomics/metabolomics) using dried blood spots: advantages and pitfalls.* Bioanalysis, 2011. **3**(20): p. 2255-2257.
38. John, H.M.I., A.P. Philip, C.R. Sarah, and K.H. Thomas, *An on-card Approach for Assessment of Hematocrit on Dried Blood Spots which Allows for Correction of Sample Volume.* Analytical and Bioanalytical Techniques, 2013. **4**(162).
39. Bassaganyas, L., G. Freedman, D. Vaka, E. Wan, R. Lao, F. Chen, M. Kvale, R.J. Currier, J.M. Puck, and P.-Y. Kwok, *Whole exome and whole genome sequencing with dried blood spot DNA without whole genome amplification.* Human Mutation, 2017. **39**(1): p. 167-171.
40. Smit, P.W., I. Elliot, R.W. Peeling, D. Mabey, and P.N. Newton, *An overview of the clinical use of filter paper in the diagnosis of tropical diseases.* The American journal of tropical medicine and hygiene, 2014. **90**(2): p. 195-210.
41. Lawson, A.J., L. Bernstone, and S.K. Hall, *Newborn screening blood spot analysis in the UK: influence of spot size, punch location and haematocrit.* Journal of Medical Screening, 2015. **23**(1): p. 7-16.
42. Rufual, M.L., L.J. McCloskey, and D.F. Stickle, *Estimation of hematocrit in filter paper dried blood spots by potassium measurments: advantage of use of perimeter ring samples over circular center sub-punch samples.* Clinical Chemistry and Laboratory Medicine, 2017. **55**(1): p. 53-57.
43. Koster, R., D. Touw, and J.-W. Alffenaar, *Dried blood spot analysis; facing new challenges.* Journal of applied bioanalysis, 2015. **1**(2): p. 38-41.
44. Prentice, P., C. Turner, M.C. Wong, and R.N. Dalton, *Stability of metabolites in dried blood spots stored at different temperature over a 2-year period.* Bioanalysis, 2013. **5**(12): p. 1507-14.
45. Fingerhut, R., R. Ensenauer, W. Roschinger, R. Arnecke, B. Olgemoller, and A.A. Roscher, *Stability of Acylcarnitines and Free Carnitine in Dried Blood Spot Samples: Implications for Retrospective Diagnosis of Inborn Errors of Metabolism and Neonatal Screening for Carnitine Transporter Deficiency.* Analytical Chemistry, 2009. **81**(9).
46. Shama, N., D.C.M.D. Santos, A. Garcia, and C. Barbas, *Analytical protocols based on LC-MS, GS-MS, CE-MS for nontargeted metabolomics of biological tissue.* Bioanalysis, 2014. **6**(12): p. 1657-1677.

47. Gika, H.G., G.A. Theodoridis, R.S. Plumb, and I.D. Wilson, *Current practice of liquid chromatography–mass spectrometry in metabolomics and metabonomics.* Journal of Pharmaceutical and Biomedical Analysis, 2014. **87**: p. 12-25.
48. Dunn, W.B., I.D. Wilson, A.W. Nicholls, and D. Broadhurst, *The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans.* Bioanalysis, 2012. **4**(18): p. 2249-2264.
49. Patti, G.J., O. Yanes, and G. Siuzdak, *Metabolomics: the apogee of the omics triology.* Nature Reviews Molecular Cell Biology, 2012. **13**(4): p. 263-9.
50. Niessen, W.M.A. and D. Falck, *Introduction to Mass Spectrometry, a Tutorial*, in *Analyzing Biomolecular Interactions by Mass Spectrometry*. 1st ed. 2015, Wiley: Weinheim, Germany.
51. Emwas, A., *The Strengths and Weaknesses of NMR Spectroscopy and Mass Spectrometry with Particular Focus on Metabolomics Research.* Methods in Molecular Biology, 2015. **1277**: p. 161-193.
52. Wishart, D.S., *Advances in metabolite identification.* Bioanalysis, 2011. **3**(15): p. 1769-1782.
53. Goldansaz, S.A., A.C. Guo, T. Sajed, M.A. Steele, G.S. Plastow, and D.S. Wishart, *Livestock metabolomics and the livestock metabolome: A systematic review.* Public Library of Science one, 2017. **12**(5).
54. Stringer, K., R. McKay, A. Karnovsky, B. Quemerais, and P. Lacy, *Metabolomics and Its Application to Acute Lung Diseases.* 2016. **7**: p. 44.
55. Harris, D.C., *Quantitative Chemical Analysis*. 8th ed. 2010, New York: W. H. Freeman and Company. p. 516-521, 596-628.
56. Lopes, A.S., E.C. Cruz, A. Sussulini, and A. Klassen, *Metabolomic Strategies Involving Mass Spectrometry Combined with Liquid and Gas Chromatography.* Advances in Experimental Medicine and Biology, 2017. **965**: p. 77-98.
57. Wikoff, W.R., A.T. Anfora, J. Liu, P.G. Schults, S.A. Lesley, and G. Siuzdak, *Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolitess.* Proceedings of the National Academy of Sciences, 2009. **106**(10): p. 3698-3703.
58. Wang, C., *The Ionization Technology of LC-MS, Advantages of APPI on Detection of PPCPs and Hormones.* Austin Chromatography, 2015. **2**(2): p. 1032.
59. Cox, D.G., J. Oh, A. Keasling, K. Colson, and M. Hamann, *The Utility of Metabolomics in Natural Product and Biomarker Characterization.* Molecular and Cell Biology of Lipids, 2014. **1840**(12): p. 3460-3474.
60. Awad, H., M.M. Khamis, and A. El-Aneed, *Mass spectrometry, Review of the Basics: Ionization.* Applied Spectroscopy Reviews, 2015. **50**(2): p. 158-175.
61. Wilm, M., *Principles of Electrospray Ionization.* Molecular & Cellular Proteomics, 2011. **10**(7): p. M111.009407.
62. Wang, Y., S. Liu, Y.-J. Hu, P. Li, and J. Wan, *Current state of the art of mass spectrometry-based metabolomics studies-A review focusing on wide coverage, high throughput and easy identification.* Royal Society of Chemistry Advances, 2015. **5**(96): p. 78728-78737.
63. Banerjee, S. and S. Mazumdar, *Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte.* International Journal of Analytical Chemistry, 2012. **2012**: p. 40.
64. Toffali, K., A. Zamboni, A. Anesi, M. Stocchero, M. Pezzotti, M. Levi, and F. Guzzo, *Novel aspects of grape berry ripening and post-harvest withering revealed by untargeted LC-ESI-MS metabolomics analysis.* Metabolomics, 2011. **7**(3): p. 424-436.

65. Lu, W., B.D. Bennett, and J.D. Rabinowitz, *Analytical strategies for LC–MS-based targeted metabolomics.* Journal of Chromatography B, 2008. **871**(2): p. 236-242.
66. Becker, S., L. Kortz, C. Helmschrodt, J. Thiery, and U. Ceglarek, *LC–MS-based metabolomics in the clinical laboratory.* Journal of Chromatography B, 2012. **883-884**: p. 68-75.
67. Greaves, J. and J. Roboz, *Mass Spectrometry for the Novice*. 1st ed. 2014, Boca Raton: CRC Press p. 1-26, 83-87.
68. Gross, J.H., *Instrumentation*, in *Mass spectrometry*. 2 ed. 2017, Springer, Cham: Berlin.
69. Cho, R., Y. Huang, J.C. Schwartz, Y. Chen, T.J. Carlson, and J. Ma, *MS(M), an efficient workflow for metabolite identification using hybrid linear ion trap Orbitrap mass spectrometer.* Journal of the American Society for Mass Spectrometry, 2012. **23**(5): p. 880-888.
70. Perry, R.H., R.G. Cooks, and R.J. Noll, *Orbitrap mass spectrometry: instrumentation, ion motion and applications.* Mass Spectrometric Reviews, 2008. **27**(6): p. 661-99.
71. Makarov, A. and S. Eliuk, *Evolution of Orbitrap Mass Spectrometry Instrumentation.* Annual Review of Analytical Chemistry, 2015. **8**: p. 61-80.
72. Scigelova, M., M. Hornshaw, A. Giannakopulos, and A. Makarov, *Fourier transform mass spectrometry.* Molecular and Cellular Proteomics, 2011. **10**(7): p. M111.009431.
73. Cech, N.B. and C.G. Enke, *Practical implications of some recent studies in electrospray ionization fundamentals.* Mass Spectrometry Reviews, 2001. **20**(6): p. 362-387.
74. Lundanes, E., L. Reubsaet, and T. Greibrokk, *Chromatography: Basic principles, Sample preparations and Related methods.* 1st ed. 2013, Weinheim, Germany: Wiley. p. 2-71.
75. Neue, U.D., *Theory of peak capacity in gradient elution.* Journal of Chromatography A, 2005. **1079**(1): p. 153-161.
76. Fenaille, F., P. Barbier Saint-Hilaire, K. Rousseau, and C. Junot, *Data acquisition workflows in liquid chromatography coupled to high resolution mass spectrometry-based metabolomics: Where do we stand?* Journal of chromatography. A, 2017. **1526**: p. 1-12.
77. Schrimpe-Rutledge, A.C., S.G. Codreanu, S.D. Sherrod, and J.A. McLean, *Untargeted metabolomics strategies - Challenges and Emerging Directions.* Journal of the American Society for Mass Spectrometry, 2016. **27**(12): p. 1897-1905.
78. Broadhurst, D., R. Goodacre, S.N. Reinke, J. Kuligowski, I.D. Wilson, M.R. Lewis, and W.B. Dunn, *Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies.* Metabolomics, 2018. **14**(6): p. 72.
79. Franceschi, P., R. Mylonas, N. Shahaf, M. Scholz, P. Arapitsas, D. Masuero, G. Weingart, S. Carlin, U. Vrhovsek, F. Mattivi, and R. Wehrens, *MetaDB a Data Processing Workflow in Untargeted MS-Based Metabolomics Experiments.* Frontiers in bioengineering and biotechnology, 2014. **2**: p. 72.
80. Tan, A., N. Boudreau, and A. Lévesque, *Internal Standards for Quantitative LC-MS Bioanalysis*, in *LC-MS in Drug Bioanalysis*, Q.A. Xu and T.L. Madden, Editors. 2012, Springer US: Boston, MA DOI: 10.1007/978-1-4614-3828-1_1
81. Sysi-Aho, M., M. Katajamaa, L. Yetukuri, and M. Oresic, *Normalization method for metabolomics data using optimal selection of multiple internal standards.* Bioinformatics, 2007. **8**(93).

82. Watson, D.G., *A rough guide to metabolite identification using high resolution liquid chromatography mass spectrometry in metabolomic profiling in metazoans.* Computational and Structure Biotechnology Journal, 2013. **15**: p. 4.
83. Petrick, L., W. Edmands, C. Schiffman, H. Grigoryan, K. Perttula, Y. Yano, S. Dudoit, T. Whitehead, C. Metayer, and S. Rappaport, *An untargeted metabolomics method for archived newborn dried blood spots in epidemiologic studies.* Metabolomics, 2017. **13**(3): p. 27.
84. Østeby, A. *Global metabolomics of urine profiles relevant to Inborn Errors of Metabolism: Liquid chromatography positive electrospray ionization tandem mass spectrometry analysis* [*Master thesis]*, in Department of Chemistry. 2013, Oslo: University of Oslo.
85. Løvoll, S.M. *Untargeted metabolomics in urine related to Inborn errors of metabolism using liquid chromatography coupled to tandem mass spectrometry with electrospray ionization* [*Master thesis]*, in Department of Chemistry. 2013, Oslo: University of Oslo.
86. Arnesen, C.E. and H.B. Skogvold. *Utvikling av metabolomikkmetode for blood spot som diagnostisk verktøy* [*Bachelor thesis]*, in Institutt for Industriell Utvikling. 2015, Oslo: Høgskolen i Oslo og Akershus.
87. Arnesen, C.E. *Optimization of liquid chromatographic parameters for untargeted metabolomics of dried blood spots* [*Master thesis]*, in Department of Chemistry. 2017, Oslo: University of Oslo
88. Skogvold, H.B. *Laboratory diagnostics: Maximizing sensitivity of a Q-Exactive Orbitrap mass spectrometer for untargeted metabolomics of dried blood spots* [*Master thesis]*, in Department of Chemistry. 2017, Oslo: University of Oslo.
89. Breindl, A., B. Beck, T. Clark, and R.C. Glen, *Prediction of the n-Octanol/Water Partition Coefficient, logP, Using a Combination of Semiempirical MO-Calculations and a Neural Network.* Molecular modeling annual, 1997. **3**(3): p. 142-155.
90. Pascual, J.A., M. Thevis, P. van Eenoo, and T. Wan *Minimum criteria for chromatographic-mass spectrometric confirmation of the identity of analytes for doping control purposes*. WADA, 2015.
91. Dunn, W.B., A. Erban, R.J.M. Weber, D.J. Creek, M. Brown, R. Breitling, T. Hankemeier, R. Goodacre, S. Neumann, J. Kopka, and M.R. Viant, *Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics.* Metabolomics, 2013. **9**(1): p. 44-66.
92. Theodoridis, G.A., H.G. Gika, E.J. Want, and I.D. Wilson, *Liquid chromatography-mass spectrometry based global metabolite profiling: A review.* Analytica Chimica Acta, 2012. **711**: p. 7-16.
93. Capiau, S., L.S. Wilk, M.C.G. Aalders, and C.P. Stove, *A Novel, Nondestructive, Dried Blood Spot-Based Hematocrit Prediction Method Using Noncontact Diffuse Reflectance Spectroscopy.* Analytical Chemistry, 2016. **88**(12): p. 6538-6568.
94. Miller, J.N. and J. Miller, C, *Statistics and Chemometrics for Analytical Chemistry*. 6th ed. 2010, Edinburgh, England: Pearson Education Limited. p. 114-121.
95. Zhu, X., Y. Chen, and R. Subramanian, *Comparison of information-dependent acquisition, SWATH, and MS(all) techniques in metabolite identification study employing ultrahigh-performance liquid chromatography-quadrupole time-of-flight mass spectrometry.* Analytical Chemistry, 2014. **86**(2): p. 1202-9.
96. Junot, C., F. Fenaille, B. Colsch, and F. Bécher, *High resolution mass spectrometry based techniques at the crossroads of metabolic pathways.* Mass Spectrometry Reviews, 2014. **33**(6): p. 471-500.

97. Johansson, M., S. Ekroth, O. Scheibner, and M. Bromirski. *How to screen and identify unexpected and unwanted compounds in food*. 2018 [cited 2018 May 6th]; Available from: https://assets.thermofischer.com/an-65144-lc-ms-unexpected-unwanted-compounds-food-an65144-en.pdf.

6 Appendix

6.1 Average and relative standard deviation of peak area and retention time of each compound in the experiments

The average and % RSD of all peak areas and retention times for all compounds observed were calculated in all experiments. See **Tables 13-24**.

Table 13: Spotted standard solution: Average, standard deviation (STD) and relative standard deviation in % (% RSD) of peak areas (PA) and retention times (Rt) of the initial compounds in five spotted standard solution (N=5) analyzed with the initial method.

