Marie Austdal

Biomarkers for prediction and characterization of preeclampsia using magnetic resonance metabolomics

Thesis for the degree of Philosophiae Doctor

Trondheim, August 2015

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

NTNU
Norwegian University of
Science and Technology

Biomarkører for prediksjon og karakterisering av svangerskapsforgiftning ved bruk av MR metabolomics

Hypertensive svangerskapskomplikasjoner er en ledende årsak til maternell sykelighet og død. Svangerskapsforgiftning, eller preeklampsi, oppstår i omtrent 3% av svangerskap i Norge og kan medfølge for tidlig fødsel, lav fødselsvekt, alvorlige komplikasjoner hos mor, og høyere risiko for hjerte- og karsykdom senere i livet for både mor og barn. Svangerskapsforgiftning kjennetegnes av nyoppstått høyt blodtrykk og protein i urinen etter 20. svangerskapsuke.

Til tross for intensiv forskning er årsakssammenhengene for utviklingen av svangerskapsforgiftning ikke fullt ut forstått. En mangelfull utvikling av blodårene som fører til morkaken (placenta) spiller en vesentlig rolle i mange, men ikke alle tilfeller. Når fosteret vokser til utover i svangerskapet, blir tilførselen av blod og næringsstoffer til morkaken etter hvert utilstrekkelig. Den stressede morkaken sender ut faktorer til kvinnens blodsirkulasjon, og en inflammasjonsreaksjon oppstår som forårsaker de kliniske symptomer på svangerskapsforgiftning.

Arbeidet i denne avhandlingen omhandler bruken av magnetisk resonans (MR)-spektroskopi og fagfeltet metabolomics for å undersøke den metabolske profilen under utviklingen av svangerskapsforgiftning, både tidlig og sent i svangerskapet, før og etter de kliniske tegnene på sykdom har oppstått. Metabolomics er en systematisk analyse av metabolitter, små molekylære forbindelser som er bestanddeler eller produkter av metabolismen. Sammensetningen av metabolitter forteller om de biologiske prosessene som foregår i kroppsvæsker eller vev. I denne avhandlingen ble blodprøver, urin og morkakebiopsier fra kvinner med svangerskapsforgiftning sammenlignet med friske kvinner, for å undersøke hvilke endringer i metabolismen som finner sted ved sykdom.

Resultatene fra disse analysene viste at metabolittsammensetningen er vesentlig endret i serum og urin fra kvinner med aktiv svangerskapsforgiftning sammenlignet med kvinner med normale svangerskap, og flere av de samme metabolittene var endret også tidlig i svangerskapet hos kvinner som senere utviklet svangerskapsforgiftning. Blant endringene var lavere glysin og p-cresol sulfat, og hippurat i urinen hos kvinner med svangerskapsforgiftning, samt en sammensetning av lipoproteiner i blodet som også ses hos kvinner med hjerte-karsykdom: høyere nivåer av "very-low

density lipoprotein" (VLDL) og lavere nivåer av "high-density lipoprotein" (HDL), også kalt en aterogen lipidprofil (artikkel 1). Når urin og blod fra kvinner i første trimester av svangerskapet ble analysert, viste det seg at kvinner som senere utviklet svangerskapsforgiftning hadde den samme lipidprofilen tidlig i svangerskapet. I tillegg hadde kvinnene lavere utskillelse av hippurat i urinen og høyere utskillelse av kreatinin. Denne sammensetningen kunne brukes til å predikere svangerskapsforgiftning i disse damene like godt som en ultralydmåling av blodstrømningen til morkaken, når målingene ble kombinert med kvinnens kroppsmasseindeks og alder (artikkel 2). Morkakebiopsier analysert med MR-metabolomics viste tegn til endring i flere metabolske funksjoner som tidligere er vist å kunne være affisert i svangerskapsforgiftning, blant annet i metabolismen av taurin, glutamin og fosfolipider (artikkel 3). I tillegg ble morkakeforandringene relatert til endringer i triglyserid, urat og sykdomsmarkøren sFlt-1 målt i blodet til de samme kvinnene. Denne måten å analysere vevbiopsier er ny i forskning på morkaken i svangerskapsforgiftning, og kan bidra til å finne nye subgrupperinger av morkakesykdom i fremtiden.

Samlet sett har arbeidet i denne avhandlingen bidratt til å kartlegge de metabolske endringene som skjer under utviklingen av svangerskapsforgiftning, og som kan være med på å forutsi hvilke damer som senere i svangerskapet utvikler sykdommen. I tillegg danner avhandlingen et solid grunnlag for videre studier på metabolske forandringer i svangerskapsforgiftning.

Kandidat: Marie Austdal

Institutt: Institutt for sirkulasjon og bildediagnostikk

Veiledere: Professor Tone F. Bathen, Førsteamanuensis Ann-Charlotte Iversen og Professor Rigmor Austgulen

Ovennevnte avhandling er funnet verdig til å forsvares offentlig for graden Philosophiae Doctor i molekylær medisin.

Disputas finner sted i Auditoriet, Medisinsk-Teknisk Forskningssenter, mandag 21. september 2015 kl. 12:15.

Acknowledgement

want to thank my dear Knut with all my heart for being my rock during these years. Thank you for being so patient and supportive.

Marie Austdal
Trondheim, May 2015

Summary

Hypertensive diseases of pregnancy are major contributors to maternal and fetal morbidity worldwide. Preeclampsia, characterized by gestational hypertension and proteinuria occurring after mid gestation, occurs in approximately 3% of pregnancies and may lead to preterm birth, low birth weight, severe maternal organ involvement, and increased risk of cardiovascular diseases in later life for both mother and child.

Despite a wealth of research and investigations, the causative agents for the development of preeclampsia are as yet unknown. A staged model is proposed, wherein the placenta is insufficiently developed at the start of pregnancy. Later in pregnancy the stressed placenta releases soluble factors into the maternal circulation, eventually causing the clinical symptoms of the disease. Preeclampsia varies in severity, time of onset, and degree of placental and fetal involvement. There is no current treatment or prevention, and no sensitive way of predicting the disease. In addition, todays diagnostic criteria are based on blood pressure and proteinuria cutoffs, which poorly reflect the root causes of the disease nor correlate with maternal and fetal outcomes.

Nuclear magnetic resonance (NMR) metabolomics analyses assess the downstream products of gene and protein expressions, called metabolites, and have been shown to provide predictive and prognostic information for several types of disease. Proton high resolution magnetic resonance spectroscopy is a nondestructive and high throughput technique for analyzing biofluids and tissues. In this thesis, multivariate data analyses and metabolite quantification of ^{1}H NMR spectra were performed to investigate the potential of metabolomics in prediction and characterization of preeclampsia.

In the first paper, the urine and serum metabolic profiles of women with active preeclampsia were compared to those of normotensive pregnant women and non-pregnant women. The differences in urine excretion showed an alteration in glycine flux for women with preeclampsia, attributed to increased oxidative stress or disturbance in the methionine-homocysteine metabolism. There was also retention of uremic solutes with possible toxic effects. Serum measurements showed a significant increase in lipid levels for women with preeclampsia, and a shift towards an atherogenic lipid profile with increased triglyceride-containing VLDL and decreased HDL.

In the second paper, urine and serum metabolic profiles of women at the end of the first trimester (gestational weeks 11-13) were analyzed for potential discriminative and predictive metabolites for preeclampsia and gestational hypertension. A metabolite profile including urinary hippurate, proline betaine, creatinine, 4-deoxythreonic acid and dimethylamine was found predictive of preeclampsia. When combined with maternal clinical characteristics, the urinary hippurate/creatinine ratio was found to improve prediction sensitivity compared to using the uterine artery pulsatility index. Serum metabolites were not found to be sensitive in prediction of preeclampsia, but revealed a significantly different serum lipid profile in women who later developed preeclampsia.

In the third paper, the placental metabolic profile was evaluated using high resolution magic angle spinning spectroscopy, a nondestructive method of analyzing intact tissue biopsies which is novel in placenta research. Several metabolic pathways were found to be affected in the placenta of preeclamptic pregnancies compared to normotensive pregnancies. These were related to taurine metabolism, glutamate and glutamine metabolism, phospholipid metabolism and the homocysteine-methionine metabolism. Metabolite profiles in placenta correlated to maternal serum lipids and markers of disease severity.

Collectively this thesis presents a thorough investigation into the metabolic state of women with active or developing preeclampsia. The results highlight the metabolic alterations that appear before and during the clinical symptom stages of preeclampsia, and the differences and similarities of gestational hypertension and preeclampsia. Both the placental and maternal component of the disease were examined and found significantly affected. New biomarkers of the disease were discovered, and suggestions for future research are posited. This thesis serves as a contribution to understanding the maternal and placental components of preeclampsia.

Symbols & Abbreviations

Symbol	Description	Page
γ	Magnetogyric ratio	15
μ	Magnetic moment of a precessing nucleus	15
B_0	External magnetic field in an NMR spectrometer	15
BMI	Body mass index [kg m^{-2}]	36
CARS	Competitive adaptive reweighted sampling	28
CPMG	Carr Purcell Meiboom Gill pulse sequence	18
CV	Cross validation	29
CVD	Cardiovascular disease	2
CS	Caesaerian section	38
ER	Endoplasmic reticulum	8
FID	Free induction decay	16
FPR	False positive rate	30
FGR	Fetal growth restriction	38
GFR	Glomerular filtration rate	10
GH	Gestational hypertension	39
HDL	High density lipoprotein	9
HMDB	Human metabolome database	45
HR-MAS	High resolution magic angle spinning	19
HSQC	Heteronuclear single quantum coherence spectroscopy	19
I	Nuclear spin number	15
J	Coupling constant between peaks in a multiplet	17
JRES	J-resolved spectroscopy	19
LV	Latent variable	23
M_0	Macroscopic magnetization	16
MAP	Mean arterial pressure	36
MS	Mass spectrometry	11
MSEA	Metabolite set enrichment analysis	48
(N)MR	(Nuclear) Magnetic resonance	15
NOESY	Nuclear Overhauser effect spectroscopy	18
P	Angular momentum for a precessing nucleus	15
PAPP-A	Pregnancy associated plasma protein A	68
PC	Principal component	23

List of Papers

Paper I

Metabolomic Biomarkers in Serum and Urine in Women with Preeclampsia. <u>Austdal M</u>, Skråstad RB, Gundersen AS, Austgulen R, Iversen A-C, Bathen TF (2014)
PLoS ONE 9(3): e91923. doi: 10.1371/journal.pone.0091923

Paper II

First Trimester Urine and Serum Metabolomics for Prediction of Preeclampsia and Gestational Hypertension: A Prospective Screening Study. <u>Austdal M</u>, Tangerås LH, Skråstad RB, Salvesen KÅ, Austgulen R, Iversen A-C, Bathen TF
Submitted to Journal of Proteome Research 2015.

Paper III

Metabolic profiles of placenta in preeclampsia using HR-MAS MRS metabolomics.
<u>Austdal M</u>, Thomsen LCV, Tangerås LH, Skei B, Mathew S, Bjørge L, Austgulen R, Bathen TF, Iversen A-C
Submitted to Placenta 2015

Contents

List of Figures

List of Tables

1   Introduction

Preeclampsia is a syndrome of pregnancy which affects families across the world. Without proper care, the syndrome can be deadly for both mother and baby. The origin of the syndrome has long been a mystery and the knowledge about its development limited. This knowledge has rapidly increased in recent years, but there is still much left to learn. Modern analysis of gene, protein and metabolite expression produce increasing amounts of data, which enable holistic profiling of complex diseases. This leaves large amounts of information to be processed and analyzed into biologically meaningful results. The 'Omics' branches of research - genomics, proteomics, metabolomics - are often termed 'Hypothesis generating', where differences in expression can form the basis of new hypotheses. Preeclampsia was recognized as early as the time of Hippocrates, as an unfavorable state of headache and convulsions during pregnancy.[1] Even today the root causes of preeclampsia remain unclear and the syndrome can neither be predicted or prevented effectively, nor cured except by terminating the pregnancy. This thesis serves to look further into the development and pathogenesis of preeclampsia using state of the art analytical methods and pattern recognition models.

1.1   Preeclampsia

1.1.1   Epidemiology and risk factors

Primary indicators of preeclampsia are newly onset high blood pressure (hypertension), and an excess of proteins in the urine (proteinuria) occurring in the second half of pregnancy.[2] Preeclampsia affects about 2-7% of women worldwide, in Norway 3.0%.[3,4] Hypertensive diseases of pregnancy cause 16.1% of maternal deaths in the developed world, and are a major contributor to maternal deaths in Norway.[5,6] Currently, preeclampsia can not effectively be prevented or predicted, and the only effective cure is delivery of the baby with the placenta.[7] Preeclampsia is the main cause of preterm birth in the developed world.[5] Several risk factors mediate the incidence of preeclampsia (See Table 1.1). Left untreated, preeclampsia can progress to eclampsia, characterized by seizures, or lethal complications such as cerebrovascular events, liver rupture, or acute renal failure.[7,8] For the fetus, complications associated with preeclampsia include fetal growth restriction (FGR) and premature delivery.[9]

Table 1.1: Risk factors for preeclampsia[3, 4, 8, 10]

Nulliparity
Multifetal gestation
Obesity
Family history of preeclampsia (mother or sister)
Preeclampsia in a previous pregnancy
Abnormal uterine Doppler studies at 18 and 24 weeks gestation
Pregestational diabetes mellitus
Presence of thrombophilias
Chronic hypertension or renal disease
Hyperglycemia
Nonsmoking
Extreme maternal age (≤ 20 or ≥ 35 years)
'Dangerous father' or limited sperm exposure

The syndrome recurs in families, suggesting genetic components present in maternal and/or fetal (maternal or paternal) genes.[11, 12] Familial preeclampsia is often more severe.[13] Preeclampsia shares many risk factors with cardiovascular disease, and indeed has been described as a 'stress test' for the cardiovascular system.[14] It is still debated whether preeclampsia causes irreversible changes to the maternal metabolic and vascular system, or if both syndromes are an expression of the same metabolic abnormalities or risk factors.[14, 15] Women with a history of preeclampsia have a two- to eightfold increased risk for cardiovascular diseases later in life.[14, 16–18] Additionally, FGR is a major risk factor for cardiovascular disease (CVD).[2] FGR is the failure of a fetus to grow to its genetically determined potential.[19]

Increasing evidence shows that preeclampsia is not a single disease.[20, 21] Several subdivisions have been suggested. One of them is the designation of early and late onset preeclampsia, with a gestational age cutoff at 34 weeks.[21] Early onset preeclampsia is associated with worse maternal and fetal outcomes such as FGR, maternal hepatic, renal, and pulmonary injury, and increased risk of maternal and fetal death.[2, 17, 21, 22] Preeclampsia occurring closer to term is more often associated with excessive fetal demands, such as in multiple gestation and fetal macrosomia. It often has a milder presentation and is the most common variant of the syndrome.[23] Preeclampsia can also be classified by severity, where any of several features in-

cluding severe hypertension or proteinuria, evidence of renal, hepatic or pulmonary involvement, visual disturbances or epigastric pain will classify the disease as severe.[2] In addition, presence or absence of fetal involvement evident by FGR is a possible disease subgrouping.[24] Subphenotypes of preeclampsia are thus based on end stage maternal and fetal symptoms.

Hypertensive diseases of pregnancy include preeclampsia, gestational hypertension, chronic hypertension in pregnancy, and preeclampsia superimposed on chronic hypertension. These are a major cause of maternal and fetal morbidity and mortality, sharing risk factors with cardiovascular disease, a global major cause of death. Preeclampsia varies in maternal and fetal outcome, time of onset, and probably by yet undefined underlying causes, making the syndrome a challenging subject of study.

1.1.2 Pregnancy and the placenta

The placenta is essential to pregnancy. It is the interface between mother and fetus and represents an immunological barrier as well as the site of life sustaining activities for the fetus.[23,25] The placenta develops in part from the fertilized egg and consists of the fetal component (*chorion*) and the maternal component (*decidua basalis*), the thickened inner lining of the uterus which is discarded during delivery.[25,26]

The development of the human placenta is a delicate and tightly regulated process involving an extensive remodeling of the spiral arteries in the uterine wall in order to supply the developing placenta with sufficient maternal blood.[23,28] The main cell type of the placenta is the fetal trophoblast, an embryonic stem cell which forms the outer layer of the blastocyst (fertilized ovum) and develops into a large part of the placenta including the chorionic villi. Villi are tree-like structures through which metabolic exchange occurs between maternal and fetal blood(Figure 1.1).[29]

Trophoblasts enter two main differentiation pathways. The villous pathway involves fusion into the multinucleate syncytiotrophoblast which covers the villi and comes into direct contact with the maternal blood, and anchor into maternal decidua.[30] The extravillous trophoblast pathway develops an invasive cell type which burrows into the maternal decidual tissue and uterine wall, and replaces maternal endothelial cells in the narrow spiral arteries.[23,28] Endovascular trophoblasts then replace the maternal endothelium and alter the vessel characteristics. The artery
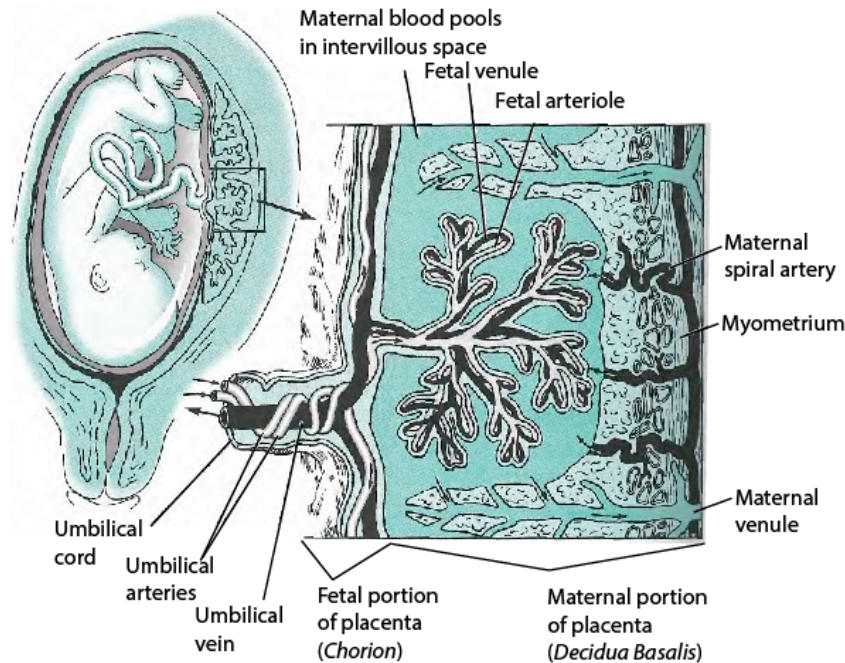
Figure 1.1: The human placenta. The placenta is composed of fetal tissue (chorion) and maternal tissue (decidua basalis). Fetal trophoblasts develop into the chorionic plate and villi. The villi (tree-like structures) are composed of fetal venules and arterioles, and are perfused with maternal blood in the intervillous space. Maternal blood is supplied from maternal arterioles (spiral arteries), exchanges nutrients and waste with the fetal circulation through the villi, and exits through maternal venules. Reproduced with permission from McGraw-Hill.[27]

becomes wider and less resistant, enabling increased blood flow to the intervillous space and ensuring that the fetus is well supplied (Figure 1.2).[25]

From the time of implantation, the embryo is supplied with nutrients from endometrial glands in a low oxygen environment.[31,32] At approximately week 10-14 of gestation, blood flow into the placenta starts.[26,33] The placental trophoblast responds to the increased oxygen tension with reduced proliferation and differentiation to an invasive phenotype.[33]

Decidual tissue contains about 40% maternal immune cells including uterine natural killer (NK) cells, T cells, macrophages and dendritic cells.[29] Fetal trophoblasts are foreign to the maternal immune system and come into direct contact with ma-
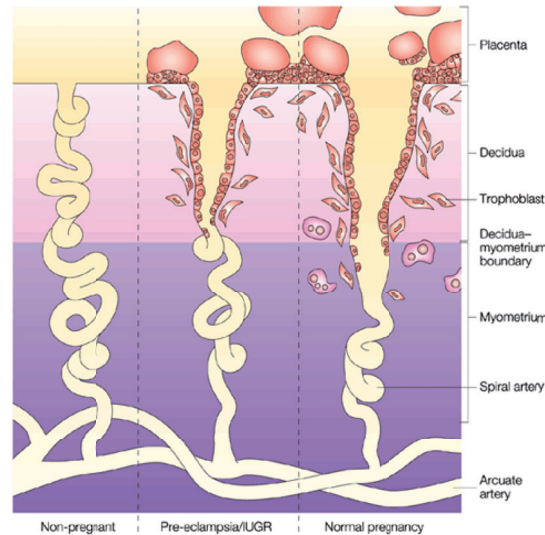
Figure 1.2: Uterine spiral artery remodeling is incomplete in placentas of FGR and preeclampsia, causing intermittent hypoxia. Reprinted by permission from Macmillan Publishers Ltd: © A. Moffet-King. *Nat. Rev. Immunol.* **2**, p 656-663, 2002.[30]

ternal blood and decidual tissue. Syncytiotrophoblast on the villous surface are in contact with maternal blood. These do not express the classical human leukocyte antigen (HLA) I and II molecules, and are protected against maternal immune response.[29] Endovascular trophoblasts lining the spiral arteries express no HLA II and an unusual array of HLA-I: HLA-C, -G and -E which protect against lysis by maternal immune cells in the decidua.[29,30] In addition, trophoblast cell surface signaling also stimulates maternal immune cells to produce cytokines, chemokines and angiogenic factors, which facilitates trophoblast invasion.[8,23] This interplay between maternal cells and fetal trophoblasts creates an immune tolerance between the two individuals involving a certain level of local inflammation in the decidua important for normal placentation.[34]

1.1.3 The staged model of preeclampsia

Human pregnancy is a challenge to the maternal immune and cardiovascular system. Normal pregnancy is characterized by a mild systemic inflammation which develops as the pregnancy progresses. In preeclampsia, an exaggerated inflammatory
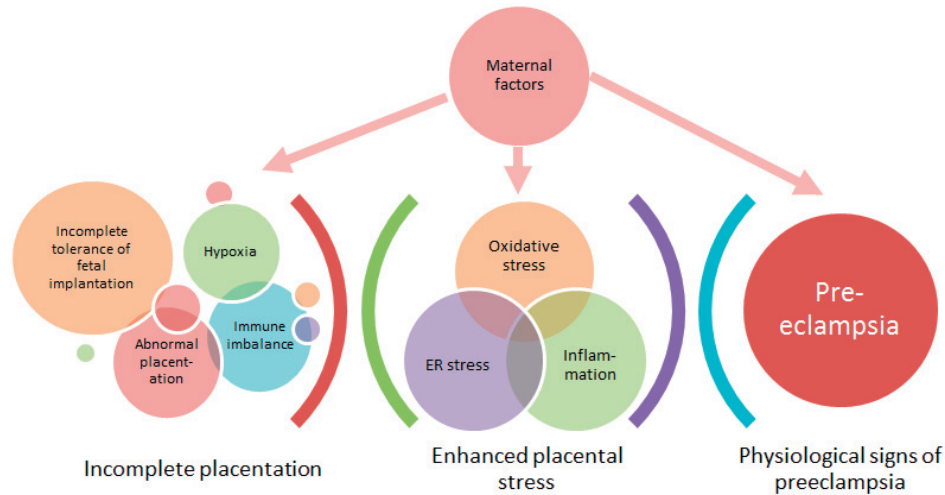
Figure 1.3: The stepwise model of preeclampsia as proposed by Redman.[7,37] Due to immune or angiogenic imbalance, hypoxia, or undetermined causes, the development of the placenta is abnormal at the start of pregnancy. This leads to a poorly vascularized placenta which exhibits signs of oxidative, endoplasmic reticulum and inflammatory stress, and sends stress signals to the maternal circulation. Together with maternal constitutional factors (genetic, environmental, metabolic) this leads to the overt condition of preeclampsia.

response occurs with resulting endothelial dysfunction causing the clinical symptoms.[8,35,36]

The origins of preeclampsia start in the early stages of pregnancy with the development of the placenta.[23,24,38] A staged model has been proposed to explain the disease development (as summarized in Figure 1.3). In the first stage, there is insufficient trophoblast invasion, and the maternal spiral arteries fail to remodel properly during the development of the placenta, leading to reduced placental perfusion and placental hypoxia (Figure 1.2).[39] Later in pregnancy, the stressed placenta releases signals into the maternal circulation, causing a widespread inflammatory response with endothelial dysfunction.[34,38] The clinical manifestation of preeclampsia is not evident until this later stage in pregnancy, when fetal demands begin to exceed the compromised placental supply.[23,38]

It is thought that defective invasion of fetal trophoblasts into the maternal spiral arteries is the main defect in preeclampsia.[30] The cause of the defective placentation may result from an inflammatory imbalance locally in the placenta,[2,34] from reduced
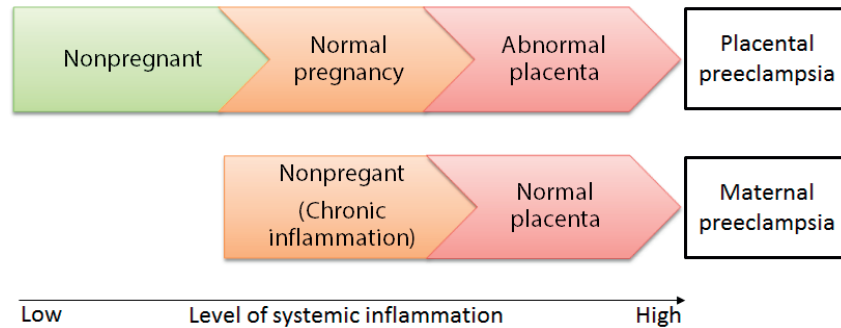
Figure 1.4: The increasing inflammatory response in pregnancy and preeclampsia. Normal pregnancy is associated with an increase in systemic inflammation, which is further increased in preeclamptic pregnancies with a stressed placenta. In a woman with chronic increase in systemic inflammation, even a normal placenta might stimulate a sufficient systemic response to cause preeclampsia. Both maternal constitution and placental ischemia may contribute.[34] Preeclampsia is not intrinsically different from normal pregnancy, but an extreme end of a continuous spectrum of responses common to pregnancy.[36] Figure adapted from Borzychowski *et al*, Semin Fetal Neonatal Med, 2006.[34]

differentiation to invasive trophoblastic phenotype due to hypoxia,[8,40] or from an ill-conditioned uterus.[41] Defective placental invasion can also cause other complications such as FGR, recurring miscarriage and preterm labor.[24,42,43] As a result, the spiral arteries are narrower and more elastic, and remain responsive to hormones and blood pressure signaling. This situation causes intermittent ischemia and reperfusion of the placenta, with oxidative stress caused by higher flow velocity of blood into the placenta[18,33,38]

An inadequately developed placenta is not enough to cause preeclampsia on its own, nor is it required for the syndrome to develop.[2] In addition, maternal factors such as underlying inflammation status, diseases such as clotting disorders, ischaemic heart disease, hypertension and renal disease predispose to preeclampsia.[34] These diseases may provide a higher baseline inflammatory state, and the mild increase of inflammation in a normal pregnancy may then be enough to cause the systemic response of preeclampsia. The increase in systemic inflammation during pregnancy and preeclampsia is illustrated in Figure 1.4. There are likely several routes linking the first stage (insufficient placentation) with the second stage (maternal systemic response).[38]

The compromised placenta sends signals into the maternal circulation, which the maternal immune system reacts to in preeclampsia. The signals are though to be either cellular signals like growth factors and cytokines, or products of placental oxidative stress, like placental debris from apoptotic cells, or anti-angiogenic factors. The soluble receptor for Vascular Endothelial Growth Factor (VEGF), also called soluble FMS-like tyrosine kinase receptor 1 (sFlt-1), is a candidate anti-angiogenic factor, as it is increased early in preeclamptic pregnancies with concomitant decrease in pro-angiogenic factors VEGF and Placental Growth factor (PlGF) in maternal serum.[38,44,45] Increased placental stress and apoptosis causes increased release of vesicular particles containing fetal DNA to the maternal circulation.[46–48] The released microparticles may be recognized as foreign by the maternal immune system, causing a systemic inflammatory response. Alternatively, the trophoblasts in preeclamptic pregnancies may themselves release danger signals such as pro-inflammatory cytokines in excess.[2,34]

Regardless of the pathway to preeclampsia, the clinical symptoms are similar. A systemic increase in inflammation causes widespread endothelial dysfunction and hypertension with renal injury leading to proteinuria. Preeclampsia can be seen as a spectrum of disorders ranging from a more maternal disease to a more placental disease. "Maternal preeclampsia" constitutes a normal placenta in a women who is predisposed in some way, and in "placental" preeclampsia there is an abnormal placenta in a normal woman. Most women will exist between these two extremes.[34]

1.1.4 Altered metabolism in the preeclamptic placenta

Several pathophysiological changes occur in the placenta during its development preceding the clinical symptoms of preeclampsia.[24] Intermittent ischemia-reperfusion causes oxidative and endoplasmic reticulum (ER) stress.[8,49] Oxidative stress is an imbalance between systematic generation of reactive oxygen species by immune or other cells, and the capability of antioxidant defenses to prevent oxidative damage.[34,50] Reactive oxygen species cause genomic damage, create inflammatory oxidized lipid species, and denature enzymes.[34,50] Oxidative stress is also a necessary component of normal placental development.[50] ER stress is caused by a demand for ER function that exceeds its capacity, leading to a backlog of unfolded or misfolded proteins.[51] ER and oxidative stress may cause apoptosis or necrosis in the placenta, or cause a chronic state of stress tolerance.[51] Oxidative stress in placenta induces release of proinflammatory cytokines, chemokines, and cellular debris.[8] Preeclampsia

is associated with reduced antioxidant capacity in the placenta.[8] Oxidative and ER stress are powerfully inflammatory, and the interaction between the two stresses and the inflammatory response in the placenta drive the development of preeclampsia.[37]

1.1.5 Altered maternal metabolism in preeclampsia

The widespread endothelial dysfunction and inflammatory response in preeclampsia induces changes in the maternal metabolism. Acute phase response, oxidative stress, hyperlipidaemia, and insulin resistance are all exaggerated in preeclampsia, though these also happen to some degree in normal pregnancy due to the mild inflammatory state.[34] The metabolic changes of preeclampsia are similar to those seen in patients with metabolic syndrome, obesity, diabetes, and chronic hypertension. Systemic inflammation is common to all these diseases.[34,52]

Normal pregnancy induces a marked increase in circulating triglycerides and lipoprotein cholesterol.[53,54] In the second half of pregnancy, there is a shift towards using lipids as a maternal energy source, preserving glucose and amino acids for fetal nutrition.[54,55] The use of lipids for energy production produces ketone bodies which can also be taken up by the fetus and used in energy production or lipid construction.[54] In preeclampsia there are proatherogenic changes in the serum lipid profile, including increased triglycerides and very low density lipoprotein (VLDL) levels, and decreased high density lipoprotein (HDL) cholesterol compared to normal pregnancies.[56,57] The increased lifetime risk of CVD in women who experienced a hypertensive disorder in pregnancy is related to the common CVD risk factors in the diseases.[42]

Systemic inflammatory responses cause increased insulin resistance and hyperlipidaemia.[34] Inflammatory cytokines released from immune cells induce insulin resistance and lipolysis, and inhibit lipogenesis, causing an increase in circulating free fatty acids.[34] Systemic inflammation is also associated with increases in triglyceride-rich lipoproteins, decreases in HDL and impaired cholesterol transport. These metabolic changes promote atherosclerosis, and may provide a link between inflammation and the development of CVD.[18,34,58]

Another link between preeclampsia and cardiovascular disease is elevated homocystein levels, which is found in women with preeclampsia and is a risk factor for CVD.[59] Homocysteine levels can accumulate through deficiency of B vitamins and folate.[59] Women with a history of hypertensive disorders have elevated homocysteine compared to women with a history of uncomplicated pregnancies.[60]

In normal pregnancy, glomerular filtration rate (GFR) increases by 40-60% during the first trimester.[61] In preeclamptic pregnancies, both GFR and renal plasma flow decrease by 30-40% compared to normal pregnancies; though the common serum markers of renal function may remain in the normal range.[61]

1.1.6   The disease of theories

The causes of the disease continuum of preeclampsia have been widely discussed and researched in the last decades, and the description as a disease of theories is still apt.[62] In summary, preeclampsia is an extreme end of an inflammatory continuum common to all pregnancies. The underlying reasons for developing preeclampsia are still largely unknown, but thought to be a staged process in which a placental insufficiency develops early in pregnancy. Later in the pregnancy the compromised placenta releases signals to which the maternal system reacts. Both the first and second stage may be mediated by genetic and environmental factors. Preeclampsia can be seen as a stress test which may reveal propensity to develop cardiovascular disease later in life, due to similarities in risk factors and symptoms. The diagnosis and the most common subdivisions of the disease reflect the maternal end-stage disease phenotype, and poorly reflect the root cause of the disease. More specific disease phenotyping and diagnostic criteria are necessary for the future of research in preeclampsia. The multifactorial causes and subdivisions of preeclampsia are part of the reason that effective prediction, prevention and cure of the syndrome has not yet been reached.
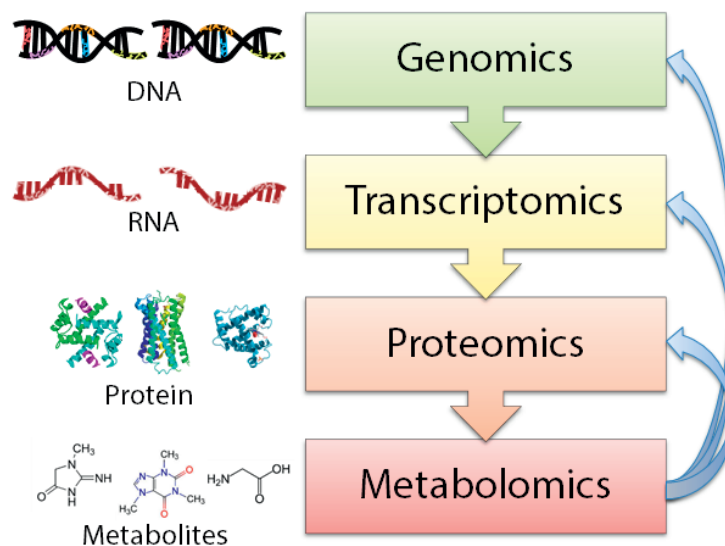
Figure 1.5: The 'Omics' cascade. All biological processes affect each other. Small changes in gene or protein expression may induce large fold changes in metabolite expression. Metabolites may also have regulatory effects in the genome and proteome.