	Creatinine	**Creatine**	**Uric acid**	**C2**	**C12**	**C16**
Average PA (a.u.) N=5	1.66E+09	3.10E+09	3.29E+08	2.30E+07	8.69E+06	7.59E+04
STD PA (a.u.) N=5	2.91E+07	3.74E+07	6.25E+06	1.08E+06	2.51E+05	3.68E+04
% RSD PA (a.u.) N=5	2 %	1 %	2 %	5 %	3 %	49 % *
Average Rt (min) N=5	2.08	2.38	3.94	3.04	12.85	13.80
STD Rt (min) N=5	0.01	0.00	0.00	0.01	0.00	0.01
% RSD Rt (min) N=5	0.3 %	0.2 %	0.1 %	0.4 %	0.0 %	0.1 %

* 7 points across the chromatographic peak.

Table 14: Dried blood spot samples: Average, standard deviation (STD) and relative standard deviation in % (% RSD) of peak areas (PA) and retention times (Rt) of the initial compounds in a dried blood spot sample analyzed 20 times (N=20) with the initial method.

	Creatinine	**Creatine**	**Uric acid**	**C2**	**C12**	**C16**
Average PA (a.u.) N=20	4.26E+07	1.83E+08	2.52E+07	2.34E+07	ND *	ND *
STD PA (a.u.) N=20	8.85E+05	2.87E+06	4.32E+05	3.32E+05	ND *	ND *
% RSD PA (a.u.) N=20	2 %	2 %	2 %	1 %	ND *	ND *
Average Rt (min) N=20	2.03	2.38	3.94	2.96	ND *	ND *
STD Rt (min) N=20	0.00	0.00	0.01	0.01	ND *	ND *
% RSD Rt (min) N=20	0.2 %	0.2 %	0.2 %	0.2 %	ND *	ND *

* ND = Not detectable.

Table 15: Assay reproducibility: Average, standard deviation (STD) and relative standard deviation in % (% RSD) of peak areas (PA) and retention times (Rt) of the initial compounds in 20 dried blood spot samples (N=20) analyzed with the initial method.

	Creatinine	**Creatine**	**Uric acid**	**C2**	**C12**	**C16**
Average PA (a.u.) N=20	3.79E+07	1.72E+08	2.36E+07	2.12E+07	ND *	ND *
STD PA (a.u.) N=20	3.55E+06	1.75E+07	1.41E+06	2.28E+06	ND *	ND *
% RSD PA (a.u.) N=20	10 %	10 %	6 %	11 %	ND *	ND *
Average Rt (min) N=20	2.02	2.38	3.95	2.95	ND *	ND *
STD Rt (min) N=20	0.00	0.00	0.00	0.01	ND *	ND *
% RSD Rt (min) N=20	0.2 %	0.2 %	0.1 %	0.2 %	ND *	ND *

* ND = Not detectable.

Table 16: Specific punch locations: Average and relative standard deviation in % (% RSD) of peak areas (PA) and retention times (Rt) of the initial compounds in 10 dried blood spot samples from each punch location A, B, C and D (A was in the center, B and C had increased distance to the center, while D was in the perimeter) analyzed with the initial method.

Position	**Creatinine**	**Creatine**	**Uric acid**	**C2**	**C12**	**C16**
A, center						
Average PA (a.u.) N=10	3.64E+07	1.74E+08	2.46E+07	2.10E+07	ND *	ND *
% RSD PA (a.u.) N=10	3 %	3 %	2 %	3 %	ND *	ND *
B						
Average PA (a.u.) N=10	3.51E+07	1.57E+08	2.32E+07	1.91E+07	ND *	ND *
% RSD PA (a.u.) N=10	9 %	9 %	9 %	10 %	ND *	ND *
C						
Average PA (a.u.) N=10	3.82E+07	1.69E+08	2.24E+07	2.13E+07	ND *	ND *
% RSD PA (a.u.) N=10	5 %	8 %	2 %	7 %	ND *	ND *
D, perimeter						
Average PA (a.u.) N=10	4.20E+07	1.87E+08	2.42E+07	2.36E+07	ND *	ND *
% RSD PA (a.u.) N=10	8 %	11 %	3 %	11 %	ND *	ND *
Retention time						
Average Rt (min) N=40	2.03	2.38	3.95	2.95	ND *	ND *
STD Rt (min) N=40	0.00	0.01	0.01	0.01	ND *	ND *
% RSD Rt (min) N=40	0.2 %	0.3 %	0.2 %	0.3 %	ND *	ND *

* ND = Not detectable.

Table 17: Dried blood spots and dried plasma spots samples: Average and relative standard deviation in % (% RSD) of peak areas (PA) of initial compounds in three dried blood spot samples (N=3) and three dried plasma samples (N=3) with center and perimeter punches analyzed with the initial method, with the ratio of peak areas from center/perimeter and blood/plasma.

	Creatinine	Creatine	Uric acid	C2	C12	C16
Blood center						
Average PA (a.u.) N=10	4.65E+07	1.69E+08	4.28E+07	1.74E+07	ND *	ND *
% RSD PA (a.u.) N=10	5 %	1 %	5 %	2 %	ND *	ND *
Blood perimeter						
Average PA (a.u.) N=10	5.09E+07	1.80E+08	4.47E+07	1.87E+07	ND *	ND *
% RSD PA (a.u.) N=10	5 %	4 %	4 %	6 %	ND *	ND *
Plasma center						
Average PA (a.u.) N=10	4.01E+07	1.47E+07	5.32E+07	3.92E+06	ND *	ND *
% RSD PA (a.u.) N=10	3 %	1 %	7 %	0 %	ND *	ND *
Plasma perimeter						
Average PA (a.u.) N=10	4.38E+07	1.62E+07	4.70E+07	4.67E+06	ND *	ND *
% RSD PA (a.u.) N=10	4 %	5 %	9 %	6 %	ND *	ND *
Retention time						
Average Rt (min) N=40	2.03	2.38	3.95	2.95	ND *	ND *
STD Rt (min) N=40	0	0.01	0.01	0.01	ND *	ND *
% RSD Rt (min) N=40	0.20 %	0.30 %	0.20 %	0.30 %	ND *	ND *

* ND = Not detectable.

Table 18: Organic solvent amount: Average and relative standard deviation in % (% RSD) of peak areas (PA) and retention times (Rt) of the initial compounds in three dried blood spot samples (N=3) with 2 %, 40 %, 80 % and extract 80 % (extraction without evaporation to dryness and reconstitution steps) organic solvent amount in the reconstitution solution analyzed with the initial method.

Methanol amount	**Creatinine**	**Creatine**	**Uric acid**	**C2**	**C12**	**C16**
2 %						
Average PA (a.u.) N=3	2.82E+07	1.22E+08	2.48E+07	8.97E+06	ND *	ND *
% RSD PA (a.u.) N=3	9 %	7 %	5 %	3 %	ND *	ND *
Average Rt (min) N=3	2.10	2.40	3.90	3.09	ND *	ND *
% RSD Rt (min) N=3	0.3 %	0.2 %	0.3 %	0.0 %	ND *	ND *
40 %						
Average PA (a.u.) N=3	2.09E+07	1.29E+08	2.73E+07	8.53E+06	ND *	2.64E+05*
% RSD PA (a.u.) N=3	5 %	6 %	0.6 %	2 %	ND *	ND *
Average Rt (min) N=3	2.10	2.40	3.87	2.73	ND *	13.77
% RSD Rt (min) N=3	0.3 %	0.2 %	0.2 %	0.4 %	ND *	0.1 %
80 %						
Average PA (a.u.) N=3	3.12E+07	1.29E+08	2.58E+07	1.03E+07	ND *	9.32E+05*
% RSD PA (a.u.) N=3	10 %	6 %	1 %	8 %	ND *	12 %
Average Rt (min) N=3	2.10	2.39	3.86	2.69	ND *	13.77
% RSD Rt (min) N=3	0.3 %	0.4 %	0.4 %	0.2 %	ND *	0.1 %
Extract 80 %						
Average PA (a.u.) N=3	3.95E+07	1.56E+08	3.51E+07	1.28E+07	ND *	1.03E+06*
% RSD PA (a.u.) N=3	4 %	2 %	2 %	4 %	ND *	13 %
Average Rt (min) N=3	2.09	2.39	3.87	2.69	ND *	13.77
% RSD Rt (min) N=3	0.3 %	0.0 %	0.8 %	0.2 %	ND *	0.1 %

* 7 points across the chromatographic peak.

* ND = Not detectable.

Table 19: Acylcarnitine standard: Average and relative standard deviation in % (% RSD) of peak areas (PA) of acylcarnitines in the acylcarnitine standard solution analyzed three times (N=5) with the initial method.

Acylcarnitine	Peak area (a.u.)	%RSD Peak area (a.u) N=3
C3	3.17E+07	0.6 %
C4	1.13E+07	0.2 %
C5	1.06E+07	8 %
C6	4.07E+06	21 %
C8	7.87E+06	20 %
C10	9.23E+06	8 %
C12	2.97E+06	27 %
C14	1.45E+06	32 %
C16	7.99E+06	8 %
C18	2.40E+06	7 %

Table 20: Scan range: Average and relative standard deviation in % (% RSD) of peak areas (PA) and retention times (Rt) of the initial compounds and acylcarnitines in a dried blood spot sample analyzed three times (N=5) with the initial method (scan ranges of *m/z* 50-750 and 750-1700) and the final method (scan range *m/z* 50-750).

	Scan ranges *m/z* 50-750 and 750-1700		**Scan range *m/z* 50-750**		**Retention time**	
	Average PA (a.u.) N=3	**% RSD PA (a.u.) N=3**	**Average PA (a.u.) N=3**	**% RSD PA (a.u.) N=3**	**Average Rt (min) N=6**	**%RSD Rt (min) N=6**
Creatinine	5.03E+07	1 %	1.05E+08	2 %	2.07	0.2 %
Creatine	2.24E+08	3 %	4.39E+08	1 %	2.29	0.1 %
Uric acid	3.07E+07	5 %	5.90E+07	2 %	3.85	0.1 %
C0	3.66E+07	2 %	7.26E+07	2 %	2.08	0.1 %
C2	2.42E+07	3 %	4.80E+07	2 %	2.61	0.1 %
C3	8.27E+05	9 %	1.65E+06	12 %	4.95	0.1 %
C4	7.28E+05	8 %	1.42E+06	7 %	8.15	0.1 %
C5	4.92E+05	3 %	9.58E+05	3 %	11.19	0.1 %
C6	ND *	ND *	5.27E+05	0.5 %	11.67	0.1 %
C8	ND *	ND *	3.30E+05	2 %	12.10	0.1 %
C10	2.21E+05	8 %	4.67E+05	4 %	12.40	0.1 %
C12	ND *	ND *	1.37E+05	4 %	12.73	0.1 %
C14	ND *	ND *	4.03E+05	3 %	13.13	0.1 %
C16	1.92E+06	3 %	3.70E+06	2 %	13.66	0.1 %
C18	7.20E+05	4 %	1.38E+06	5 %	14.35	0.1 %

* ND = Not detectable.

Table 21: Final method: Average and relative standard deviation in % (% RSD) of peak areas (PA) and retention times (Rt) of the final compounds in a dried blood spot sample analyzed 5 times (N=5) with the initial method.

Compound	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5	Average Rt (min) N=5	% RSD Rt (min) N=5
Ornithine	3.07E+07	2 %	1.62	0.2 %
Arginine	9.87E+06	7 %	1.77	0.08 %
Glycine	6.20E+07	3 %	1.99	0.2 %
Alanine	1.15E+08	2 %	2.05	0.2 %
Citrulline	3.49E+07	1 %	2.19	0.1 %
Valine	3.40E+08	4 %	2.52	0.1 %
Methionine	1.85E+07	3 %	3.10	0.2 %
Leucine	2.70E+08	3 %	3.43	0.5 %
Tyrosine	8.47E+07	1 %	4.25	0.1 %
Phenylalanine	1.23E+08	2 %	6.97	0.1 %
C0	7.23E+07	1 %	2.08	0.2 %
C2	4.68E+07	1 %	2.61	0.1 %
C3	1.47E+06	7 %	4.85	0.3 %
C4	9.04E+05	6 %	8.03	0.3 %
C5	5.45E+05	3 %	11.17	0.03 %
C8	3.13E+05	1 %	12.09	0.04 %
C12	1.13E+05	3 %	12.72	0.1 %
C14	3.94E+05	3 %	13.12	0.1 %
C16	3.65E+06	4 %	13.66	0.04 %
C18	1.28E+06	4 %	14.36	0.1 %

Table 22: Quantitativity: Average and relative standard deviation in % (% RSD) of peak areas (PA) the final compounds five dried blood spot samples (N=5) prepared with 1, 2, 3 and 4 punches and analyzed with the final method.

Ornithine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	3.07E+07	2 %
2	3.57E+07	3 %
3	3.84E+07	2 %
4	3.99E+07	3 %

C0

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	7.23E+07	1 %
2	1.28E+08	6 %
3	1.80E+08	1 %
4	2.38E+08	1 %

Arginine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	9.87E+06	7 %
2	1.28E+07	4 %
3	1.55E+07	4 %
4	1.64E+07	9 %

C2

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	4.68E+07	0.8 %
2	8.21E+07	6 %
3	1.16E+08	0.6 %
4	1.56E+08	2 %

Glycine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	6.20E+07	3 %
2	9.60E+07	5 %
3	1.26E+08	2 %
4	1.59E+08	2 %

C3

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	1.47E+06	7 %
2	2.94E+06	13 %
3	4.63E+06	6 %
4	7.26E+06	1 %

Alanine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	1.15E+08	2 %
2	1.77E+08	5 %
3	2.37E+08	1 %
4	2.93E+08	1 %

C4

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	9.04E+05	6 %
2	2.40E+06	11 %
3	3.57E+06	4 %
4	5.20E+06	1 %

Citrulline

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	3.49E+07	21 %
2	5.13E+07	5 %
3	6.98E+07	1 %
4	9.13E+07	2 %

C5

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	5.45E+05	3 %
2	1.04E+06	9 %
3	1.52E+06	1 %
4	2.06E+06	2 %

Valine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	3.40E+08	4 %
2	5.81E+08	6 %
3	7.78E+08	0.8 %
4	1.00E+09	1.5 %

C8

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	3.13E+05	0.8 %
2	6.37E+05	8 %
3	9.50E+05	6 %
4	1.35E+06	4 %

Methionine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	1.85E+07	3 %
2	3.33E+07	8 %
3	4.94E+07	10 %
4	6.55E+07	6 %

C12

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	1.13E+05	17 %
2	2.41E+05	12 %
3	3.87E+05	8 %
4	5.62E+05	2 %

Leucine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	2.70E+08	3 %
2	5.17E+08	13 %
3	7.68E+08	2 %
4	1.04E+09	3 %

C14

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	3.94E+05	3 %
2	7.81E+05	10 %
3	1.17E+06	7 %
4	1.73E+06	5 %

Tyrosine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	8.74E+07	0.9 %
2	1.27E+08	6 %
3	1.96E+08	2 %
4	2.69E+08	2 %

C16

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	3.65E+06	4 %
2	6.99E+06	8 %
3	1.05E+07	2 %
4	1.51E+07	1 %

Phenylalanine

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	1.23E+08	2 %
2	2.24E+08	7 %
3	3.31E+08	0.9 %
4	4.56E+08	1 %

C18

Punches	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5
1	1.28E+06	4 %
2	2.44E+06	8 %
3	3.89E+06	5 %
4	5.59E+06	1 %

Table 23: Matrix: Average and relative standard deviation in % (% RSD) of peak areas (PA) of the internal standard compounds in the internal standard solution with the absence of DBS matrix analyzed 15 times and in internal standard solution with the presence of DBS matrix analyzed 60 times with the final method. With the ratio of peak areas of IS in the presence of DBS matrix (ISmatrix) over IS in the absence of DBS matrix (ISsolution).

IS Compound	Presence of DBS matrix		Absence of DBS matrix		
	Average PA (a.u.) N=15	% RSD PA (a.u.) N=15	Average PA (a.u.) N=60	% RSD PA (a.u.) N=60	IS matrix/IS solution
Ornithine	5.94E+08	0.9 %	3.57E+08	3 %	166 %
Arginine	8.98E+08	1 %	3.23E+08	2 %	278 %
Glycine	5.59E+08	1 %	5.83E+08	2 %	96 %
Alanine	2.97E+08	1 %	3.04E+08	2 %	97 %
Citrulline	6.74E+08	1 %	7.36E+08	3 %	92 %
Valine	1.29E+09	1 %	1.38E+09	2 %	94 %
Methionine	9.11E+08	13 %	6.51E+08	4 %	140 %
Leucine	1.42E+09	1 %	1.59E+09	2 %	89 %
Tyrosine	6.19E+08	3 %	6.29E+08	5 %	98 %
Phenylalanine	1.42E+09	3 %	1.57E+09	4 %	91 %
C0	6.21E+08	1 %	6.74E+08	2 %	92 %
C2	2.33E+07	1 %	2.43E+07	4 %	96 %
C3	1.39E+06	8 %	1.76E+06	10 %	79 %
C4	1.28E+07	11 %	1.54E+07	8 %	83 %
C5	7.60E+07	1 %	8.24E+07	5 %	92 %
C8	2.79E+07	3 %	3.02E+07	6 %	93 %
C12	6.39E+07	3 %	6.79E+07	8 %	94 %
C14	5.98E+07	4 %	6.34E+07	7 %	94 %
C16	1.18E+08	2 %	1.24E+08	9 %	95 %
C18	1.06E+08	5 %	1.10E+08	10 %	96 %

Table 24: Extraction efficiency: Average and relative standard deviation in % (% RSD) of peak areas (PA) of the internal standard compounds in spotted and diluted IS solution analyzed five times each with the final method. With ratio of peak areas of IS in the diluted solution over IS in the spotted sample.

IS Compound	IS diluted		IS spotted		
	Average Pa (a.u.) N=5	% RSD PA (a.u.) N=5	Average PA (a.u.) N=5	% RSD PA (a.u.) N=5	IS diluted/ IS spotted
Ornithine	1.58E+07	4 %	6.36E+06	6 %	40 %
Arginine	2.54E+07	8 %	1.08E+07	8 %	43 %
Alanine	1.13E+07	6 %	4.71E+06	5 %	42 %
Citrulline	2.01E+07	7 %	9.50E+06	5 %	47 %
Valine	4.86E+07	7 %	1.75E+07	3 %	36 %
Methionine	3.07E+07	8 %	5.59E+06	8 %	18 %
Leucine	3.30E+07	8 %	1.14E+07	8 %	34 %
Tyrosine	1.15E+07	9 %	4.66E+06	6 %	40 %
Phenylalanine	8.74E+05	10 %	3.20E+05	9 %	37 %
C0	2.33E+07	6 %	8.82E+06	4 %	38 %
C2	4.58E+05	8 %	1.44E+05	5 %	31 %
C5	1.41E+06	6 %	4.25E+05	5 %	30 %
C8	4.42E+05	7 %	1.00E+05	8 %	23 %
C12	1.02E+06	6 %	2.01E+05	8 %	20 %
C14	9.24E+05	8 %	1.80E+05	9 %	19 %
C16	1.76E+06	7 %	4.72E+05	9 %	27 %
C18	1.42E+06	5 %	4.23E+05	8 %	30 %

6.2 Dried blood spots and dried plasma spots

Whole blood from a healthy volunteer was collected in an EDTA tube and three droplets were spotted on to a filter paper card. The rest of the blood was centrifuged so that the red blood cells gathered in the bottom of the tube and plasma was pipetted out, and three droplets were spotted. Three DBS and DPS samples were prepared with punches from the center and the perimeter. Peak areas of the detected compounds in all samples are shown in **Figure 37**.

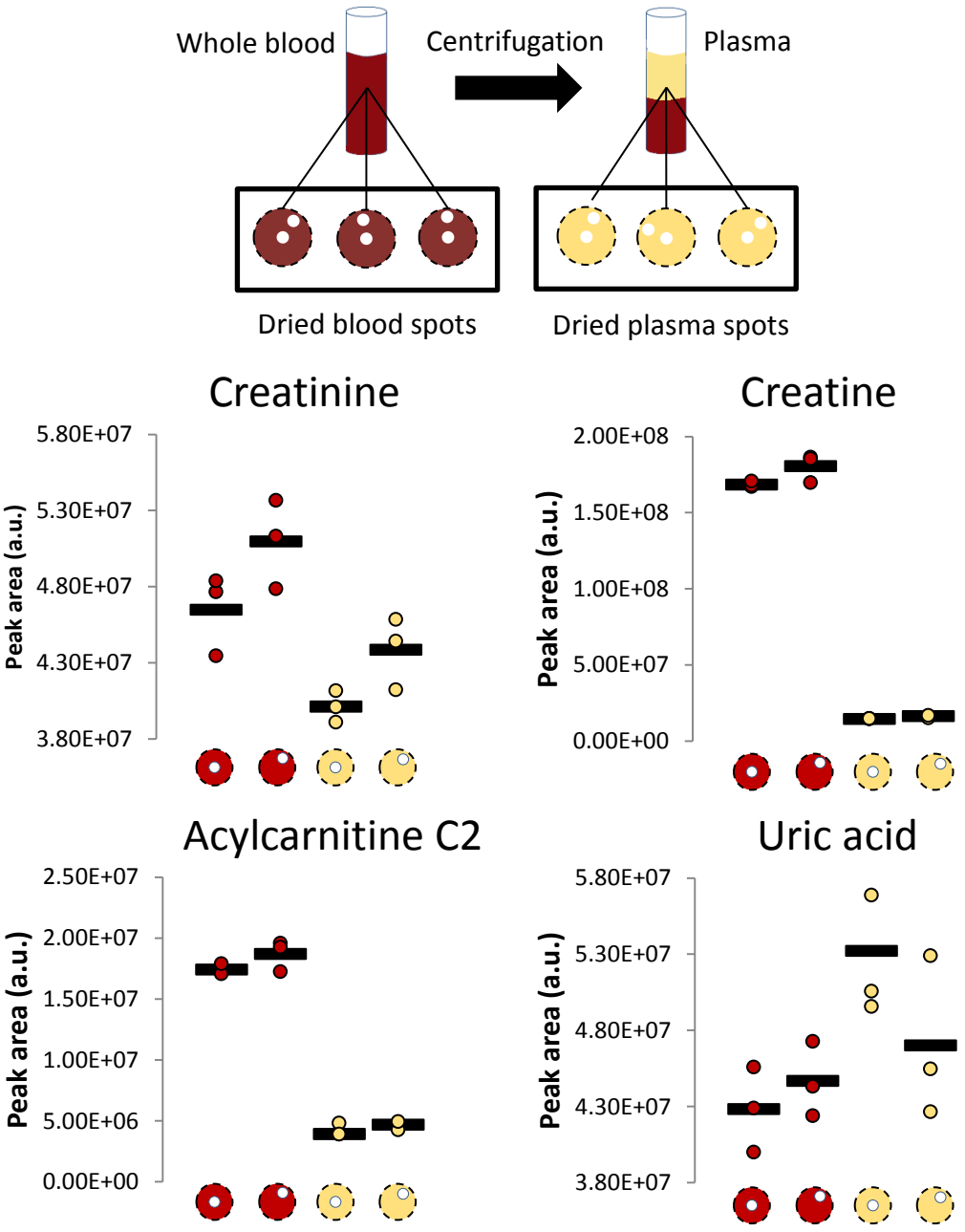


Figure 37: Whole blood spotted to make dried blood spots (DBS) and centrifuged to make dried plasma spots (DPS). Peak area of detected compounds in DBS with center punches, in DBS with perimeter punches, in DPS with center punches and DPS with perimeter punches (left to right, respectively).

Creatinine, creatine and C2 seemed to be present in higher concentrations in DBS samples than in DPS samples, but seemed to not be affected by the presence or absence of red blood cells. Uric acid seemed to be more presented in DPS than DBS and more was observed in the center of the DPS. This indicates that compounds in plasma were more presented in the center, supported by a study that observed more red blood cells in the perimeter of the punch compared to the center [42]. Their explanation was that plasma occupies a greater fraction of volume of the interior of the filter paper card than the blood cells, thus the red blood cells are spread to the perimeter. The red blood cells being more presented in the perimeter than in the center of the DBS can be seen by watching a DBS, as there is a darker perimeter than center of the DBS, see **Figure 38.**

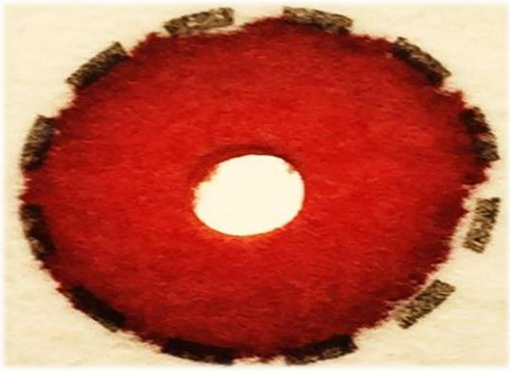


Figure 38: Close up photo of a dried blood spot illustrating that the perimeter is darker than the center.

When red blood cells are present, more compounds seemed to be present in the perimeter of the dried blood spot.

6.3 Detection of hydrophobic and low-abundant compounds

6.3.1 Mid-long acylcarnitines gave non repeatable measurements

The acylcarnitine standard solution was analyzed three times to observe the instrument method's ability to detect other hydrophobic compounds than the initial ones. The % RSD of the peak areas of the acylcarnitines plotted against its concentration in the acylcarnitine standard solution is shown in **Figure 39**.

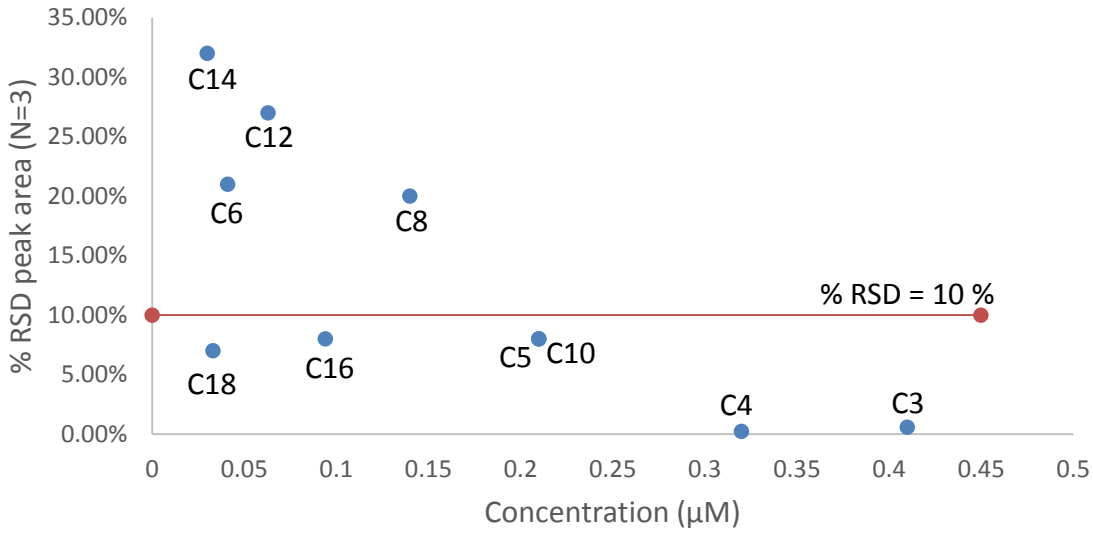


Figure 39: The % RSD of the peak areas of the acylcarnitines plotted against their concentrations in the acylcarnitine standard solution, analyzed three times. The line represents % RSD at 10 %.

Acylcarnitines C3, C4, C5, C10, C16 and C18 had a % RSD of the peak areas <10 % for the three analyses of the acylcarnitine standard solution, while the C6, C8, C12 and C14 peaks had >10 %. The latter ones were the mid-long acylcarnitines (C6-C14), except from C10 with the highest concentration of the mid-long acylcarnitines. The longest acylcarnitines (C16 and C18) were also present in low concentrations, but they were detected by the initial method with low % RSD of peak areas.

The initial method uses ESI, which is an ionization technique more suitable for polar than for non-polar compounds, thus, the poor detection of non-polar compounds may be due to poor ionization effect. Commonly, derivatization of the acylcarnitines in to butyl esters is done prior to analysis by LC-ESI-MS^2 to improve the electrospray ionization efficiency [23]. This is an example which illustrates that additional analyses with another ionization technique (e.g., APCI, which is more suitable for non-polar compounds [59]) may be useful to increase the coverage of the DBS metabolome.

6.3.2 Extraction efficiency affected the detection of low-abundant compounds

IS compounds were used to observe the extraction efficiency of compounds by comparing spotted IS solution and injected diluted IS solution with nearly the same concentrations of compounds. IS solution was spotted on a filter paper card and five samples were prepared out of these. 3 µL of IS solution was diluted in 100 µL extraction solution to obtain close to the same concentration in one punch, see **Figure 40**.