1.2 Metabolomics

Hypothesis-generating approaches such as genomics, proteomics and metabolomics identify common patterns of disease in large data compilations.[63] Systematic changes in gene, protein or metabolite expression may predict or explain disease. Metabolomics is defined as *"the quantitative measurement of the multiparametric metabolic response to pathophysiological stimuli or genetic modification"*.[64] Commonly used analytical methods in metabolomics are mass spectrometry (MS) and nuclear magnetic resonance (NMR).

Metabolites are intermediates and end products of the metabolism, chemical reactions that are necessary for life. The metabolism provides energy, structural building blocks, signaling molecules, and detoxification of exogenous and endogenous molecules. Metabolites are typically small molecular weight (≤ 1500KDa) compounds, and are connected through metabolic pathways such as the citric acid cycle, glycolysis and gluconeogenesis, protein and amino acid synthesis, and lipid metabolism. The metabolome is generally referred to as the complete set of metabolites found in a cell, organ, tissue, biofluid or organism.[65]

Metabolites are the downstream products of gene and protein processes, and in essence are a closer measure of the phenotype of an organism.[66] Small changes in genes or protein activity can have large consequences in metabolite concentrations.[65] At the same time, metabolites may have regulatory effects on gene or protein expression. An illustration of the "Omics" cascade is shown in Figure 1.5. The total count of human metabolites is unknown, but the most comprehensive databases currently contain 20-40 000 annotated metabolites of endogenous or microbial origin, toxins or pollutants, food or drug derived, or a combination.[67,68]

There is a constant flux of metabolites in and out of cells and tissues. Consequently, disturbed processes due to genetic abnormalities, exogenous chemicals, errors in protein translation or function, or inflammation will be reflected in the metabolite composition of biofluids and tissues.[64]

Due to the limited sensitivity range of analytical instruments and the multiple fold range of metabolite concentrations, no single analytical method can detect and quantify all metabolites. Comprehensive profiling thus requires multiple platforms.[66] Metabolic profiling research is limited to the metabolites that may be detected on the instrument in use.

1.2.1 The human metabolome

Human metabolism is influenced by diet, stress, disease, gender, age, body mass index (BMI), and many other processes.[69–71] Metabolomic studies of human subjects have investigated biofluids and tissues such as urine, blood serum or plasma, cancer tissue, ocular tissue, and even hair.[69,72–74] Metabolomics has been employed to assess risk, diagnosis and subgrouping of CVD, diabetes and cancer, amongst others.[75] This thesis concerns the analysis of urine, serum and placental tissue samples.

Urine is a result of soluble waste products and water from the bloodstream being filtered through the kidneys. The main components of urine are urea from amino acid metabolism, inorganic salts, creatinine, ammonia, organic acids, and a variety of breakdown products from endogenous and exogenous metabolism.[66] The urine metabolome is the end product of human systemic and cellular metabolism, gut microbial metabolism, kidney function, exogenous compounds from diet and medication, and hydration status. Metabolic alterations in urine have been studied for many conditions, including blood pressure and cardiovascular risk, cancer, and pregnancy disorders.[71,76,77]

Blood is a favored subject of study because of its relatively noninvasive sampling and proximity to many potential biological sites of disease. Serum is the liquid remaining when a whole blood sample is allowed to clot, and subsequently centrifuged to remove red and white blood cells, platelets, and clotting proteins. Serum contains nutrients, electrolytes, lipids and lipoproteins, amino acids, and a variety of small organic compounds. The role of blood is to transport signaling compounds and waste products to various organs, defend against pathogens and blood loss, and regulate body temperature.[66,78] Serum is a common subject of metabolomics studies, spanning almost any human disease subject.[79]

The placenta plays a crucial role in pregnancy and preeclampsia, and also itself has metabolic activity. Published analyses of the placental metabolome are limited.[80–85] Placental extracts, cellular explants, or live imaging of placenta have been used to understand the metabolic changes occurring in the placenta in response to hypoxia, FGR, and preeclampsia.[80,86–89] However, specific changes in placental and fetal metabolism in response to preeclampsia and FGR is a wide area of research, as is genomic and proteomic analyses of placental tissue.[24,49,90] Metabolic research in placental tissue and explants is expected to provide a deeper understanding of the placental dysfunction in preeclampsia.[80]

1.2.2 Experimental design of metabolomics experiments

Metabolomics experiments are hypothesis-generating: they do not require a specific research question, but rather measure a complement of compounds to identify possible biomarkers. The results of the experiment then create hypotheses which may be tested in further studies. Experimental design is an important facet of metabolomics studies, as these are very prone to bias and confounding.

Subtle differences in sample collection, analyses or storage of the case and control samples may influence the analysis results in a manner that is impossible to differentiate from real disease differences. Typical experimental design of metabolomics studies have several stages, as summarized in Figure 1.6. These are 1) Generation of a biological question and study design, 2) Sample acquisition and analysis, 3) Data preprocessing, 4) Statistical analysis and metabolite identification, 5) Biological interpretation and hypothesis generation.[91] Each step is vital for the final results.

During the study design phase the number of samples required must be determined. Sample size estimation in metabolomic analysis is not as simple as a power
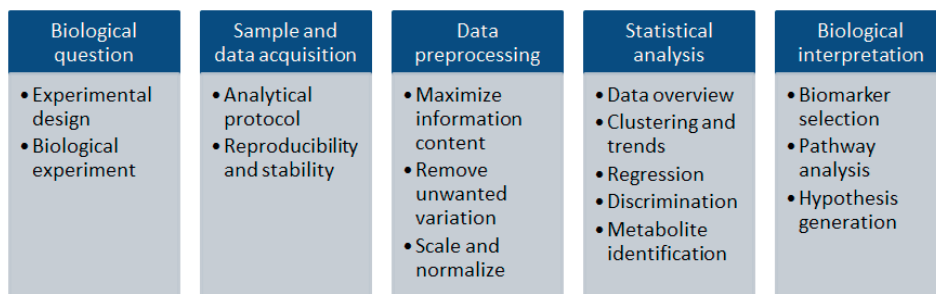
Figure 1.6: Work flow of a typical metabolomics study.

analysis for univariate studies. The collinearity of the measured variables compli-
cates this analysis.[91] Sample acquisition, analysis, preprocessing and identification
will be described in detail in the following chapters. Final biological interpreta-
tion depends on online databases of molecular pathways and interactions, previous
research, and intimate knowledge of the system and disease under study.

1.3 Magnetic resonance spectroscopy

Nuclear magnetic resonance (NMR) signals were first observed in 1946, and have developed into a valuable tool for chemists, biochemists, physicists and later medical scientists as a method of structural analysis of molecules, quantitative analysis of complex mixtures, and imaging of live organisms.[92]

Atomic nuclei with non-zero spin number ($I \neq 0$) precess about their axis with an intrinsic and quantized angular momentum $\boldsymbol{P}$, which is associated with a magnetic moment μ as follows: $\mu = \gamma \mathbf{P}$.[92] The proportionality constant γ is called the magnetogyric ratio, and is constant for each element and isotope. The nuclides ^{1}H, ^{13}C, ^{19}F, and ^{31}P are sensitive to NMR spectroscopy because of their spin numbers of $\frac{1}{2}$, with ^{1}H being the most commonly used in metabolomic studies as it has the highest natural abundance. However, the nuclides ^{16}O and ^{12}C, which are most common in organic molecules, are invisible to NMR spectroscopy. The sensitivity of nuclei in NMR spectroscopy is determined by a combination of natural abundance, the size of γ, and the magnetic field strength.

1.3.1 Nuclear magnetic resonance

When placed in an external magnetic field $\mathbf{B}_0$, individual nuclei with spin number $\frac{1}{2}$ will orient with spin axes parallel (α state) or anti-parallel (β state) to the field, as illustrated in Figure 1.7. The momentum with which the nuclei spins is proportional to the strength of $\mathbf{B}_0$, and is called the Larmour frequency ν_L. The α state is slightly more energetically favorable. Therefore the population of nuclei in the α state (N_α) will be slightly higher.
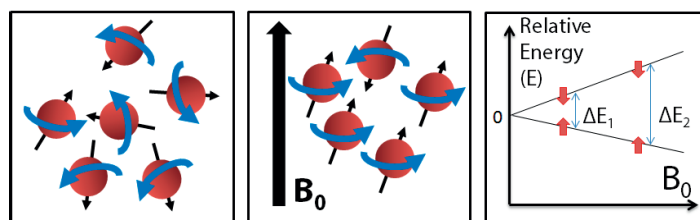
Figure 1.7: Atomic nuclei spin about their axis. When placed within an external magnetic field $\mathbf{B}_0$, atomic nuclei with spin number $\frac{1}{2}$ will orient with axes parallel or antiparallel to the field. The antiparallel direction has a higher energy state. The difference in energy ΔE between the parallel and antiparallel spin states increases with the strength of the external magnetic field.[92]

The energy difference between two adjacent energy levels is $\Delta E = \gamma \hbar B_0$,* and thus increases with magnetic field strength. The distribution between the higher and lower energy states is provided by the Boltzmann Equation (Equation 1). For protons the energy difference ΔE is very small compared to the average energy $K_B T$ of thermal motion, and consequently N_β/N_α is typically in the region of parts per million (ppm).

$$\frac{N_\beta}{N_\alpha} = e^{\frac{-\Delta E}{k_B T}} \approx 1 - \frac{\Delta E}{k_B T} = 1 - \frac{\gamma \hbar B_0}{k_B T} \tag{1}$$

In Equation 1, N_α/N_β is the ratio of nuclei in spin states α and β respectively, k_B is the Boltzmann constant ($=1.3805$ x 10^{-23} J K^{-1}), and T is the absolute temperature in K. The combined vector of the magnetized nuclei in α and β states result in a macroscopic magnetization $\mathbf{M}_0$ of the sample along the field direction.

In NMR experiments, transitions between energy states are induced by irradiating the nuclei with a radiofrequency pulse which matches the Larmour frequency of the nuclei. As a result, $\mathbf{M}_0$ is tipped away from the direction of $\mathbf{B}_0$ at an angle determined by the duration and amplitude of the pulse. A 90° pulse will flip $\mathbf{M}_0$ with an angle of 90° from $\mathbf{B}_0$. $\mathbf{M}_0$ returns to the equilibrium state through longitudinal (T_1) and transverse (T_2) relaxation (Figure 1.8). Longitudinal relaxation (T_1), also called spin-lattice relaxation, transfers energy to the environment through molecular movement to regain alignment to $\mathbf{B}_0$.
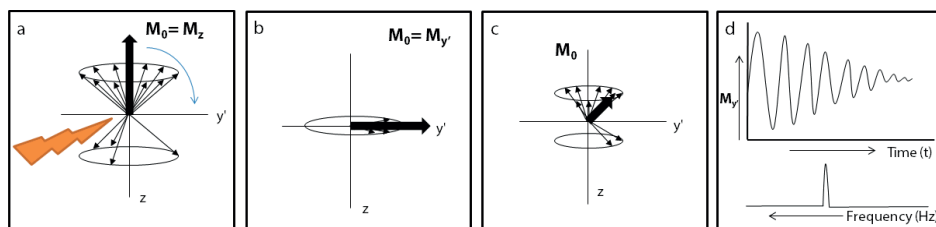
Figure 1.8: The macroscopic magnetization ($\mathbf{M}_0$) of the sample is tilted 90° by a radio-frequency pulse of appropriate power and length (panels a and b). Equilibrium is regained through longitudinal (T_1) and transverse (T_2) relaxation (panel c). The signal recorded in the y' direction is the free induction decay (FID), which when Fourier transformed gives an NMR spectrum (panel d).[92]

After a 90° pulse is applied, a fraction of the nuclear spins are bunched together in phases as they precess around the z axis. Transverse relaxation (T_2),

---

*$\hbar$ is $\frac{h}{2\pi}$ where h is the Planck's constant $= 6.6256 \times 10^{-34}$ J s

also called spin-spin relaxation, is caused by the dephasing of the individual magnetic moments of the nuclei, losing the x'y' component of the net magnetization. Short T_2-relaxation times give broad NMR signals, which increase with viscosity and decreasing temperature, presence of paramagnetic impurities, and magnetic field inhomogeneities. The real transverse relaxation time T_2^* is shorter than T_2 due to inhomogeneities in the magnetic field. The decay of the transverse magnetization as detected in the receiver is called the free induction decay (FID). The transverse magnetization is recorded as the raw NMR signal. The FID is Fourier transformed from time domain (s) to frequency domain (Hz) for interpretation as an NMR spectrum.

The resonance frequencies of nuclei are influenced by their environments. Because of shielding from surrounding electrons, each nucleus experiences a slightly different effective magnetic field. Nuclei in different chemical arrangements thus have different resonance frequencies ν. This difference compared to a standard reference frequency is expressed as the chemical shift and given in parts per million (ppm).

$$\delta[ppm] = \frac{\nu_{signal} - \nu_{reference}}{\nu_{reference}} \frac{[Hz]}{[MHz]} \qquad (2)$$

The chemical shift is independent of the field strength $\mathbf{B}_0$. The resonance signal is also affected by neighboring nuclei which results in a splitting of a peak into multiplets, through spin-spin coupling. The interaction stems from the two possible states of the nucleus under observation; α or β. The distance between peaks in a multiplet is called the coupling constant and denoted J (Hz). The shielding and splitting is characteristic to the position and bonding of the nucleus, and so the coupling pattern and chemical shift of a resonance signal give structural information about the molecule to which it belongs. The area under the signal curve is proportional to the number of nuclei producing the signal, which makes quantitative analysis possible.

1.3.2 NMR spectroscopy of biofluids and tissues

High resolution NMR spectroscopy can be applied to biofluids such as urine and serum,[64,93] and with some modifications to semisolid tissue biopsies such as from breast, brain, prostate, etc.[94] NMR spectroscopy requires little pretreatment and is capable of detecting a range of biochemical components simultaneously in an

unbiased manner. The analysis is quantitative or semi-quantitative depending on the acquisition parameters. The metabolic profile represented in a spectrum can be optimized by choice of acquisition parameters and pulse sequences, to suppress or enhance specific molecular subtypes.

Standard 1D spectra are acquired using the Nuclear Overhauser Effect Spectroscopy (NOESY) pulse sequence (D1-90°-t_1-90°-t_m-90°-AQ) with presaturation of the water frequency. NOESY spectra of urine samples produce high resolution spectra with narrow linewidths.

Macromolecules such as proteins, lipids and lipoproteins have less freedom of rotation and give broad signals in the NOESY spectrum because of their short T_2. It is possible to filter out signals from these molecules by spin-echo or T_2-editing the acquisition parameters, leaving spectra consisting of only signals from small metabolites.[95] A useful pulse sequence for lipid suppression in spectra of blood serum and tissues is the Carr-Purcell-Meiboom-Gill (CPMG) sequence in which a series of 180° pulses follow a 90° pulse after a delay τ (illustrated in Figure 1.9).[95] The 180° pulses refocus the signals which are dephased by T_2*, giving an echo of the 90° signal. Molecules with long T_2 have better preserved signals, and signals from molecules with short T_2 are reduced.
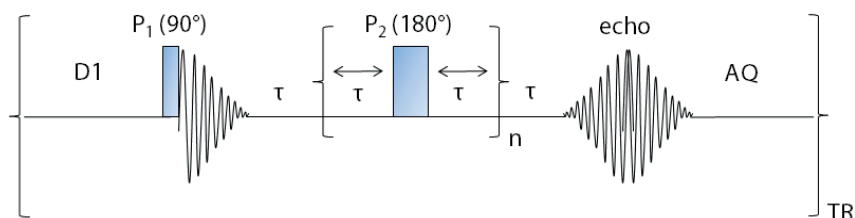
Figure 1.9: The Carr-Purcell-Meiboom-Gill pulse sequence. After a presaturation of the water resonance frequency for a time D1, a 90° pulse is followed by n repeated 180° pulses spaced by a time τ, giving a total echo time of TE=n(2τ+P_2)+2τ. The 180° pulses refocus the spins, and an echo is observed for a time AQ. TR is the total repetition time.

A spectrum focusing on the macromolecules can be made with a diffusion-edited pulse sequence (RD-90°-G_1-180°-G_1-90°-G_2-T-90°-G_1-180°-G_1-90°G_2-τ-90°-AQ). G_1 and G_2 are gradients that allow diffusion editing.[96] Signals from small and rapidly moving molecules are reduced, and a spectrum of only slower diffusing molecules is acquired.

J-resolved experiments (JRES) generate a simpler NMR spectrum. The pulse sequence is in the form (RD-90°-t_1-180°-t_1-AQ) where t_1 is an incremented time period. The spin-spin coupled multiplets are spread out over a second dimension, and the skyline projection gives a singlet for every multiplet in the recorded spectrum.[96]

Most biological samples contain large quantities of water, which may obscure the smaller metabolite signals. The water signal may be suppressed by irradiating the resonance frequency of water with a weak radiofrequency pulse, equalizing the population of the energy levels so that the water signal is almost completely attenuated before a spectrum is acquired.

In analyzing semisolid tissue samples, a method called high resolution magic angle spinning, or HR-MAS is used. Normal spectroscopy methods will result in very broad peaks due to anisotropic interactions of the less mobile molecules. These interactions are usually averaged in liquid state NMR by rapid molecular motion. The averaging in solution is mimicked by rapidly spinning the sample (4-6kHz) at the magic angle, $\theta = 54.74°$ relative to $\mathbf{B}_0$, which produces high resolution spectra approaching the liquid state spectra in linewidth.[94,97,98] Additionally, the tissue samples remain intact after the spectral acquisition and can be further analyzed by histopathology, proteomics or gene expression.[99]

Commonly, a standard water-presaturated NOESY experiment, a T_2-weighted CPMG experiment, J-resolved and sometimes diffusion edited sequences (DIFF) are used in metabolomics experiments.[96,100] Representative spectra of urine, serum and placental tissue are shown in Figure 1.10.

Two-dimensional (2D) spectroscopy such as heteronuclear single quantum coherence (HSQC), where proton NMR resonances are correlated to ^{13}C resonances, are useful for unambiguous metabolite identification, and can even be used directly in multivariate analysis.[101,102] ^{13}C-^{1}H spectroscopy approaches can identify more metabolites than one-dimensional ^{1}H spectroscopy alone, but the spectra take longer time to acquire because of the low natural abundance of NMR-visible ^{13}C nuclei (1%) and because of spectrum acquisition in two dimensions.

In HSQC, a polarization of ^{1}H ($\mathbf{M}_H$) is transferred to the much less NMR-sensitive ^{13}C nuclei, allowed to develop for a time t_1, then the ^{13}C polarization ($\mathbf{M}_C$) is transferred back to the protons, and the ^{1}H spectrum is recorded. A FID is recorded for each incremental t. In this way, information about the carbon attached to each recorded proton is revealed. A 2D spectrum is obtained with the directly observed ^{1}H spectrum along the top edge, and the indirectly observed ^{13}C

spectrum, along the left edge. An example HSQC spectrum of human urine is shown in Figure 1.11
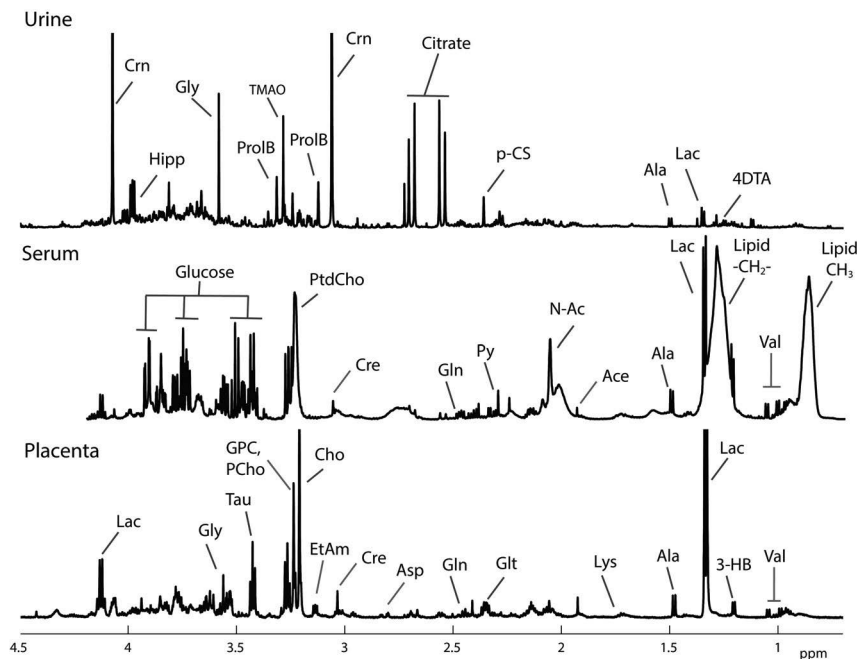
Figure 1.10: NMR spectra of urine, serum and placental tissue showing some of the metabolites that are detectable in NMR.

Abbreviations: 4DTA, 4-deoxythreonic acid; Ace, acetate; Ala, alanine; Asp, aspartate; Cho, choline; Cre, creatine, Crn, creatinine; EtAm, ethanolamine; Gln, glutamine; Glt, glutamate; Gly, glycine; GPC, glycerophosphocholine; Hipp, hippurate; Lac, lactate; Lys, lysine; N-Ac, N-acetylated carbohydrate side chains of glycoproteins; NMR, nuclear magnetic resonance; p-Cr, p-cresol sulfate; PCho, phosphocholine; ppm, parts per million; ProlB, proline betaine; Py, pyruvate, Tau, taurine; TMAO, trimethylamine-N-Oxide; Val, valine
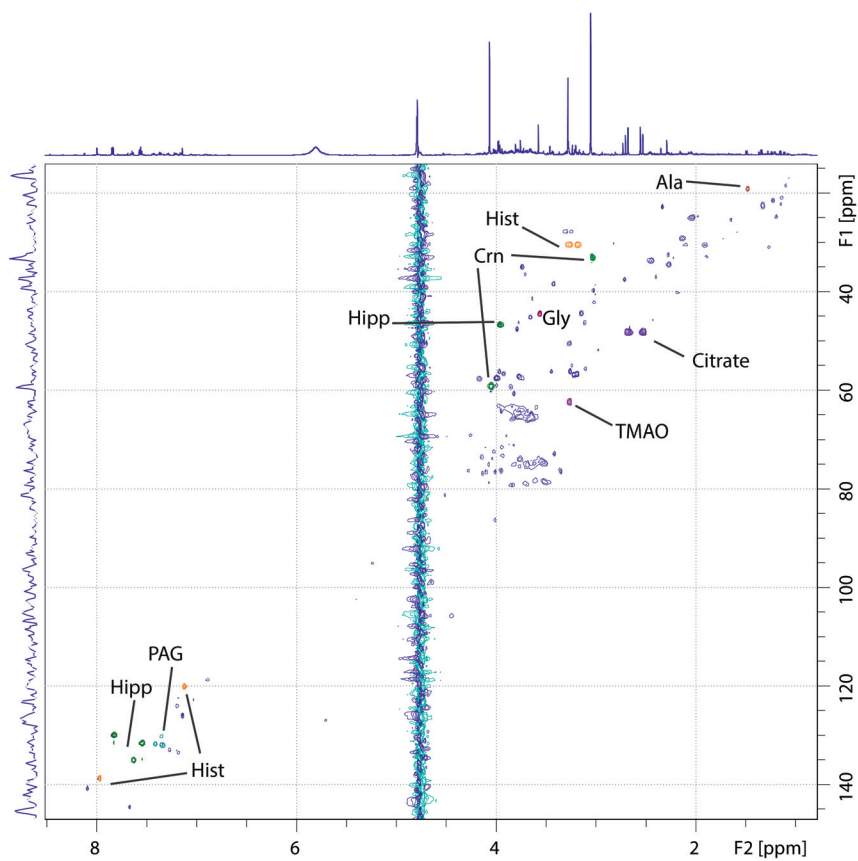
Figure 1.11: Heteronuclear Single Quantum Coherence spectroscopy of a urine sample showing the proton-carbon bonds. Signals from the same molecules are colored for illustration purposes.

Abbreviations: Ala, alanine; Crn, creatinine; Gly, glycine; Hipp, hippurate; Hist, histidine; TMAO, trimethylamine-N-oxide; PAG, phenylacetylglutamine; ppm, parts per million.

1.4   Multivariate analysis

Multivariate analyses is a set of methods for analyzing multiple, covarying measurements. Common goals of multivariate analysis are data description, discrimination and classification, and regression and prediction. Data description methods including principal component analysis (PCA) are unsupervised methods in that no *a priori* class information is input to the models. These methods are used for visualizing and exploring data sets. Discrimination and classification methods including partial least squares discriminant analysis (PLS-DA) are supervised methods, where the goal is to find differences between groups or classes that allow objects of unknown class to be assigned based on a set of measurements. Regression and prediction models such as partial least squares (PLS) regression aim to relate two sets of continuous variables to each other, e.g. to predict a property from a set of measurements.[103]

NMR spectra contain thousands of data points, many of them describing approximately the same information. The variation of metabolites in a biofluid or tissue may also be highly correlated due to their connection in metabolic pathways. In metabolomic experiments the number of variables measured commonly exceed the number of samples analyzed. All these considerations together mean that conventional statistics are not ideal to analyze metabolomic data sets. Dimension-reducing multivariate analysis applied to NMR spectra of biofluids may produce reliable metabolic profiles of physiological states.

1.4.1   Principal component analysis

A typical NMR spectrum of a biofluid or tissue consists of over 25 000 variables, with information about typically 20-50 compounds. The essential information within the spectra often lies in how the metabolites vary with respect to each other. This information must be separated from noise (e.g., regions of the spectra without signals) and redundancy (e.g., several signals from the same metabolite).

A typical method for visualizing highly multivariate data is PCA, which finds combinations of variables that describe major trends in the data. The multidimensional data is projected onto a lower dimensional space, typically two or three dimensions, for visual inspection. To do this, a new set of variables called latent variables (LVs) are created which are linear combinations of the original variables. In PCA LVs are called principal components (PCs). The PCs are mutually orthogonal,

that is, they describe independent variation, and they are ordered by the amount of variance captured. By selecting the first two or three principal components, most of the data variation can be plotted for visual inspection.[104]
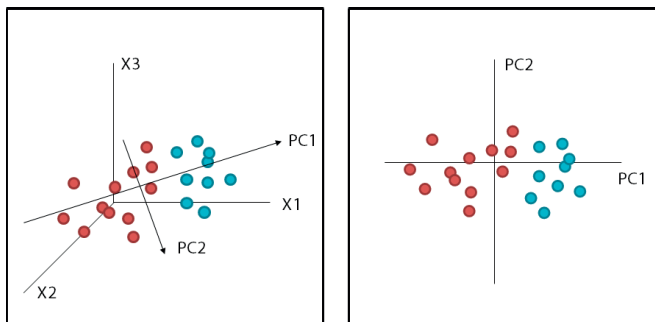
Figure 1.12: Principal Component Analysis for variable reduction, illustrated for three original variables transformed to two latent variables or principal components (PCs). The original data is projected into a lower dimensional space.

PCA gives the new coordinate system with a set of *scores* and *loadings*. The scores give the position of an object in the new coordinate system, and describes how the samples relate to each other. Scores can be useful to detect groupings, trends and outliers in the samples. Loadings define how the variables contribute to each principal component. The loadings also describe how the variables are related to each other and can be used to interpret the biological meaning of the explained variance.

Scores and loadings are related to the original data matrix as shown in Equation 3. The data matrix with samples in rows and variables in columns is called $\mathbf{X}$. $\mathbf{T}$ is the scores matrix giving the coordinates of the samples in the new coordinate system, $\mathbf{P}$ is the loadings matrix giving the contribution of each variable, and $\mathbf{E}$ is the residual matrix of unexplained variation.[104]

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \qquad (3)$$

A metabolomic dataset can usually be described adequately by far fewer PCs than the number of original variables. The combinations of variables found by PCA can be useful descriptors or even predictors of underlying phenomena such as physiological changes in the metabolome.

Figure 1.12 shows the PCA graphically. In this example, three variables are measured for each sample. When these are plotted in three dimensional space, a

pattern emerges that can be adequately described in two dimensions, where one axis describes most of the variation (PC1). In this case, the first PC also happens to describe a difference between two classes.

1.4.2 Partial Least Squares

Regular linear regression involves predicting a response variable y from one or more independent measurements x_1, x_2, (...), x_n. When a large number of variables are measured such methods are no longer valid. In addition, the variables measured by NMR are noisy and highly collinear. In this case, the partial least squares (PLS) regression method - also called projection to latent structures - is useful.[105]

In PLS, LVs are estimated so that they give directions of highest variance in a data set $\mathbf{X}$ with regards to a response vector or matrix $\mathbf{Y}$. The LVs are directly relevant for prediction of the response variables and are, like PCs, orthogonal. The $\mathbf{X}$ scores, denoted $\mathbf{T}$, are predictors of $\mathbf{Y}$ and are linear combinations of the original variables $\mathbf{X}$ with coefficient weights $\mathbf{W}$ (Equation 4).

$$\mathbf{T} = \mathbf{XW} \tag{4}$$

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \tag{5}$$

$$\mathbf{Y} = \mathbf{TC}^T + \mathbf{F} \tag{6}$$

The $\mathbf{X}$ scores, $\mathbf{T}$, are good summaries of the $\mathbf{X}$ matrix when multiplied with the loadings, $\mathbf{P}$ (Equation 5), and good predictors of $\mathbf{Y}$, when multiplied with the transpose of $\mathbf{C}$, the weights on $\mathbf{Y}$ (Equation 6). The residual matrices of unexplained variance, $\mathbf{E}$ and $\mathbf{F}$, are minimized. Equations 5 and 6 can be rewritten to give a multiple regression model (Equation 7) where $\mathbf{B}$ is a matrix of regression coefficients.

$$\mathbf{Y} = \mathbf{XB} + \mathbf{F}$$
$$where \tag{7}$$
$$\mathbf{B} = \mathbf{WC}^T$$

The PLS parameters can be estimated using non-linear iterative partial least squares (NIPALS) or the faster SIMPLS algorithm.[105,106] An assumption of the PLS model is that the modeled process is influenced by only a few underlying LVs, the number of which is unknown and can be estimated by PLS.[105]
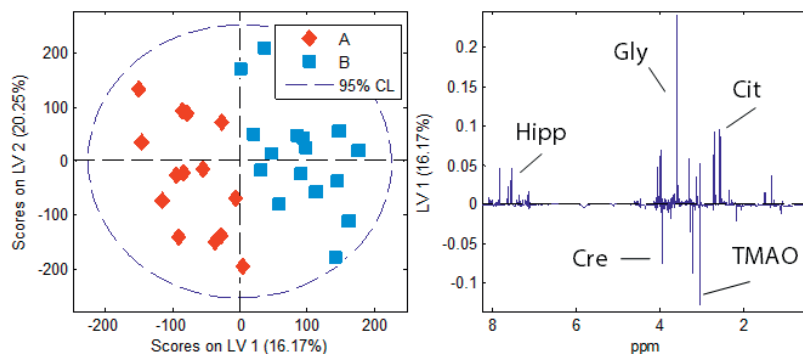
Figure 1.13: A constructed PLS-DA example with a scores plot (left) and a loading plot (right) showing a difference between group A and group B on LV1. From the plot of LV1 it is seen that group B has increased excretion of glycine (Gly), citrate (Cit) and hippurate (Hipp), and decreased excretion of creatine (Cre) and trimethylamine-N-Oxide (TMAO).

If the response variable is categorical, PLS can be used to find a threshold value of scores which discriminates between classes. In this case the analysis is PLS discriminant analysis, PLS-DA. A constructed example of a PLS-DA model discriminating between urine samples from group A and B is shown in Figure 1.13

1.4.3   Preprocessing

Before data is analyzed with multivariate analysis, it must be preprocessed to focus on the biologically relevant information content. Typical preprocessing of NMR spectra involves excluding uninformative regions, peak alignment, normalization, scaling and mean centering. Variable selection may also be applied.

Baseline correction. Distortions of the baseline may occur due to broad macromolecule signals, alterations in the FID, or instrument instability.[107] This might introduce systematic differences which are detected by multivariate methods. Several baseline correction algorithms exist, involving fitting a polynomial to the nonsignal region of the baseline or correcting in the time domain. Baseline offset corrections may be done by shifting the minimum intensity of the spectrum to zero.

Peak alignment. The position of a metabolite peak in the NMR spectrum may be shifted slightly due to variations in pH, temperature, ion concentrations and protein interaction. Multivariate methods require that corresponding data points in every spectra contain information from the same compound. This is seldom

26

the case in raw data, so the spectra must be aligned. In biofluids or tissue where concentrations and pH does not vary much, aligning the entire spectra by a single peak may be sufficient. For urine, where there is higher variability in the sample composition and pH, a more complex algorithm is required. The icoshift method independently aligns user defined segments of the spectrum by optimizing within-segment correlation.[108,109]

Normalization. Normalization multiplies each sample by a factor to account for dilution or sample weight differences. Dividing by total spectral area approximates the relative metabolite content of the sample.[110] There are several ways to account for dilution factors in urine samples. Creatinine normalization is common in clinical chemistry, and assumes a constant creatinine clearance rate.[110,111] Probabilistic Quotient Normalization (PQN) estimates a probable dilution factor by finding the median difference from a reference spectrum.[112]

Mean centering. Centering is a procedure typically carried out prior to scaling in which the mean intensity of each variable or column is subtracted from each individual intensity value. This centers the variation around zero, adjusting for differences between high and low abundance.

Scaling. Scaling methods divide each variable by a unique factor to account for fold differences between metabolites. Autoscaling uses the standard deviation as the scaling factor, so that all variables have equal variances and thus contribute equally to the model.[104] This will result in inflation of small, noisy values like the non-signal regions of NMR spectra.[104,113] Pareto scaling uses the square root of the standard deviation, so that large fold changes are reduced more than small fold changes.[113]

1.4.4 Variable selection

Several thousand variables are recorded for each sample in metabolomics. Given the redundancy and the required computational power, often a subset of the best variables for prediction can be selected. Variables with little biological meaning (such as the no signal regions of NMR spectra) can lead to worse predictions as they may cancel important information.[114] Several methods exist for variable selection.

Variable Importance in Projection (VIP). VIP variable selection uses a measure of variable importance to the prediction as a selection criteria.[115,116] The VIP score of a variable j is found by summing its influence in every latent variable

k (Equation 8). The average of the squared VIP scores is 1, and variables with VIP$\geq$1 can be assumed to be important in the model.

$$VIP_j = \sqrt{\frac{p \sum\limits_{k=1}^{h} b_k^2 \mathbf{t}_k^T \mathbf{t}_k \frac{\mathbf{w}_{jk}}{|\mathbf{w}_k|^2}}{\sum\limits_{k=1}^{h} b_k^2 \mathbf{t}_k^T \mathbf{t}_k}} \qquad (8)$$

In Equation 8, $\frac{\mathbf{w}_{jk}}{|\mathbf{w}_k|^2}$ is the weight of variable j in latent variable k, b_k is the regression coefficient of k, and $\mathbf{t}_k$ is the scores vector for k. The number of variables is p.

Competitive Adaptive Reweighted Sampling (CARS) CARS variable selection selects the variables with largest absolute coefficients in a multivariate regression model. In each run, a subset of samples are selected to create a calibration model. Next, important variables are selected by applying an exponentially decreasing function (EDF) and adaptive reweighting sampling of regions.[117] Four successive steps which are repeated several times comprise the CARS variable selection algorithm:

1. A PLS model is built on a subset of samples

2. Variables are selected by weights using EDF. The PLS variables with lowest regression weights are removed in every repeat. A higher ratio of variables are removed in the first repeats, according to the EDF.

3. After reducing the number of variables, adaptive reweighted sampling is used to further filter out variables selected in the previous step in a competitive way. The variables with the strongest weights are selected as most fit for prediction.

4. Finally, the remaining variables are evaluated using cross validation (Section 1.4.5). After all repetitions have been performed, the variable set with the lowest cross validation error is selected.

1.4.5 Assessment and validation of multivariate models

Multivariate prediction models are prone to overfitting data. An overfitted model is one that has modeled so much variation that it no longer applies to samples that were not used to build the model. In data sets with thousands of variables, it is nearly

guaranteed that some variables will be statistically different at the 95% confidence limit, creating a discriminative model. To evaluate whether a real predictive model can be made, proper validation should be performed. Typical approaches include cross validation and permutation testing.

Cross validation. Ideally a subset of samples is set aside completely as the model is developed, then used to validate the prediction accuracy of the final model. However, lack of sufficient numbers of samples is usually a problem. Cross validation is used to validate classification models when fewer samples are available. The model is built using a subset of samples, known as the training set, and tested on the remaining samples, comprising the test set. The procedure is repeated until all the samples have been included in the test set. The number of misclassified samples and other model diagnostics can then be found based on the test samples.[118]

In datasets with sparse numbers of cases, leave-one-out or leave-n%-out validation can be used. One or several samples are left out of the model building. Then the classes are predicted on the left out samples, and the process is repeated. The number of PLS LVs are selected as the number giving the lowest cross validated error of classification. This method tends to overestimate prediction accuracy as the test set is not completely independent; it has been used to estimate the parameters for the model.[118]
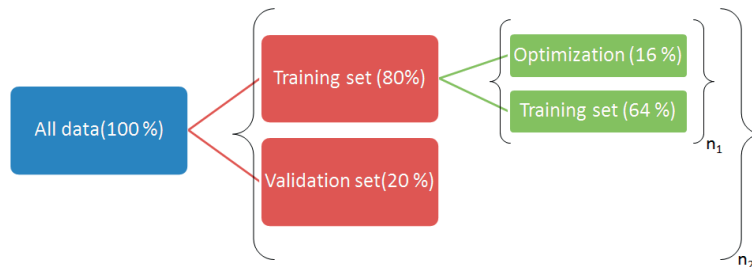
Figure 1.14: Double cross validation as detailed by Westerhuis *et al.*[118] A test set is set aside, and the remaining data is divided into an optimization and training set. In this way, the model parameters are optimized without information from the test set. The inner (n_1) and outer (n_2) loops are repeated several times. A typical single cross validation procedure only contains the outer loop n_2)

If number of samples permits, a double cross validation procedure can be implemented. The procedure is illustrated in Figure 1.14. A subset of samples is set aside as a validation set, and the remaining samples are divided into a training and

optimization set. Model parameters such as the number of LVs are found in the inner optimization set, while the final model is used to predict the test set. Now the model is tested on completely independent samples and the correct prediction errors are found.[118] Assessing the predictive or discriminative ability of a PLS model should be done on cross validated parameters.

Permutation testing. A permutation test evaluates whether a given regression or discrimination model is significantly better than models built on permuted classes; where the y vector is shuffled and a new model built on this random data. By building many permuted models, an H_0 distribution of classification or regression results of random models is found. The classification results from the real model should then be outside the 95% confidence limit of the H_0 distribution to be considered valid.[118]

When using variable selection there is an additional challenge in validating models with permutation testing and cross validation. An information leakage about class will happen when class information is used to select variable subsets for prediction. One way to alleviate this is to allow the permutation analysis to also select subsets of variables the same way after the classes have been shuffled. This extra step ensures that a true relationship has been found between class-selected variables and cross-validated response.

Sensitivity, specificity and classification accuracy. The sensitivity of a prediction model is the percentage of true positive classifications (TP) out of the sum of true positives and false negative (FN) classifications.

$$Sensitivity = \frac{TP}{TP + FN} \tag{9}$$

The specificity is the percentage of true negative (TN) classifications. Often in prediction research the false positive rate is used, which is 1-specificity.

$$Specificity = \frac{TN}{TN + FP} \tag{10}$$

Classification error is the total number of incorrectly classified objects, while model accuracy is 1-error, the number of correctly classified samples.

$$Error = \frac{FP + FN}{TP + TN + FP + FN} \tag{11}$$

The receiver-operating characteristic curve (ROC) is a plot of 1-specificity vs. sensitivity, and can be used to find the best cutoff for a prediction or to calibrate

a predictive test to a given specificity e.g. 90%. An often used measure is the area
under the curve (AUROC), ranging from 0.5 to 1, as an evaluation of a predictive
test.

2 Aims

2.1 Overall objective

The aim of the thesis work was to find new metabolic biomarkers for development of preeclampsia that may predict and characterize the disease, aid in risk stratification, and assess metabolic differences in its early and late stages.

2.2 Specific objectives

- To determine the metabolic profiles in urine and serum of women diagnosed with preeclampsia in comparison to normotensive pregnant and nonpregnant women, for possible clues to the etiology and pathogenesis of the disease.

- To identify predictive biomarkers for preeclampsia and gestational hypertension in urine and serum in early pregnancy, and assess the metabolic changes related to the diseases.

- To characterize the placental metabolic profile in preeclampsia in comparison to normotensive pregnancies, and assess metabolic differences according to established subgroups of preeclampsia.

3 Materials and Methods

3.1 Patients and Data Sets

The papers included in this thesis are based on data from three different sources. Paper I uses data from the Pilot Study for Identification of Biomarkers in Preeclampsia (The Pilot Study), paper II uses data from the ScreenTox study, and paper III uses data from the Preeclampsia Biobank). These data sources are briefly summarized below.

3.1.1 Pilot Study for Identification of Biomarkers of Preeclampsia

The Pilot Study was a small study conducted at the Medical Faculty, NTNU. Women with diagnosed preeclampsia (n=10), women with healthy pregnancies (n=10), and nonpregnant women (n=10) were recruited from the local environment of St. Olavs Hospital, Trondheim University Hospital, Norway and at Røros Medical Center, Røros, Norway. Patients admitted with suspected preeclampsia to the Maternity Ward at St. Olavs Hospital were asked to contribute blood and urine samples at the time of or soon after their diagnosis was set. Normotensive pregnant controls were matched by maternal age and gestational age to the preeclamptic women, and maternal blood and urine samples were collected. The nonpregnant women were matched by age and blood and urine samples were collected. All the women were of Scandinavian ethnicity. Women were included in the study between September 1st 2011 and December 15th 2012. Previous and current hypertensive diagnoses, medication and time of last meal were recorded at the time of sample donation. Gestational diagnoses, highest blood pressure, degree of proteinuria, and length, weight and gestational age at birth of the baby were obtained from hospital records. All women gave written informed consent and the study was approved by the Regional Committee for Medical and Health Research Ethics (REC), Central Norway, reference number REK 2011/761.

3.1.2 The ScreenTox Study

The ScreenTox Study was a prospective screening study conducted at the National Center for Fetal Medicine, St. Olavs Hospital in Trondheim, Norway. The cohort has been described in detail in previous publications.[119,120] Women who were nulliparous or had previously had preeclampsia or gestational hypertension, and who

were referred to routine ultrasound at 18 weeks of gestation, were sent a letter with information about the study, and the study was advertized on the Internet at the hospital web page and via Google AdWords. Women were included in the study between September 1st 2010 and March 31st 2012. During the study inclusion period, 5769 women attended a second trimester routine ultrasound at the National Center for Fetal Medicine and 2602 (45%) of these women were nulliparous. All study participants were scheduled for a visit between gestational week 11+0 and 13+6 based on previous ultrasound scans or the last menstrual period.[121] Women who were para ≤ 1 without previous preeclampsia or gestational hypertension, or were using aspirin or low molecular heparin were not eligible for inclusion.

At the time of inclusion between gestational week 11+0 and 13+6, women were interviewed about chronic diseases, medication, ethnic origin, smoking status, method of conception, previous or familial history of preeclampsia, and height. All women had their weight and blood pressure measured. Information about pregnancy outcomes was obtained from hospital records. Estimated date of delivery and gestational age at delivery was based on a second trimester routine ultrasound examination.[121] If spontaneous abortion or termination of pregnancy occurred after the study visit and before the second trimester routine ultrasound, gestational age was based on crown-rump length measurements. For estimation of weight deviation at birth, the normal growth curves from Maršál *et al.* were applied.[122] The definition of small for gestational age (SGA) was birth weight with a mean less than -22%. Stillbirth was defined as a dead child born with a birth weight $\geq$500 g or, if information about birth weight was unavailable, a gestational age $\geq$22 weeks.[123] If birth weight was <500g, the birth was classified as a spontaneous abortion. Blood pressure was measured with a CAS 740 MAX NIBP automated device (CAS Medical systems Inc, CT, USA).[124] Mean arterial pressure (MAP) was calculated as $MAP = \frac{1}{3}SP + \frac{2}{3}DP$ where DP is diastolic blood pressure and SP is systolic blood pressure. MAP from the arm with the highest MAP was used. BMI was calculated in kg/m^2 from maternal weight in the first trimester. Participants in study were examined with transabdominal ultrasound with a Siemens ACUSON Antares^{TM} machine (Siemens Medical Solutions Inc, CA, USA).[125] The UtAPI was measured three times on each side, and the average of three measurements on each side was used. The average of the two sides was used in calculations. All scans were carried out by specialized trained midwifes who were certified by the Fetal Medicine Foundation (`http://www.fetalmedicine.com`).

At the end of the study inclusion period, 640 women had attended a study visit of which 585 were nullipara. Of these, 38 were excluded for reasons appearing at or after the study visit: missed abortion, multiple pregnancies and severe congenital anomalies. Eight serum samples and 12 urine samples were excluded due to missing samples or due to failed NMR acquisitions. A flow chart of study participants is shown in Figure 3.1. Of the remaining cohort, 26 (4.5%) women developed preeclampsia (25 late onset preeclampsia), and 21 (3.6%) developed gestational hypertension. Ninety-eight percent were of European, Middle Eastern or North African ethnic origin. All women gave written informed consent at study entry. The study was approved by the REC, reference numbers REK 2010/102 and 2013/386
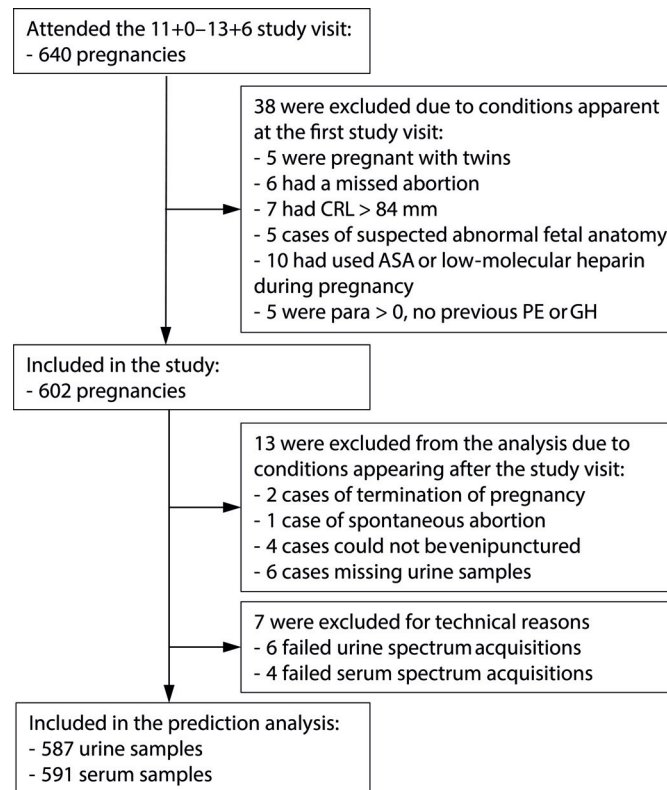
Figure 3.1: Flowchart of study participants in paper II.
Abbreviations: ASA, acetylsalicylic acid (aspirin); CRL, crown rump length; GH, gestational hypertension; PE, preeclampsia.

3.1.3 The Preeclampsia Biobank

The Preeclampsia Biobank consists of biological materials and health information collected from pregnant women at St. Olavs Hospital, Trondheim, Norway, and Haukeland University Hospital, Bergen, Norway. This NTNU project has collected pregnancy related material in three phases named Trondheim 1 (2002-2006), Bergen 1 (2003-2006), and Bergen 2 (2010-2012). Placental tissue and serum samples from Bergen 2 were used in Paper III.

Women with preeclampsia (n=19) and normotensive controls with otherwise healthy pregnancies (n=15) were included. Women with preeclampsia were included in the study if they were undergoing caesarian section (CS) for complications of preeclampsia. Normotensive women were included if they were undergoing CS for reasons considered not relevant for the study: maternal request, fetal breech position, or previous CS. Only singleton pregnancies without labor activity were included. All women were of Scandinavian ethnicity. Pregnancies with chromosomal alterations, fetal and placental structural abnormalities or perinatal infections were excluded. Information about the current and previous pregnancies, diagnoses, height and weight, blood pressure and proteinuria were recorded by interview and by reviewing medical journals. Maternal venous blood was collected prior to CS, and placental tissue samples were collected after delivery.

Fetal growth restriction (FGR) in the PE biobank was diagnosed when serial ultrasound measurements showed reduced intrauterine growth. In absence of serial ultrasound measurements, neonates were defined with FGR if their birth weights were $\leq$5th percentile for gestational age according to Norwegian fetal weight reference curves.[126] Large for gestational age (LGA) was defined as birth weight $\geq$95th percentiles. The study was approved by REC (REK 2012/1040) and informed, written consent was obtained from all participants.

3.2 Clinical Diagnostics

The Norwegian Association for Obstetrics and Gynecologists' definition of preeclampsia and severe preeclampsia from 2008, a slightly modified version of the 2002 guidelines of the American College of Obstetricians and Gynecologists,[127] was used in the Pilot study and ScreenTox biobanks.[128] In the "PE biobank", the 2014 Norwegian guidelines were used.[129] However, the diagnostic criteria were equal in the Norwegian guidelines.

For all the biobanks used in the scope of this thesis, preeclampsia was defined as persistent systolic/diastolic blood pressure $\geq$140/90mmHg and proteinuria $\geq$0.3g/24h occurring after gestational week 20. Gestational hypertension was defined as persistent systolic/diastolic blood pressure $\geq$140/90mmHg without proteinuria, occurring after gestational week 20.[127] Superimposed preeclampsia was defined as hypertension pre-existing the pregnancy plus proteinuria developing after 20 weeks of gestation.[129] Severe preeclampsia was defined as preeclampsia with presence of one or more severe clinical signs or symptoms: epigastric pain, severe headache and/or other cerebral symptoms (visual disturbances, hyperreflexia, edema, pulmonary edema (dyspnea, cyanosis), eclampsia, persistent blood pressure $\geq$160/110mmHg, proteinuria $\geq$3g per 24h, oliguria $\leq$500mL/24h, thrombocytopenia, microangiopathic hemolytic anemia, or elevated liver enzymes.[2] FGR was not a criteria for severe preeclampsia. Early onset preeclampsia was defined as preeclampsia diagnosed before 34 weeks gestational age.[3]

3.3 Sample Handling

3.3.1 Sample collection and storage

Urine and serum samples in the Pilot Study and ScreenTox biobank (papers I and II): peripheral venous blood (5mL) and spot urine samples (20mL) were collected from nonfasting women at their study visit. Venous blood was drawn into a non-heparinized tube, and blood samples were left to clot for $\leq$30 minutes and centrifuged at 1800G for 10 minutes. Aliquots (1.8mL urine and 1/0.5 mL serum (papers I/II) were stored at -80°C until analysis. The serum for paper II was thawed once for further splitting into aliquots for metabolomic analyses.

Placenta and serum samples from the PE biobank (paper III): a tangential section (100mg) from the maternal central side of the placenta was taken as soon as possible after CS (mean $\pm$ SD, 101 $\pm$ 49 minutes). The biopsies were placed directly in cryotubes and snap frozen either in liquid nitrogen or directly at -80°C until analysis. Maternal venous blood was collected prior to CS, left to clot for $\leq$30 minutes, centrifuged at 1800G for 10 minutes, and serum aliquots (1mL) stored at -80° until analyses.

3.3.2 Clinical chemistry serum analyses

Serum samples from the PE biobank (paper III) were analyzed for common clinical chemistry measurements to complement the HR-MAS placental analysis. Serum sFlt-1 was measured in duplicate using a quantitative sandwich ELISA according to the manufacturer's instructions (#DVR100B, R&D Systems, Abingdon, UK). High sensitivity C-reactive protein (hsCRP) (turbidimetric assay, Modular P analyzer, Roche, Burgess Hill, UK), total cholesterol, HDL, triglyceride and creatinine (enzymatic colorimetric assays, Modular P analyzer) were measured at the Department of Clinical Chemistry, St. Olavs Hospital, Trondheim, Norway.

3.3.3 Sample preparation for NMR analysis

Urine samples were thawed at room temperature (paper I) or over ice (paper II) and centrifuged at 6000RPM. The supernatant(540μL) was mixed with a bacteriostatic buffer containing Trimethylsilylpropionic acid (TSP) as a reference (6mM TSP, 1.5M KH_2PO_4, 2mM NaN_3, pH 7.4 in D_2O), transferred to a 5mm NMR tube, and kept at 5 °C until analysis.

Serum samples were thawed at room temperature (paper I) or over ice (paper II). A bacteriostatic buffer containing TSP as a reference (4.6mM TSP, 0.075M $Na_2HPO_4 \bullet 7H_2O$, 6.2mM NaN_3, pH 7.4, in 20% D_2O/H_2O) was mixed with serum (90μL), transferred to a 3mm NMR tube, and kept at 5°C until analysis.

Placental biopsy samples were prepared on a metal plate bathed in liquid nitrogen within a short time period in order to minimize the effect of tissue degradation on the metabolic profiles. Biopsies (mean ± SD, 7.5 ± 1.4mg) were cut while still frozen to fit 30μL disposable inserts (Bruker Biospin Corp, USA), and filled with 3μL cold D_2O containing 25mM formate for shimming. The inserts were placed in 4mm zirconium MAS rotors and transferred immediately to the magnet.

3.4 NMR Analysis

3.4.1 Optimization of NMR parameters

Analytical reproducibility. Reproducibility was tested throughout the acquisition period of spectra in papers I and II. Five to eight samples were independently prepared and analyzed. A PCA score plot of the samples (Figure 3.2) show the interindividual variation to be much larger than the analytical and sample handling differences,

which consist mainly of small peak position shifts due to minute inconsistencies in pH after buffer addition.
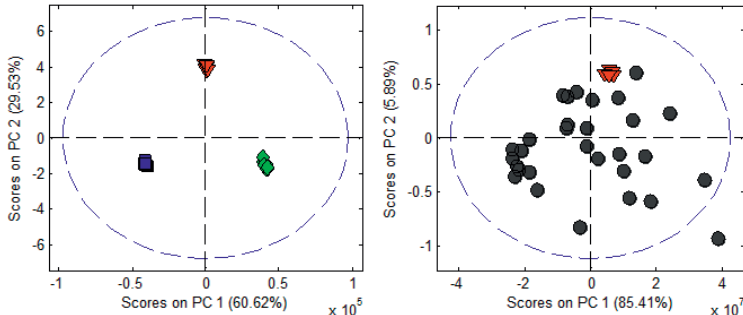
Figure 3.2: Reproducibility of NMR measurements. Left: Urine samples from three individuals prepared in 6-8 replicates and analyzed separately by NMR, then compared in a PCA. Right: a PCA of serum from participants in paper I with samples from one individual prepared in 5 replicates (in red triangles). Sample handling and analytical variation produced smaller differences than individual variation.

Temperature calibration. Sample temperature was calibrated for every new experiment setup. A 99.8% deuterated methanol sample equilibrated in the sample probe for 5 minutes before spectrum acquisition. The exact temperature was calculated by measuring the distance Δ between the two peaks, and applying Equation 12. The temperature in the probe was then adjusted to correct for the offset.

$$T(K) = -23.832\Delta^2 - 29.46\Delta + 403 \tag{12}$$

Shimming and water suppression. For biofluid spectroscopy, a thorough optimization of the automated shimming and water suppression was performed on a standard 2mM sucrose in D_2O solution. Ideal parameters produces a linewidth $\leq$1Hz and irradiation of the center of the water peak, suppressing the water signal in a urine sample to below the CH_3 peak of creatinine. A perfectly shimmed, water suppressed and pulse calibrated urine sample is shown in Figure 3.3

Tuning, matching, pulse and magic angle calibration. Automatic tuning and matching of the ^{1}H resonance frequency as well as calibration of the 90° pulse were used for biofluid spectroscopy. For the HR-MAS analysis, tuning, matching, shimming and pulse calibration were manually adjusted while in gs mode (continuous acquisition). Linewidth of the formate peak close to 1Hz were typically obtained
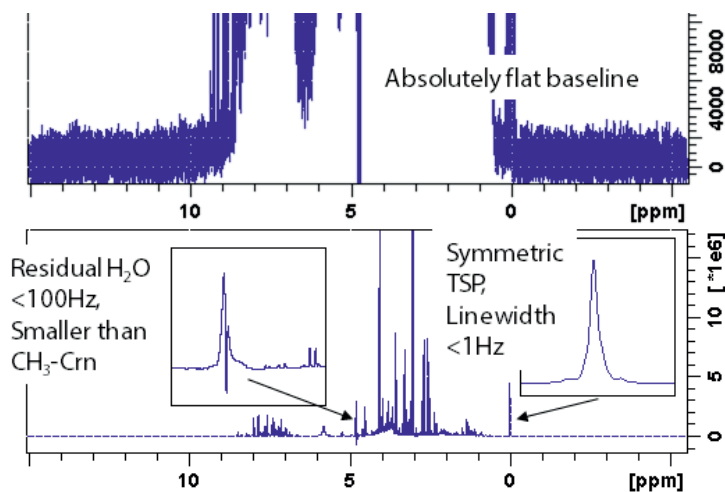
Figure 3.3: A perfectly shimmed, baseline corrected and water suppressed urine spectrum obtained by automated procedures. The linewidth is $\leq$1Hz, the baseline is flat and the residual H_2O peak is smaller than the highest metabolite peak in the spectrum.

on placental samples (mean $\pm$ SD, 1.29 $\pm$ 0.36Hz). For HR-MAS, additional adjustment of the magic angle was necessary to obtain maximum spectral resolution. This was done by continuously observing the ^{79}Br frequency of a KBr sample spun at 5KHz, while adjusting the angle under the magnet. At the magic angle, the spinning sidebands of the ^{79}Br signal are maximized at 10-12% of the main signal.

3.4.2   NMR equipment and settings

The NMR spectroscopy parameters for the analysis protocols are described in Table 3.1. Urine and serum samples were analyzed on a Bruker Avance III Ultrashielded Plus 600MHz spectrometer (Bruker Biospin GmbH, Rheinstetten, Germany) equipped with 5mm QCI Cryoprobe with integrated, cooled preamplifiers for ^{1}H, ^{2}H and ^{13}C. Experiments were fully automated using the SampleJet^{TM} sample changer in combination with Icon-NMR on TopSpin 3.1 software (Bruker Biospin).

Urine samples were analyzed at 300K with NOESY (noesygppr1d; Bruker) and JRES (jresgpprqf; Bruker) pulse programs, with total experiment time of 15 minutes including sample transfer and temperature equilibration time. Serum samples were analyzed at 310K to decrease viscosity, with NOESY, CPMG (cpmgpr1d; Bruker) and JRES pulse programs. For paper I a DIFF spectrum (ledbpgppr2s1d; Bruker)

Table 3.1: NMR spectroscopy parameters used in this thesis.

Analyte	Spectrum	NS	TD (Hz)	AQ (s)	D1 (s)	TE (ms)
Urine	NOESY	32	65k	2.65	4	
	JRES	2	8k×40	0.41×0.26	2	
Serum	NOESY	32	98k	2.74	4	
	CPMG	64	65k	2.73	4	78
	JRES	1	8k×40	0.41×0.26	2	
	DIFF	64	65k	1.83	4	
Placenta	NOESY	128	98k	2.74	4	
	CPMG	256	73k	3.07	4	78
	JRES	1	8k×40	0.41×0.26	2	

Abbreviations: AQ; acquisition time; CPMG, Carr-Purcell-Meiboom-Gill pulse sequence; D1, relaxation delay with water suppression; DIFF; diffusion-edited spectroscopy; JRES, J-resolved spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; NS, number of scans; SW, sweep width; TE, total echo time; TD, spectrum width in the time domain

was also acquired. Total experiment time for serum samples was 23 minutes (31 with DIFF) including sample transfer and heating time for the samples.

Placental biopsies were analyzed by HR MAS NMR on a Bruker Avance DRX600 spectrometer (Bruker Biospin) equipped with a ^{1}H/^{13}C MAS probe with gradient using 5KHz spin rate. Samples were analyzed at 5ºC to minimize tissue degradation. NOESY, CPMG and JRES spectra were acquired. Total experiment time was 43 minutes, with an additional 15-20 minutes for sample preparation and optimization of NMR acquisition parameters.

Additional 2D spectra were obtained for selected urine and placental samples for aid in metabolite identification: Urine was analyzed with HSQC (hsqcedetgpisp.2; Bruker), COSY (cosygpprqf; Bruker), HSQC-TOCSY (hsqcdietgpisp.2; Bruker) and HMBC (hmbcetgpl3ndpr; Bruker). Placenta was analyzed with HSQC and HMBC.

3.5 Data analysis

3.5.1 Multivariate analyses

Following spectral acquisition, the spectra were automatically phased and baseline corrected, and a 0.3Hz line broadening was applied in TopSpin 3.1 (Bruker Biospin). Spectra were imported to Matlab R2010 / R2013b (The Mathworks Inc., Natick, USA) using a script provided by Bruker. Regions containing water signal residuals and nonsignal regions at the edges of the spectra were removed.

Spectral preprocessing and validation methods used in the thesis are summarized in Table 3.2. Urine spectra were aligned using the iCoshift algorithm[108] and normalized to unit area in paper I and with PQN normalization[112] in paper II. Serum NOESY spectra were aligned by TSP (paper I) or by alanine (paper II) and normalized to unit area (paper II). In paper I the diffusion-edited serum spectra were limited to the lipid regions and normalized to unit area. Placental spectra were aligned by iCoshift and normalized to unit area. Mean centering was performed by subtracting the sample mean from every variable. VIP and CARS variable selection was performed prior to PLS-DA in paper II. VIP selection was performed by creating a five-fold cross validated model, then selecting variables with VIP$\geq$1 to use in a further model. CARS variable selection was performed with five fold cross validation and 50 Monte Carlo samplings.

PCA, PLS and PLS-DA were performed in Matlab using PLS Toolbox v7.1.1 (Eigenvector Research, Wenatchee, USA). PLS and PLS-DA were performed using the SIMPLS algorithm. In paper I, full cross validation (leave one out) was performed on the PLS-DA models due to the low number of samples. In paper II, a double cross validation procedure was performed. A set of samples (20%) was set aside for independent validation (outer loop). The remaining samples were then split into an inner calibration (80%) and test (20%) set for determining the optimal number of LVs (inner loop). The inner and outer loops were both repeated 20 times, and the model diagnostics (classification error, sensitivity, specificity) calculated from the validation results. In paper III, a single loop cross validation was performed. One fifth (20%) of the samples were set aside and used for validation until all of the samples had been in the test set once, and the procedure was repeated 20 times. The number of LVs giving the lowest cross validated error of prediction was used. Permutation testing was performed by randomly shuffling class or continuous variable assignments, and building a model using the same parameters as for

the true class labels. Where variable selection was performed, a variable selection step was included in the permutation test. The permuted results were compared to the true results, and where p≤0.05, the models were considered valid.

Table 3.2: Preprocessing of NMR spectra before multivariate analyses

	Paper I		**Paper II**		**Paper III**
	Urine	Serum	Urine	Serum	Placenta
Spectrum	NOESY	CPMG	NOESY	CPMG	CPMG
Preprocessing					
Normalization	Area	-	PQN	Area	Area
Scaling	Pareto	-	Pareto	-	Autoscaling
Alignment	Icoshift	TSP	Icoshift	Alanine	Icoshift
Mean centering	yes	yes	yes	yes	yes
Variable selection	-	-	VIP, CARS	VIP, CARS	-
Validation					
Cross validation	LOO	LOO	$2CV^a$	$2CV^a$	$1CV^b$
Permutations	1000	1000	100	100	1000

Abbreviations: CARS, competitive adaptive reweighted sampling; CPMG, Carr-Purcell-Meiboom-Gill pulse sequence; LOO, leave-one-out; NOESY, nuclear Overhauser effect spectroscopy; PQN, probabilistic quotient normalization; TSP, trimethyl-silyl propionic acid; VIP, variable importance in projection.