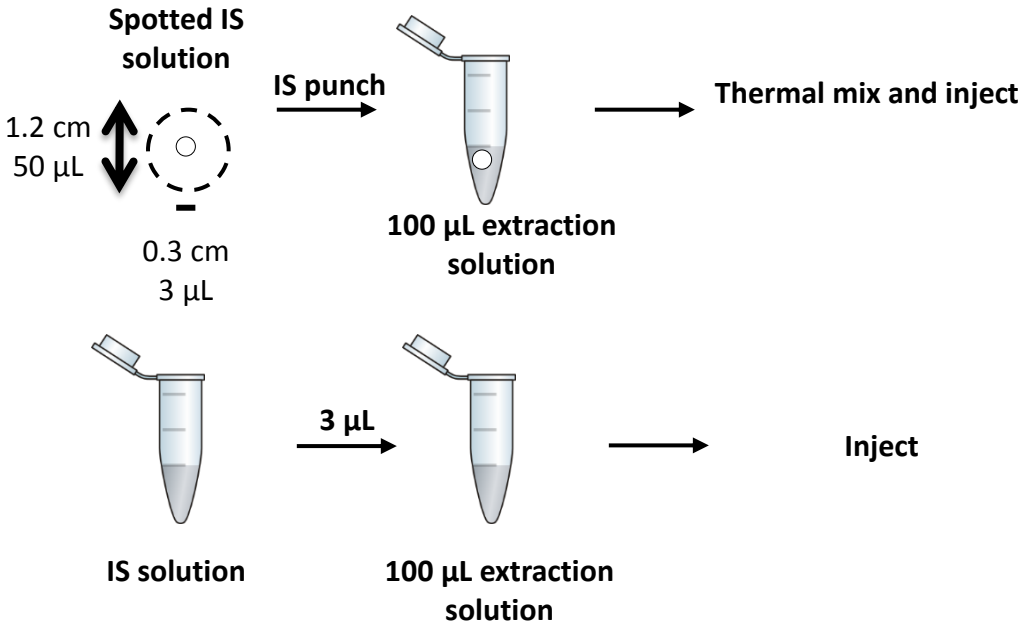


Figure 40: The preparation of spotted ISs and diluted IS to obtain about the same concentration used to investigate the extraction efficiency of the IS compounds.

The peak areas of the compounds detected in the IS solution were compared to peak areas of the compounds in the spotted ISs. **Figure 41** shows that methionine and acylcarnitines C8, C12 and C14 had the poorest extraction efficiencies. The peak shapes of acylcarnitine C3 and

C4 made it difficult to integrate the peak and were therefore not included in these results. C6 and C10 were not added to the IS solution.

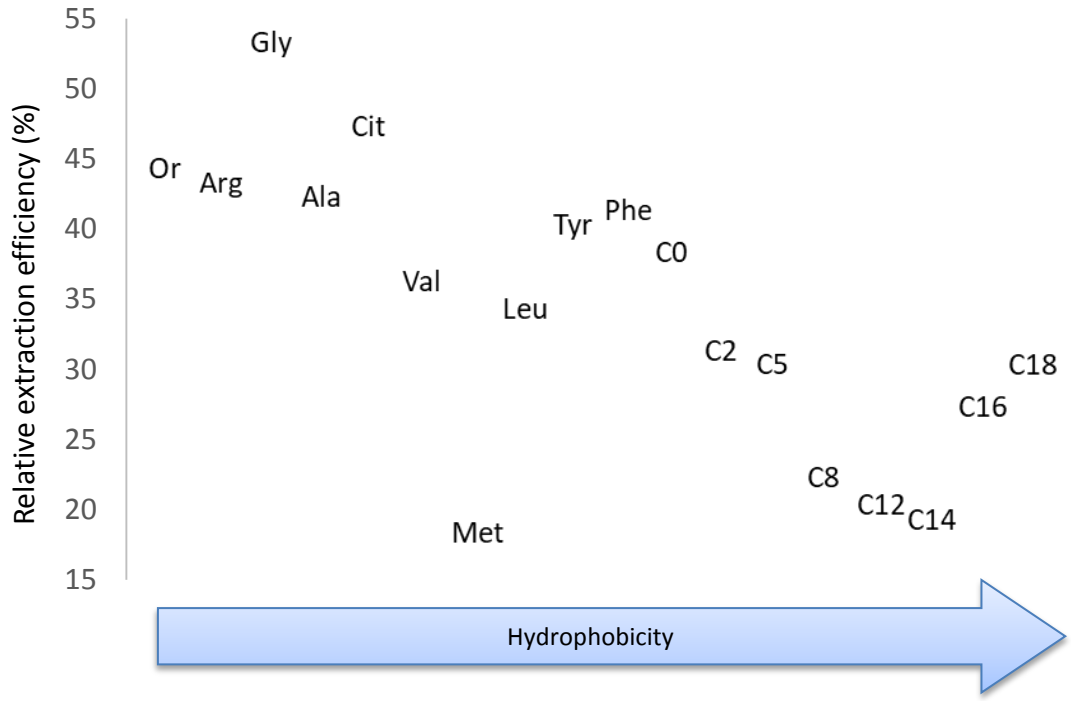


Figure 41: Relative extraction efficiency in % of the chosen compounds using internal standards.

The poor extraction efficiency will affect the detection of low abundant compounds in whole blood (e.g., acylcarnitines C8-C14), but not for the high abundant compounds (e.g., methionine) using the final method. The poor extraction of some compounds may be explained by the extraction solution not being suitable for these compounds, or that they are more presented in red blood cells (more difficult to get into solution). The most important factor is that the extraction efficiency is repeatable to obtain confidential quantifications. The % RSD (N=5) was satisfactory, <10 % for all final compounds.

6.4 Final sample preparation method

Figure 42 shows the final sample preparation method in pictures.

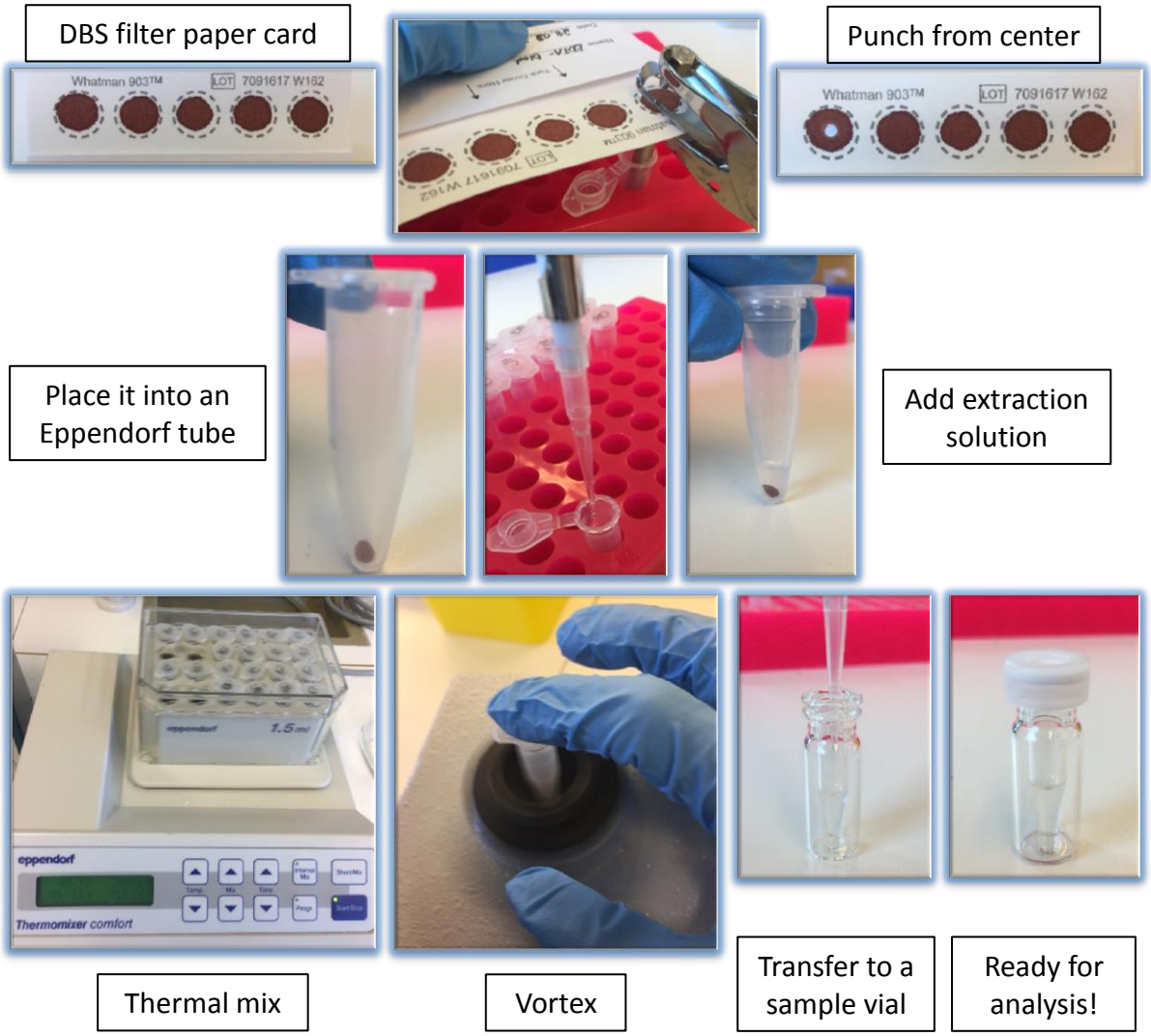
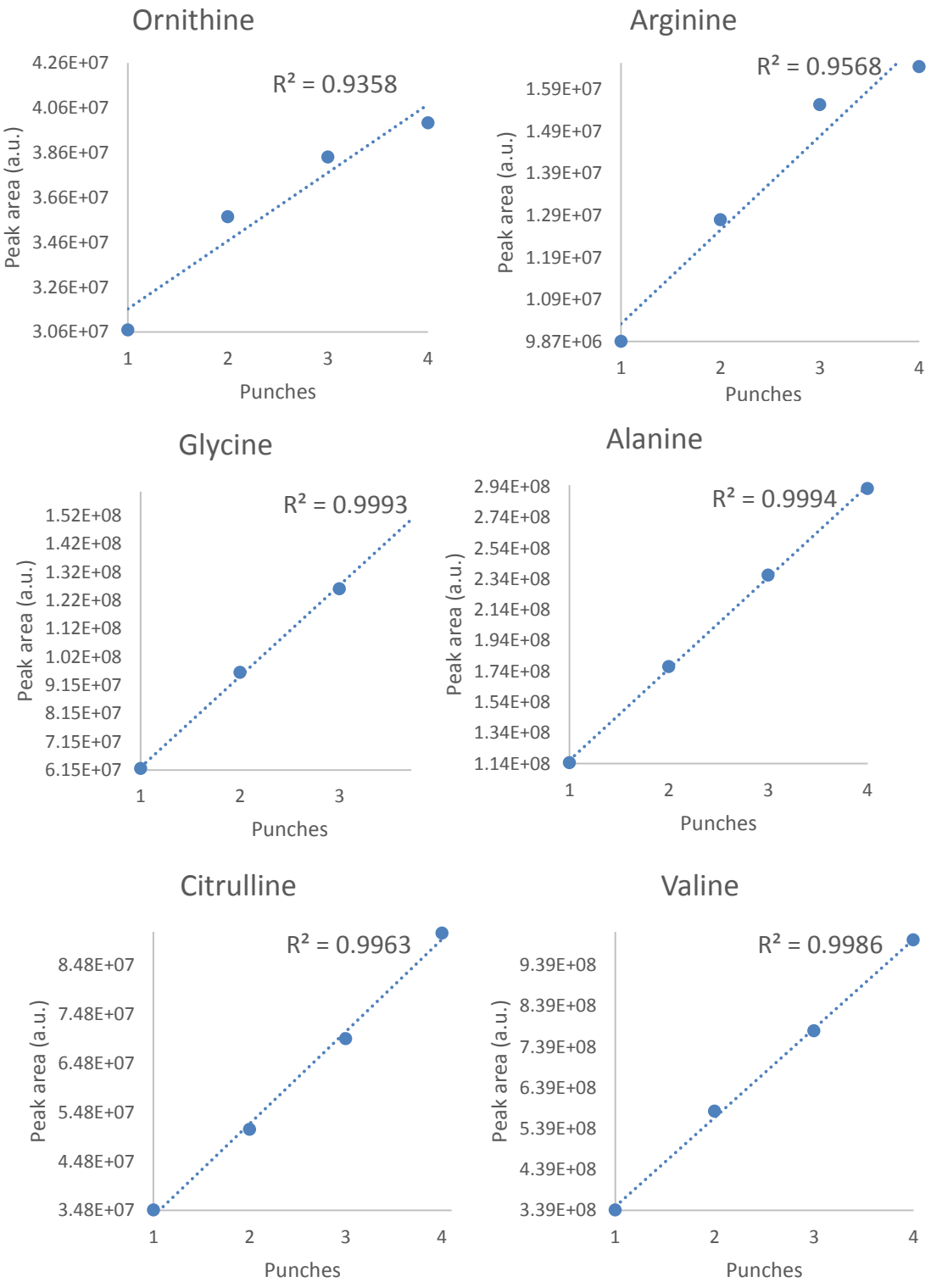
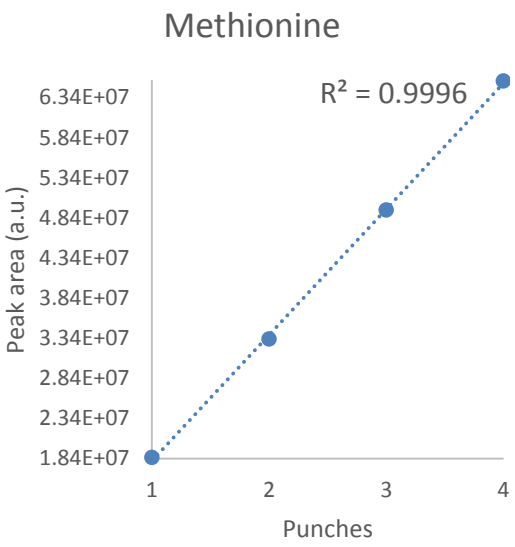


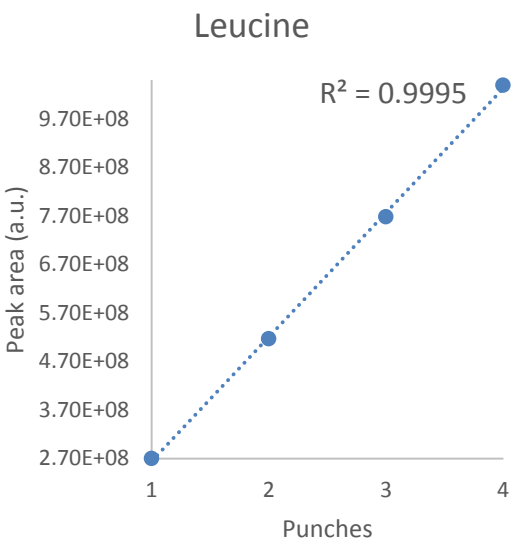
Figure 42: Final sample preparation method.

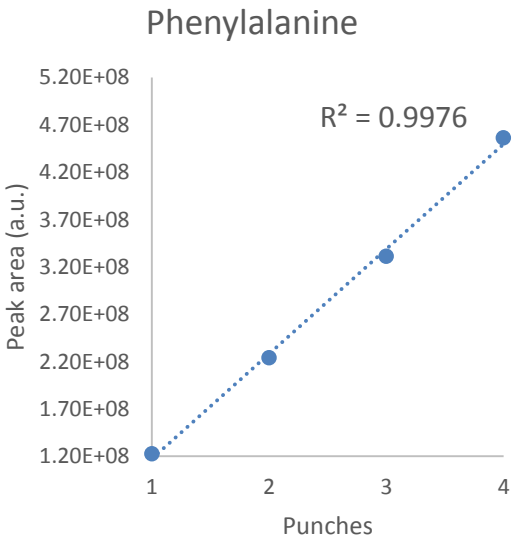
6.5 Peak area over punches plots

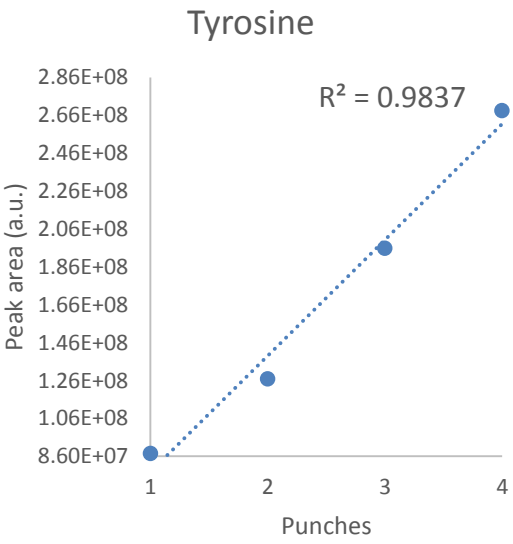
Figure 43 shows the peak area over punches plots with R^2 value for all final compounds obtained by analyzing 1, 2, 3 and 4 punches by the final method.