[a] Double loop cross validation

[b] Single loop cross validation (leave 20% out)

3.5.2   Metabolite identification and quantification

Metabolite peaks in the NMR spectra were compared to online databases of ^{1}H and ^{13}C metabolite resonances: the Human Metabolome Database (HMDB)[67], the Madison-Qingdao Metabolomics Consortium Database[68] and software databases of reference spectra AMIX v2.5 (Analysis of MIXtures software, Bruker Biospin GmbH) and Chenomx v.7.11 (Chenomx Inc., Alberta, Canada). Further confirmation of metabolite identities were made using Statistical Total Correlation Spectroscopy (STOCSY) to obtain statistically correlated peaks in NOESY, CPMG and HSQC spectra resembling pure metabolite spectra.

In paper I, metabolites were quantified using Chenomx based on the visible TSP concentration. TSP was quantified in Topspin using a synthetic signal based on a reference creatine (14.43mM) spectrum recorded at equal NMR acquisition parameters.[130] Serum metabolites were quantified from CPMG spectra. Urine metabolites were quantified from NOESY spectra and normalized to creatinine. In paper II, regions of metabolites found important to prediction (hippurate, 4-deoxythreonic acid, lactate, dimethylamine, creatinine) were integrated from urine spectra and normalized to the creatinine value.

Lipoproteins and lipids were not quantified directly from NMR in this thesis, but the contributions of signals arising from the lipid moieties within the lipoproteins were considered in the multivariate models in papers I and II. Water-insoluble lipids in blood are transported by lipoproteins in the form of triglycerides and esterified cholesterol.[131] The position of the signals from the lipids depends on the size of the lipoprotein carrying it; smaller, denser lipoproteins like HDL have lower ppm values while larger LDL and VLDL lipoproteins, containing triglycerides and cholesterol, have higher ppm values.[131] The methyl ($\approx \delta 0.9$ppm) and methylene ($\approx \delta 1.3$ppm) regions of serum NMR spectra contain overlapping resonances from VLDL and HDL, as well as contributions from low- and intermediate density lipoprotein. The VLDL/triglyceride and HDL regions of paper I were identified by comparison to reports based on purified lipoprotein subfractions.[132] In paper II the VLDL and HDL regions were additionally identified by PLS regression of the spectra to lipoprotein concentrations obtained by methods described in Section 3.3.2, and by STOCSY, showing separate contributions to the broad lipid resonances from the left (VLDL) and right (HDL) side of the peak (Figure 3.4).

In paper III, relative metabolite intensities were calculated in by integrating the area under the metabolite signal in the normalized spectra after removal of the residual water signal. The metabolite regions with the least signal overlap were chosen for quantification.

3.5.3 Univariate data analysis

Univariate tests, logistic regression and linear regression models were performed in SPSS Statistics 20.0.0 (IBM Corp, NY, USA). The Kolmogorov-Smirnov test was used to evaluate normality of the data. Nonparametric metabolite data was compared between groups using the Mann-Whitney U test for two groups or Kruskal-
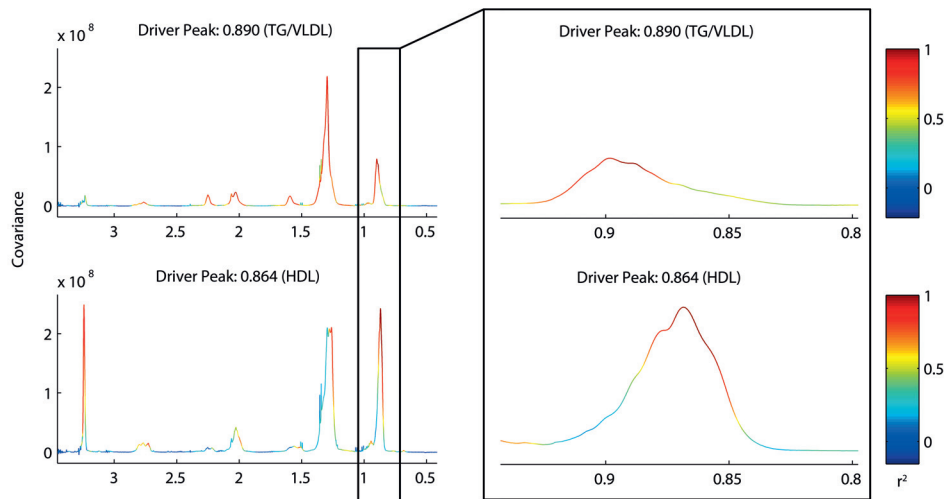
Figure 3.4: Lipoprotein contributions to serum NMR spectra. Statistical total correlation spectroscopy (STOCSY) of CPMG ^{1}H spectra of serum from pregnant women. The height expresses the variable covariation to the driver peak, and the color scale represents the Pearson correlation to the driver peak. Due to particle-size dependent differences in susceptibility, NMR signals from VLDL triglycerides appear at higher ppm values than HDL cholesterol.[131]
Abbreviations: CPMG, Carr-Purcell-Meiboom-Gill pulse sequence; HDL, high density lipoprotein; NMR, nuclear magnetic resonance; STOCSY, statistical total correlation spectroscopy; VLDL, very low density lipoprotein.

Wallis test for three groups. T-tests were used to compare normally distributed groups.

In paper II, metabolite to creatinine ratios were combined with clinical parameters MAP, UtAPI and a dummy variable signifying maternal age (≥ 34 or ≤ 20) in logistic regression models to predict development of preeclampsia.[133] ROC curve analysis was performed to find prediction sensitivity at 90% specificity.

In paper III, linear regression models with log transformed metabolites as dependent variable and preeclampsia and gestational age as independent variables were generated. The interaction term between preeclampsia and gestational age was not included in the models as it was not significant (p$\geq$0.05) for any metabolite $\times$ gestational age interactions.

For all analyses, p values $\leq$ 0.05 were considered significant. Adjustment for multiple testing was performed with the Benjamini-Hochberg false discovery rate correction.[134]

3.5.4 Metabolite set enrichment analyses

Metabolite levels from the placenta HR-MAS spectra were analyzed with Metabo-Analyst online (www.metaboanalyst.ca)[135] by metabolite set enrichment analysis (MSEA) for inferring the main metabolic pathways associated with disease.[136] Quantitative enrichment analysis of metabolites based on metabolic pathway associated metabolite sets was performed. In quantitative enrichment analysis, a generalized linear model is used to create a Q statistic for each metabolite set, describing correlation between a metabolite and the clinical outcomes. The Q statistics are then compared to expected Q values for the metabolite sets. The metabolite sets are curated from the Small Molecular Pathway Database (SMPDB).[137,138]

3.6   Ethical considerations

Prediction of gestational disease carries special ethical issues. Stratifying women who are of higher risk for disease in the current pregnancy, after pregnancy and for subsequent pregnancies, may cause more stress and anxiety than is offset by improved outcomes. As of today there are few or no successful ways of preventing the disease of preeclampsia of occurring. However, in order to find a useful prophylaxis, a reasonable selection of high risk women must be found which may benefit from treatment. Meta-analyses suggest that starting treatment of high-risk women with aspirin before 16 weeks of gestation has preventive effect against preterm or severe preeclampsia, but not term and moderate preeclampsia.[139–141] In order to find successful methods of preventing preeclampsia, research in early pregnancy is necessary.

Early fetal diagnostics are a controversial topic in Norway. An important consideration was whether any information could be found about the health of the baby before it was born. For this reason, no analyses were started before all babies in the project had been born. The studies were approved considering that the research focused on preeclampsia as a maternal and fetal disease, not as fetal diagnostics. However, during data collection in paper II it was impossible to measure the uterine artery pulsatility index (UtAPI) without also observing the baby. The women were advised of this before they consented to take part in the study.

Placental tissue is comprised of fetal cells, but in Norway the placenta is not defined as part of the fetus. Biobanking of placental tissue is therefore not considered a controversial ethical issue, as for example cord blood biobanking, though both procedures involve using cells without consent from their owner. The placental samples are taken at birth with no additional inconvenience to the mother.

All papers in this thesis use and generate much information about relatively few people. An important ethical consideration was to not disseminate enough information that the participants in any way become identifiable. Additionally, study results are published at group level and never returned individually to study participants. All studies in this thesis were approved by the regional ethic committee (REC) and signed consent was obtained for all participants.

4   Summary of papers

4.1   Paper I

**Metabolomic Biomarkers in Serum and Urine in Women
with Preeclampsia**

The aim of paper I was to determine the metabolic profiles in urine and serum of women diagnosed with preeclampsia in comparison to normotensive pregnant and nonpregnant women, for possible clues to the etiology and pathogenesis of the disease.

Serum and urine from women with preeclampsia, healthy pregnant controls and healthy nonpregnant controls were analyzed by NMR metabolomics. Metabolomic analysis showed a significant difference between women with preeclampsia, women with healthy pregnancies, and nonpregnant women in both urine and serum samples. PLS-DA analyses of the healthy and preeclamptic pregnant groups revealed that several metabolites differed in both biofluids. In urine samples, women with preeclampsia had increased levels of choline and dimethylamine, and decreased levels of glycine, p-cresol sulphate, and hippurate. In serum, the levels of histidine were higher in women with preeclampsia, and the loading plots of the PLS-DA analysis showed an increase of total lipids and the VLDL region of the methyl region in particular. The lower levels of urinary glycine were attributed to a possible increased demand for glutathione in antioxidant processes, as glycine is a precursor in the glutathione pathway. Glycine has also been shown to be decreased in mothers with small for gestational age babies. The decreased excretion of the uremic toxin p-cresol sulfate was considered an effect of the renal injury, and retention may contribute to endothelial dysfunction. The loading plot for the urine spectra revealed a grouping of the early onset preeclampsia cases, suggesting a possible subgrouping of the disease in metabolomic analyses. The serum PLS-DA plots of the diffusion-edited spectra showed a lipoprotein signal pattern of increased VLDL and decreased HDL, a proatherogenic pattern which has been related to cardiovascular disease, suggesting a connection in the metabolic profiles of the two diseases.
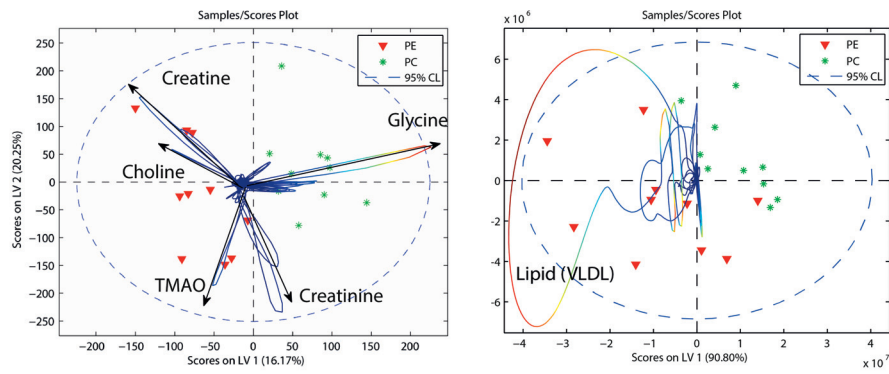
Figure 4.1: Paper I: Two PLS-DA models discriminating urine (left) and serum (right) samples from women with preeclampsia (PE) or healthy pregnancies (PC) from the Pilot Study. The 2D loading plots are superimposed on the score plots, and colored by variable importance in projection score. Abbreviations: CL, confidence limit; TMAO, trimethylamine-N-oxide; VLDL, very low density lipoprotein.

4.2 Paper II

First Trimester Urine and Serum Metabolomics for Prediction of Preeclampsia and Gestational Hypertension: A Prospective Screening Study

The aim of this study was to evaluate the potential of metabolomics to predict preeclampsia and gestational hypertension from urine and serum samples in early pregnancy, and to elucidate the metabolic changes related to the diseases. Current risk stratification methods for preeclampsia and gestational hypertension in first trimester pregnancy have low sensitivity, and new predictive biomarkers are warranted. Conventional "prior risk" methods using maternal characteristics can predict 30-40% of preeclampsia cases at a 10% false positive rate (FPR).

Nuclear magnetic resonance spectra (Standard ^{1}H NOESY (urine) and CPMG (serum)) were acquired on samples from 599 women at medium to high risk of preeclampsia. The data were explored using PCA, and a predictive model created using PLS-DA with variable selection. The models were optimized and evaluated with a double cross validation procedure. Areas of the spectra that were important for discrimination were identified using HSQC, HMBC and TOCSY spectra of selected samples.

A total of 599 women with 587 urine samples and 591 serum samples were included in the analysis. Twenty six women (4.3%) developed preeclampsia, and 21 women (3.5%) developed gestational hypertension.

Urinary metabolic profiles predicted preeclampsia and gestational hypertension at 51% and 40% sensitivity respectively, at 10% FPR. Metabolite differences included higher urinary creatinine and 3-hydroxyisobutyrate, and lower urinary hippurate, and proline betaine in women who later developed preeclampsia. Serum metabolic profiles predicted preeclampsia at 15% sensitivity and gestational hypertension at 33% sensitivity at 90% specificity with increased lipid levels and an atherogenic lipid profile as most important. Serum spectra from the preeclampsia group contained higher levels of triglycerides and VLDL, and lower levels of phosphatidylcholine and HDL.

Early urine changes may result from metabolic disturbances from hypertension, diet and low level renal involvement in the first trimester of pregnancies which later develop preeclampsia. The lipid profile preceding the clinical onset of disease

may contribute to the pathogenesis or be reflective of a cardiovascular risk profile contributing to hypertensive diseases of pregnancy.

Combining MAP, maternal age and UtAPI gave an area under the ROC curve (AUROC) of 0.42 (95% CI 0.64-0.84) for prediction of preeclampsia in a logostic regression model. Combining maternal characteristics (MAP and age) with the urinary hippurate to creatinine level gave an AUROC of 0.78 (95% CI 0.70 - 0.86). Combining the hippurate/creatinine ratio with MAP, age and UtAPI gave an AUROC of 0.81 (95% CI 0.72-0.89).

These results show a potential clinical importance for metabolomic analysis of urine samples to predict preeclampsia. Preeclampsia could be predicted from both urine and serum spectra, but with higher sensitivity from the urine spectra.
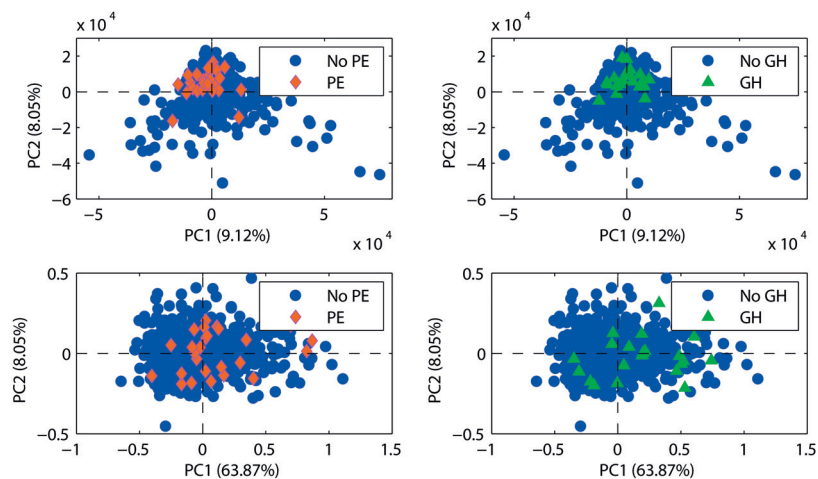
Figure 4.2: Paper II: Results from principal component analyses of NMR spectra of urine (top) and serum (bottom) from first trimester pregnant women in the ScreenTox Study. The first two principal components (PCs) are plotted. Samples from women with pre-eclampsia (PE, pink diamonds) and gestational hypertension (GH, green triangles) cluster together in the urine analysis, but not in the serum analyses. Samples from normotensive pregnancies are shown in blue circles.

4.3 Paper III

Metabolic profiles of placenta in preeclampsia by HR-MAS MRS metabolomics

Our aim in this paper was to phenotype the preeclamptic placenta using High-resolution Magic Angle Spinning Nuclear Magnetic Resonance spectroscopy (HR-MAS MRS). Preeclampsia is a heterogeneous gestational disease initiated by insufficient placental development, but studies characterizing the placental metabolic disease components are lacking.

Placental samples were collected after delivery by CS from women with preeclampsia (n=19) and normotensive controls (n=15). The samples were analyzed for metabolic biomarkers including amino acids, osmolytes, components of energy and phospholipid metabolism. The metabolic biomarkers were correlated to clinical characteristics and maternal serum inflammatory biomarkers. Principal component analysis showed an inherent difference in placental metabolic profile between preeclamptic and normotensive pregnancies. There was an overlap in placental metabolic expression between established subtypes of preeclampsia, such as disease severity and presence of FGR. Using PLS-DA, significant differences in metabolic profiles were found between normotensive and preeclamptic placentas, and between placentas from severe and non-severe preeclampsia. Several metabolic pathways were found to be involved using metabolite set enrichment analysis: taurine, glutamate and phospholipid biosynthesis in particular. Placental metabolites, especially the cell membrane breakdown product glycerophosphocholine, correlated with the placental stress marker sFlt-1 in serum. A different metabolic profile correlated with maternal serum triglycerides, suggesting a variation in placental stress signaling in different placental phenotypes.

Gestational age differences between preeclamptic and normal placentas may be considered part of the disease, as preeclampsia necessitates early delivery. However, differences in metabolites which also vary due to gestational age variation might confound the results. Variation due to gestational age was adjusted for using a linear regression model, and metabolites related to taurine and glutamate metabolism remained significant after this correction.

HR-MAS MRS is in this study shown to be a sensitive method for defining the placental disease component of preeclampsia, identifying several altered metabolic pathways. Placental HR-MAS MRS analysis may provide improved insights into

which processes are disturbed in placental subtypes of preeclampsia, and introduce a new long-required tool for more detailed and sensitive placental phenotyping of this heterogeneous disease.
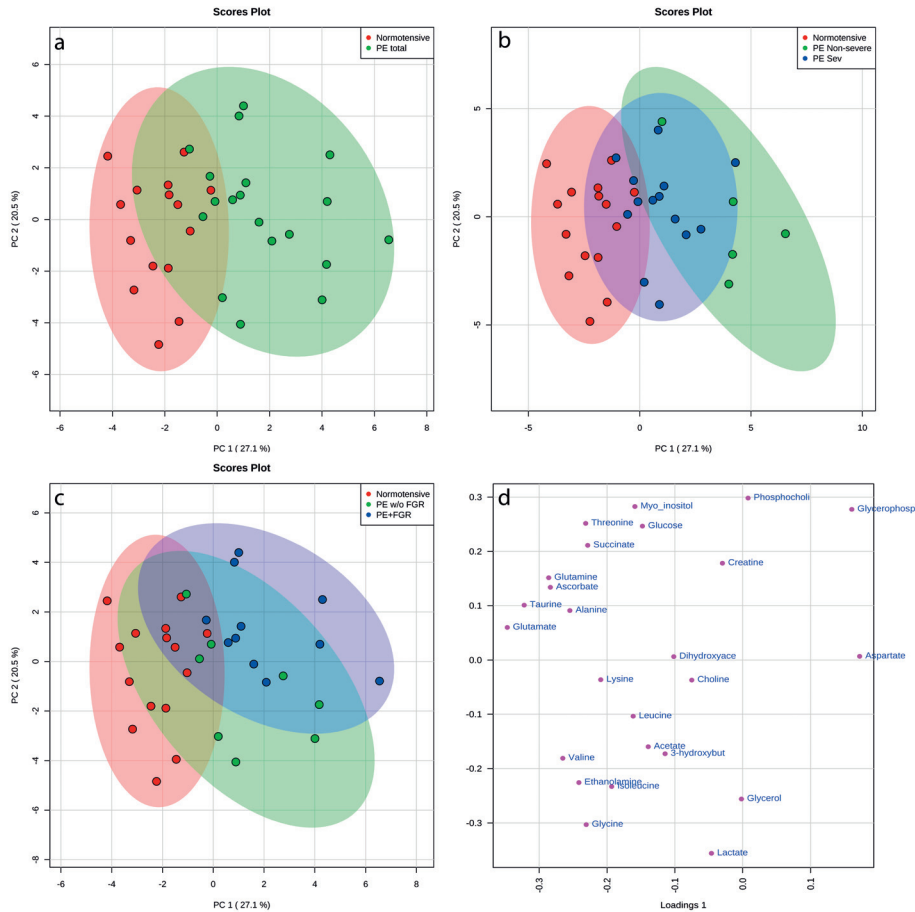
Figure 4.3: Paper III: Principal component analysis of the HR-MAS MRS spectra from paper III. a) Placentas from all preeclamptic pregnancies (PE total, green) and from normotensive pregnancies (Normotensive, red) have clearly different metabolic profiles. b) Placentas from severe preeclampsia (PE Sev, blue) have a closer phenotype to the normotensive group than the non-severe preeclampsia group (PE non-severe, green). c) Placenta with PE and FGR (PE+FGR, blue) and PE without FGR (PE w/o FGR, green) have more overlapping phenotypes. c) The loading plot shows the metabolite contributions to the sample variation.

5   Discussion

The main aim of this thesis was to identify new metabolic biomarkers of pre-eclampsia, in order to characterize the disease and predict its manifestation. The subjects in the thesis ranged from nonpregnant controls, to pregnant women with or without preeclampsia in early and late stages of pregnancy, and at time of delivery.

In paper I and II, maternal biofluids were analyzed with high throughput, automated metabolic profiling methods. Paper I analyzed urine and serum from women with manifest disease compared to normotensive women, and proved a clear preeclampsia-related disturbance in the metabolic profiles. Comparison between healthy pregnancy and nonpregnant women reflected a normal systemic pregnancy related metabolic response. Paper II continued the use of these methods in a larger cohort, aiming to identify metabolic predictors of preeclampsia and gestational hypertension in first trimester pregnancy. The urinary metabolic profile was found to be significantly altered at this early stage of disease, with predictive value for preeclampsia. The serum metabolic profiles had less predictive sensitivity, but showed an early pregnancy evidence of an atherogenic lipid profile in preeclampsia. Paper III assessed the metabolic profile of intact placental tissue from women with preeclamptic and with normotensive pregnancies. Several metabolic pathways were altered in the preeclamptic placenta, highlighting the importance of dysfunctional metabolism in disease related processes, and revealing novel biomarkers of the placental disease.

5.1   Metabolic profiles of preeclampsia

New metabolic biomarkers of disease were found in all three papers presented in the thesis. The results are shortly summarized in Table 5.1, and set in context of the stepwise model for the development of preeclampsia in Figure 5.1.

5.1.1   Redefining biomarkers of preeclampsia

The multimodal nature of preeclampsia is not fully described by its diagnostic hall-marks; hypertension and proteinuria. Therefore sets of biomarkers which would aid in the correct diagnosis or prediction of preeclampsia were sought in this thesis. The multivariate methods applied make use of the entire metabolic profile, which means that metabolites that are not significantly different alone can be included

Table 5.1: Summary of the biomarkers of disease identified in this thesis

Case group	Paper	Sample type	Metabolites ↑	Metabolites ↓
Active PE	I	Urine	Isobutyrate, DMA	Gly, p-CS, hippurate, histidine
	I	Serum	Histidine, TG / VLDL	
	III	Placenta	GPC, aspartate	Taurine, Glt, Gln, ethanolamine, Gly, lysine
Risk of PE	II	Urine	Creatinine, 4-DTA, Gly, DMA	Hippurate, lactate proline betaine
	II	Serum	3-HB, TG	Pyruvate, PtdCho, lactate
Risk of GH	II	Urine	Creatinine, α-HIB, DMA	Hippurate, lactate, proline betaine
	II	Serum	TG	PtdCho, HDL, lactate, N-Ac

Symbols and Abbreviations: ↑, increased expression in preeclampsia; ↓, decreased expression in preeclampsia; α-HIB, α-hydroxyisobutyrate; 3-HB, 3-hydroxybutyrate; 4-DTA, 4-deoxythreonic acid; DMA, Dimethylamine GH, gestational hypertension; Gln, glutamine; Glt, glutamate; GPC; glycerophosphocholine; Gly, glycine; HDL, high-density lipoprotein; N-Ac, N-acetylated residues of glycoproteins; p-CS, p-cresol sulfate; PE, preeclampsia; PtdCho, phosphatidylcholines; TG, triglyceride; VLDL, very low density lipoprotein
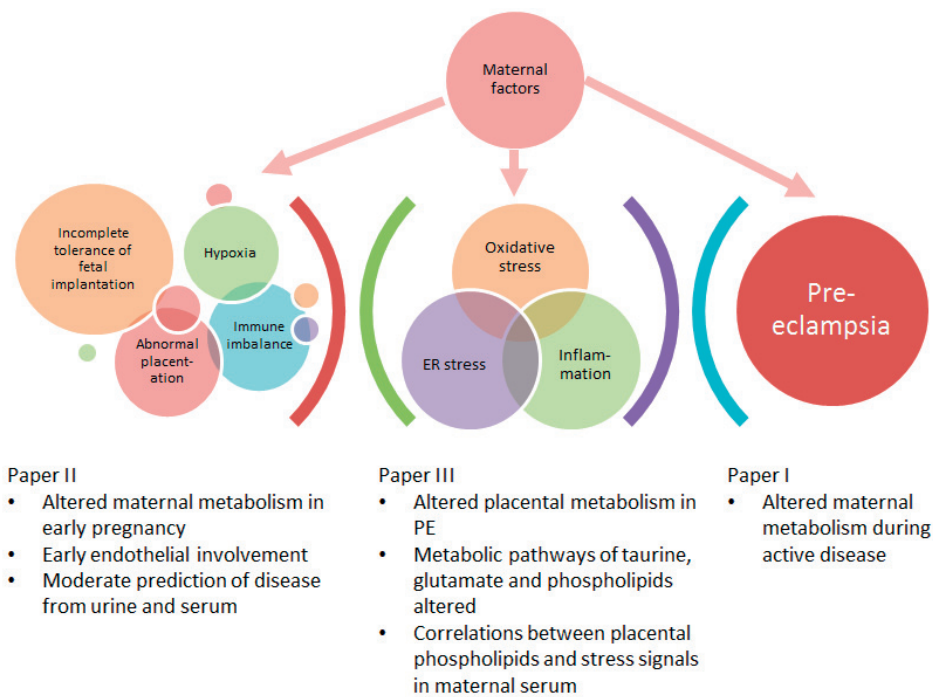
Figure 5.1: Results from the thesis in the context of the staged model of preeclampsia as detailed previously (Figure 1.3, page 6). Early metabolic alterations in maternal biofluids predate preeclampsia and gestational hypertension (paper II) and are highly apparent in active preeclampsia (paper I). Metabolic pathways of phospholipid biosynthesis, taurine metabolism and protein biosynthesis are altered in placentas from women with preeclampsia (Paper III).

in discriminative models. In order to move predictive metabolic biomarkers to a clinical setting it may be necessary to move towards panels of metabolites, measurable with simpler tests.[75] However, the first steps are identifying the possible biomarkers and inferring their role in disease development or pathophysiology. In this thesis, metabolic profiles of active disease were found both in urine and serum, and in placental tissue corresponding to serum markers. Predictive biofluid profiles for preeclampsia were also established.

5.1.2 Maternal metabolism in early and late pregnancy

Maternal urine and serum were profoundly changed in preeclampsia compared to healthy pregnancy (paper I). Some metabolites in active preeclampsia were also reflected in early pregnancy for women later developing the disease (paper II). These were the atherogenic lipid profile in serum and metabolites associated with high blood pressure in urine. In paper III, serum markers of renal function and placental stress were significantly altered for women with preeclampsia, and the serum lipid levels measured by clinical chemistry methods shared the same trend.

Women with active preeclampsia had significantly lower urinary excretion of glycine, p-cresol sulfate and hippurate (paper I). Glycine excretion increased almost fourfold from nonpregnant to the third trimester normal pregnant state, but was reduced to half in preeclamptic pregnancies. Increased glycine excretion in normal pregnancy has been previously described,[142] as well as a relative decrease in preeclamptic pregnancies, FGR and preterm birth.[77,143–145] Glycine as a precursor in glutathione synthesis may be depleted due to maternal oxidative stress.[146,147] Glycine is also involved in the one-carbon metabolism involving tetrahydrofolate, methionine and homocysteine (Figure 5.2).[77] Methionine is required for many biosynthetic reactions, particularly in pregnancy.[146] Dysregulation of key enzymes in the pathway has been suggested in preeclampsia, with increased homocysteine and decreased methionine linked to low birth weight and future CVD.[148–150] Decreased glycine excretion could be due to reduced methylation of homocysteine, also explaining the increased choline excretion. Decreased glycine and increased glycerophosphocholine was also evident in placental biopsies from women with preeclampsia compared to normotensive (paper III), giving additional evidence that the folate-related metabolism may be a pathogenic factor in preeclampsia.[151] Urinary glycine was not amongst the most predictive metabolites for preeclampsia (paper II) and was in fact slightly increased in women who later developed the disease, in-
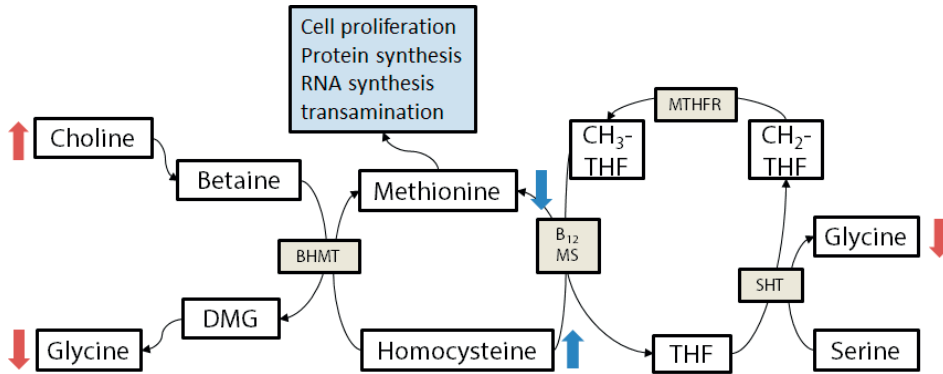
Figure 5.2: Folate-methionine metabolism. Methionine is generated by transfer of a methyl group from tetrahydrofolate (THF) via the enzyme methionine synthase (MS) and the cofactor vitamin B_{12}. The primary source for one-carbon units for THF is the conversion of serine to glycine via serine hydroxymethyltransferase (SHT). An alternative source of methyl is from choline via betaine, producing dimethylglycine (DMG) and glycine.[146,153] Red arrows, findings in paper I and III; Blue arrows, findings in previous research;[148–150] BHMT, betaine homocysteine methyltransferase; MTHFR: methylene tetrahydrofolate reductase

dicating this metabolite as an effect of disease rather than connected to risk profile. Indeed, plasma homocysteine is not elevated before clinical signs of preeclampsia, but increases once the signs develop.[152]

Similarly, the decrease in urinary p-cresol sulfate for women with preeclampsia (paper I) was not reflected in early pregnancy (paper II). This metabolite is a gut microbial byproduct,[154] and a uremic retention solute which has shown proinflammatory effects on endothelium and kidney cells, and may induce shedding of endothelial microparticles.[155] There is a link between p-cresol sulfate retention in patients with chronic kidney disease and later CVD[154,155] suggesting a possible mechanism for CVD risk for women who have been diagnosed with preeclampsia. However, p-cresol sulfate was not identified in the serum spectra of the preeclamptic women (paper I), probably due to low concentrations, and thus any concomitant increase in serum concentration could not be established.

Fewer specific metabolite differences were observed in serum metabolic profiles than urine in active preeclampsia (paper I). However, the lipid profile difference between the normotensive and preeclamptic group in paper I was highly significant and alone discriminated between cases and controls. This lipid profile was also found

in first trimester in women developing preeclampsia (paper II), suggesting that the atherogenic lipid profile may be part of the pathogenesis of preeclampsia. Increase in triglyceride-rich VLDL and decrease in HDL have been associated with increased risk of cardiovascular events[156] and preeclampsia.[157] HDL levels remain low in the years after pregnancy for mild and moderate preeclampsia, but not severe,[158] suggesting that the pro-atherogenic lipid profile persists even after preeclampsia. In paper III, serum levels of lipids in women with preeclampsia measured by clinical chemistry methods showed a trend to decreased HDL levels and increased total cholesterol and triglyceride levels. However, the differences were not significant. A reason for this could be the earlier and more severe phenotype of the women with preeclampsia in paper III, which may be associated more with inflammatory and angiogenic imbalances in the placenta than with maternal constitutional factors, thus, representing the more "placental" side of the disease continuum.[158]

5.1.3 Placental metabolism in active preeclampsia

The placenta is in most cases the causative organ in preeclampsia. Close to term, it is affected by the inadequate placentation originating at its development and by adaptations to continuous stressors. In turn it releases stress signals and causes endothelial dysfunction in the maternal circulation. Metabolic changes are thus expected to be evident of oxidative, ER and inflammatory stress, and to be associated with the maternal disease.

Paper III explored the metabolic effects of preeclampsia in the placenta. The placental metabolic profile in preeclampsia showed a highly significant altered metabolic state. The metabolites were connected to taurine metabolism, glutamate and glutamine metabolism, phospholipid metabolism, and methionine-folate metabolism.

Taurine depletion in preeclamptic placenta and cord blood is a known event in preeclampsia and was confirmed in paper III.[159] Likewise, the decreases of glycine, glutamate and ascorbate may be connected to the state of increased oxidative stress in the placenta of preeclamptic women; glycine and glutamate as precursors to glutathione, and ascorbate (Vitamin C) as an antioxidant.

Metabolites involved in phospholipid metabolism appeared highly disturbed in preeclampsia. Glycerophosphocholine is a cell membrane breakdown product which was posited to be decreased in preeclamptic placenta due to either increased apoptosis or necrosis, or to increased cell membrane catabolism for replenishing the cellular choline pool, or a combination of the two effects (paper III). The latter cause of the

decrease is interesting because it is associated with the methionine-folate cycle, where choline may be used as a methyl donor (Figure 5.2).[160] Folate is a vitamin that is essential for reproductive health, and a well known prenatal supplement.[152] Choline levels were equal in placenta from preeclamptic and normotensive pregnancies (paper III), suggesting a compensatory effect on placental choline levels by increasing cell membrane catabolism. Comparing severe with non-severe preeclampsia, an additional increase in choline was found in tissue from the severe group (paper III). Accumulation of cellular choline in severe PE may be caused by reduced remethylation of homocysteine to methionine as illustrated in Figure 5.2.

The possible involvement of the enzyme phospholipase A_2 has also called to attention in preeclampsia.[161,162] This enzyme releases glycerophosphocholine from cell membrane phosphatidylcholines with production of arachidonic acid,[163] a contributor to immune reactions.[146] Increased activity of PLA_2 might contribute to increased glycerophospholipids in placental tissue, and to increased inflammatory stress as registered by increased serum sFlt-1 (paper III). Non-steroidal anti-inflammatory agents such as aspirin inhibit conversion of arachidonic acid to prostaglandins and thromboxanes which may contribute to inflammation, vasoconstriction and platelet aggregation.[146] This is thought to be the mechanism for low-dose aspirin treatment in prevention of preeclampsia in a subset of cases.[141]

Of special interest was the correlation of placental metabolite profiles and maternal serum markers of preeclampsia. Glycerophosphocholine in particular correlated to increasing maternal serum sFlt-1 levels, which established a link between maternal and placental disease (paper III). If such a link from preeclampsia-specific phenotypes could be established earlier in pregnancy, a more targeted treatment and expectant management could be developed. There was also a significant correlation between maternal serum triglyceride levels and placental metabolites (paper III): increased placental choline, glutamine and glycine and decreased 3-hydroxybutyrate. This is a different placental profile compared to that associated with placental stress (sFlt-1) and of preeclampsia in general. The findings might suggest a separate placental disease in those women who have high triglycerides, associated with an atherogenic lipid profile - such as the women in paper II. In preeclamptic decidua from the Bergen 1 and Trondheim 1 collections, a downregulated gene, ACOX1, has been shown inversely correlated with triglyceride levels in a separate cohort, suggesting a shared genetic risk factor between preeclampsia and cardiovascular disease.[164] These findings are definitely worth exploring further.

No difference was found in NMR-visible metabolites between placentas from preeclamptic pregnancies with or without FGR. Both isolated FGR and preeclampsia are independently associated with placentation disorder, but only in preeclampsia is there a maternal pathophysiological syndrome.[38] There are few studies on the differences between preeclampsia with and without FGR, and further research is required for elucidation of the mechanisms behind fetal growth involvement in preeclampsia.

Paper III highlights the importance of the placenta in the disease, illustrating a wide range of placental metabolic changes. There are no previous publications detailing HR-MAS MRS analyses of placental tissue, although the method has many advantages in sample preparation and the range of metabolites measured. Paper III represents a new method for characterizing the placental component of preeclampsia compared to placentas from healthy pregnancies and in stratifying maternal and fetal outcomes. Different placental phenotypes have been hypothesized to give rise to the same diffuse end stage maternal presentation, hypertension and proteinuria.[51]

Only the end stage of the disease is examined in placental analysis, and causative factors may have been masked by the overreaching consequences of placental hypoxia and inflammation.[165] It might be possible to obtain samples of chorionic villi as is used in detection of aneuploidy at the end of the first trimester.[166] This method would allow research into causative factors. However it suffers from the same problem as other first trimester research; the incidence rate of the disease demands a very large biobank which would also contain a large degree of chromosomal disorders due to the indications for chorionic villi sampling.[165]

5.1.4 Metabolic profiles of preeclampsia and gestational hypertension in early pregnancy

Investigation of the first trimester metabolic profile of women developing preeclampsia marks a change in focus from effects to causes or risk factors for the disease. Significant differences were found preceding preeclampsia and gestational hypertension in urine and serum, and urine samples clearly proved most sensitive (paper II). The altered metabolite profiles found at the end of the first trimester were not directly associated with placental disease, but were more evident of a differing maternal constitution early in pregnancy. Of the 47 women who developed hypertensive diseases of pregnancy, three delivered before term ($\leq$37 weeks gestation) focusing paper II primarily on late onset disease.

A similar metabolic profile in urine and serum preceded both preeclampsia and gestational hypertension (paper II). The findings indicate that a similar risk profile visible to NMR precedes the two hypertensive diseases. The late onset version of preeclampsia is largely associated with maternal factors (metabolic syndrome, hypertension, diabetes[8,167]), and with a less affected placenta.[38] The vascular and inflammatory early placental stress are observed less often in late onset disease.[168] A subset of late onset preeclampsia may be difficult to predict because the pathogenic changes in the placenta have not yet happened at the study visit in the first trimester, rather occurring closer to term with crowding of placental villi causing reduced perfusion.[51] A bimodal skewing of birth weights towards increased SGA and LGA, suggests a dual etiology of late onset preeclampsia, one closer to early onset with signs of underperfusion, and one associated with maternal inability to meet placental demand, causing hypoxia.[169] Of preeclampsia births in paper II, four births (15%) were >90th percentile of weight, and three (12%) were <10th percentile of weight, suggesting that this may also be the case in the ScreenTox cohort, though the numbers are too small for definite conclusions. Contrarily, in paper II the uterine artery flow resistance measured by UtAPI was significantly higher at the 11-13 week study visit for those later developing preeclampsia, but not gestational hypertension (paper II), proving placental involvement in the first trimester for this group. A comparison between metabolic profiles from the preeclampsia and gestational hypertension groups found no significant difference. A conclusion that may be drawn from this is that differences in placental perfusion could not be estimated from early maternal biofluid profiles with NMR, although the maternal constitutional factors increasing the risk of hypertensive diseases could be detected. A cytokine profiling of the ScreenTox cohort demonstrated that the women who later developed gestational hypertension had increased levels of cytokines (interleukin (IL) 1β, IL-5, IL-7, IL-8, IL-13, basic fibroblast growth factor and VEGF) compared to the women who later developed preeclampsia, pointing to early systemic inflammation and stress, accompanied by compensatory mechanisms (Tangerås *et al., submitted*). NMR-visible metabolites are at higher abundance than cytokines in serum, and it may be that differences between the two hypertensive diseases are seen primarily in low-abundance components of the immune system at this early stage.

The decrease of urinary hippurate in women who later developed preeclampsia was attributed to increased MAP in early pregnancy.[71] However, because the dif-

ference remained significant after correction for MAP in a linear regression model, blood pressure alone can not be the cause. Hippurate and proline betaine have been associated with fruit intake,[170,171] and the decrease of these metabolites in urine of women who went on to develop preeclampsia might suggest a healthier diet and lifestyle within the normotensive control group. High fruit intake was included in prediction models for preeclampsia in the Screening for pegnancy end-points (SCOPE) study, indicating an association with either diet or general lifestyle and preeclampsia,[172,173] and adherence to the "New Nordic diet" of Nordic fruits, root vegetables, cabbage etc. was found to be associated with reduced risk of pre-eclampsia in the Norwegian mother-child (MoBa) study.[174] Together this suggests that healthy diet and disturbed placentation are inversely correlated, but the details of this interaction are unknown. Without detailed information about the diet of the study participants, no further associations can be made in our cohort. The interplay between urinary metabolites, blood pressure, diet and gut microbial activity is a complex network, which seems to be affected in pregnant women who go on to develop preeclampsia. Hippurate decrease was also observed in active pre-eclampsia (paper I). It would be interesting to examine the predictive sensitivity of the hippurate/creatinine ratio in mid-gestation or closer to term. Metabolic profiling in mid-gestation, before disease onset, would also enable comparisons to first- and third trimester samples, for inferring which changes precede and follow onset of preeclampsia.

Changes in creatinine excretion in first trimester urine suggest an early, subtle renal involvement, before any serum markers of renal involvement can be observed.[175–177] Both preeclampsia and gestational hypertension had early increases in creatinine excretion (paper II). However only in the preeclampsia group did the renal involvement progress to proteinuria. Increased levels of sFlt-1 appear close to the onset of preeclampsia[178] but not gestational hypertension.[179] This must be consolidated with the increased UtAPI seen in the preeclampsia group and not the gestational hypertension group (paper II). To speculate, both groups have increased renal filtration in early pregnancy due to constitutional factors, but only in preeclampsia is the placenta also compromised, later in pregnancy releasing antiangiogenic factors affecting the susceptible glomerulis.

The serum analyses identified metabolic alterations that were only slightly predictive of preeclampsia, but that reflect an altered lipid composition which may contribute to the exaggerated inflammatory state of the women developing pre-

eclampsia (paper II). We did not adjust for confounding factors; for example, BMI and maternal obesity might influence the lipid levels. However the preeclampsia group was not significantly heavier than the normotensive group. The results reflect previous studies of lipoprotein, cholesterol and triglyceride at early gestation.[180] A result of elevated lipids and VLDL lipoproteins is oxidative stress and endothelial dysfunction, which may increase the risk of preeclampsia.[181,182] Dyslipidemia may be a subpart of a more general metabolic syndrome, including insulin resistance, obesity, hypertension and hyperglycemia, which together predispose to preeclampsia and gestational hypertension. Dyslipidemia precedes pregnancies with preeclampsia or gestational diabetes,[182] indicating that dyslipidemia may contribute to development of preeclampsia - possibly by altering functional characteristics of endothelial cells.[183] However, NMR has limited ability to differentiate lipid and fatty acid species. An interesting approach would be to use lipidomic MS-methods to further assess the variation in lipid components in this cohort.[184]

The connection between preeclampsia and CVD is reflected in all studies in this thesis. Serum profiles reflect an atherogenic lipid profile in both early and late pregnancies where preeclampsia develops (papers I and II). Urine profiles reflect both a maternal metabolic risk profile (paper II) and potential pathogenic processes increasing the risk of later CVD (paper I). In placenta, metabolic profiles reflect both the association to maternal triglycerides and local inflammation, and methionine-folate metabolism; all of which are indicated in metabolic processes leading to CVD. The question of whether preeclampsia is a risk factor or a stress test for future CVD remains unanswered, but it is clear that hypertensive diseases of pregnancy are closely related to cardiovascular diseases. This association also requires further research.

5.1.5 Predicting preeclampsia using metabolic biomarkers

Current Norwegian recommendations for hypertensive disease of pregnancy include advising women at high risk for developing preeclampsia to start a low dose aspirin treatment from 12 weeks gestation.[129] High risk is defined as previous preeclampsia at ≤34-36 weeks gestation. However, primiparas have a fourfold risk of preeclampsia compared to multipara.[4] Thus there is a need for better risk markers for prediction of preeclampsia in nulliparous as well as parous populations.

It has become clear that the early onset preeclampsia has better prediction rates, and benefits more from prophylactic treatment with aspirin.[139,140,168] Ma-

ternal risk factor profiles differ between early onset preeclampsia and late onset preeclampsia,[168] and commonly cited factors (maternal MAP, UtAPI, PlGF and pregnancy associated plasma protein A (PAPP-A)) better predict the early onset disease.[168] Late onset preeclampsia is often referred to as the milder disease, but is in fact responsible for the majority of preeclampsia-related maternal deaths worldwide.[165] Severe gestational hypertension has higher rates of adverse perinatal outcomes than moderate preeclampsia.[185] Although much research has focused on prediction of early onset preeclampsia, there is also a critical need for prediction of late onset disease and gestational hypertension.

In the ScreenTox cohort, the best combination of maternal markers were found to be MAP, UtAPI and maternal age, which predicted 39% of preeclampsia cases at 10% FPR.[120] Gestational hypertension was best predicted by MAP alone, giving a sensitivity of 43%.[120] The metabolomics analyses in urine provided a better prediction rate for preeclampsia at 51% sensitivity (paper II). When metabolite regions were quantified and used in logistic regression models, some of the multivariate power of prediction was lost, including contributions from regions of unresolved metabolite signals, and metabolite covariance. Using only the hippurate decrease relative to creatinine a better prediction of preeclampsia was found compared to using the UtAPI (both in combination with MAP and maternal age). The results suggest that for this cohort, maternal factors measured in urine by NMR were more associated with preeclampsia than reduced placental flow measured by UtAPI.

Preeclampsia is a heterogeneous and multifactorial syndrome, and different risk profiles may necessitate different prevention or treatment approaches.[186] Aspirin treatment targets a prothrombotic risk profile associated with early onset preeclampsia.[186] Women with metabolic or cardiovascular risk profiles might benefit from statin treatment or antihypertensive treatment, options which are currently under investigation in clinical trials.[186, 187] These preventive treatments also require that appropriate risk populations are selected.

Placental metabolite profiles may also enable identification of differing metabolic profiles or subtypes of placental preeclampsia (paper III). For example, identification of a placental subtype in which the methionine-folate metabolism is affected, could enable targeted treatment with high-dose folate supplementation.[186, 188] Especially if the metabolic subtype could be associated with maternal serum markers. This is a focus of our continued work in placental metabolism.

Based on the results from Skråstad *et al.*,[119,120] and from paper II in this thesis, first trimester screening for hypertensive diseases of pregnancy were not sensitive enough for introduction into routine prenatal care. However, compared to the current strategy of risk assessment, these results represent an improvement in identifying women who are at high risk for developing preeclampsia or gestational hypertension. Additionally, the results demonstrate an early change in metabolites in both urine and serum from women who later develop preeclampsia, which may guide future research. Importantly, if a set of biomarkers measurable by simpler methods in urine could effectively predict preeclampsia, the time consuming and certification-requiring UtAPI measurement could be avoided, which would be very helpful in low resource areas.

5.2 Methodological considerations

The validity of research findings are dependent on rigid methodology and well defined sample collection and analytical methods. The aim of this section is to discuss the methodological strengths and limitations of the studies included in this thesis, and how they may have influenced the results.

5.2.1 Hypothesis-generating research

NMR based metabolomics is generally thought to be an untargeted approach wherein all NMR-visible metabolites are measured and included in pattern recognition models. In this way, novel candidate biomarkers and pathways for disease can be identified. Untargeted studies can identify expression patterns which can be targeted in later approaches; generating further studies into a disease phenotype. Metabolites that are found to be differentially expressed between cases and controls are matched to knowledge databases, and in this way generate hypotheses about their function or causative effect.

A limitation to hypothesis generating research is the generation of false positive results. As the number of measurements increase, so does the chance of finding spurious correlations. Univariate comparisons in this thesis were corrected using the Benjamini-Hochberg False discovery rate. Metabolites may interact in combinations and networks that are significantly associated with a disease, even when individual metabolites are not significantly different. In this thesis, this interaction was dealt with by using pattern recognition models PCA, PLS-DA, and MSEA. All these methods include variables that may not be individually different between classes, but together represent a significant pattern.

5.2.2 Diagnostic criteria and phenotypes

Preeclampsia is a syndrome varying in degree of fetal and maternal outcome, for which the diagnostic criteria have been set by consensus, not by an understanding of the underlying pathogenesis. The strict diagnostic limits of blood pressure and proteinuria are not sensitive or specific for maternal or fetal outcomes[189] and may mask the phenotypic diversity of the disease in research settings. There is a current debate on whether the inclusion of proteinuria should be a requirement for diagnosing preeclampsia when evidence of systemic involvement is present.[8] Nev-

ertheless, specific description of diagnostic guidelines are essential for comparing research results.

The criteria for diagnosis were equal in all studies in this thesis, though they are sourced from the Norwegian Association for Obstetrics and Gynecologists' guidelines from 2008 and 2014 for papers I/II and paper III respectively. All case women included in paper III delivered by CS due to complications of preeclampsia, indicating a more severe version of the disease. Therefore the results from this study might not generalize well to milder and later onset forms of preeclampsia.

In all papers, the obstetrician responsible for the patient made the diagnosis at the time the woman was admitted at the hospital. The diagnoses were later reviewed and confirmed from hospital records to be in adherence to the current diagnostic criteria, including multiple measures of blood pressure and proteinuria, fetal birth weight estimations, and prenatal ultrasound measurements. Sufficient information was recorded from the medical history so that if the diagnostic criteria change, it would be possible to reevaluate the diagnoses made in the studies. This was done for paper III, where all diagnoses were reviewed to be in accordance with the 2014 guidelines.

In paper III a stringent SGA definition ($\leq$5th weight percentile) was used as a proxy for FGR in one pregnancy where serial ultrasound measurements were missing. The estimated fetal weights were based on Norwegian population curves and considered gestational age and fetal sex.[126] SGA was therefore considered a good approximation of FGR in this study.

In the studies included in this thesis, the preeclampsia and gestational hypertension cases were compared to controls without signs of preeclampsia, termed "pregnant controls" in paper I, and "normotensive controls" in paper II and III. In case-control studies such as papers I and III, it is important that members of the control or "healthy" group have no related diseases; specifically for preeclampsia, no cardiovascular disease, renal or hypertensive diseases, or inflammatory diseases. For paper I, pregnant controls were specified to have "normal pregnancies", while in paper III, the controls were specified to be "exclusively healthy women with no prior FGR or preeclampsia". The inclusion criteria were more stringent for paper III than for paper I, which may affect the results. For example, two of the nonpregnant control women in paper I had gestational hypertension in a previous pregnancy, which might give a persistant risk profile reflected in metabolic profile. However,

none of the pregnant controls in paper I or paper III had previous preeclampsia or fetal growth retardation.

In paper II a cohort of women were followed from the first trimester of pregnancy until birth. In this paper the aim was to identify biomarkers which may identify women at risk of preeclampsia from information available in the first trimester, thus, the "normotensive controls" in each prediction model include all women who did not develop the diagnosis in question. In studies of predictors, comparing women with preeclampsia with completely healthy women is not appropriate, as this will falsely enhance the predictive power.[189] A predictive test must separate women with preeclampsia from all other outcomes. Inclusion criteria for the cohort were for women with medium to high risk of preeclampsia, which included both nulliparous women and those with a history of preeclampsia or gestational hypertension. The combination of risk groups may have complicated the biomarker search. Recurrent preeclampsia may have different mechanisms than preeclampsia occurring in primiparas, and gestational hypertension also has different mechanisms than preeclampsia. Due to the low numbers of cases, subdivision of the case groups (e.g into nulliparous only, or severe preeclampsia only) was not explored. The strength of this study was the complete follow up and the prospective design allowing evaluation of prediction in a cohort resembling the general pregnant population.

5.2.3 Confounding

A confounder in statistics is a variable or factor which correlates with the variables in the study. Confounders may create false positive relationships between the dependent and the independent variable in the study, or obscure a real relationship, if they are not accounted for. Confounders can be dealt with by adequate matching of cases and controls in a study, or by taking into account known confounders and including them as covariates in a statistical model. Explorative research is especially sensitive to confounders because unknown factors may be different between cases and controls, possibly skewing the analysis results.

Some known or possible confounders in preeclampsia research are obesity, blood pressure, smoking status, sex of the fetus, and ethnicity.[189,190] In metabolomics, additional confounders may be sample storage and handling, medication, and dietary differences. In tissue samples specifically, difference in time passed before quenching of metabolic activity, or total time thawed before analysis may introduce bias.[87,100]

Incidence of preeclampsia increases with BMI and the increased incidence of preeclampsia in recent years may be partly due to increase in obesity rates.[191] In paper I, no pre-pregnancy weight information was recorded for the women with preeclampsia. In paper II, participants were weighed at the study visit and in paper III, maternal weight at the first pregnancy visit, before week 12, was recorded from hospital journals. BMI was not higher in the preeclamptic women in the first trimester in papers II or III, but the women with gestational hypertension in paper II had higher BMI than the normotensive and preeclamptic groups, possibly affecting their metabolic profiles.[192] Smoking status was recorded only in paper II. There were no differences between hypertensive and normotensive groups. However, smoking status in papers I and III could potentially confound the results as this information was not available. Sex of the fetus was recorded for all pregnancies. Fetal sex is suggested to have an impact on placental inflammation, apoptosis and mitochondrial function in preeclampsia.[193] No differences due to fetal sex were observed in any of the studies in cursory discriminative models.

Gestational age differences could impact both maternal and placental metabolic profiles. In paper I, pregnancies were matched by gestational age in order to mitigate gestational age bias between cases and controls. In the cohort of paper II, gestational age at sampling was not different between the groups. In paper III, a lower gestational age was found in the preeclamptic group (average 31 weeks, range 25-39 weeks) than the normotensive group (average 39 weeks, range 38-40 weeks). This is expected due to the need for early delivery in patients with severe and early onset preeclampsia. Metabolite expression in the placenta may be influenced by gestational age, but there are no published reports on this within the third trimester. Difference in gestational age was investigated in paper III by applying a linear regression model. After adjustment for gestational age differences, differences in levels remained significant for half of the metabolites indicating that at least within the third trimester, disease induced differences dominate the placental metabolic profile.

5.2.4 Confounding in metabolomics studies

The hypothesis free nature of metabolomics is ideal to analyze biofluids and tissues for both subtle and widespread metabolic alterations in disease. However, several factors must be taken into account which may specifically bias metabolomic studies.

Metabolic profiles of body fluids vary by diet, stress, drug use, physical activity, and many other variables.[75] Recommended sample collection routines for serum

and urine include collecting samples from fasted participants.[194] However, dietary restrictions can not be applied to pregnancy study participants. Stability of blood lipid analysis and CVD risk to fasting status has been investigated, finding that also postprandial triglyceride levels correspond to risk of CVD.[195] In papers I and II, participants were asked to fast for one hour before their study visit. Nonfasting conditions ($\geq$2h) do not have strong effects on the overall characteristics of study groups in NMR of plasma[196] but the time passed since the last meal, and its contents, could still have an effect on the results.

Metabolites are stable at -80°C,[197] and prepared biofluid samples were stored at 5°C for no longer than 12 hours before NMR analysis, a time period for which metabolic profiles of urine and serum are shown to be stable.[198] Based on cited studies and our own unpublished results, the differences in storage times within studies should have negligible impact on the results. The total time from the start of tissue preparation for HR MAS to spectral acquisition was recorded, and a cursory PLS model found no significant correlation in metabolite spectra to these times. The continuous cold environment during preparation of tissue samples is probably an important factor for this stability.

Preparation of samples was done strictly adhering to procedures provided by Bruker. However, human error is always possible by minute differences in sample handling, buffer addition, or even sample mixups. Reproducibility testing for biofluid analyses showed that differences in sample preparation and instrument stability were smaller than individual differences. However, it is possible that automated preparation would reduce technical variation.

Collection of placental tissue for the PE Biobank was not done with metabolomic profiling in mind, and due to limitations in tissue availability, only one sample per placenta was analyzed. In retrospect, a more systematic analysis of tissue from several sites in the placenta would be preferred, to infer the amount of individual variation between regions of placental tissue. Since all samples were obtained from pregnancies without signs of labor, any effect of labor could not bias the results.[199]

The placental samples were snap frozen as soon as possible after delivery by CS. Time between sample biopsy and freezing of tissue samples is an important consideration in metabolomic studies. Samples should ideally have been snap frozen immediately after delivery to quench any metabolic activity. However, due to patient priority in a CS situation, often at least one hour passed before biopsies were frozen. Delay of sample quenching may cause changes in the metabolic profile,[87]

but there was no difference in collection times between placentas collected from normotensive or preeclamptic pregnancies. Detailed recommendations for collection of placental samples for research were recently published.[190] Adherence to these recommendations, particularly to sample collection times, are vital to high impact placental research in the future.

5.2.5 Choice of analytical method

Both MS and NMR are common analytical choices in metabolomics studies. Although MS has higher analytical sensitivity both in lower concentration ranges and number of visible metabolites, the analysis is destructive to the sample, and requires pretreatment and separation steps, and is more expensive.[75] In contrast, NMR requires only addition of buffer to the sample before analysis. NMR is unbiased, has fast acquisition times, is quantitative and highly reproducible across equipment, and allows safe identification of unknown substances. In addition, NMR has the advantage of nondestructive sampling even of whole tissue samples, and has the possibility of translational value to *in-vivo* spectroscopy.[100,200] A key advantage is high throughput and reproducible analysis. However, the low sensitivity means that low abundance markers of oxidative and inflammatory stress may go unnoticed.[66]

5.2.6 Metabolite quantification

Metabolites were quantified in different ways throughout this work. In paper I a software package (Chenomx) was employed which matches a database of line shapes to the spectra. This method gives accurate quantification in NOESY spectra, avoiding signal overlap and taking into account proton multiplicities. Chenomx quantification was also applied to CPMG spectra in paper I, which is less accurate due to the T_2 filtering of metabolic signals, causing signals from macromolecules and metabolites bound to them to be filtered out. However, as the same spectral acquisition parameters were applied to all spectra, the metabolite concentrations were still comparable. Chenomx quantification has advantages, but it is manually adjusted and thus time consuming and possibly prone to human bias, and finally does not work optimally with baseline distortions by residual macromolecule signals.

In paper II, quantification by integration of metabolite peaks was done only for metabolites important to prediction. Integration of area under signal curve is accurate if there is low signal overlap. The metabolites were normalized to creatinine

in order to correct for dilution effects, and in order to conform to clinical practice and ease comparison to other studies.

In paper III integrated metabolite regions of the area normalized spectra were used as approximations of concentration. Absolute quantification in tissue samples is possible by integration of regions or line fitting metabolite signals, with appropriate normalization for number of protons contributing to the signal, and sample weight. However, for exploratory analysis, area normalized metabolite ratios were considered sufficient.

5.2.7 Multivariate analyses

Multivariate analysis methods are commonly applied metabolomics measurements, to compress the multivariate data into meaningful and useful profiles of diseases or states. Single biomarkers may have limited sensitivity, and combining sets or patterns of biomarkers may increase predictive ability. The advantage of multivariate analysis is the ability to simultaneously model many biomarkers. Multivariate analysis can be applied to concentration data from metabolomic experiments, or to whole spectra without the need for prior quantification. When applied to whole spectra, multivariate analyses may point to important regions of the spectra for which unknown metabolites may be identified *post hoc*. Proper validation is essential to avoid false positive results.

Sample size. Sample size is important for assessing the power of metabolomic studies, but is not straightforward as the expected differences between classes are difficult to predict *a priori*. At least 600 samples has been suggested as necessary for serum metabolomics studies.[69] The studies in this thesis are much smaller than this, representing 10, 26 and 24 preeclampsia cases in papers I, II and III respectively. Small sample sizes increase the chance of overfitting data. In paper I, even with the small sample size, reproducible results could be found compared to similar studies.[77,142,201] In paper II, although there were relatively few hypertensive cases, the large number of normotensive pregnancies established a robust baseline to which the hypertensive groups could be compared. Sample collection is labor intensive and it is difficult to obtain sufficient samples for high powered metabolomic studies. Collecting samples over too long time periods may introduce temporal shifts in the data. In the future tightly controlled multicenter studies as well as an increased culture of data sharing could improve the power of metabolomics studies.

Preprocessing.   Preprocessing of spectra before multivariate analysis must be critically considered, as every method has potential drawbacks as well as advantages.

Spectral alignment is necessary for urine spectra, but the peak shifts themselves may contain information if there is a systematic difference in pH or ion chelation between the considered classes. This is true for tumor samples expressing a pronounced Warburg effect, and may be true for placental samples enduring hypoxia.[109] This information may be lost due to spectral alignment.

Normalization and scaling procedures have different effects on the information content. Normalization is a row- or sample wise procedure affecting all variables in a sample equally, while scaling is a column- or variable wise procedure employed to enhance or reduce contributions from different types of variables.

Normalization procedures have large effects on the results. Area under curve normalization is a fair approximation of sample size or dilution effects in many cases;[100,202] however if disease or medication causes a large signal in the NMR spectra, the total spectral area will be increased and the remaining metabolites levels decreased correspondingly. Area normalization was used on urine spectra in paper I, and on serum and placental samples in papers I, II, and III. In paper II we used a median dilution factor (PQN) on urine spectra, which is more robust to individual metabolite fluctuations. Creatinine normalization in urine is common in clinical chemistry, but is not ideal because metabolic processes and individual factors may alter creatinine excretion.[110,111,203] In paper I, creatinine normalization was used in the quantified urinary metabolite levels, and was not significantly different between preeclamptic and normal pregnant women. In paper II, creatinine excretion was found to be increased in preeclampsia, and although it was not significant alone (p=0.09), it contributed to the multivariate prediction model and also to the significance of the hippurate / creatinine marker.

Pareto scaling of urine spectra in paper I and II gave the best multivariate models, suppressing nonsignal regions of the spectra and causing the multivariate models to make use of metabolite signals rather than noise regions. There was still a tendency for high concentration metabolites to dominate the model; e.g., creatinine, hippurate and glycine. Other scaling procedures (variable stability scaling, range scaling)[113,204] were explored on the data and though they might increase the participation from lower abundance metabolites, the classification was less effective. In paper III integrated metabolite areas were used directly in multivariate models. Autoscaling could then be used since there were nonsignal variables to suppress.

This method gave equally good discrimination models to using the entire spectra (data not shown), but had the advantage of easier interpretation and the ability to assess each metabolite contribution alone to the model.

Model optimization and validation The optimal number of LVs adequately captures the variation without overfitting the data. Due to low numbers of samples, the number of LVs in papers I and III were found by leave-one-out and leave-20%-out cross validation. Although usually quite robust, this method may overfit the data because it uses the same samples to optimize the model and then "predict" samples based on that same model. In paper II, a double cross validation procedure was used which enables optimization of model parameters on a subset of data, before the final model is tested on a held out validation set. In this way, model accuracy and sensitivity is found on a completely independent set. Permutation testing performs an additional validation to the cross validation and was performed for all models.

Variable Selection. In paper II, variable selection was applied to find subsets of metabolites predicting preeclampsia, and to reduce contribution from noise regions. Both VIP and CARS variable selection found reliable subsets of metabolites which increased the prediction sensitivity of the PLS-DA models. However, both were applied on the full datasets. VIP and CARS employ cross validation within their variable selection algorithms, however there is a danger of information leakage when the chosen subsets are employed for prediction on the same dataset. For this reason, variable selection introduces an optimistic bias. In retrospect, the variable selection could have been implemented in the double cross validation procedure to ameliorate these problems. However, with a limited amount of case samples available, cross-validation optimized variable selection would utilize few initial samples for variable selection. The risk of overfitting was deemed acceptable in this explorative study.

The bias in variable selection was partly corrected for by incorporating a variable selection step into the final permutation testing of the models. This ensures the validity of the models even with variable selection, though the specificity, sensitivity and error rate may still be overstated.

6 Concluding remarks and future perspectives

In this thesis, MR metabolic profiling of biofluids and tissues from preeclamptic pregnancies were used to assess the metabolic profile of maternal and placental disease, and to predict the onset of preeclampsia and gestational hypertension.

Specific metabolic markers of preeclampsia were found in urine of women with active preeclampsia, suggesting both depletions of required nutrients and retention of toxic uremic solutes, which may contribute to the inflammatory state. Lipid levels were significantly higher in the preeclamptic group, and a proatherogenic shift towards higher VLDL levels and lower HDL levels was observed.

Metabolite profiles consisting of urinary markers of renal function, blood pressure, gut microbiota and diet were found predictive of preeclampsia in early pregnancy. Serum metabolic profiles were also significantly altered in early pregnancy, showing evidence of atherogenic lipid changes. The metabolites are associated with maternal risk factors and metabolic state at the start of pregnancy, leading to higher risk of developing the disease.

The placental component of preeclampsia was examined by metabolic profiling using a novel method in placental research. The findings indicated that several metabolic pathways are altered in the preeclamptic placenta, and that specific markers of cell membrane catabolism correlated to a maternal serum marker of placental stress. Some evidence of involvement of the methionine-homocysteine pathways was also found in the preeclamptic placentas. An interesting direction of future research could be the clustering of preeclamptic placental metabolite profiles into groups of differentially expressed phenotypes, for elucidation of placental subtypes of preeclampsia. A larger study where metabolic clustering is the main aim is currently in the planning stages.