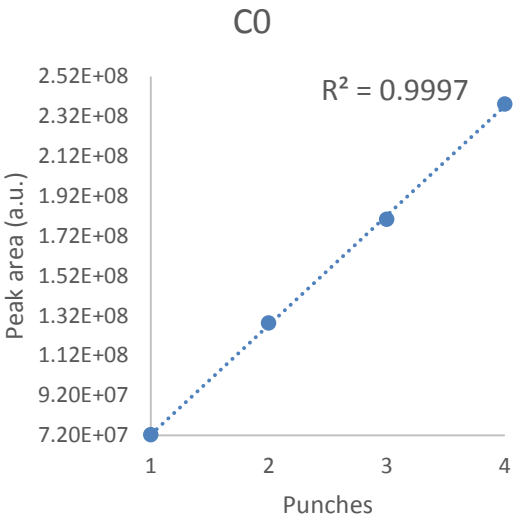


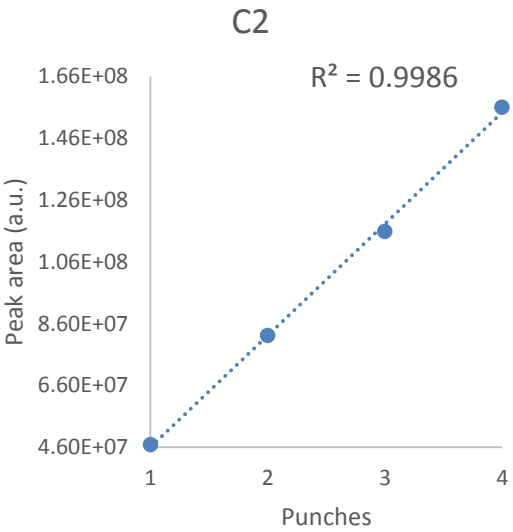

Methionine
R² = 0.9996
6.34E+07
5.84E+07
5.34E+07
4.84E+07
4.34E+07
3.84E+07
3.34E+07
2.84E+07
2.34E+07
1.84E+07
Peak area (a.u.)
1
2
3
4
Punches

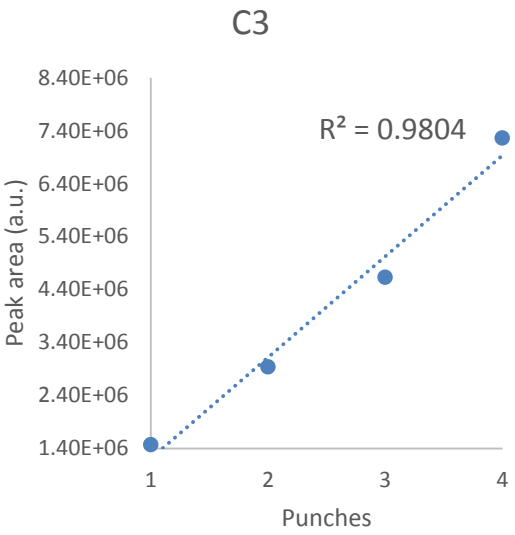

Leucine
R² = 0.9995
9.70E+08
8.70E+08
7.70E+08
6.70E+08
5.70E+08
4.70E+08
3.70E+08
2.70E+08
Peak area (a.u.)
1
2
3
4
Punches

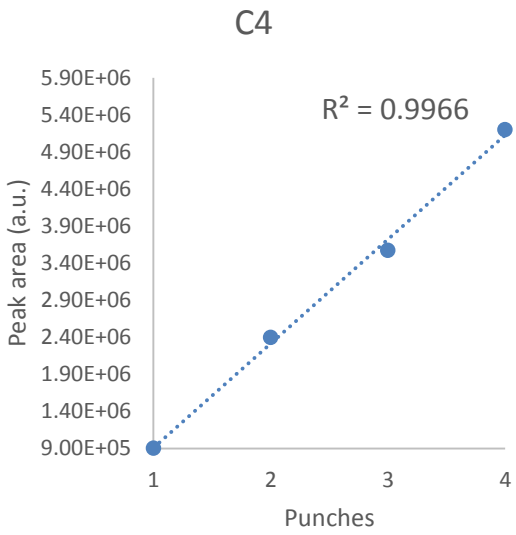

Phenylalanine
R² = 0.9976
5.20E+08
4.70E+08
4.20E+08
3.70E+08
3.20E+08
2.70E+08
2.20E+08
1.70E+08
1.20E+08
Peak area (a.u.)
1
2
3
4
Punches

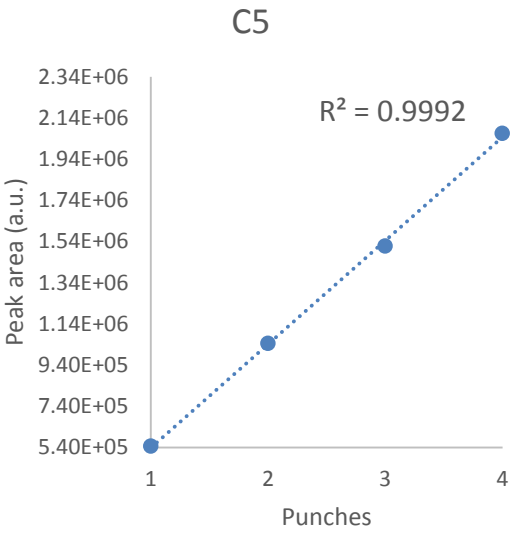

Tyrosine
R² = 0.9837
2.86E+08
2.66E+08
2.46E+08
2.26E+08
2.06E+08
1.86E+08
1.66E+08
1.46E+08
1.26E+08
1.06E+08
8.60E+07
Peak area (a.u.)
1
2
3
4
Punches

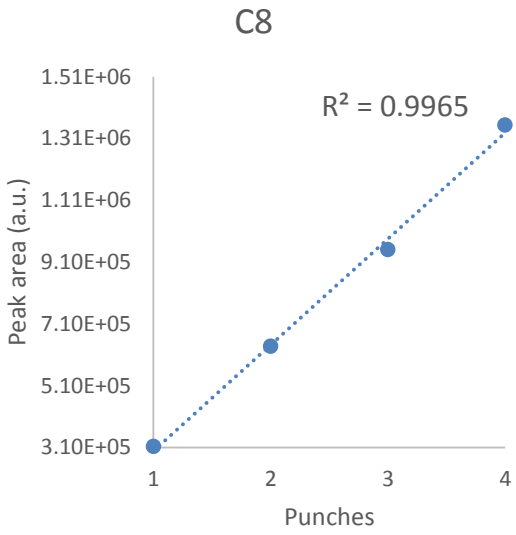

C0
R² = 0.9997
2.52E+08
2.32E+08
2.12E+08
1.92E+08
1.72E+08
1.52E+08
1.32E+08
1.12E+08
9.20E+07
7.20E+07
Peak area (a.u.)
1
2
3
4
Punches

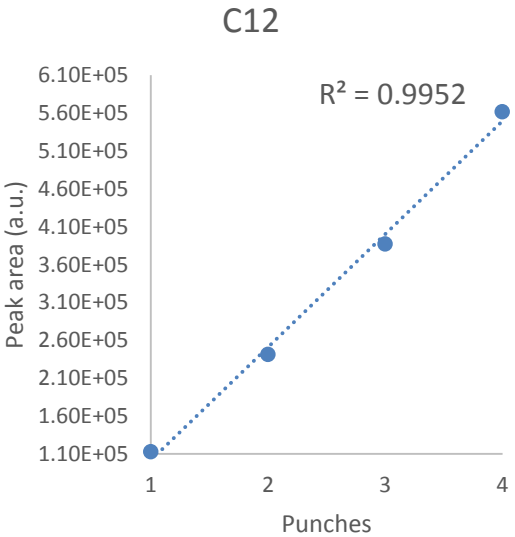

C2
R² = 0.9986
1.66E+08
1.46E+08
1.26E+08
1.06E+08
8.60E+07
6.60E+07
4.60E+07
Peak area (a.u.)
1
2
3
4
Punches

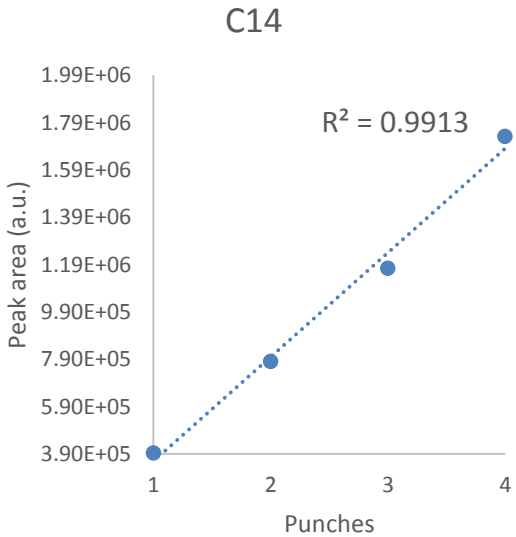

C3
R² = 0.9804
Peak area (a.u.)
8.40E+06
7.40E+06
6.40E+06
5.40E+06
4.40E+06
3.40E+06
2.40E+06
1.40E+06
1
2
3
4
Punches


C4
R² = 0.9966
Peak area (a.u.)
5.90E+06
5.40E+06
4.90E+06
4.40E+06
3.90E+06
3.40E+06
2.90E+06
2.40E+06
1.90E+06
1.40E+06
9.00E+05
1
2
3
4
Punches


C5
R² = 0.9992
Peak area (a.u.)
2.34E+06
2.14E+06
1.94E+06
1.74E+06
1.54E+06
1.34E+06
1.14E+06
9.40E+05
7.40E+05
5.40E+05
1
2
3
4
Punches


C8
R² = 0.9965
Peak area (a.u.)
1.51E+06
1.31E+06
1.11E+06
9.10E+05
7.10E+05
5.10E+05
3.10E+05
1
2
3
4
Punches


C12
R² = 0.9952
Peak area (a.u.)
6.10E+05
5.60E+05
5.10E+05
4.60E+05
4.10E+05
3.60E+05
3.10E+05
2.60E+05
2.10E+05
1.60E+05
1.10E+05
1
2
3
4
Punches


C14
R² = 0.9913
Peak area (a.u.)
1.99E+06
1.79E+06
1.59E+06
1.39E+06
1.19E+06
9.90E+05
7.90E+05
5.90E+05
3.90E+05
1
2
3
4
Punches

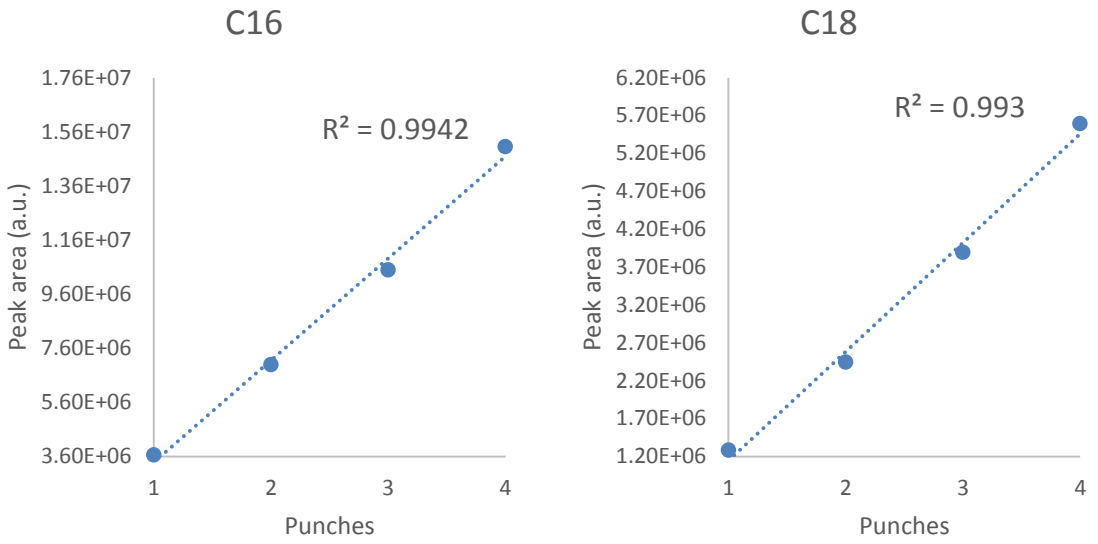


Figure 43: Peak area over punches for all final compounds with R^2 value for all final compounds obtained by analyzing 1, 2, 3 and 4 punches by the final method.

6.5.1 The most polar compounds were affected by matrix effects

Other compounds can affect the quantification of specific compounds in a biological sample. The ions may compete for space in the ion source resulting in suppression or enhancement of ion signal, which are examples of matrix effects. The isotopically labeled ISs have the same chemical and physical properties as the endogenous compounds, thus the analysis of the ISs can represent the impact on compounds by the instrument. The ISs were used to observe the matrix effects of the chosen compounds by analyzing the IS solution in the absence and presence of the DBS matrix. The shape and area of the chromatographic peaks after analysis of the samples were compared. A DBS sample was prepared with ISs in the extraction solution and analyzed 60 times, and the IS solution was analyzed 15 times.

All the peaks of the 60 injections of the DBS sample with IS and the 15 injections of the IS solution were investigated. The peak area of the polar compounds, ornithine, arginine and methionine were enhanced with ratios of 278, 166 and 140 %, respectively, and the peak area of C3 and C4 were slightly suppressed with ratios of 79 and 83 %, respectively. Ornithine, arginine and C3 showed poorer quantitativity than the other compounds and may be explained by the matrix effect. **Figure 44** shows the chromatographic peaks with and without the DBS matrix of the compounds that were most affected by the DBS matrix and two compounds that were not affected: alanine and C18. The rest of the chromatographic peaks of the final compounds were not affected and are shown in **Figure 45**.

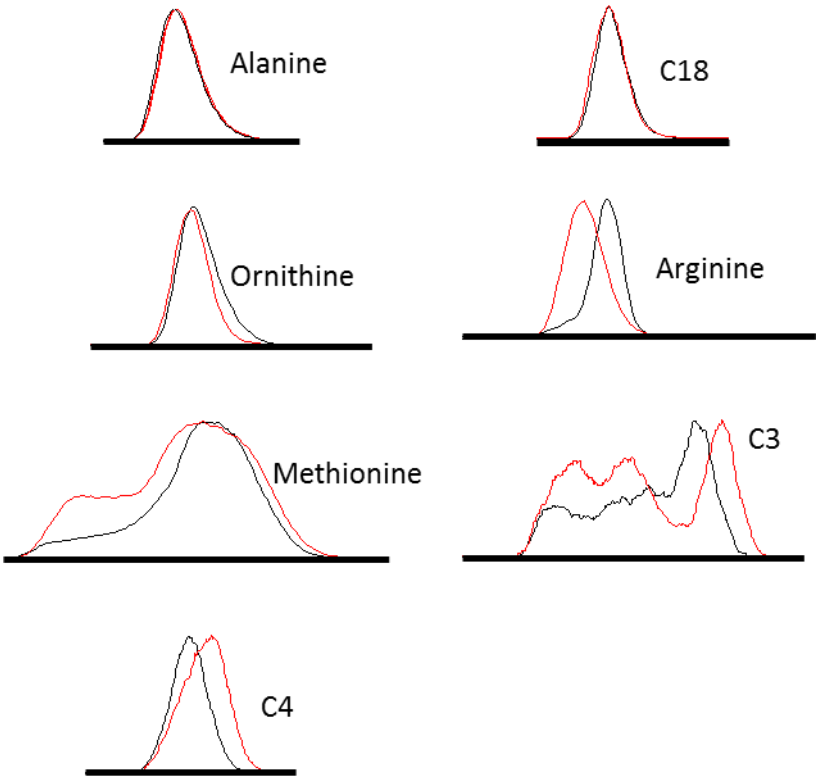


Figure 44: Comparison of the chromatographic peaks of the IS compounds in the presence (red) and in the absence (black) of the DBS matrix. Alanine and C18 and the rest of the final compounds were not affected by the DBS matrix, except from ornithine, arginine, methionine, C3 and C4.

Polar compounds are retained poorly using reverse phase chromatography. HILIC and normal phase chromatography are complementary separation techniques to reverse phase chromatography. The use of one of these separation techniques may solve the matrix effect to an extent by separating the polar compounds better than reverse phase chromatography, reducing the number of polar compounds entering the ion source simultaneously [47].

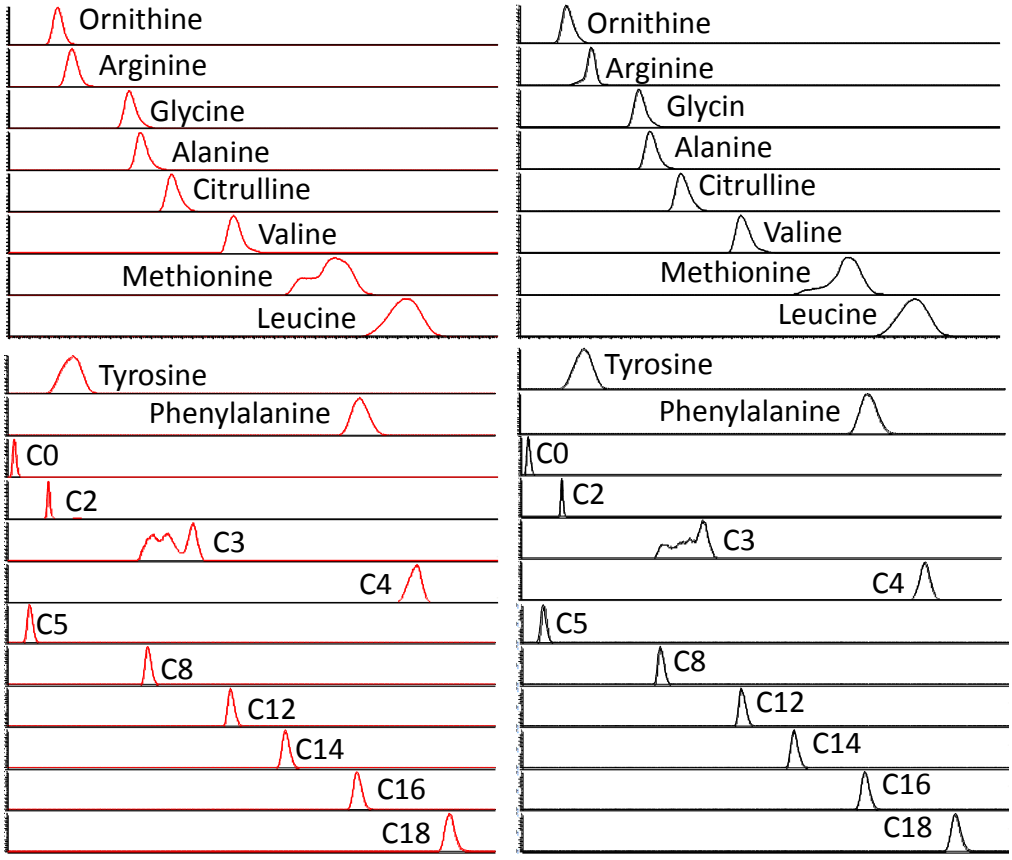


Figure 45: Chromatographic peaks of the final compounds in the presence (red) and absence (black) of the DBS matrix.

6.6 Fasting project

6.6.1 Regional Committee for Medical and Health Research Ethics informed consent form

An application for a research project to investigate the normal metabolome was approved by REC. The informed consent form for the healthy volunteers of the project is described below.

Forespørsel om å avgi biologisk materiale til

KARTLEGGING AV NORMALMETABOLOMET

Bakgrunn og hensikt

Formålet med biobanken er å undersøke sammensetningen av stoffene som kan påvises i kroppsvæsker i en normalbefolkning. Metabolismen er summen av alle kjemiske reaksjoner i kroppen og innebærer nedbrytning og oppbygging av ulike stoffer som inngår i prosessene som skjer i kroppen. Stoffene som deltar i metabolismen kalles metabolitter, og sammensetningen av metabolitter utgjør det som kalles metabolomet. Metabolismen er et dynamisk system, det vil si det er i endring hele tiden, og påvirkes av en rekke naturlige biologiske faktorer slik som alder, kjønn, tid på døgnet, og ytre påvirkninger slik som matinntak/faste, fysisk aktivitet, sykdom, inntak av legemidler m.m.

For å kunne skille variasjoner i metabolomet som skyldes normale prosesser fra avvik som skyldes sykdom er det nødvendig å undersøke metabolomet fra et stort antall friske personer under ulike betingelser (f.eks. tid på døgnet, fødeinntak, fysisk aktivitet).

Forskningsansvarlig er Oslo Universitetssykehus, Avdeling for medisinsk biokjemi ved avdelingsleder. Prosjektleder er Katja B. Prestø Elgstøen ved samme avdeling.

Det biologiske materialet blir oppbevart på ubestemt tid og skal brukes i fremtidig forskning til ulike studier av normalmetabolomet.

Hvilket biologisk materiale skal innsamles?

Vi undersøker i hovedsak blod og/eller urin. I sjeldne tilfeller kan det være aktuelt med hår, spytt, tårevæske, svette eller vev fra deg. Det vil bli et stikk (blodprøve fra armen og/eller stikk i fingeren) i forbindelse med blodprøvetaking. Dersom du gir fibroblaster (hudprøve), vil

du få lokalbedøvelse før vi tar en knappenålshode stor prøve fra huden din. Dersom du gir vev, vil dette være i forbindelse med en annen operasjon og du vil ikke oppleve ekstra ubehag pga dette.

Bredt samtykke

Prøvene som du avgir lagres i den generelle forskningsbiobanken 2018/787 «Kartlegging av normalmetabolomet». Når du avgir biologisk materiale til denne forskningsbiobanken gir du også et bredt samtykke til at materiale og relevante helseopplysninger kan brukes til fremtidig forskning som har til hensikt å kartlegge normalmetabolomet, og som normalkontroll i studie av ulike sykdommer.

Innsamling og bruk av helseopplysninger

Biobanken vil inneholde noen opplysninger om deg. Disse er kun tilgjengelige gjennom en koblingsnøkkel som skal beskytte din identitet, men samtidig gjøre det mulig å knytte opplysningene om deg til ditt materiale gjennom en kodeliste dersom det skulle være nødvendig. Oslo Universitetssykehus er ansvarlig for at koblingsnøkkel oppbevares og forvaltes forsvarlig. Materiale og opplysningene om deg lagres permanent og vil analyseres i forbindelse med spesifiserte forskningsprosjekter. Det kan også være ønskelig å registrere annen relevant informasjon enn bare alder og kjønn (kosthold, medisinbruk, kjent sykdom eller liknende). Du står i så fall helt fritt til å oppgi dette.

Det kan være aktuelt å gjøre relevante undersøkelser på gennivå for å bidra til forståelsen av variasjoner i metabolomet.

Genetiske undersøkelser

Det kan være aktuelt å gjøre genetiske analyser på det materialet som er samlet inn for å se på sammenhengen mellom metabolomet og underliggende gener og genvarianter (sekvensering av enkeltgen, grupper av gener eller hele genomet). Genomsekvensen til hvert enkelt menneske er så unik at ingen prøver som inneholder DNA i teorien kan være anonyme. I praksis blir imidlertid alle prøver avidentifiserte og ingen (verken forskere eller personer som deltar i biobanken) informeres om resultater og funn hos enkeltpersoner.

Godkjenning av fremtidige forskningsprosjekter

Alle fremtidige forskningsprosjekter som benytter materialet fra deg skal forhåndsgodkjennes av en regional komité for medisinsk og helsefaglig forskningsetikk, men du vil kun unntaksvis bli spurt på nytt om slik bruk. Det vil derfor kun unntaksvis kunne være behov for å spørre deg på ny om slik bruk av prøvematerialet.

Informasjon om fremtidige prosjekter

Informasjon om forskningsbiobanken 2018/787 kartlegging av normalmetabolomet finnes på Biobankportalen til Oslo Universitetssykehus: https://oslo-universitetssykehus.no/forskningsbiobanker-ved-oslo-universitetssykehus

På denne nettsiden er det lenke videre til forskningsprosjektets nettside hvor informasjon om pågående og framtidige prosjekter er beskrevet.

Utlevering av prøvemateriale

Det kan være aktuelt å sende prøver ut av landet (innad i EU og til USA, data som overføres til USA behandles i henhold til EUs Personverndirektiv), som et ledd i internasjonalt samarbeid. Materialet vil i så fall kun utleveres uten navn, fødselsnummer eller andre personidentifiserbare opplysninger.

Det er frivillig å delta

Å avgi biologisk materiale til Kartlegging av normalmetabolomet er frivillig og krever samtykke. Informasjonen som registreres om deg skal kun brukes slik som beskrevet i hensikten med studien.

Dersom du ønsker å avgi biologisk materiale og godkjenner at prøvene dine lagres for eventuell bruk i fremtidige forskningsprosjekter, undertegner du samtykkeerklæringen på siste side. Der vil det bli spurt spesifikt om a) oppbevaring av din prøve i biobanken til fremtidige studier av normalmetabolomet, b) bruk av prøvene som normalkontroll i andre forskningsprosjekter, og c) om det kan utføres relevante analyser på gennivå knyttet til kartlegging av normalmetabolomet.

Mulighet for å trekke sitt samtykke, innsynsrett, endring og sletting av opplysninger

Du kan til enhver tid få innsyn i hvilket materiale som er lagret fra deg. Du kan når som helst kreve at materialet blir destruert, uten at du må oppgi noen grunn. Destruksjon av materialet vil imidlertid ikke innebære sletting av utledete opplysninger som allerede har inngått i sammenstilling eller analyser.

Kontakt

Biobanken er lokalisert ved Avdeling for medisinsk biokjemi, Rikshospitalet, Oslo Universitetssykehus. Katja B Prestø Elgstøen (kelgstoe@ous-hf.no, tlf. 23073079) er ansvarshavende for biobanken. Dette er en prospektiv biobank, og prøvene vil bli lagret til de eventuelt ønskes fjernet av deg.

OPPLYSNINGER OM DEG

Kode for avidentifisering:_____

Alder:

Kjønn:

Andre opplysninger (frivillig å oppgi):

Faste medisiner:

Andre medisiner tatt siste døgn:

Eventuelle sykdommer:

Kosttilskudd:

Eventuelt spesielt kosthold (f.eks. vegetarianer, laktoseintolerant):

Har du spist i dag? I så fall: hvor lenge er det siden, og hva spiste du?

Røyker eller snuser du?

Annen relevant informasjon:

Samtykke til lagring av biologisk materiale

a) Jeg avgir bredt samtykke til at mitt biologiske materiale kan oppbevares varig i biobanken 2018/787 Kartlegging av normalmetabolomet og kan bli benyttet i fremtidig forskning til ulike studier som omhandler kartlegging av normalmetabolomet

Sted og dato Deltakers signatur

b) Jeg godkjenner at mine prøver kan bli brukt som normalkontroll i andre forskningsprosjekter som er godkjent av Regional komité for medisinsk og helsefaglig forskningsetikk

Sted og dato Deltakers signatur

c) Jeg godkjenner at det kan utføres relevante analyser av mine prøver på gennivå knyttet til kartlegging av normalmetabolomet

Sted og dato Deltakers signatur

Deltakers navn med trykte bokstaver

Jeg bekrefter å ha gitt informasjon om prosjektet.

Sted og dato Signatur

Rolle i prosjektet

6.6.2 Fasting programme

Figure 46 shows the sampling programme during the fasting project and **Table 25** shows the programme with comments.

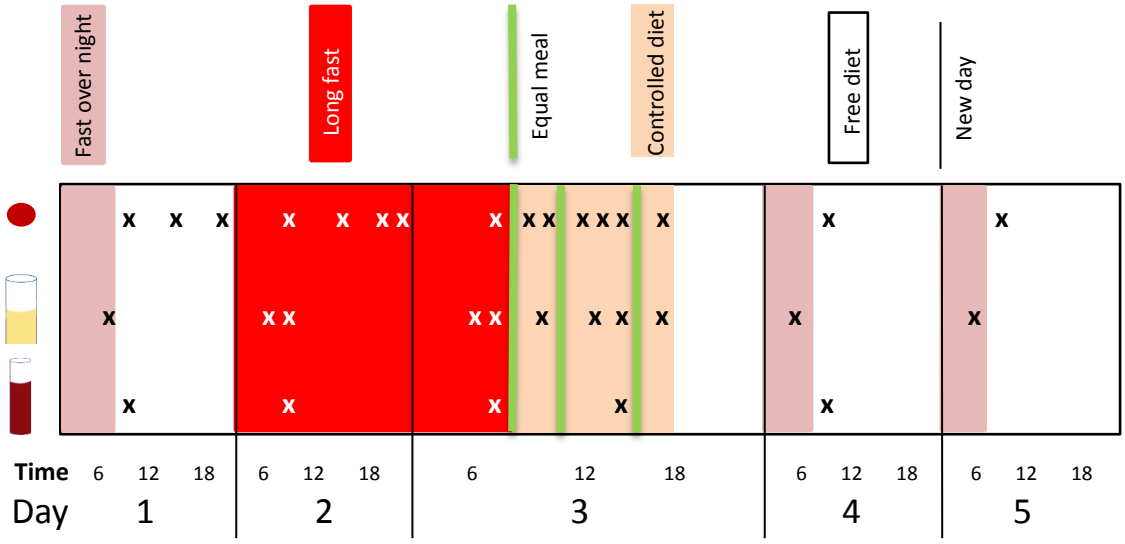


Figure 46: Sampling collection program during the fasting for 36 hours project.