Together, the results highlight the metabolic alterations that appear before and during the clinical symptom stages of preeclampsia, and the differences and similarities of gestational hypertension and preeclampsia. Both the placental and maternal component of the disease were examined and found significantly affected. The current diagnosis of preeclampsia relies on criteria set without full understanding of the disease. The results in this thesis may aid in better stratification of women with preeclampsia based on biomarkers reflecting the metabolic state, a closer measure of the phenotype. Connecting maternal serum markers to placental metabolic profiles may aid in the search for treatment targets and in differential management of women

with preeclampsia. A combination of metabolic biomarkers and maternal clinical characteristics provided a prediction higher sensitivity for prediction than uterine artery flow resistance measurement, a result which holds promise for prediction of preeclampsia in low resource areas.

The work presented in this thesis provides a foundation for continued studies on the metabolic alterations in preeclampsia. The predictive biomarkers need validation in separate cohorts, and should also be further explored in second trimester urine samples to see which metabolic features are common to active disease or to earlier pregnancy timepoints. The metabolic profiling of placenta opens a possibility for subgrouping the placental component of preeclampsia based on metabolic subgroups, a subphenotyping which would closer represent the cause of the disease rather than its effect. Phenotypic subgroups could be found which may require different treatment options or correspond to separate maternal and fetal outcome groups. Combining NMR with MS-based metabolomics could improve identification of lipid classes, which could be useful to identify possibly inflammatory lipid subclasses.

Applications of systems biology methods have begun to unravel the complex pathways resulting in preeclampsia. Still a great deal of work remains before the full picture is revealed. Genetic variation, gene transcription, proteomic function, and metabolic interactions contribute in complex pathways resulting in preeclampsia. Combinations of the 'Omics' platforms in systems biology may provide a more complete understanding of the disease. MR metabolomics can assist in identification of biomarkers of preeclamptic disease: maternal active and predictive biomarkers, and phenotyping of the preeclamptic placenta. This thesis has presented novel investigations that have significantly contributed to the current body of knowledge regarding preeclampsia, and outlined potential directions for future research.

References

[1] M. J. Bell, "A historical overview of preeclampsia-eclampsia," *J Obstet Gynecol Neonatal Nurs*, vol. 39, no. 5, pp. 510–8, 2010.

[2] B. Sibai, G. Dekker, and M. Kupferminc, "Pre-eclampsia," *Lancet*, vol. 365, no. 9461, pp. 785 – 799, 2005.

[3] B. M. Sibai, "Diagnosis and management of gestational hypertension and preeclampsia," *Obstet Gynecol*, vol. 102, no. 1, pp. 181–92, 2003.

[4] K. Klungsoyr, N. H. Morken, L. Irgens, S. E. Vollset, and R. Skjaerven, "Secular trends in the epidemiology of pre-eclampsia throughout 40 years in Norway: prevalence, risk factors and perinatal survival," *Paediatr Perinat Epidemiol*, vol. 26, no. 3, pp. 190–8, 2012.

[5] K. S. Khan, D. Wojdyla, L. Say, A. M. Gulmezoglu, and P. F. Van Look, "WHO analysis of causes of maternal death: a systematic review," *Lancet*, vol. 367, no. 9516, pp. 1066–74, 2006.

[6] S. Vangen, L. Ellingsen, A. B. Andersgaard, A. F. Jacobsen, B. Lorentzen, L. T. Nyfløt, A. B. Rygh, S. M. Skulstad, C. Tappert, and P. Øian, "Mødredødsfall i Norge 2005–09," *Tidsskrift for Den norske legeforening*, vol. 134, no. 8, pp. 836–9, 2014.

[7] C. W. Redman and I. L. Sargent, "Latest advances in understanding preeclampsia," *Science*, vol. 308, no. 5728, pp. 1592–1594, 2005.

[8] T. Chaiworapongsa, P. Chaemsaithong, L. Yeo, and R. Romero, "Pre-eclampsia part 1: current understanding of its pathophysiology," *Nat Rev Nephrol*, vol. 10, no. 8, pp. 466–80, 2014.

[9] L. Duley, "The global impact of pre-eclampsia and eclampsia," *Semin Perinatol*, vol. 33, no. 3, pp. 130–7, 2009.

[10] L. Trogstad, P. Magnus, and C. Stoltenberg, "Pre-eclampsia: Risk factors and causal models," *Best Pract Res Cl Ob*, pp. 1–14, 2011.

[11] M. P. Johnson, S. P. Brennecke, C. E. East, T. D. Dyer, L. T. Roten, J. M. Proffitt, P. E. Melton, M. H. Fenstad, T. Aalto-Viljakainen, K. Mäkikallio, S. Heinonen, E. Kajantie, J. Kere, H. Laivuori (for the FINNPEC Study Group), R. Austgulen, J. Blangero, and E. K. Moses, "Genetic dissection of the pre-eclampsia susceptibility locus on chromosome 2q22 reveals shared novel risk factors for cardiovascular disease," *Mol Hum Reprod*, vol. 19, no. 7, pp. 423–437, 2013.

[12] J. Roberts and D. Cooper, "Pathogenesis and genetics of pre-eclampsia," *Lancet*, vol. 357, no. 9249, pp. 53 – 56, 2001.

[13] R. Skjærven, L. J. Vatten, A. J. Wilcox, T. Rønning, L. M. Irgens, and R. T. Lie, "Recurrence of pre-eclampsia across generations: exploring fetal and maternal genetic components in a population based cohort," *Brit Med J*, vol. 331, no. 877, 2005.

[14] I. Craici, S. Wagner, and V. D. Garovic, "Preeclampsia and future cardiovascular risk: formal risk factor or failed stress test?," *Ther Adv Cardiovasc Dis*, no. 4, pp. 249–259, 2008.

[15] P. R. Romundstad, E. B. Magnussen, G. D. Smith, and L. J. Vatten, "Hypertension in pregnancy and later cardiovascular risk: common antecedents?," *Circulation*, vol. 122, no. 6, pp. 579–84, 2010.

[16] M. C. Brown, K. E. Best, M. S. Pearce, J. Waugh, S. C. Robson, and R. Bell, "Cardiovascular disease risk in women with pre-eclampsia: systematic review and meta-analysis," *Eur J Epidemiol*, vol. 28, no. 1, pp. 1–19, 2013.

[17] H. U. Irgens, L. Reisaeter, L. M. Irgens, and R. T. Lie, "Long term mortality of mothers and fathers after pre-eclampsia: population based cohort study," *BMJ (Clinical research ed.)*, vol. 323, no. 7323, pp. 1213–7, 2001.

[18] A. C. Staff, R. Dechend, and R. Pijnenborg, "Learning from the placenta: acute atherosis and vascular remodeling in preeclampsia-novel aspects for atherosclerosis and future cardiovascular health," *Hypertension*, vol. 56, no. 6, pp. 1026–34, 2010.

[19] J. Gardosi, A. Chang, B. Kalyan, D. Sahota, and E. M. Symonds, "Customised antenatal growth charts," *Lancet*, vol. 339, no. 8788, pp. 283–7, 1992.

[20] L. J. Vatten and R. Skjaerven, "Is pre-eclampsia more than one disease?," *Br J Obstet Gynecol*, vol. 111, no. 4, pp. 298–302, 2004.

[21] P. von Dadelszen, L. A. Magee, and J. M. Roberts, "Subclassification of preeclampsia," *Hypertens Pregnancy*, vol. 22, no. 2, pp. 143–8, 2003.

[22] F. Mattar and B. M. Sibai, "Eclampsia. VIII. risk factors for maternal morbidity," *Am J Obstet Gynecol*, vol. 182, no. 2, pp. 307–12, 2000.

[23] G. Eastabrook, M. Brown, and I. Sargent, "The origins and end-organ consequences of pre-eclampsia," *Best Pract Res Cl Ob*, pp. 1–13, 2011.

[24] J. M. Roberts and C. Escudero, "The placenta in preeclampsia," *Pregnancy Hypertens*, vol. 2, no. 2, pp. 72–83, 2012.

[25] A. Moffett and C. Loke, "Immunology of placentation in eutherian mammals," *Nat Rev Immunol*, vol. 6, pp. 584–594, 2006.

[26] R. Pijnenborg, L. Vercruysse, and M. Hanssens, "The uterine spiral arteries in human pregnancy: facts and controversies," *Placenta*, vol. 27, no. 9-10, pp. 939–58, 2006.

[27] G. E. Morgan, M. S. Mikhail, and M. J. Murray, *Clinical Anesthesiology*. McGraw-Hill Companies, 4th ed., 2006.

[28] R. Pijnenborg, J. Bland, W. Robertson, G. Dixon, and I. Brosens, "The pattern of interstitial trophoblastic invasion of the myometrium in early human pregnancy," *Placenta*, vol. 2, no. 4, pp. 303–316, 1981.

[29] A. Trundley and A. Moffett, "Human uterine leukocytes and pregnancy," *Tissue Antigens*, vol. 63, no. 1, pp. 1–12, 2004.

[30] A. Moffet-King, "Natural killer cells and pregnancy," *Nat Rev Immunol*, vol. 2, pp. 656–663, 2002.

[31] G. J. Burton, E. Jauniaux, and D. S. Charnock-Jones, "The influence of the intrauterine environment on human placental development," *Int J Dev Biol*, vol. 54, no. 2-3, pp. 303–12, 2010.

[32] G. J. Burton, A. L. Watson, J. Hempstock, J. N. Skepper, and E. Jauniaux, "Uterine glands provide histiotrophic nutrition for the human fetus during the first trimester of pregnancy," *J Clin Endocrinol Metab*, vol. 87, no. 6, pp. 2954–9, 2002.

[33] G. J. Burton, A. W. Woods, E. Jauniaux, and J. C. Kingdom, "Rheological and physiological consequences of conversion of the maternal spiral arteries for uteroplacental blood flow during human pregnancy," *Placenta*, vol. 30, no. 6, pp. 473–82, 2009.

[34] A. M. Borzychowski, I. L. Sargent, and C. W. G. Redman, "Inflammation and pre-eclampsia," *Semin Fetal Neonat M*, vol. 11, pp. 309–316, 2006.

[35] J. M. Roberts, R. N. Taylor, T. J. Musci, G. M. Rodgers, C. A. Hubel, and M. K. McLaughlin, "Preeclampsia: An endothelial cell disorder," *Am J Obst Gynecol*, vol. 161, no. 5, pp. 1200–1204, 1989.

[36] C. W. G. Redman, G. P. Sacks, and I. L. Sargent, "Preeclampsia: An excessive maternal inflammatory response to pregnancy," *Am J Obstet Gynecol*, vol. 180, pp. 499–506, 1999.

[37] C. W. Redman and I. L. Sargent, "Immunology of pre-eclampsia," *Am J Reprod Immunol*, vol. 63, no. 6, pp. 534–43, 2010.

[38] J. M. Roberts and C. A. Hubel, "The two stage model of preeclampsia: variations on the theme," *Placenta*, vol. 30 Suppl A, pp. S32–7, 2009.

[39] I. A. Brosens, W. B. Robertson, and H. G. Dixon, "The role of the spiral arteries in the pathogenesis of preeclampsia," *Obstet Gynecol Annu*, vol. 1, pp. 177–91, 1972.

[40] I. Caniggia, H. Mostachfi, J. Winter, M. Gassmann, S. J. Lye, M. Kuliszewski, and M. Post, "Hypoxia-inducible factor-1 mediates the biological effects of oxygen on human trophoblast differentiation through tgfbeta(3)," *J Clin Invest*, vol. 105, no. 5, pp. 577–87, 2000.

[41] J. J. Brosens, M. G. Parker, A. McIndoe, R. Pijnenborg, and I. A. Brosens, "A role for menstruation in preconditioning the uterus for successful pregnancy," *Am J Obstet Gynecol*, vol. 200, no. 6, pp. 615 e1–6, 2009.

[42] J. H. Veerbeek, P. G. Nikkels, H. L. Torrance, J. Gravesteijn, E. D. Post Uiterweer, J. B. Derks, S. V. Koenen, G. H. Visser, B. B. Van Rijn, and A. Franx, "Placental pathology in early intrauterine growth restriction associated with maternal hypertension," *Placenta*, vol. 35, no. 9, pp. 696–701, 2014.

[43] M. Kovo, L. Schreiber, and J. Bar, "Placental vascular pathology as a mechanism of disease in pregnancy complications," *Thromb Res*, vol. 131 Suppl 1, pp. S18–21, 2013.

[44] R. J. Levine, S. E. Maynard, C. Qian, K.-H. Lim, L. J. England, K. F. Yu, E. F. Schisterman, R. Thadhani, B. P. Sachs, F. H. Epstein, B. M. Sibai, V. P. Sukhatme, and S. A. Karumanchi, "Circulating angiogenic factors and the risk of preeclampsia," *N Eng J Med*, vol. 350, no. 7, pp. 672–683, 2004.

[45] E. Shibata, A. Rajakumar, R. W. Powers, R. W. Larkin, C. Gilmour, L. M. Bodnar, W. R. Crombleholme, R. B. Ness, J. M. Roberts, and C. A. Hubel, "Soluble fms-like tyrosine kinase 1 is increased in preeclampsia but not in normotensive pregnancies with small-for-gestational-age neonates: relationship to circulating placental growth factor," *J Clin Endocrinol Metab*, vol. 90, no. 8, pp. 4895–903, 2005.

[46] B. Huppertz, J. Kingdom, I. Caniggia, G. Desoye, S. Black, H. Korr, and P. Kaufmann, "Hypoxia favours necrotic versus apoptotic shedding of placental syncytiotrophoblast into the maternal circulation," *Placenta*, vol. 24, no. 2-3, pp. 181–90, 2003.

[47] A. E. Heazell, A. N. Sharp, P. N. Baker, and I. P. Crocker, "Intra-uterine growth restriction is associated with increased apoptosis and altered expression of proteins in the p53 pathway in villous trophoblast," *Apoptosis*, vol. 16, no. 2, pp. 135–44, 2011.

[48] M. S. Alberry, D. G. Maddocks, M. A. Hadi, H. Metawi, L. P. Hunt, S. A. Abdel-Fattah, N. D. Avent, and P. W. Soothill, "Quantification of cell free fetal DNA in maternal plasma in normal pregnancies and in pregnancies with placental dysfunction," *Am J Obstet Gynecol*, vol. 200, no. 1, pp. 98 e1–6, 2009.

[49] I. A. Lian, M. Loset, S. B. Mundal, M. H. Fenstad, M. P. Johnson, I. P. Eide, L. Bjorge, K. A. Freed, E. K. Moses, and R. Austgulen, "Increased endoplasmic reticulum stress in decidual tissue from pregnancies complicated by fetal growth restriction with and without pre-eclampsia," *Placenta*, vol. 32, no. 11, pp. 823–9, 2011.

[50] G. J. Burton and E. Jauniaux, "Oxidative stress," *Best Pract Res Clin Obstet Gynaecol*, vol. 25, no. 3, pp. 287–99, 2011.

[51] C. W. Redman, I. L. Sargent, and A. C. Staff, "IFPA senior award lecture: making sense of pre-eclampsia - two placental causes of preeclampsia?," *Placenta*, vol. 35 Suppl, pp. S20–5, 2014.

[52] B. Isomaa, "A major health hazard: The metabolic syndrome," *Life Sciences*, vol. 73, no. 19, pp. 2395–2411, 2003.

[53] A. Basaran, "Pregnancy-induced hyperlipoproteinemia: Review of the literature," *Reprod Sci*, vol. 16, no. 5, pp. 431–437, 2009.

[54] F. M. von Versen-Hoeynck and R. W. Powers, "Maternal-fetal metabolism in normal pregnancy and preeclampsia," *Front Biosci*, vol. 12, pp. 2457–70, 2007.

[55] E. Herrera, "Lipid metabolism in pregnancy," *Endocrine*, vol. 19, pp. 43–55, October 2002.

[56] N. Sattar, A. Bendomir, C. Berry, J. Shepherd, I. A. Greer, and C. J. Packard, "Lipoprotein subfraction concentrations in preeclampsia: Pathogenic parallels to atherosclerosis," *Obstet Gynecol*, vol. 89, no. 3, pp. 403 – 408, 1997.

[57] A. M. Baker, R. L. Klein, K. L. Moss, S. Haeri, and K. Boggess, "Maternal serum dyslipidemia occurs early in pregnancy in women with mild but not severe preeclampsia," *Am J Obstet Gynecol*, vol. 201, no. 3, pp. 293.e1–293.e4, 2009.

[58] R. Akolekar, A. Syngelaki, L. Poon, D. Wright, and K. H. Nicolaides, "Competing risks model in early screening for preeclampsia by biophysical and biochemical markers," *Fetal Diagn Ther*, vol. 33, no. 1, pp. 8–15, 2013.

[59] V. Shenoy, K. Kanasaki, and R. Kalluri, "Pre-eclampsia: connecting angiogenic and metabolic pathways," *Trends Endocrinol Metab*, vol. 21, no. 9, pp. 529–36, 2010.

[60] S. Visser, W. Hermes, J. C. Ket, R. H. Otten, M. G. van Pampus, K. W. Bloemenkamp, A. Franx, B. W. Mol, and C. J. de Groot, "Systematic review and metaanalysis on nonclassic cardiovascular biomarkers after hypertensive pregnancy disorders," *Am J Obstet Gynecol*, vol. 211, no. 4, pp. 373 e1–9, 2014.

[61] S. A. Karumanchi, S. E. Maynard, I. E. Stillman, F. H. Epstein, and V. P. Sukhatme, "Pre-eclampsia: A renal perspective," *Kidney Int*, vol. 67, no. 6, pp. 2101–2113, 2005.

[62] T. N. Jeffcoate, "Pre-eclampsia and eclampsia: the disease of theories," *Proc R Soc Med*, vol. 59, no. 5, pp. 397–404, 1966.

[63] M. Mohaupt, "Molecular aspects of preeclampsia," *Mol Aspects Med*, vol. 28, no. 2, pp. 169 – 191, 2007.

[64] J. C. Lindon, J. K. Nicholson, E. Holmes, and J. R. Everett, "Metabonomics: Metabolic processes studied by NMR spectroscopy of biofluids," *Concept Magnetic Res*, vol. 12, no. 5, pp. 289–320, 2000.

[65] H. Pearson, "Meet the human metabolome," *Nature*, vol. 446, no. 7131, pp. 8–8, 2007.

[66] S. Bouatra, F. Aziat, R. Mandal, A. C. Guo, M. R. Wilson, C. Knox, T. C. Bjorndahl, R. Krishnamurthy, F. Saleem, P. Liu, Z. T. Dame, J. Poelzer, J. Huynh, F. S. Yallou, N. Psychogios, E. Dong, R. Bogumil, C. Roehring, and D. S. Wishart, "The human urine metabolome," *PLoS ONE*, vol. 8, p. e73076, 09 2013.

[67] D. S. Wishart, T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, E. Dong, S. Bouatra, I. Sinelnikov, D. Arndt, J. Xia, P. Liu, F. Yallou, T. Bjorndahl, R. Perez-Pineiro, R. Eisner, F. Allen, V. Neveu, R. Greiner, and A. Scalbert, "HMDB 3.0 - the human metabolome database in 2013," *Nucl Acids Res*, vol. 41, no. D1, pp. D801–D807, 2013.

[68] Q. Cui, I. A. Lewis, A. D. Hegeman, M. E. Anderson, J. Li, C. F. Schulte, W. M. Westler, H. R. Eghbalnia, M. R. Sussman, and J. L. Markley, "Metabolite identification via the Madison Metabolomics Consortium Database," *Nat Biotech*, vol. 26, no. 2, pp. 162–164, 2008.

[69] W. B. Dunn, W. Lin, D. Broadhurst, P. Begley, M. Brown, E. Zelena, A. A. Vaughan, A. Halsall, N. Harding, J. D. Knowles, S. Francis-McIntyre, A. Tseng, D. I. Ellis, S. O'Hagan, G. Aarons, B. Benjamin, S. Chew-Graham, C. Moseley, P. Potter, C. L. Winder, C. Potts, P. Thornton, C. McWhirter, M. Zubair, M. Pan, A. Burns, J. K. Cruickshank, G. C. Jayson, N. Purandare, F. C. Wu, J. D. Finn, J. N. Haselden, A. W. Nicholls, I. D. Wilson, R. Goodacre, and D. B. Kell, "Molecular phenotyping of a UK population: defining the human serum metabolome," *Metabolomics*, vol. 11, pp. 9–26, 2015.

[70] C. Stella, B. Beckwith-Hall, O. Cloarec, E. Holmes, J. C. Lindon, J. Powell, F. van der Ouderaa, S. Bingham, A. J. Cross, and J. K. Nicholson, "Susceptibility of human metabolic phenotypes to dietary modulation," *J Proteome Res*, vol. 5, no. 10, pp. 2780–8, 2006.

[71] E. Holmes, R. L. Loo, J. Stamler, M. Bictash, I. K. Yap, Q. Chan, T. Ebbels, M. De Iorio, I. J. Brown, K. A. Veselkov, M. L. Daviglus, H. Kesteloot, H. Ueshima, L. Zhao, J. K. Nicholson, and P. Elliott, "Human metabolic phenotype diversity and its association with diet and blood pressure," *Nature*, vol. 453, no. 7193, pp. 396–400, 2008.

[72] J. T. Bjerrum, Y. Wang, F. Hao, M. Coskun, C. Ludwig, U. Gunther, and O. H. Nielsen, "Metabonomics of human fecal extracts characterize ulcerative colitis, crohn's disease and healthy individuals," *Metabolomics*, vol. 11, pp. 122–133, 2015.

[73] K. Sulek, T. L. Han, S. G. Villas-Boas, D. S. Wishart, S. E. Soh, K. Kwek, P. D. Gluckman, Y. S. Chong, L. C. Kenny, and P. N. Baker, "Hair metabolomics: identification of fetal compromise provides proof of concept for biomarker discovery," *Theranostics*, vol. 4, no. 9, pp. 953–9, 2014.

[74] M. B. Tessem, T. F. Bathen, S. Lofgren, O. Saether, V. Mody, L. Meyer, X. Dong, P. G. Soderberg, and A. Midelfart, "Biological response in various compartments of the rat lens after in vivo exposure to UVR-B analyzed by HR-MAS ^{1}H NMR spectroscopy," *Invest Ophthalmol Vis Sci*, vol. 47, no. 12, pp. 5404–11, 2006.

[75] P. Soininen, A. J. Kangas, P. Würtz, T. Suna, and M. Ala-Korpela, "Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics," *Circ Cardiovasc Genet*, vol. 8, no. 1, pp. 192–206, 2015.

[76] I. K. S. Yap, I. J. Brown, Q. Chan, A. Wijeyesekera, I. Garcia-Perez, M. Bictash, R. L. Loo, M. Chadeau-Hyam, T. Ebbels, M. De Iori, E. Maibaum, L. Zhao, H. Kesteloot, M. L. Daviglus, J. Stamler, J. K. Nicholson, P. Elliot, and E. Holmes, "Metabolome-wide association study identifies multiple biomarkers that discriminate North and South Chinese populations at differing risks of cardiovascular disease: INTERMAP study," *J Proteome Res*, vol. 9, pp. 6647–6654, 2010.

[77] S. O. Diaz, A. S. Barros, B. J. Goodfellow, I. F. Duarte, I. M. Carreira, E. Galhano, C. Pita, C. Almeida Mdo, and A. M. Gil, "Following healthy pregnancy by nuclear magnetic resonance (NMR) metabolic profiling of human urine," *J Proteome Res*, vol. 12, no. 2, pp. 969–79, 2013.

[78] N. Psychogios, D. D. Hau, J. Peng, A. C. Guo, R. Mandal, S. Bouatra, I. Sinelnikov, R. Krishnamurthy, R. Eisner, B. Gautam, N. Young, J. Xia, C. Knox, E. Dong, P. Huang, Z. Hollander, T. L. Pedersen, S. R. Smith, F. Bamforth, R. Greiner, B. McManus, J. W. Newman, T. Goodfriend, and D. S. Wishart, "The Human Serum Metabolome," *PloS One*, vol. 6, pp. 1–18, 2011.

[79] M. Ala-Korpela, "Critical evaluation of 1H NMR metabonomics of serum as a methodology for disease risk assessment and diagnostics," *Clin Chem Lab Med*, vol. 46, pp. 27–42, 2008.

[80] A. E. Heazell, M. Brown, S. A. Worton, and W. B. Dunn, "Review: The effects of oxygen on normal and pre-eclamptic placental tissue–insights from metabolomics," *Placenta*, vol. 32 Suppl 2, pp. S119–24, 2011.

[81] A. E. Heazell, M. Brown, W. B. Dunn, S. A. Worton, I. P. Crocker, P. N. Baker, and D. B. Kell, "Analysis of the metabolic footprint and tissue metabolome of placental villous explants cultured at different oxygen tensions reveals novel redox biomarkers," *Placenta*, vol. 29, no. 8, pp. 691–8, 2008.

[82] W. B. Dunn, M. Brown, S. A. Worton, I. P. Crocker, D. Broadhurst, R. Horgan, L. C. Kenny, P. N. Baker, D. B. Kell, and A. E. Heazell, "Changes in the metabolic footprint of placental explant-conditioned culture medium identifies metabolic disturbances related to hypoxia and pre-eclampsia," *Placenta*, vol. 30, no. 11, pp. 974–80, 2009.

[83] R. P. Horgan, D. I. Broadhurst, W. B. Dunn, M. Brown, A. E. P. Heazell, D. B. Kell, P. N. Baker, and L. C. Kenny, "Changes in the metabolic footprint of placental explant-conditioned medium cultured in different oxygen tensions from placentas of small for gestational age and normal pregnancies," *Placenta*, vol. 31, no. 10, pp. 893–901, 2010.

[84] M. C. Tissot van Patot, A. J. Murray, V. Beckey, T. Cindrova-Davies, J. Johns, L. Zwerdlinger, E. Jauniaux, G. J. Burton, and N. J. Serkova, "Human placental metabolic adaptation to chronic hypoxia, high altitude: hypoxic preconditioning," *Am J Physiol Regul Integr Comp Physiol*, vol. 298, no. 1, pp. R166–72, 2010.

[85] W. B. Dunn, M. Brown, S. A. Worton, K. Davies, R. L. Jones, D. B. Kell, and A. E. P. Heazell, "The metabolome of human placental tissue: investigation of first trimester tissue and changes related to preeclampsia in late pregnancy," *Metabolomics*, vol. 8, no. 4, pp. 579–597, 2011.

[86] M. O. Pulkkinen, S. Nyman, M. M. Hamalainen, J. Mattinen, M. O. Pulkkinen, M. M. Hamalainen, S. Nyman, K. Pihlaja, and J. Mattinen, "Proton NMR spectroscopy of the phospholipids in human uterine smooth muscle and placenta," *Gynecol Obstet Invest*, vol. 46, no. 4, pp. 220–4, 1998.

[87] N. Serkova, J. Bendrick-Peart, B. Alexander, and M. C. Tissot van Patot, "Metabolite concentrations in human term placentae and their changes due to delayed collection after delivery," *Placenta*, vol. 24, no. 2-3, pp. 227–35, 2003.

[88] H. H. Kay, J. D. Gordon, A. A. Ribeiro, and L. D. Spicer, "Phosphorus 31 magnetic resonance spectroscopy of human placenta and quantitation with perchloric acid extracts," *Am J Obstet Gynecol*, vol. 164, no. 1.1, pp. 80–7, 1991.

[89] H. H. Kay, S. R. Hawkins, J. D. Gordon, Y. Wang, A. A. Ribeiro, and L. D. Spicer, "Comparative analysis of normal and growth-retarded placentas with phosphorus nuclear magnetic resonance spectroscopy," *Am J Obstet Gynecol*, vol. 167, no. 2, pp. 548–53, 1992.

[90] J. H. Kang, H. Song, J. A. Yoon, D. Y. Park, S. H. Kim, K. J. Lee, A. Farina, Y. K. Cho, Y. N. Kim, S. W. Park, G. J. Kim, S. H. Shim, and D. H. Cha, "Preeclampsia leads to dysregulation of various signaling pathways in placenta," *J Hypertens*, vol. 29, no. 5, pp. 928–36, 2011.

[91] M. M. W. B. Hendriks, F. A. v. Eeuwijk, R. H. Jellema, J. A. Westerhuis, T. H. Reijmers, H. C. J. Hoefsloot, and A. K. Smilde, "Data-processing strategies for metabolomics studies," *Trends Anal Chem*, vol. 30, no. 10, pp. 1685–1698, 2011.

[92] H. Friebolin, *Basic One- and Two-Dimensional NMR Spectroscopy.* Weinheim: Wiley-VCH-Verl., 5th ed., 2011.

[93] J. K. Nicholson and P. J. D. Foxall, "750 MHz ^{1}H and ^{1}H-^{13}C NMR Spectroscopy of Human Blood Plasma," *Anal Chem*, vol. 67, pp. 793–811, 1995.

[94] S. Moestue, B. Sitter, T. F. Bathen, M. B. Tessem, and I. S. Gribbestad, "HR MAS MR spectroscopy in metabolic characterization of cancer," *Curr Top Med Chem*, vol. 11, no. 1, pp. 2–26, 2011.

[95] S. Meiboom and D. Gill, "Modified spin-echo method for measuring nuclear relaxation times," *Rev Sci Instrum*, vol. 29, no. 8, pp. 688–691, 1958.

[96] O. Beckonert, H. C. Keun, T. M. Ebbels, J. Bundy, E. Holmes, J. C. Lindon, and J. K. Nicholson, "Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts," *Nat Prot*, vol. 2, no. 11, pp. 2692–703, 2007.

[97] A. E.R and N. R.A, "The narrowing of nuclear magnetic resonance spectra by molecular rotation in solids," *Proc Phys Soc*, vol. 72, no. 6, pp. 959–972, 1958.

[98] I. J. Lowe, "Free induction decays of rotating solids," *Phys Rev Lett*, vol. 2, no. 7, pp. 285–287, 1959.

[99] H. Lyng, B. Sitter, T. F. Bathen, L. R. Jensen, K. Sundfor, G. B. Kristensen, and I. S. Gribbestad, "Metabolic mapping by use of high-resolution magic angle spinning ^{1}H MR spectroscopy for assessment of apoptosis in cervical carcinomas," *BMC Cancer*, vol. 7, p. 11, 2007.

[100] O. Beckonert, M. Coen, H. C. Keun, Y. Wang, T. M. Ebbels, E. Holmes, J. C. Lindon, and J. K. Nicholson, "High-resolution magic-angle-spinning NMR spectroscopy for metabolic profiling of intact tissues," *Nat Prot*, vol. 5, no. 6, pp. 1019–32, 2010.

[101] W. Gronwald, M. S. Klein, H. Kaspar, S. R. Fagerer, N. Nurnberger, K. Dettmer, T. Bertsch, and P. J. Oefner, "Urinary metabolite quantification employing 2D NMR spectroscopy," *Anal Chem*, vol. 80, no. 23, pp. 9288–97, 2008.

[102] M. Hedenstrom, S. Wiklund-Lindstrom, T. Oman, F. Lu, L. Gerber, P. Schatz, B. Sundberg, and J. Ralph, "Identification of lignin and polysaccharide modifications in populus wood by chemometric analysis of 2D NMR spectra from dissolved cell walls," *Molecular Plant*, vol. 2, no. 5, pp. 933–42, 2009.

[103] K. H. Esbensen, *Multivariate Data Analysis - In Practice.* Camo Process AS, 5th ed., 2002.

[104] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometr Intell Lab*, vol. 2, pp. 37–52, 1987.

[105] S. Wold, M. Sjöström, and L. Eriksson, "PLS-Regression: a basic tool of chemometrics," *Chemometr Intell Lab*, vol. 58, pp. 109–130, 2001.

[106] S. de Jong, "SIMPLS: An alternative approach to partial least squares regression," *Chemometr Intell Lab*, vol. 18, no. 3, pp. 251 – 263, 1993.

[107] L. R. Euceda, G. F. Giskeodegard, and T. F. Bathen, "Preprocessing of NMR metabolomics data," *Scand J Clin Lab Invest*, vol. 75, no. 3, pp. 193–203, 2015.

[108] F. Savorani, G. Tomasi, and S. Engelsen, "icoshift: A versatile tool for the rapid alignment of 1D NMR spectra," *J Magn Reson*, vol. 2, pp. 190–202, 2010.

[109] G. F. Giskeodegard, T. G. Bloemberg, G. Postma, B. Sitter, M. B. Tessem, I. S. Gribbestad, T. F. Bathen, and L. M. Buydens, "Alignment of high resolution magic angle spinning magnetic resonance spectra using warping methods," *Anal Chim Acta*, vol. 683, no. 1, pp. 1–11, 2010.

[110] A. Craig, O. Cloarec, E. Holmes, J. K. Nicholson, and J. C. Lindon, "Scaling and normalization effects in NMR spectroscopic metabonomic data sets," *Anal Chem*, vol. 78, no. 7, pp. 2262–7, 2006.

[111] S. S. Waikar, V. S. Sabbisetti, and J. V. Bonventre, "Normalization of urinary biomarkers to creatinine during changes in glomerular filtration rate," *Kidney Int*, vol. 78, no. 5, pp. 486–94, 2010.

[112] F. Dieterle, A. Ross, G. Schlotterbeck, and H. Senn, "Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics," *Anal Chem*, vol. 78, no. 13, pp. 4281–90, 2006.