Table 25: Sampling collection programme during the fasting for 36 hours, including comments.

Sample	Day	Time	Period	Comment
01	1	10.30	Free diet	After breakfast
02	1	15.00	Free diet	After ½ day of free diet
03	1	22.00	Free diet	After one day of free diet
04	2	09.00	Long fasting	No food intake, 11 h
05	2	15.00	Long fasting	No food intake, 17 h
06	2	18.00	Long fasting	No food intake, 20 h
07	2	22.00	Long fasting	No food intake, 24 h
08	3	08.00	Long fasting	No food intake, 36 h
09	3	09.45	Controlled diet	After meal number 1
10	3	10.30	Controlled diet	45 min after meal number 1
11	3	11.30	Controlled diet	Before meal number 2
12	3	13.30	Controlled diet	1.5 h after meal number 2
13	3	15.00	Controlled diet	Before meal number 3
14	3	18.30	Controlled diet	3.5 h after meal number 3
15	4	09.00	Free diet	After breakfast
16	5	09.00	Free diet	After breakfast

Individual F did not follow the programme completely. This individual did not eat the controlled diet like the others, thus, sample 9-12 were not collected. Samples 4, 15 and 16 for this individual were taken before breakfast (4 hours), unlike the others.

6.6.3 Control of the instrumental drift

Samples taken during free diet and fasting from all individuals were analyzed by the method in one analysis run with ISs in the extraction solution. Another analysis run analyzed all 16 DBS samples with control samples. The peak area % RSD of IS compounds in samples from all six individuals are shown in **Table 26.** All IS compounds in all samples had peak area % RSD <10 %, except for methionine, tyrosine, C3 and C4. There was not a single sample or group where the amount of all IS compounds changed significantly, thus, the instrumental variation was considered to not affect the measurements, see **Figure 47**.

Table 26: Relative standard deviation in % (% RSD) in peak area of internal standard compounds in samples 1-8 from all healthy volunteers in the fasting project (N=6).

Sample →	1	2	3	4	5	6	7	8
IS compound	**% RSD PA (a.u.) (N=6)**							
Ornithine	5 %	2 %	6 %	6 %	3 %	3 %	3 %	6 %
Arginine	9 %	3 %	4 %	6 %	9 %	8 %	3 %	10 %
Glycine	1 %	2 %	4 %	3 %	2 %	4 %	4 %	3 %
Alanine	2 %	4 %	2 %	5 %	4 %	4 %	3 %	3 %
Citrulline	3 %	2 %	4 %	2 %	4 %	5 %	5 %	4 %
Valine	1 %	2 %	2 %	2 %	5 %	4 %	3 %	4 %
Methionine	8 %	7 %	18 %	9 %	32 %	28 %	10 %	9 %
Leucine	6 %	2 %	4 %	3 %	2 %	4 %	4 %	3 %
Tyrosine	25 %	17 %	27 %	24 %	34 %	33 %	41 %	24 %
Phenylalanine	10 %	3 %	14 %	3 %	6 %	7 %	5 %	3 %
C0	3 %	2 %	2 %	2 %	4 %	5 %	3 %	4 %
C2	2 %	3 %	5 %	5 %	3 %	5 %	5 %	5 %
C3	27 %	17 %	20 %	30 %	16 %	10 %	26 %	18 %
C4	20 %	11 %	35 %	15 %	25 %	8 %	3 %	13 %
C5	2 %	2 %	2 %	3 %	4 %	5 %	4 %	4 %
C8	2 %	1 %	4 %	4 %	5 %	6 %	9 %	4 %
C12	4 %	4 %	6 %	3 %	3 %	8 %	11 %	8 %
C14	3 %	6 %	7 %	4 %	9 %	4 %	6 %	9 %
C16	2 %	5 %	4 %	4 %	6 %	2 %	4 %	9 %
C18	2 %	4 %	5 %	3 %	5 %	9 %	8 %	3 %

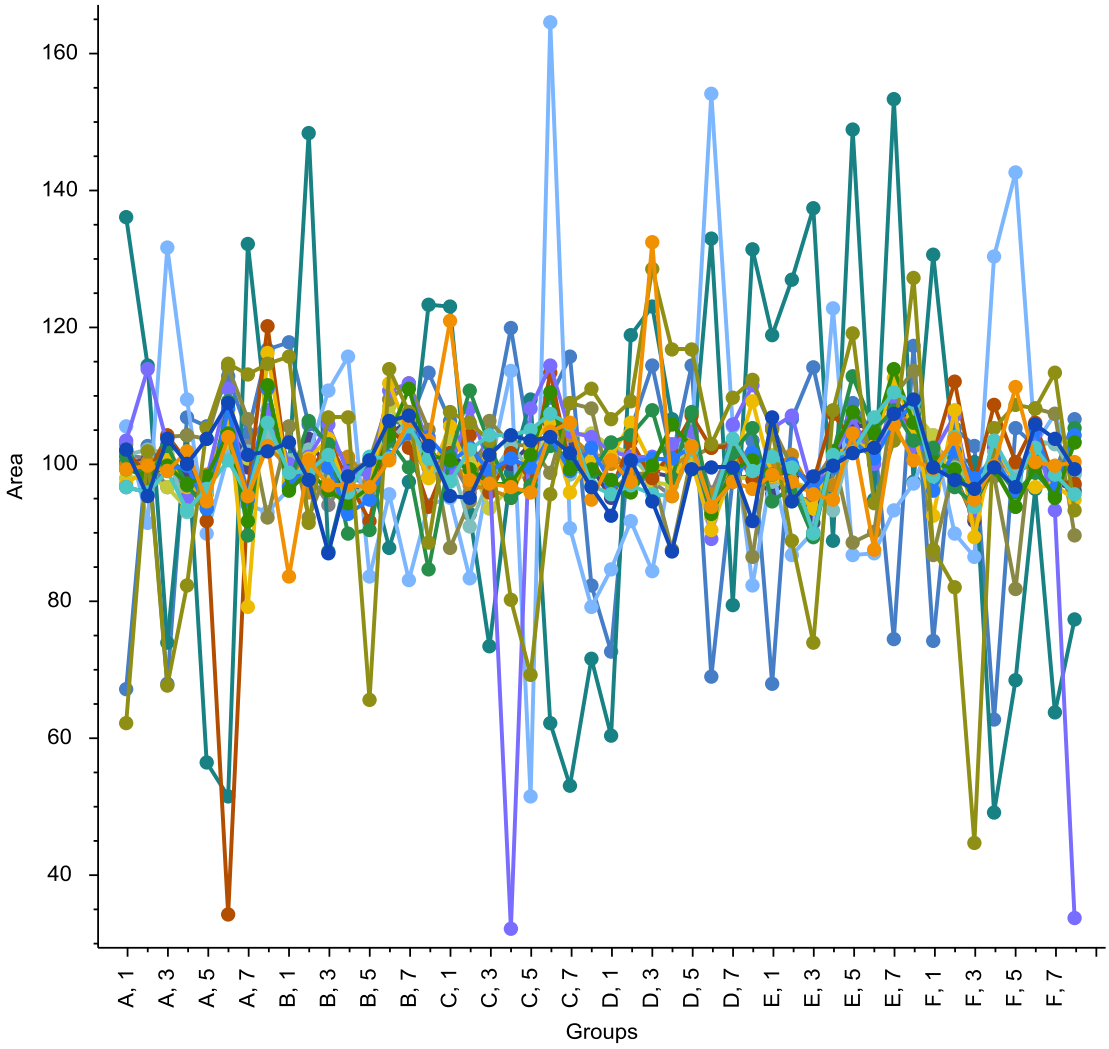


Figure 47: Internal standard peak area in all samples (1-8), from all individuals (A-F).

A DBS sample was used to control the instrument drifts during the sample analyses. A sample was injected every day (about 24 hours in between), from the same vial, to control signal variation of the MS and HPLC. **Table 27** shows the % RSD of the retention time and peak area of some final compounds. The software was not able to align the retention times of samples 1, 2 and 5, thus, these samples were not included when the data was processed.

Table 27: Relative standard deviation (% RSD) of retention time and peak area of final compounds in a control sample (DBS) injected three times every day during an analysis run for 11 days.

Compound	Average Pa (a.u.) N=30	% RSD PA (a.u.) N=30	Average Rt (min) N=30	% RSD Rt (min) N=30
Ornithine	1.81E+07	4 %	1.62	0.3 %
Arginine	6.88E+06	3 %	1.75	0.4 %
Citrulline	1.75E+07	3 %	2.19	0.2 %
Valine	1.71E+08	9 %	2.50	0.1 %
Methionine	8.17E+06	7 %	3.05	0.2 %
Leucine	1.37E+08	6 %	3.33	0.3 %
Tyrosine	3.34E+07	5 %	4.10	0.3 %
Phenylalanine	6.73E+07	3 %	6.73	0.2 %
C0	5.16E+07	2 %	2.11	0.3 %
C2	3.10E+07	4 %	2.61	0.2 %
C3	1.02E+06	8 %	4.97	0.4 %
C8	2.09E+05	8 %	12.09	0.1 %
C14	3.35E+05	10 %	13.14	0.1 %
C16	2.90E+06	4 %	13.69	0.1 %
C18	1.14E+06	4 %	14.42	0.1 %

The % RSD (N=30) of both the retention time and peak area of the final compounds were satisfactory. Retention times varied less than <1 % and peak area <10 % during 11 days of analysis. The injections were taken from the same vial, meaning that samples can be placed in the autosampler for at least 11 days with little effect on variation in peak area for the final compounds. The MeOH did not vaporize even though the concentration was high (80 %). 30 injections were performed and an article describing system suitability recommends a maximum of three injections of a QC sample from one vial [78].

ISs can determine the instrumental drift in a specific sample and were chosen to be used in further experiments.

6.6.4 Metabolites identified in the fasting samples

Figures 48-52 show the measured peak area of identified compounds that changed significantly in the fasting samples from the analysis with IS (1-8) and control samples (1-16). The figures are box and whisker plots where the ends of the whiskers represent the highest and lowest point. The whisker below the box contains about 25 % of the lowest data. The line in the box represents the median, and below this line are approximately 25% of the lowest data inside the box and above the line are approximately 25 % of the highest data in the box. The whisker above the bow contains approximately 25 % of the highest data.

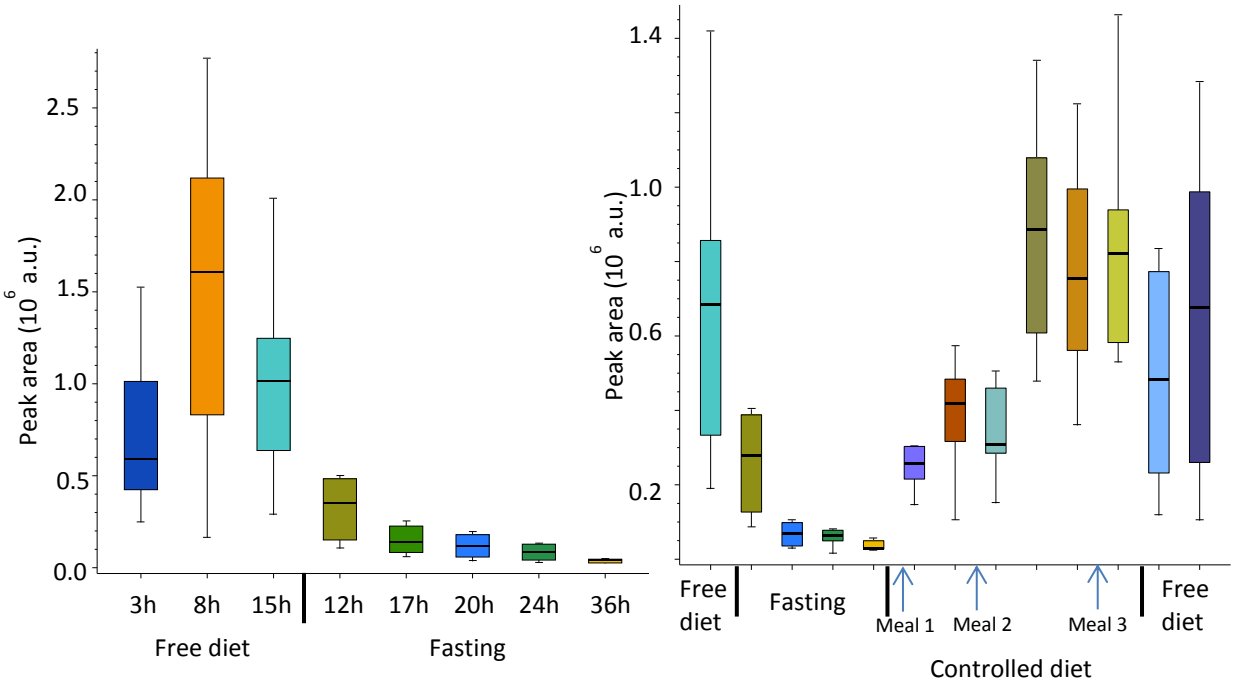


Figure 48: Caffeine (from coffee): Measured peak area of caffeine in the fasting samples from six healthy volunteers. Analyses with ISs of samples taken during free diet and fasting to the left. Analyses with control samples of DBS samples taken during free diet, fasting and controlled diet to the right.

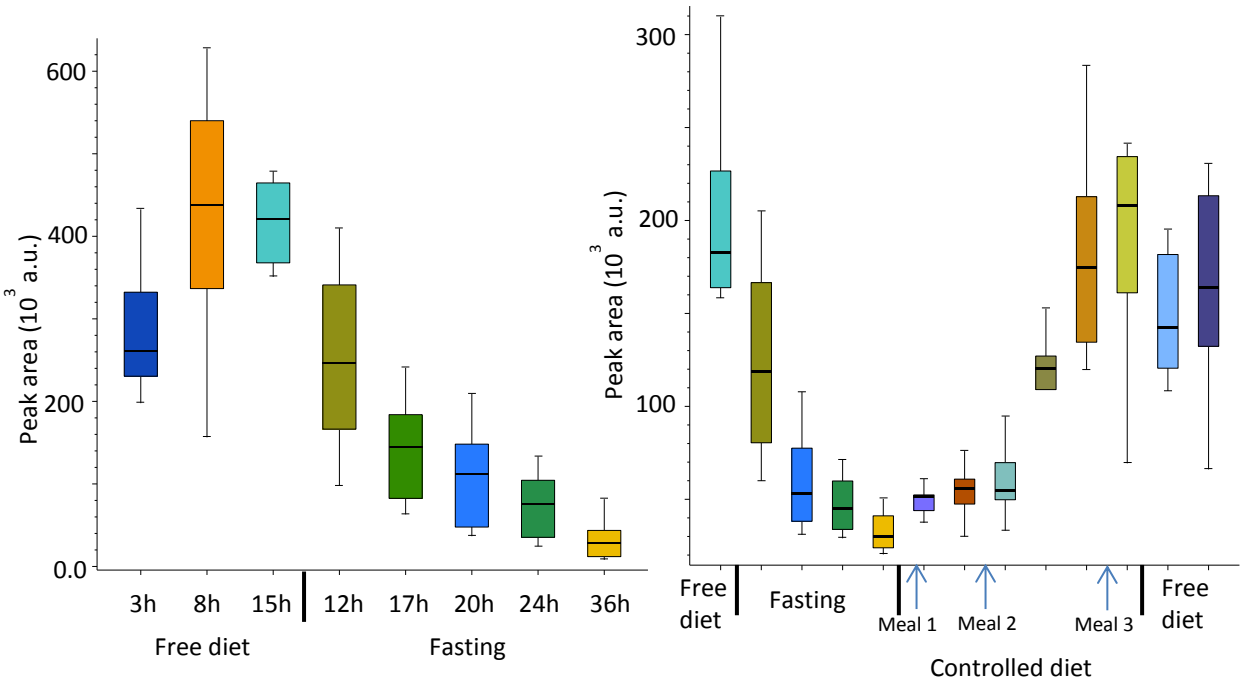


Figure 49: Paraxanthine (coffee metabolite): Measured peak area of paraxanthine in the fasting samples from six healthy volunteers. Analyses with control samples of DBS samples taken during free diet, fasting and controlled diet to the right.