[113] R. van den Berg, H. Hoefsloot, J. Westerhuis, A. Smilde, and M. van der Werf, "Centering, scaling, and transformations: improving the biological information content of metabolomics data," *BMC Genomics*, vol. 7, no. 1, p. 142, 2006.

[114] D. E. Axelson, *Data Preprocessing For Chemometric and Metabonomic Analysis*. CreateSpace Independent Publishing Platform, 2012.

[115] G. Palermo, P. Piraino, and H. D. Zucht, "Performance of PLS regression coefficients in selecting variables for each response of a multivariate PLS for omics-type data," *Adv Appl Bioinform Chem*, vol. 2, pp. 57–70, 2009.

[116] I.-G. Chong and C.-H. Jun, "Performance of some variable selection methods when multi-collinearity is present," *Chemometr Intell Lab*, vol. 78, no. 1–2, pp. 103 – 112, 2005.

[117] H. Li, Y. Liang, Q. Xu, and D. Cao, "Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration," *Anal Chim Acta*, vol. 648, no. 1, pp. 77–84, 2009.

[118] J. A. Westerhuis, H. C. J. Hoefsloot, S. Smit, D. J. Vis, A. K. Smilde, E. J. J. Velzen, J. P. M. Duijnhoven, and F. A. Dorsten, "Assessment of PLSDA cross validation," *Metabolomics*, vol. 4, no. 1, pp. 81–89, 2008.

[119] R. Skrastad, G. Hov, H. G. Blaas, P. Romundstad, and K. Salvesen, "Risk assessment for preeclampsia in nulliparous women at 11-13 weeks gestational age: prospective evaluation of two algorithms," *Br J Obstet Gynaecol*, 2014.

[120] R. B. Skrastad, G. G. Hov, H. G. Blaas, P. R. Romundstad, and K. A. Salvesen, "A prospective study of screening for hypertensive disorders of pregnancy at 11-13 weeks in a scandinavian population," *Acta Obstet Gynecol Scand*, 2014.

[121] H. K. Gjessing, P. Grottum, and S. H. Eik-Nes, "A direct method for ultrasound prediction of day of delivery: a new, population-based approach," *Ultrasound Obstet Gynecol*, vol. 30, no. 1, pp. 19–27, 2007.

[122] K. Marsal, P.-H. Persson, T. Larsen, H. Lilja, A. Selbing, and B. Sultan, "Intrauterine growth curves based on ultrasonically estimated foetal weights," *Acta Paediatr*, vol. 85, pp. 843–848, Jul 1996.

[123] J. S. Lawson and P. Mayberry, "How can infant and perinatal mortality rates be compared internationally?," *World Health Forum*, vol. 15, no. 1, pp. 85–88, 1994.

[124] G. Mancia, R. Fagard, K. Narkiewicz, J. Redon, A. Zanchetti, M. Böhm, T. Christiaens, R. Cifkova, G. De Backer, A. Dominiczak, M. Galderisi, D. E. Grobbee, T. Jaarsma, P. Kirchhof, S. E. Kjeldsen, S. Laurent, A. J. Manolis, P. M. Nilsson, L. M. Ruilope, R. E. Schmieder, P. A. Sirnes, P. Sleight, M. Viigimaa, B. Waeber, and F. Zannad, "2013 ESH/ESC guidelines for the management of arterial hypertension," *Blood Pressure*, vol. 22, no. 4, pp. 193–278, 2013.

[125] A. Khalil and K. H. Nicolaides, "How to record uterine artery doppler in the first trimester," *Ultrasound Obstet Gynecol*, vol. 42, no. 4, pp. 478–9, 2013.

[126] S. L. Johnsen, S. Rasmussen, T. Wilsgaard, R. Sollien, and T. Kiserud, "Longitudinal reference ranges for estimated fetal weight," *Acta Obstet Gynecol Scand*, vol. 85, no. 3, pp. 286–97, 2006.

[127] ACOG Committee on Obstetric Practice, "ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. American College of Obstetricians and Gynecologists," *Int J Gynaecol Obstet*, vol. 77, no. 1, pp. 67–75, 2002.

[128] B. Lorentzen, P. Øian, A. Staff, and J. H. S. Mortensen, "Hypertensive svangerskapskomplikasjoner; veileder i fødselshjelp 2008." Norsk gynekologisk forening, November 2008.

[129] A. Staff, A. B. Andersgaard, T. Henriksen, E. Langesæter, E. Magnusson, T. M. Michelsen, L. C. Thomsen, and P. Øian, *Hypertensive svangerskapskomplikasjoner og eklampsi*, ch. 28. Oslo: Legeforeningen, 2014. In Norwegian.

[130] S. Akoka, L. Barantin, and M. Trierweiler, "Concentration measurement by proton NMR using the ERETIC method," *Anal Chem*, vol. 71, pp. 2554–2557, 1999.

[131] J. Lounila, M. Ala-Korpela, J. Jokisaari, M. J. Savolainen, and Y. A. Kesaniemi, "Effects of orientational order and particle size on the NMR line positions of lipoproteins," *Phys Rev Lett*, vol. 72, no. 25, pp. 4049–4052, 1994.

[132] M. Ala-Korpela, A. Korhonen, J. Keisala, S. Hörkkö, P. Korpi, L. P. Ingman, J. Jokisaari, M. J. Savolainen, and Y. A. Kesäniemi, "^{1}H NMR-based absolute quantitation of human lipoproteins and their lipid contents directly from plasma," *J Lipid Res*, vol. 35, 1994.

[133] A. F. Saftlas, D. R. Olson, A. L. Franks, H. K. Atrash, and R. Pokras, "Epidemiology of preeclampsia and eclampsia in the United States, 1979-1986," *Am J Obstet Gynecol*, vol. 163, no. 2, pp. 460–5, 1990.

[134] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," *Ann Stat*, vol. 29, no. 4, pp. 1165–1188, 2001.

[135] J. Xia, R. Mandal, I. V. Sinelnikov, D. Broadhurst, and D. S. Wishart, "MetaboAnalyst 2.0 - a comprehensive server for metabolomic data analysis," *Nucl Acids Res*, 2012.

[136] J. Xia and D. S. Wishart, "MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data," *Nucl Acids Res*, vol. 38, pp. W71–7, 2010.

[137] A. Frolkis, C. Knox, E. Lim, T. Jewison, V. Law, D. D. Hau, P. Liu, B. Gautam, S. Ly, A. C. Guo, J. Xia, Y. Liang, S. Shrivastava, and D. S. Wishart, "SMPDB: The small molecule pathway database," *Nucl Acids Res*, vol. 38, pp. D480–7, 2010.

[138] T. Jewison, Y. Su, F. M. Disfany, Y. Liang, C. Knox, A. Maciejewski, J. Poelzer, J. Huynh, Y. Zhou, D. Arndt, Y. Djoumbou, Y. Liu, L. Deng, A. C. Guo, B. Han, A. Pon, M. Wilson, S. Rafatnia, P. Liu, and D. S. Wishart, "SMPDB 2.0: big improvements to the small molecule pathway database," *Nucl Acids Res*, vol. 42, pp. D478–84, 2014.

[139] S. Roberge, Y. Giguere, P. Villa, K. Nicolaides, M. Vainio, J. C. Forest, P. von Dadelszen, D. Vaiman, S. Tapp, and E. Bujold, "Early administration of low-dose aspirin for the prevention of severe and mild preeclampsia: a systematic review and meta-analysis," *Am J Perinatol*, vol. 29, no. 7, pp. 551–6, 2012.

[140] S. Roberge, P. Villa, K. Nicolaides, Y. Giguere, M. Vainio, A. Bakthi, A. Ebrashy, and E. Bujold, "Early administration of low-dose aspirin for the prevention of preterm and term preeclampsia: a systematic review and meta-analysis," *Fetal Diagn Ther*, vol. 31, no. 3, pp. 141–6, 2012.

[141] P. M. Villa, E. Kajantie, K. Raikkonen, A. K. Pesonen, E. Hamalainen, M. Vainio, P. Taipale, H. Laivuori, and P. S. group, "Aspirin in the prevention of pre-eclampsia in high-risk women: a randomised placebo-controlled PREDO trial and a meta-analysis of randomised trials," *Br J Obstet Gynecol*, vol. 120, no. 1, pp. 64–74, 2013.

[142] D. Sachse, L. Sletner, K. Morkrid, A. K. Jenum, K. I. Birkeland, F. Rise, A. P. Piehler, and J. P. Berg, "Metabolic changes in urine during and after pregnancy in a large, multiethnic population-based cohort study of gestational diabetes," *PloS ONE*, vol. 7, no. 12, p. e52399, 2012.

[143] L. Maitre, E. Fthenou, T. Athersuch, M. Coen, M. B. Toledano, E. Holmes, M. Kogevinas, L. Chatzi, and H. C. Keun, "Urinary metabolic profiles in early pregnancy are associated with preterm birth and fetal growth restriction in the Rhea mother-child cohort study," *BMC Medicine*, vol. 12, no. 1, 2014.

[144] R. H. Glew, G. Melah, A. I. El-Nafaty, Y. Brandt, D. Morris, and D. J. VanderJagt, "Plasma and urinary free amino acid concentrations in preeclamptic women in northern Nigeria," *Clin Chim Acta*, vol. 342, no. 12, pp. 179 – 185, 2004.

[145] T. Hauet, M. Thame, H. Fletcher, T. Baker, and F. Jahoor, "Comparing the in vivo glycine fluxes of adolescent girls and adult women during early and late pregnancy," *Brit J Nutr*, vol. 104, no. 04, pp. 498–502, 2010.

[146] D. L. Nelson and M. M. Cox, *Principles of Biochemistry*. W.H. Freeman and Company, 2008.

[147] R. W. Friesen, E. M. Novak, D. Hasman, and S. M. Innis, "Relationship of dimethylglycine, choline, and betaine with oxoproline in plasma of pregnant women and their newborn infants," *J Nutr*, vol. 137, no. 12, pp. 2641–2646, 2007.

[148] A. Perez-Sepulveda, P. P. Espana-Perrot, X. B. Fernandez, V. Ahumada, V. Bustos, J. A. Arraztoa, A. Dobierzewska, H. Figueroa-Diesel, G. E. Rice, and S. E. Illanes, "Levels of key enzymes of methionine-homocysteine metabolism in preeclampsia," *Biomed Res Int*, vol. 2013, p. 731962, 2013.

[149] R. O. Bahado-Singh, R. Akolekar, R. Mandal, E. Dong, J. Xia, M. Kruger, D. S. Wishart, and K. H. Nicolaides, "Metabolomics and first-trimester prediction of early-onset preeclampsia," *J Matern Fetal Neonatal Med*, vol. 25, no. 10, pp. 1840–7, 2012.

[150] K. Braekke, P. M. Ueland, N. K. Harsem, A. Karlsen, R. Blomhoff, and A. C. Staff, "Homocysteine, cysteine, and related metabolites in maternal and fetal plasma in preeclampsia," *Pediatr Res*, vol. 62, no. 3, pp. 319–324, 2007.

[151] C. Mislanova, O. Martsenyuk, B. Huppertz, and M. Obolenskaya, "Placental markers of folate-related metabolism in preeclampsia," *Reproduction*, vol. 142, no. 3, pp. 467–76, 2011.

[152] T. Tamura and M. F. Picciano, "Folate and human reproduction," *Am J Clin Nutr*, vol. 83, no. 5, pp. 993–1016, 2006.

[153] S. A. Craig, "Betaine in human nutrition," *Am J Clin Nutr*, vol. 80, no. 3, pp. 539–49, 2004.

[154] R. Vanholder, E. Schepers, A. Pletinck, E. V. Nagler, and G. Glorieux, "The uremic toxicity of indoxyl sulfate and p-cresyl sulfate: a systematic review," *J Am Soc Nephrol*, vol. 25, no. 9, pp. 1897–907, 2014.

[155] B. K. Meijers, S. Van Kerckhoven, K. Verbeke, W. Dehaen, Y. Vanrenterghem, M. F. Hoylaerts, and P. Evenepoel, "The uremic retention solute p-cresyl sulfate and markers of endothelial damage," *Am J Kidney Dis*, vol. 54, no. 5, pp. 891–901, 2009.

[156] P. Wurtz, J. R. Raiko, C. G. Magnussen, P. Soininen, A. J. Kangas, T. Tynkkynen, R. Thomson, R. Laatikainen, M. J. Savolainen, J. Laurikka, P. Kuukasjarvi, M. Tarkka, P. J. Karhunen, A. Jula, J. S. Viikari, M. Kahonen, T. Lehtimaki, M. Juonala, M. Ala-Korpela, and O. T. Raitakari, "High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis," *Eur Heart J*, vol. 33, no. 18, pp. 2307–16, 2012.

[157] C. N. Spracklen, C. J. Smith, A. F. Saftlas, J. G. Robinson, and K. K. Ryckman, "Maternal hyperlipidemia and the risk of preeclampsia: a meta-analysis," *Am J Epidemiol*, vol. 180, no. 4, pp. 346–58, 2014.

[158] I. V. Alsnes, I. Janszky, M. R. Forman, L. J. Vatten, and I. Okland, "A population-based study of associations between preeclampsia and later cardiovascular risk factors," *Am J Obstet Gynecol*, vol. 211, no. 6, pp. 657 e1–7, 2014.

[159] M. Desforges, A. Ditchfield, C. Hirst, C. Pegorie, K. Martyn-Smith, C. Sibley, and S. Greenwood, "Reduced placental taurine transporter (TauT) activity in pregnancies complicated by pre-eclampsia and maternal obesity," *Adv Exp Med Biol*, vol. 776, pp. 81–91, 2013.

[160] L. G. Mikael, J. Pancer, X. Jiang, Q. Wu, M. Caudill, and R. Rozen, "Low dietary folate and methylenetetrahydrofolate reductase deficiency may lead to pregnancy complications through modulation of ApoAI and IFN-gamma in spleen and placenta, and through reduction of methylation potential," *Mol Nutr Food Res*, vol. 57, no. 4, pp. 661–70, 2013.

[161] A. C. Staff, T. Ranheim, and B. Halvorsen, "Augmented pla2 activity in pre-eclamptic decidual tissue–a key player in the pathophysiology of 'acute atherosis' in pre-eclampsia?," *Placenta*, vol. 24, no. 10, pp. 965–73, 2003.

[162] R. S. Bowen, Y. Zhang, Y. Gu, D. F. Lewis, and Y. Wang, "Increased phospholipase A2 and thromboxane but not prostacyclin production by placental trophoblast cells from normal and preeclamptic pregnancies cultured under hypoxia condition," *Placenta*, vol. 26, no. 5, pp. 402–9, 2005.

[163] E. Ackerstaff, K. Glunde, and Z. M. Bhujwalla, "Choline phospholipid metabolism: a target in cancer cells?," *J Cell Biochem*, vol. 90, no. 3, pp. 525–33, 2003.

[164] A. Johansson, J. E. Curran, M. P. Johnson, K. A. Freed, M. H. Fenstad, L. Bjorge, I. P. Eide, M. A. Carless, D. L. Rainwater, H. H. Goring, R. Austgulen, E. K. Moses, and J. Blangero, "Identification of acox2 as a shared genetic risk factor for preeclampsia and cardiovascular disease," *Eur J Hum Gen*, vol. 19, no. 7, pp. 796–800, 2011.

[165] S. Hahn, O. Lapaire, and N. G. Than, "Biomarker development for presymptomatic molecular diagnosis of preeclampsia: feasible, useful or even unnecessary?," *Exp Rev Mol Diagn*, vol. 15, no. 5, pp. 617–629, 2015.

[166] S. A. Founds, Y. P. Conley, J. F. Lyons-Weiler, A. Jeyabalan, W. A. Hogge, and K. P. Conrad, "Altered global gene expression in first trimester placentas of women destined to develop preeclampsia," *Placenta*, vol. 30, no. 1, pp. 15–24, 2009.

[167] S. Lisonkova and K. S. Joseph, "Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease," *Am J Obstet Gynecol*, vol. 209, no. 6, pp. 544 e1–544 e12, 2013.

[168] L. C. Poon and K. H. Nicolaides, "Early prediction of preeclampsia," *Obstet Gynecol Int*, vol. 2014, p. 297397, 2014. Poon, Leona C Nicolaides, Kypros H.

[169] S. Verlohren, K. Melchiorre, A. Khalil, and B. Thilaganathan, "Uterine artery Doppler, birth weight and timing of onset of pre-eclampsia: providing insights into the dual etiology of late-onset pre-eclampsia," *Ultrasound Obstet Gynecol*, vol. 44, no. 3, pp. 293–8, 2014.

[170] S. S. Heinzmann, C. A. Merrifield, S. Rezzi, S. Kochhar, J. C. Lindon, E. Holmes, and J. K. Nicholson, "Stability and robustness of human metabolic phenotypes in response to sequential food challenges," *J Proteome Res*, vol. 11, no. 2, pp. 643–55, 2012.

[171] S. S. Heinzmann, I. J. Brown, Q. Chan, M. Bictash, M.-E. Dumas, S. Kochhar, J. Stamler, E. Holmes, P. Elliott, and J. K. Nicholson, "Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption," *Am J Clin Nutr*, vol. 92, no. 2, pp. 436–443, 2010.

[172] L. C. Kenny, M. A. Black, L. Poston, R. Taylor, J. E. Myers, P. N. Baker, L. M. McCowan, N. A. Simpson, G. A. Dekker, C. T. Roberts, K. Rodems, B. Noland, M. Raymundo, J. J. Walker, and R. A. North, "Early pregnancy prediction of preeclampsia in nulliparous women, combining clinical risk and biomarkers: The screening for pregnancy endpoints (SCOPE) international cohort study," *Hypertension*, vol. 64, no. 3, pp. 644–652, 2014.

[173] R. A. North, L. M. McCowan, G. A. Dekker, L. Poston, E. H. Chan, A. W. Stewart, M. A. Black, R. S. Taylor, J. J. Walker, P. N. Baker, and L. C. Kenny, "Clinical risk prediction for pre-eclampsia in nulliparous women: development of model in international prospective cohort," *BMJ*, vol. 342, p. d1875, 2011.

[174] E. R. Hillesund, N. C. Overby, S. M. Engel, K. Klungsoyr, Q. E. Harmon, M. Haugen, and E. Bere, "Associations of adherence to the new nordic diet with risk of preeclampsia and preterm delivery in the Norwegian mother and child cohort study (MoBa)," *Eur J Epidemiol*, vol. 29, no. 10, pp. 753–65, 2014.

[175] P. Palatini, "Glomerular hyperfiltration: a marker of early renal damage in pre-diabetes and pre-hypertension," *Nephrol Dial Transplant*, 2012.

[176] R. Okada, Y. Yasuda, K. Tsushita, K. Wakai, N. Hamajima, and S. Matsuo, "Glomerular hyperfiltration in prediabetes and prehypertension," *Nephrol Dial Transplant*, vol. 27, no. 5, pp. 1821–5, 2012.

[177] K. Kuromoto, M. Watanabe, K. Adachi, K. Ohashi, and Y. Iwatani, "Increases in urinary creatinine and blood pressure during early pregnancy in pre-eclampsia," *Ann Clin Biochem*, vol. 47, no. 4, pp. 336–42, 2010.

[178] G. C. McKeeman, J. E. Ardill, C. M. Caldwell, A. J. Hunter, and N. McClure, "Soluble vascular endothelial growth factor receptor-1 (sflt-1) is increased throughout gestation in patients who have preeclampsia develop," *Am J Obstet Gynecol*, vol. 191, no. 4, pp. 1240–6, 2004.

[179] S. Verlohren, I. Herraiz, O. Lapaire, D. Schlembach, M. Moertl, H. Zeisler, P. Calda, W. Holzgreve, A. Galindo, T. Engels, B. Denk, and H. Stepan, "The sFlt-1/PlGF ratio in different types of hypertensive pregnancy disorders and its prognostic potential in preeclamptic patients," *Am J Obstet Gynecol*, vol. 206, no. 1, 2012.

[180] D. A. Enquobahrie, M. A. Williams, C. L. Butler, I. O. Frederick, R. S. Miller, and D. A. Luthy, "Maternal plasma lipid concentrations in early pregnancy and risk of preeclampsia," *Am J Hypertens*, vol. 17, no. 7, pp. 574–581, 2004.

[181] B. Lorentzen and T. Henriksen, "Plasma lipids and vascular dysfunction in preeclampsia," *Sem Reprod Endocrinol*, vol. 16, no. 1, pp. 33–9, 1998.

[182] A. Wiznitzer, V. Mayer A Fau Novack, E. Novack V Fau Sheiner, H. Sheiner E Fau Gilutz, A. Gilutz H Fau Malhotra, L. Malhotra A Fau Novack, and L. Novack, "Association of lipid levels during gestation with preeclampsia and gestational diabetes mellitus: a population-based study," *Am J Obstet Gynecol*, vol. 201, no. 1097-6868 (Electronic), 2009.

[183] N. J. Robinson, L. J. Minchell, J. E. Myers, C. A. Hubel, and I. P. Crocker, "A potential role for free fatty acids in the pathogenesis of preeclampsia," *J Hypertens*, vol. 27, no. 6, pp. 1293–302, 2009.

[184] H. A. Korkes, N. Sass, A. F. Moron, N. O. Camara, T. Bonetti, A. S. Cerdeira, I. D. Da Silva, and L. De Oliveira, "Lipidomic assessment of plasma and placenta of women with early-onset preeclampsia," *PloS one*, vol. 9, no. 10, p. e110747, 2014.

[185] A. Buchbinder, B. M. Sibai, S. Caritis, C. MacPherson, J. Hauth, M. D. Lindheimer, M. Klebanoff, P. VanDorsten, M. Landon, R. Paul, M. Miodovnik, P. Meis, and G. Thurnau, "Adverse perinatal outcomes are significantly higher in severe gestational hypertension than in mild pre-eclampsia," *Am J Obstet Gynecol*, vol. 186, no. 1, pp. 66 – 71, 2002.

[186] A. A. Baschat, "First-trimester screening for pre-eclampsia: moving from personalized risk prediction to prevention," *Ultrasound Obstetrics Gynecol*, vol. 45, no. 2, pp. 119–29, 2015.

[187] M. M. Costantine and K. Cleary, "Pravastatin for the prevention of preeclampsia in high-risk pregnant women," *Obstet Gynecol*, vol. 121, no. 2 Pt 1, pp. 349–53, 2013.

[188] S. W. Wen, J. Champagne, R. Rennicks White, D. Coyle, W. Fraser, G. Smith, D. Fergusson, and M. C. Walker, "Effect of folic acid supplementation in pregnancy on preeclampsia: the folic acid clinical trial study," *J Pregnancy*, vol. 2013, p. 294312, 2013.

[189] L. Myatt, C. W. Redman, A. C. Staff, S. Hansson, M. L. Wilson, H. Laivuori, L. Poston, and J. M. Roberts, "Strategy for standardization of preeclampsia research study design," *Hypertension*, vol. 63, no. 6, pp. 1293–301, 2014.

[190] G. J. Burton, N. J. Sebire, L. Myatt, D. Tannetta, Y. L. Wang, Y. Sadovsky, A. C. Staff, and C. W. Redman, "Optimising sample collection for placental research," *Placenta*, vol. 35, no. 1, pp. 9–22, 2014.

[191] A. K. Mbah, J. L. Kornosky, S. Kristensen, E. M. August, A. P. Alio, P. J. Marty, V. Belogolovkin, K. Bruder, and H. M. Salihu, "Super-obesity and risk for early and late pre-eclampsia," *Brit J Obstet Gynecol*, vol. 117, no. 8, pp. 997–1004, 2010.

[192] S. Cheng, E. P. Rhee, M. G. Larson, G. D. Lewis, E. L. McCabe, D. Shen, M. J. Palma, L. D. Roberts, A. Dejam, A. L. Souza, A. A. Deik, M. Magnusson, C. S. Fox, C. J. O'Donnell, R. S. Vasan, O. Melander, C. B. Clish, R. E. Gerszten, and T. J. Wang, "Metabolite profiling identifies pathways associated with metabolic risk in humans," *Circulation*, vol. 125, no. 18, pp. 2222–31, 2012.

[193] L. Myatt, S. Muralimanoharan, and A. Maloyan, "Effect of preeclampsia on placental function: influence of sexual dimorphism, microRNA's and mitochondria," *Adv Exp Med Biol*, vol. 814, pp. 133–46, 2014.

[194] A.-H. M. Emwas, R. M. Salek, J. L. Griffin, and J. Merzaban, "NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations," *Metabolomics*, vol. 9, no. 5, pp. 1048–1072, 2013.

[195] S. Bansal, J. E. Buring, N. Rifai, S. Mora, F. M. Sacks, and P. M. Ridker, "Fasting compared with nonfasting triglycerides and risk of cardiovascular events in women," *J Am Med Assoc*, vol. 298, no. 3, pp. 309–16, 2007.

[196] J. Pinto, M. R. Domingues, E. Galhano, C. Pita, C. Almeida Mdo, I. M. Carreira, and A. M. Gil, "Human plasma stability during handling and storage: impact on NMR metabolomics," *Analyst*, vol. 139, no. 5, pp. 1168–77, 2014.

[197] M. Lauridsen, S. H. Hansen, J. W. Jaroszewski, and C. Cornett, "Human urine as test material in 1H NMR-based metabonomics:? recommendations for sample preparation and storage," *Anal Chem*, vol. 79, no. 3, pp. 1181–1186, 2007.

[198] R. H. Barton, J. K. Nicholson, P. Elliott, and E. Holmes, "High-throughput 1H NMR-based metabolic analysis of human serum and urine for large-scale epidemiological studies: validation study," *Int J Epidemiol*, vol. 37 Suppl 1, pp. i31–40, 2008.

[199] T. Cindrova-Davies, H. W. Yung, J. Johns, O. Spasic-Boskovic, S. Korolchuk, E. Jauniaux, G. J. Burton, and D. S. Charnock-Jones, "Oxidative stress, gene expression, and protein changes induced in the human placenta during labor," *Am J Pathol*, vol. 171, no. 4, pp. 1168–79, 2007.

[200] S. Sohlberg, A. K. Wikstrom, M. Olovsson, P. Lindgren, O. Axelsson, A. Mulic-Lutvica, J. Weis, and J. Wikstrom, "In vivo ^{31}P-MR spectroscopy in normal pregnancy, early and late pre-eclampsia: a study of placental metabolism," *Placenta*, vol. 35, no. 5, pp. 318–23, 2014.

[201] S. O. Diaz, J. Pinto, G. Graca, I. F. Duarte, A. S. Barros, E. Galhano, C. Pita, M. Almeida, B. J. Goodfellow, I. M. Carreira, and A. M. Gil, "Metabolic biomarkers of prenatal disorders: An exploratory NMR metabonomics study of second trimester maternal urine and blood plasma," *J Proteome Res*, vol. 10, pp. 3732–3742, 2011.

[202] M. Cao, S. Lamichhane, S. Lundgren, A. Bofin, H. Fjosne, G. Giskeodegard, and T. Bathen, "Metabolic characterization of triple negative breast cancer," *BMC Cancer*, vol. 14, no. 1, p. 941, 2014.

[203] D. J. Greenblatt, B. J. Ransil, J. S. Harmatz, T. W. Smith, D. W. Duhme, and J. Koch-Weser, "Variability of 24-hour urinary creatinine excretion by normal subjects," *J Clin Pharm*, vol. 16, no. 7, pp. 321–8, 1976.

[204] H. C. Keun, T. M. D. Ebbels, H. Antti, M. E. Bollard, O. Beckonert, E. Holmes, J. C. Lindon, and J. K. Nicholson, "Improved analysis of multivariate data by variable stability scaling: application to NMR-based metabolic profiling," *Anal Chim Acta*, vol. 490, no. 1-2, pp. 265–276, 2003.

Paper I

PLOS ONE

Metabolomic Biomarkers in Serum and Urine in Women with Preeclampsia

Marie Austdal[1,2,3]*, Ragnhild Bergene Skråstad[4,5], Astrid Solberg Gundersen[6], Rigmor Austgulen[6,7], Ann-Charlotte Iversen[6,7], Tone Frost Bathen[1,3]

1 Department of Circulation and Medical Imaging, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, 2 Liaison Committee between the Central Norway Regional Health Authority (RHA) and Norwegian University of Science and Technology (NTNU), Trondheim, Norway, 3 St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway, 4 Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, 5 National Center for Fetal Medicine, St. Olavs Hospital, Trondheim, Norway, 6 Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, 7 Centre of Molecular Inflammation Research (CEMIR), Faculty of Medicine, NTNU, Trondheim, Norway

Abstract

Objective: To explore the potential of magnetic resonance (MR) metabolomics for study of preeclampsia, for improved phenotyping and elucidating potential clues to etiology and pathogenesis.

Methods: Urine and serum samples from pregnant women with preeclampsia (n = 10), normal pregnancies (n = 10) and non-pregnant women (n = 10) matched by age and gestational age were analyzed with MR spectroscopy and subjected to multivariate analysis. Metabolites were then quantified and compared between groups.

Results: Urine and serum samples revealed clear differences between women with preeclampsia and both control groups (normal pregnant and non-pregnant women). Nine urine metabolites were significantly different between preeclampsia and the normal pregnant group. Urine samples from women with early onset preeclampsia clustered together in the multivariate analysis. The preeclampsia serum spectra showed higher levels of low and very-low density lipoproteins and lower levels of high-density lipoproteins when compared to both non-pregnant and normal pregnant women.

Conclusion: The MR determined metabolic profiles in urine and serum from women with preeclampsia are clearly different from normal pregnant women. The observed differences represent a potential to examine mechanisms underlying different preeclampsia phenotypes in urine and serum samples in larger studies. In addition, similarities between preeclampsia and cardiovascular disease in metabolomics are demonstrated.

Introduction

Preeclampsia (PE) is a complex syndrome affecting about 3% of pregnancies [1]. It presents serious risk of both maternal and fetal morbidity and mortality [2]. PE is characterized by high blood pressure and proteinuria in the second half of pregnancy [3]. No tests accurately predict the onset of PE, and implementation of fetal delivery is the only definitive treatment for threatening manifestations of symptoms [1].

The pathogenesis of PE is still undefined. However, it is generally assumed that it starts in early pregnancy with poorly developed placental vascularization, giving rise to placental oxidative stress and imbalanced interaction between maternal and fetal cells. Later, inappropriate and exaggerated maternal responses to the placental stress are established, involving endothelial activation and systemic inflammation [4]. The

inflammation in PE shows a strong similarity to the development of cardiovascular diseases (CVD) [4], and it has been reported that women with preeclamptic pregnancies have an up to eight-fold increased risk of later cardiovascular events [5]. The shared underlying mechanisms include endothelial dysfunction, metabolic abnormalities and increased oxidative stress [6].

Metabolites are constituents of the metabolism, chemical interactions in the body necessary for life [7]. Metabolomics is the systematic study of metabolites in tissues and biofluids [8]. The concentrations of metabolites and their combinations can be used as predictive models for disease classification and progression [8]. Robust statistical methods are applied to handle the massive data outputs.

Metabolomics analysis holds potential for detailed phenotyping of the PE syndrome, but few metabolomics studies of women with active disease have so far been undertaken. Studies by Turner *et al.*

[9,10], using MR metabolomics on serum from women diagnosed with PE reported metabolic patterns attributed to oxidative stress including decreased lipid and ketone body content, and the findings indicated that this method could be useful for PE phenotyping. Schott et al. [11] used both proton and phosphorous MR spectroscopy to analyze plasma from women with PE. They found a decrease in HDL and a trend towards higher levels of VLDL2 and LDL2 in this group compared to healthy pregnancies. However, no metabolomics studies have analyzed both urine and serum from the same group of women, which represents a more comprehensive view of the metabolome. Detailed analysis of body fluids from women with preeclampsia could improve diagnostic accuracy and possibly predict perinatal outcomes and future risk for the mother.

Further research is needed to establish the role of metabolomics and the robustness of the method in preeclampsia. Furthermore, a discussion of the discriminatory metabolites in a more biological context relevant to PE is generally lacking. To this end, the aim of the present study was to establish the metabolic profiles of body fluids (urine and serum) from women with PE, normal pregnancies and from non-pregnant women by MR metabolomics. Detailed phenotyping of the PE syndrome and potential clues to etiology and pathogenesis were explored.

Materials and Methods

Ethics Statement

All participating women signed informed consents and the study was approved by the Regional Committee for Medical and Health Research Ethics (REC), Central Norway, reference number 2011/761.

Study Population

Women admitted with PE at the maternity ward at St. Olavs Hospital, Trondheim University Hospital, Norway gave samples to the study. The PE diagnosis was based on the diagnostic criteria of the Norwegian Medical Association (blood pressure $\geq$ 140/90, proteinuria $\geq$+1, measured at least twice four to six hours apart after gestational week 20) [12]. Pregnant and non-pregnant control women were recruited by appeals to environments of St. Olavs Hospital and Røros Medical Center, Røros, Norway. Control women with previous PE pregnancies were not included. Gestational age for both cases and controls were based on routine ultrasound examination between gestational week 17 and 20. Information about health status and pregnancy was collected from interviews and medical journals. All included women were of Scandinavian ethnicity.