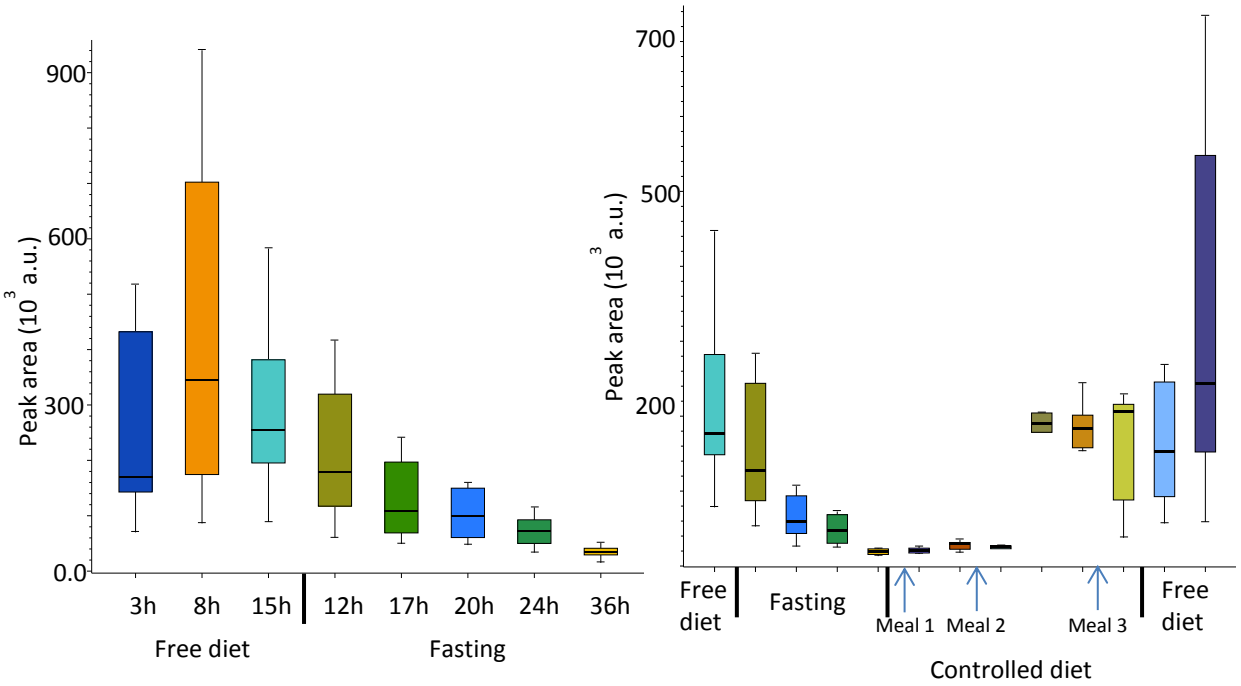


Figure 50: Theobromine (in chocolate): Measured peak area of theobromine in the fasting samples from six healthy volunteers. Analyses with control samples of DBS samples taken during free diet, fasting and controlled diet to the right.

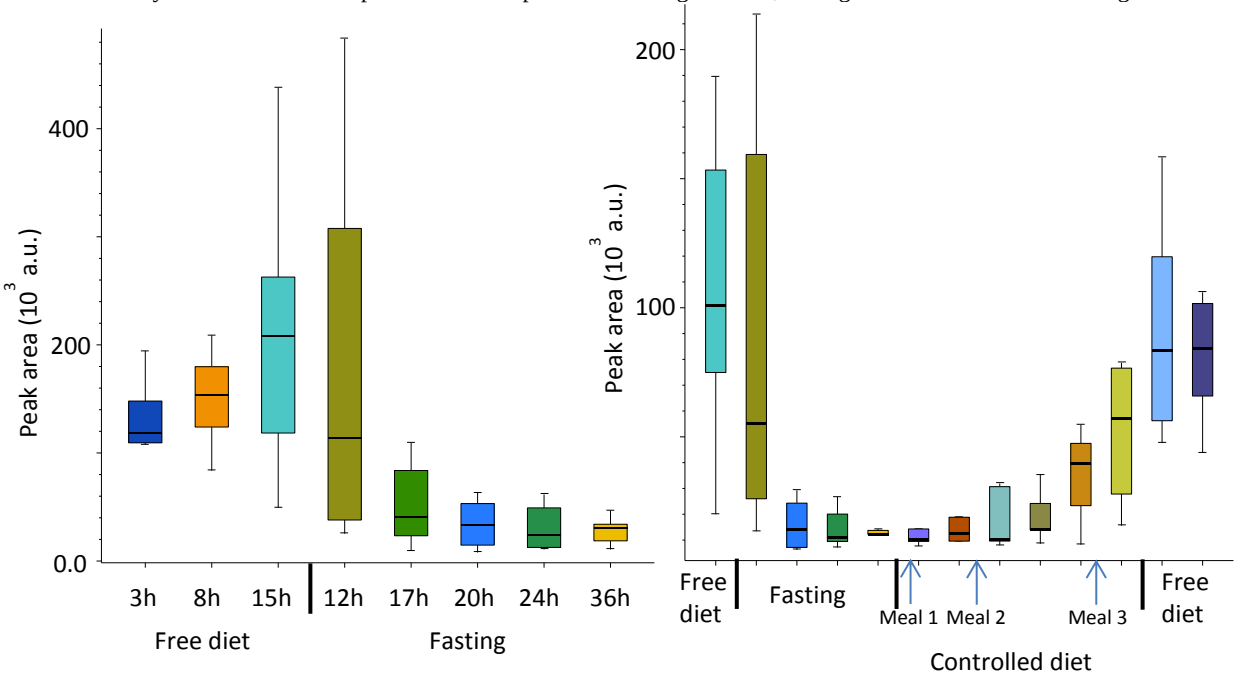


Figure 51: Hippuric acid (in orange juice and tea): Measured peak area of hippuric acid in the fasting samples from six healthy volunteers. Analyses with control samples of DBS samples taken during free diet, fasting and controlled diet to the right.

6.7 Poster presented at "Det 18. Norske seminar i væskekromatografi", Sandefjord 2018

Oslo University Hospital

UiO : University of Oslo

Evaluation of an untargeted metabolomics method of dried blood spots

E.M. Sandås[1,2], H.B. Skogvold [1], A. Østeby[1], H. Rootwelt[1], S.R. Wilson[2], K.B.P. Elgstøen[1]

[1] Department of Medical Biochemistry, Oslo University Hospital Rikshospitalet, P.O. Box 4950 Nydalen, 0424 Oslo

[2] Department of Chemistry, University of Oslo, P.O. Box 1033 Blindern, 0315 Oslo

Background

Inborn errors of metabolism (IEMs) are a group of genetic disorders causing dysfunction of specific proteins, most often enzymes, leading to abnormal levels of metabolites in biochemical pathways. Consequently, an IEM results in deviation from a normal metabolic profile. Targeted analyses of known diagnostic markers are used in laboratory diagnostics. Many IEMs are treatable if found at an early stage, putting great demand on the methods used for their diagnosis. To increase the understanding of the wider biochemical consequences of IEMs and to find better markers for diagnosis and monitoring of treatment, untargeted metabolomics (the analysis of all metabolites in a biological sample) is a powerful tool. Orbitrap mass spectrometry coupled to liquid chromatography is well suited for metabolomics, due to the high resolution and sensitivity of this technique. Earlier work by Camilla Elene Arnesen and Hanne Bendiksen Skogvold involved optimization of liquid chromatographic parameters to achieve separation of as many compounds as possible and mass spectrometric parameters to achieve better sensitivity and coverage of the metabolome in dried blood spots, respectively.

Purpose

The main purpose of this master project is to evaluate and further optimize the LC-Orbitrap MS method for dried blood spots developed by Arnesen and Skogvold by determination of instrument repeatability and assay reproducibility, and investigating the effect of punch location in dried blood spots.

Materials and methods

The liquid chromatography – mass spectrometry used was Dionex Ultimate 3000 UHPLC system coupled to a Q Exactive Orbitrap mass spectrometer from Thermo Fisher Scientific. Method parameters used in all experiments are shown in **Table 1**. Milli-Q water was used in all experiments.

Table 1: Settings used in all experiments

Parameter	Setting
Analytical column	Pursuit XRs C18-Diphenyl, 250 x 2.0 mm, 3 µm (Agilent)
Mobile phase A	Water + 0.1 % formic acid
Mobile phase B	Methanol + 0.1% formic acid
Flow gradient	**Table 2** and **Figure 1**
Injection volume	2 µL
Column temperature	30 °C
Electrospray voltage	3.5 kV
Electrospray needle position	C
Resolution	70000 a.u.
AGC target value	1e6 ion counts
Capillary temperature	250 °C
Aux gas heater temperature	300 °C
Scan ranges (*m/z*)	50-750, 750-1700
Scan type	Full MS
Flow rate	300 µL/min
Analysis time	27.5 minutes
Re-equilibration time	10 minutes

Table 2: LC Flow gradient profile. Mobile phase: A: Water + 0.1 % formic acid. B: Methanol + 0.1% formic acid

Time (min)	% B
0	2
6	10
8.5	75
25	100
27.5	100

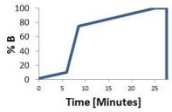


Figure 1: LC Flow gradient profile. Mobile phase: A: Water + 0.1 % formic acid. B: Methanol + 0.1% formic acid

Sample preparation of dried blood spots

To one punch (3.2 mm diameter, containing approximately 3 µL whole blood),100 µL of 80 % aqueous methanol + 0.1 % formic acid was added and extraction performed by mixing (700 rpm) for 45 minutes at 45 °C. After evaporation to dryness (N_2 40 °C), 100 µL 2 % aqueous methanol with 0.1 % formic acid was added. The sample solution was mixed and 2 µL injected. The samples were analyzed in positive and negative mode. However, only results from positive ionization is presented due to temporary technical problems with negative ionization. Blood from healthy controls were used in all experiments.

Endogenous compounds

Endogenous compounds covering a wide polarity range was chosen to represent the blood metabolome; creatinine, creatine, uric acid, glucose (only detected in negative mode) and acylcarnitines C_2, C_{12} and C_{16}.

Instrument repeatability and assay reproducibility

Instrument repeatability was determined by peak area variation in 20 injections of one prepared dried blood spot sample. Assay reproducibility was assessed by analyzing 20 samples with random punch location.

Punch location

To observe the effect of punch location, 10 dried blood spot samples were prepared from each location (A, B, C and D) shown in **Figure 2**. Peak area of creatinine, creatine, acylcarnitine C_2 and uric acid was used.

Figure 2: Punch locations of dried blood spots

Results and discussion

Instrument repeatability and assay reproducibility

Acylcarnitines C_{12} and C_{16} were not detected, probably due to very low endogenous levels in healthy controls and/or solubility issues. As a substitute, two peaks of unknown identities with high mass (*m/z* 802.79852) and long retention time (11.05 min), respectively, was included. The extracted ion chromatograms are shown in **Figure 3** and the coefficient of variation (CV) in % of peak area for each peak are shown in **Table 3**.

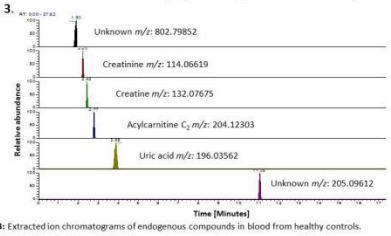


Figure 3: Extracted ion chromatograms of endogenous compounds in blood from healthy controls.

Table 3: Coefficient of variation (CV) in % of peak area for endogenous compounds (N=20).

	Creatinine	Creatine	Acylcarnitine C_2	Uric acid	*m/z*: 205.09612	*m/z*: 802.79852
Instrument repeatability CV (%)	2 %	2 %	1 %	2 %	1 %	2 %
Assay reproducibility CV (%)	9 %	10 %	11 %	6 %	12 %	9 %

Punch location

As seen in **Figure 4**, concentration varies with punch location for all compounds. In general, higher concentration was observed in punches taken from the perimeter of the spot. This may be due to the inhomogeneous nature of blood and slower diffusion of red blood cells from the center where the blood is applied compared to plasma. In addition, larger variation was observed in punches taken from the perimeter than from the center of the spot.

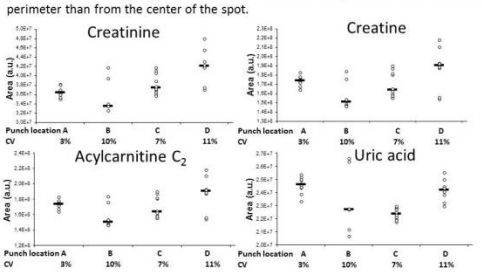


Figure 4: Peak area and CV (%) of selected endogenous compounds in samples from punch locations A, B, C and D , see **Figure 2** (N=10).

Conclusion and future work

The instrument repeatability and assay reproducibility are acceptable, with a coefficient of variation in peak area of 2% and 12% or lower, respectively, for the chosen compounds. Further work for improvement of assay reproducibility is ongoing and involves use of isotopically labelled internal standards in the extraction solvent. Concentration was found to vary with punch location, and larger variation was observed in punches taken from the perimeter than from the center of the spot. Thus, when comparing metabolomes from dried blood spots, knowledge of punch location is essential and the center area punch should be used. Some of the hydrophobic compounds were not detected. Ongoing work to improve the coverage of the hydrophobic part of the metabolome involves adjustment of the sample preparation; skipping the whole evaporation step is currently being tested and the results are promising. The long-term goal is to show the suitability of the method for detection of changes in the metabolome in health and disease. This will be achieved by studying changes in the metabolome of healthy controls with exposure to different environmental stimuli (e-g- diet and/or emotional changes), and samples from patients with known diseases.

Contact info: moerel@ous-hf.no