Sample Handling and Spectroscopy

Peripheral venous blood (5 mL) and spot urine samples (20 mL) were collected from nonfasting women with PE at time of diagnosis, from healthy pregnant women matched by age and gestational age to the PE group, and from non-pregnant women matched by age to the PE group. Aliquots (1.8 mL) were stored at $-80°C$ prior to analysis.

Samples were thawed at 20°C, mixed with bacteriostatic buffer and stored at 5°C until analysis ($\leq$15 hours). Urine was centrifuged at 6000 RPM (Sorvall RMC 14; DuPont) for five minutes. The supernatant (540 µL) was mixed with buffer (60 µL) (pH 7.4 1.5 mM KH_2PO_4 in D_2O, 0.1% Trimethyl-Silyl Propionate (TSP), 2 mM NaN_3) and analyzed in 5 mm NMR tubes (Norell Inc., NJ, USA). Serum (100 µL) was mixed with buffer (100 µL) (pH 7.4 0.075 mM Na_2HPO_4, 5 µM NaN_3, 5 µM TSP) and analyzed in 3 mm NMR tubes. MR analysis was performed at the MR Core Facility at NTNU, Trondheim, Norway using a Bruker Avance III Ultrashielded Plus 600 MHz spectrometer (Bruker Biospin GmbH, Germany) equipped with 5 mm QCI Cryoprobe with integrated, cooled preamplifiers for ^1H, ^2H and ^{13}C. Experiments were fully automated using the SampleJet^{TM} in combination with Icon-NMR on TopSpin 3.1 software (Bruker Biospin). Proton spectra were acquired using 1D NOESY with presaturation and spoil gradients on urine samples and a 1D lipid and water suppressing Carr-Purcell-Meiboom-Gill sequence (CPMG) on serum samples. Diffusion edited serum spectra (LEDBPG) for suppression of small molecular weight metabolite signals were also acquired. Additional 2D spectra of urine samples were acquired for metabolite identification: J-resolved spectroscopy (JRES), Heteronuclear Multiple Bond Correlation Spectroscopy (HMBC), Heteronuclear Single Quantum Coherence Spectroscopy (HSQC) and Total Correlation Spectroscopy (TOCSY). Additional NMR parameters are given in Tab 3. Spectra were Fourier transformed to 128 K after 0.3 Hz exponential line broadening. Chemical shifts were referenced to TSP (δ0 ppm).

Multivariate analysis

In MR Metabolomics, common statistical methods are Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA) [13]. PCA is a powerful method of data extraction, which finds combinations of variables that describe trends in large data, called principal components (PCs), visualized in scores and loading plots. The score plots show each spectrum as an object in the principal component space, and are useful for identifying clusters and outliers in the dataset. The loading plots show the contributing variables to each PC. PLS-DA models the relationship between the spectra and class information using multivariate regression methods. The metabolites responsible for the separation between classes are shown in loading variables (LVs), and may be colored by variable importance in projection (VIP) [13]. Multivariate analysis was performed using PLS_Toolbox 6.7.1 (Eigenvector Research, USA).

Spectra were imported to Matlab r2012a (The Mathworks, Inc., MA, USA). Residual water signals were removed. The urine NOESY spectra were normalized to equal area below the curve, cut to region of interest (ROI) (δ8–1 ppm) and peak aligned using icoshift [14]. The serum CPMG spectra were cut to ROI (δ4.5–0.5 ppm) and aligned by referencing the left alanine peak at δ1.50 ppm. The serum LEDBPG spectrum were cut to ROI (δ1.45–0.77 ppm) containing signals from methyl and methylene groups from lipoproteins, and normalized to equal area for additional analysis of the lipid profile.

Spectra sets were mean centered and explored by PCA with random subset cross validation for initial visualization of the data and detection of inherent trends and outliers. Using PLS-DA, a classification model was created on samples from the women with preeclampsia and healthy pregnant groups using the number of LVs giving the smallest classification error, and cross validated by "leave one out" which creates the model on all but one sample, testing it on the remaining sample. Permutation testing with 1000 repeats (random reshuffling of classes), then creating a new predictive model) was done to measure the significance of the predictive model at 95% compared to a classification in arbitrary groups.

Identification and quantification of metabolites

Metabolites were assigned using Bruker AMIX software v.2.5 (Analysis of MIXtures software, Bruker Biospin) and Chenomx v.7.11 (Chenomx Inc., Alberta, Canada), matching spectra to reference databases of metabolites. 2D NMR spectra (JRES,

Table 1. Characteristics of study participants.

Data	PE	PC	NP	p-value
n (samples)	10	10	10	-
Age (years)	29.5 (22–39)	32.6 (28–39)	30.7 (24–39)	>0.05
GA at sampling (week)	35.9 (21.7–37.9)	35.2 (18.6–37.4)	N/A	>0.05
GA at onset (week)	33.2 (21.4–36.9)	N/A	N/A	-
BP sys. (mmHg)	163 (143–174)	120(100–191)	N/A	0.000
BP dia. (mmHg)	104 (96–111)	78 (60–96)	N/A	0.000
Proteinuria[a]	3 (1–4)	0.1 (0–1)	N/A	0.000

Values are given as median (min-max). PE: Women with preeclampsia. PC: Pregnant controls. NP: Non-pregnant controls. GA: Gestational age. BP: Blood pressure. Dia: Diastolic. Sys: Systolic. N/A: Not applicable. Statistical p-values computed by Kruskal-Wallis independent samples test.
[a]Proteinuria measured with dipstick.
doi:10.1371/journal.pone.0091923.t001

COSY, HSQC, and HMBC) were reviewed to confirm assignments. Additional assignments were done with help from literature [7,15]. All identified metabolites were quantified in Chenomx, based on the visible TSP concentration. TSP was quantified in Topspin using the PULCON [16] principle based on a creatine (14.43 mM) spectrum recorded at equal parameters. Serum metabolites were quantified from CPMG spectra, and urine metabolites from NOESY spectra. Concentrations were imported to SPSS v. 20.0.0 (IBM Corp, NY, USA) and subjected to Kruskal-Wallis test of three independent samples. Urine metabolite concentrations were analyzed as [metabolite/creatinine] ratio to correct for dilution. The significance cutoff was set to $p<0.05$ after Benjamini-Hochberg correction [17] of p-value for multiple parallel tests.

Results

Details of the study groups are given in Table 1. Ten women with PE, ten healthy pregnant women and ten non-pregnant women were included. Age and gestational age were matched, and
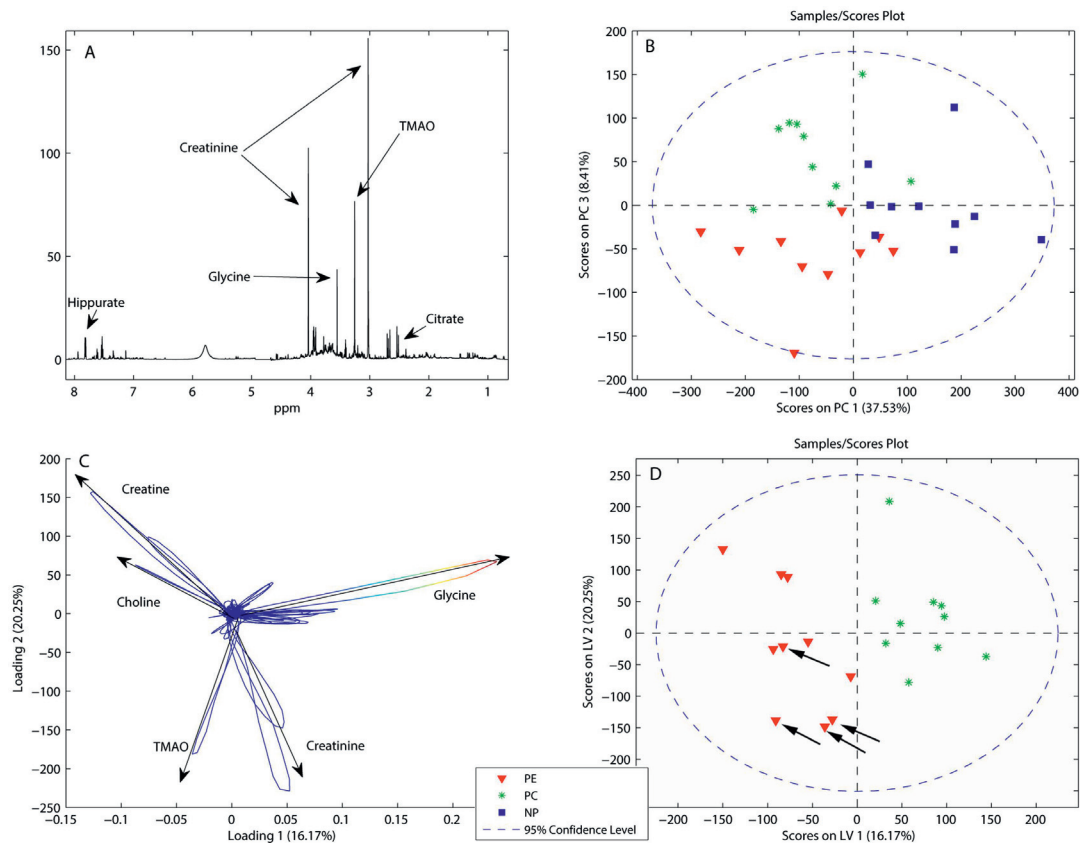
Figure 1. Results from urine analysis. Results from Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA) of urine samples from women with preeclampsia (PE, pregnant controls (PC) and non-pregnant controls (NP). **A**) Typical high resolution NMR spectrum of urine from a PE subject, most abundant metabolites annotated. **B**) PCA score plot separating all three groups in two dimensions. **C**) Loading Variables (LV) 1 and 2 of the PLS-DA used to create a model discriminating between PE and PC groups. Arrow direction indicates increased metabolite level. **D**) Scores on LV1 and LV2 showing a clustering of early onset PE samples (marked by arrows). TMAO: Trimethylamine-N-Oxide.
doi:10.1371/journal.pone.0091923.g001

Table 2. Urine metabolite concentrations.

Metabolite (μM/mM creatinine)	PE	PC	NP	p-value
Glycine*†	260±150	498±219	138±79	0.000
p-Cresol Sulfate*	6.2±4,3	39±13	30±10	0.000
Alanine†	67±57	86±51	21±5	0.000
Threonine†	84±102	94±48	20±5	0.000
Choline†	41±50	10±4	5.3±2.7	0.004
Hippurate*	88±61	265±144	280±188	0.004
Histidine*†	153±129	266±100	79±42	0.004
Asparagine*†	31±42	47±24	13±7	0.004
Isobutyrate*†	21±8	13±7	6±7	0.004
Lactate†	36±27	73±137	9.6±4.4	0.004
Citrate	203±107	473±136	371±217	0.004
Leucine†	9.4±4.8	8.0±1.6	5.0±1.3	0.010
Dimethylamine*	66±13	41±7	40±18	0.015
Trigonelline*	6.3±4.9	16±15	29±22	0.015
2-Oxoglutarate	36±23	40±9	13±10	0.016
Ethanolamine†	75±24	66±19	44±12	0.018
Isoleucine†	6.3±3.8	5.2±1.2	3.1±1.1	0.018
cis-Aconitate†	37±15	39±12	25±5	0.018
Creatine	131±118	76±138	23±25	0.041
Glutamine†	114±88	121±37	70±20	0.041
Glucose*†	38±71	75±44	35± 15	0.049
Valine†	7.6±6.7	6.1±2.0	3.8±1.4	>0.05
Tyrosine	28±19	23±12	14±6	>0.05
N-N-Dimethylglycine	6.0±4.0	5.3±3.0	3.3±1.7	>0.05
Malonate	58±49	156±187	72±105	>0.05
Uracil	8.0±3.8	10.2±2.8	7.1±3.2	>0.05
N-Phenylacetylglycine[a]	41±29	71±35	57±24	>0.05
Betaine	12±8	15±13	8.0±5.2	>0.05
2-Methylglutarate	11±7	16±9	6.9±3.2	>0.05
Guanidoacetate	70±77	77±41	63±39	>0.05
Formate	23±12	35±18	23±13	>0.05
3-Hydroxybutyrate	9.3±7.9	16±15	16±25	>0.05
Acetate	8.5±5.8	11±5	12±9	>0.05
Urea	2957±1211	2251±1077	2795±1845	>0.05
2-Hydroxybutyrate	8.0±3.5	7.4±1.2	5.9±1.3	>0.05
Pyruvate	7.8±4.9	15±15	6.3±2.1	>0.05
TMAO	67±58	54±54	66±78	>0.05
Creatinine[b]	30 ±38	21±13	26±14	>0.05
o-Acetylcholine	2.1±3.1	1.7±2.1	1.4±0.8	>0.05
π-Methylhistidine	46±55	48±67	50±56	>0.05
Phenylalanine	17±13	15±8	14±5	>0.05

Values given as mean [metabolite/creatinine]±sample standard deviation. PE: Women with preeclampsia. PC: Pregnant controls. NP: Non-pregnant controls. TMAO: Trimethylamine-N-Oxide.
[a]As suggested by Chenomx, may instead be phenylacetylglutamine.
[b]Absolute creatinine concentration – not corrected for dilution.
*Significantly different metabolite concentration between PE and PC with a cutoff value at p=0.05 after Benjamini-Hochberg correction using the Kruskal-Wallis test for nonparametric distributions of concentrations for three independent groups.
†Significantly different metabolite concentration between PC and NP.
doi:10.1371/journal.pone.0091923.t002

in accordance with the PE diagnosis, proteinuria and blood pressure were significantly different between the two pregnant groups (not measured in the nonpregnant group).

Results from the urine analyses are shown in Figure 1 and Table 2, with a typical urine spectrum shown in Figure 1A. The urine PCA (Figure 1B) score plot shows clustering of the three groups. PC1 separated samples from both the PE and healthy pregnant groups from the non-pregnant group based on a combination of higher creatinine and trimethylamine-N-oxide (TMAO) levels, and lower glycine levels for the non-pregnant group. PC3 separated preeclamptic women from healthy pregnant women similarly to the subsequent PLS-DA analysis. Creatinine levels were similar between preeclamptic and healthy pregnant women. The PLS-DA model classified urine spectra from preeclamptic and healthy pregnant groups with 95% accuracy (sensitivity = 0.9 and specificity = 1.0) using two LVs (Figure 1C and 1D) and was significant at p ≤ 0.001. The model separated the groups based on a combination of higher choline and creatine levels, and lower glycine levels for preeclamptic women compared to healthy pregnant women. Urine samples from women with early onset PE (<34 weeks) had lower scores on LV2 than the late onset women. (Fig 1D, starred). These had higher TMAO and creatinine, and lower choline and creatine compared to the late onset PE group. The urinary metabolite concentrations are shown in Table 2. The variation in concentrations was similar to the results from the multivariate analyses. Twenty-one metabolites were significantly different between all three groups at p<0.05, with nine metabolite concentrations significantly different between women with PE and healthy pregnant women, and 15 between healthy pregnant women and non-pregnant women. In summary, urine sample spectra from the PE group were clearly different from those of healthy pregnant women based on metabolite content, and a difference in metabolic profile between women with early and late onset PE may exist. Healthy pregnant women also showed a different urinary metabolic profile than non-pregnant women, with higher excretion of amino acids.

Results from the serum analyses are shown in Figure 2 and Table 3, with a typical serum spectrum in Figure 2A. The CPMG and LEDBPG spectra of all serum samples were explored using PCA. A trend of metabolite profiles showing a continuous change from non-pregnant women through healthy pregnant women to women with PE was found, mainly based on increasing total serum lipid content (Figure 2B). All pregnant women had higher serum lipid content than the non-pregnant women, and women with PE had even higher serum lipid content. The distribution of lipoproteins was also different between groups, with the PE group expressing higher signals originating from VLDL and LDL and lower signals from HDL. The signals from the lipoproteins in the NMR spectra consist of several highly overlapping peaks, arising from the lipid moieties within the various lipoproteins. The chemical shifts differ slightly between the particles due to the density differences of the lipoproteins, with lower densities at higher chemical shifts (Figure 2D) [18]. PLS-DA classified serum CPMG spectra from women with PE and healthy pregnant women with 90% accuracy (sensitivity = 0.8 and specificity = 1.0) using four LVs, with heavy reliance on lipid levels for separation between PE and PC (Figure 2C). The lipoprotein distribution was explored further in a PLS-DA on the LEDBPG spectra, as shown in Figure 2D. PE cases were discriminated from healthy pregnant controls by the lipoprotein profile alone, with increased signal in the LDL-VLDL region (the leftmost part of the lipid signal) and decreased signal in the HDL region of the serum spectra. Visible serum metabolites concentrations are shown in Table 3. Serum from women with PE had significantly lower concentrations of
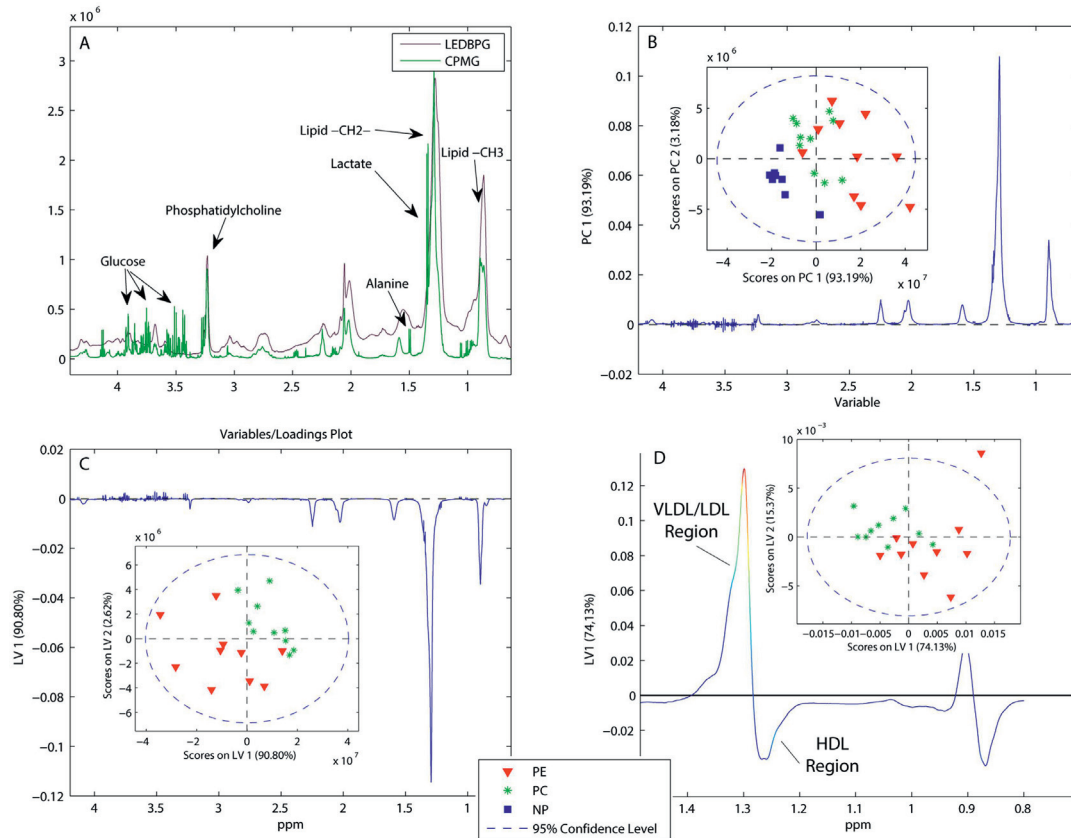
Figure 2. Results from serum analysis. Results from Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA) on the spectra of serum samples from women with preeclampsia (PE), pregnant controls (PC) and non-pregnant controls (NP). ppm: parts per million, resonance frequency of metabolite. **A**) Typical highly resolved serum CPMG (lipids suppressed) and LEDBPG (small metabolites suppressed) spectra from a woman with PE with some annotated metabolites. **B**) Scores plot and loading profile of the PCA separating CPMG spectra of the three groups. **C**) Scores plot and Loading Variable (LV) 1 from the PLS-DA of CPMG spectra showing class discrimination based on lipid level, where women with PE clearly have higher levels of total lipids in the serum compared to pregnant controls **D**) Score plot and LV1 of the LEDBPG showing distinction between PE and PC groups based on lipoprotein distribution. LV1 shows higher levels of VLDL-LDL and lower levels of HDL. VLDL: very low density lipoproteins. LDL: low density lipoproteins. HDL: high density lipoproteins.
doi:10.1371/journal.pone.0091923.g002

histidine than the healthy pregnant women, and non-significant lower levels of formate and higher levels of glycerol. Healthy pregnant women had higher alanine and lactate than the non-pregnant women.

Discussion

The present study clearly demonstrates the metabolic differences between the women with PE and those with healthy pregnancies, in both urine and in serum. The metabolomics method additionally reveals a possible way to subgroup the disease based on metabolic profiles. The metabolic profiles gave new information about possible pregnancy- and disease-induced changes.

The strength of multivariate metabolomic analysis is that the entire visible metabolome is taken into account; and metabolites with small and large variation contribute to the end result. As a

result there are limitations towards finding definite mechanisms correlating to alterations in isolated metabolites. Many small metabolites visible to MR metabolomics are involved in several pathways, and may not be comparable between the urine and serum metabolome.

Differences associated with normal pregnancy compared to non-pregnant women included increased amino acids, choline, and lactate in urine, and higher alanine and lactate in serum. A similar pattern was found in a study by Diaz *et al* [19] following healthy pregnancies with MR metabolomics on urine. The increased excretion of amino acids is suggested to be caused by impaired renal filtration in healthy pregnancies [19]. The choline increase appears as a trend increasing from non-pregnant to healthy pregnant to preeclamptic women. Although choline metabolism appears important to fetal development [20], it is difficult to pinpoint the exact cause of the increase. The increase of lactate in both urine and serum from pregnant women confirms

Table 3. Serum metabolite concentrations.

Metabolite (µM)	PE	PC	NP	p-value
Histidine*	90±26	72±21	57±12	0.036
Formate	19±4	25±4	22±2	0.059
Glycerol	146±37	105±34	92±29	0.059
Alanine†	251±39	303±82	223±31	0.059
Lactate†	907±306	1094±252	774±179	0.094
Creatine	42±14	30±7	35±7	>0.05
Glucose	2455±623	3005±656	2922±466	>0.05
Glycine	135±31	142±20	176±52	>0.05
Valine	132±26	155±31	161±26	>0.05
Acetate	22±4	26±9	32±16	>0.05
Citrate	81±18	70±19	63±13	>0.05
Phenylalanine	38±7	37±10	34±4	>0.05
Glutamine	366±95	375±81	410±60	>0.05
3-Hydroxybutyrate	80±58	47±13	51±34	>0.05
Tyrosine	34±8	39±8	39±10	>0.05
Glutamate	94±28	106±35	95±37	>0.05
Creatinine	54±6	50±7	53±7	>0.05
Leucine	116±18	114±28	109±25	>0.05
2-Methylglutarate	11±3	12±2	11±4	>0.05

Values given as mean±SD. PE: Women with preeclampsia. PC: Pregnant controls. NP: Non-pregnant controls.
*Significantly different metabolite concentration between PE and PC with a cutoff value at p = 0.05 after Benjamini-Hochberg correction using the Kruskal-Wallis test for nonparametric distributions of concentrations for three independent groups.
†Significantly different metabolite concentration between PC and NP.
doi:10.1371/journal.pone.0091923.t003

the findings by other studies [19,21], and follows an increase in prolactin linked to lactation [21].

Women with PE showed increased choline and decreased glycine, p-cresol sulfate and hippurate in urine, which may be related to increased oxidative stress and kidney dysfunction. Choline and glycine are connected through the metabolic pathway of homocysteine [22]. A previous study found elevated choline in serum of women with preeclampsia, and connected the findings to this pathway [20]. Reduced glycine has also been seen with preterm birth [20]. Urinary choline increase has been associated with fetal stress in the second trimester [23]. Glycine is a precursor to glutathione, a tripeptide important for protection against oxidative stress [20]. The decrease in glycine excretion in women with PE could be a result of increased demand for glutathione in response to oxidative stress. p-Cresol sulfate is retained in patients

with kidney damage [24], and is accordingly reduced in the PE group which suffers from kidney dysfunction. p-Cresol sulfate is known to increase oxidative stress in human kidney epithelial cells [24]. Possible consequences of retained p-cresol sulfate could be further increased kidney damage and systemic inflammation in women with PE, thus contributing to the severity of the disease. Hippurate is a metabolic conjugate of glycine, and may be reduced as a consequence of reduced glycine in PE. A relation between hippurate excretion and PE has not been reported previously. The multivariate analysis grouped the urine samples from women with early onset PE together, indicating a similarity in their metabolic profile compared with women with late onset PE, and a difference in phenotype between the two. Such a division was previously found in serum samples from women in early pregnancy [25,26]. However, as there were only four women in this group the results should be interpreted with caution and be followed up in a larger study.

Metabolic profiles in maternal serum revealed significant differences with regards to PE. The major difference detected was the higher total serum lipid content and an increase of VLDL/LDL signals for the PE group. Women with PE also had higher histidine and glycerol levels than women with normal pregnancies. Increased histidine levels were in accordance with the study by Bahado-Singh et al which looked at first-trimester serum [26]. Bolin et al [27] found disturbance in histidine metabolism, with contrarily decreased histidine-rich glycoprotein in serum throughout pregnancies which later develop PE. Although the histidine contained in glycoproteins is different from the free histidine seen in MR spectra, their metabolism may be related. Histidine-rich glycoproteins interact with the coagulation system and angiogenic pathway, and a decrease was found to predict PE in Bolins study [27]. Increased glycerol was detected in the women with PE, similar to the Bahado-Singh study [26], where it was attributed to abnormal lipid metabolism as it forms the backbone of triglycerides. The lipoprotein profiles here shown related to PE are similar to those found for people at risk of CVD [28], with increased low density lipoprotein levels. Lipid dysfunction starts early in pregnancies destined for PE development [28], suggesting that metabolomics may be used to predict the onset of PE. An increase in low- and very-low density lipoprotein has been recorded in patients with CVD and PE previously [28], underscoring the similarities between the two diseases.

The quantification of serum metabolites was done on T2-edited CPMG spectra, where lipid signals are attenuated. Therefore the concentrations of metabolites in serum are not absolute, but comparable between spectra. The multivariate analysis performed on the LEDBPG spectra, where small molecular weight metabolite signals are filtered out, showed that the lipid profile itself was sufficient to distinguish between the two groups.

The study contains relatively few samples, limiting a complete validation procedure. However, a rigorous cross validation was

Table 4. PLS-DA Classification of samples as healthy pregnant or from women with preeclampsia.

Input	LVs	Classification accuracy	Sensitivity	Specificity	AUC	p
Urine spectra	2	95%	0.9	1	0.90	0.001
Serum CPMG spectra	4	90%	0.8	1	0.86	0.002
Serum LEDBPG spectra	4	90%	1	0.8	0.70	0.001

The sensitivity is for detecting a preeclampsia sample using Partial Least Squares Discriminant Analysis. Classification accuracy, sensitivity and specificity are from the leave-one-out cross validation. The p-value is from permutation testing the model with 1000 repeats. LVs: number of loading variables in model. AUC: Area under the Receiver Operator Characteristic curve. CPMG: Lipid-suppressed. LEDBPG: Low molecular weight metabolite suppressed.
doi:10.1371/journal.pone.0091923.t004

performed to ensure that the model was valid also for samples not included in the building of the model. As this is an exploratory study highlighting main differences between groups, cross validation in combination with permutation testing is sufficient to conclude whether there is a difference between groups. Analysis of spectra using PLS-DA is prone to overfitting. However, permutation testing of the urine and serum PLS-DA models (Table 4) revealed them to be significantly different (p<0.05) from models made on random classifications. This indicates that the spectra contained sufficient information to distinguish between samples from women with PE and healthy pregnant women.

The metabolic changes found in cases compared to controls reflect the disease state of the individual. However, it is possible that some of the changes may be evident in the metabolome before the onset of the disease. This possibility must be evaluated in longitudinal studies following women earlier in their pregnancies.

Urine and serum from women with PE, normal pregnancies and non-pregnant women were effectively discriminated by MR metabolomics. Differences were observed related to disease processes and phenotypes. Samples from healthy pregnancies were clearly different from samples collected from non-pregnant women. The observed data suggest that an enlarged study is recommended to find predictive biomarkers earlier in pregnancy or to sub-phenotype the disease, and to investigate the association to later cardiovascular symptoms.

Acknowledgments

Author Contributions

Conceived and designed the experiments: MA RBS ASG RA ACI TFB. Performed the experiments: MA TFB. Analyzed the data: MA RBS ASG RA ACI TFB. Contributed reagents/materials/analysis tools: MA RBS ASG RA ACI TFB. Wrote the paper: MA RBS ASG RA ACI TFB.

References

1. Bergsjø P, Evensen SR, Steinsholt IM (2001) Preeklampsi og eklampsi. Norsk Elektronisk Legehåndbok.
2. Sibai B, Dekker G, Kupferminc M (2005) Pre-eclampsia. Lancet 365: 785–799.
3. Uzan J, Carbonnel M, Piconne O, Asmar R, Ayoubi J-M (2011) Pre-eclampsia: pathophysiology, diagnosis, and management. Vasc Health Risk Manag 7: 467–474.
4. Borzychowski AM, Sargent IL, Redman CWG (2006) Inflammation and pre-eclampsia. Semin Fetal Neonat M 11: 309–316.
5. Redman CWG, Sacks GP, Sargent IL (1999) Preeclampsia: An excessive maternal inflammatory response to pregnancy. Am J Obstet Gynecol 180: 499–506.
6. Craici I, Wagner S, Garovic VD (2008) Preeclampsia and future cardiovascular risk: formal risk factor or failed stress test? Ther Adv Cardiovasc Dis 4: 249–259.
7. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, et al. (2011) The human serum metabolome. PloS one 6: e16957.
8. Lindon JC, Nicholson JK, Holmes E, Everett JR (2000) Metabonomics: Metabolic processes studied by NMR spectroscopy of biofluids. Concept Magnetic Res 12: 289–320.
9. Turner E, Brewster JA, Simpson NAB, Walker JJ, Fisher J (2007) Plasma from women with preeclampsia has a low lipid and ketone body content - A nuclear magnetic resonance study. Hypertens Pregnancy 26: 329–342.
10. Turner E, Brewster JA, Simpson NAB, Walker JJ, Fisher J (2008) Aromatic amino acid biomarkers of preeclampsia - a nuclear magnetic resonance investigation. Hypertens Pregnancy 27: 225–235.
11. Schott S, Hahn J, Kurbacher C, Moka D (2012) 31P and 1H Nuclear Magnetic Resonance Spectroscopy of Blood Plasma in Female Patients with Preeclampsia. International Journal of Biomedical Science 8: 258.
12. Lorentzen B, Øian P, Staff A, Mortensen JHS (2008) Hypertensive svangerskapskomplikasjoner. Veileder i fødselshjelp 2008: Norsk Gynekologisk forening.
13. Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58: 109–130.
14. Savorani F, Tomasi G, Engelsen S (2010) icoshift: a versatile tool for the rapid alignment of 1D NMR spectra. J Magn Res 443: 190–202.
15. Nicholson JK, Foxall PJ, Spraul M, Farrant RD, Lindon JC (1995) 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. Analytical chemistry 67: 793–811.
16. Dreier L, Wider G (2006) Concentration measurements by PULCON using X-filtered or 2D NMR spectra. Magn Reson Chem 44: 206–212.
17. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Annals of Statistics 29: 1165–1188.
18. Bathen TF, Krane J, Engan T, Bjerve KS, Axelson D (2000) Quantification of plasma lipids and apolipoproteins by use of proton NMR spectroscopy, multivariate and neural network analysis. NMR Biomed 13: 271–288.
19. Diaz SO, Barros AS, Goodfellow BJ, Duarte IF, Carreira IM, et al. (2013) Following healthy pregnancy by nuclear magnetic resonance (NMR) metabolic profiling of human urine. Journal of proteome research 12: 969–979.
20. Friesen RW, Novak EM, Hasman D, Innis SM (2007) Relationship of Dimethylglycine, Choline, and Betaine with Oxoproline in Plasma of Pregnant Women and Their Newborn Infants. The Journal of Nutrition 137: 2641–2646.
21. Sachse D, Sletner L, Morkrid K, Jenum AK, Birkeland KI, et al. (2012) Metabolic changes in urine during and after pregnancy in a large, multiethnic population-based cohort study of gestational diabetes. PloS one 7: e52399.
22. Braekke K, Ueland PM, Harsem NK, Karlsen A, Blomhoff R, et al. (2007) Homocysteine, Cysteine, and Related Metabolites in Maternal and Fetal Plasma in Preeclampsia. Pediatr Res 62: 319–324.
23. Diaz SO, Pinto J, Graca G, Duarte IF, Barros AS, et al. (2011) Metabolic Biomarkers of Prenatal Disorders: An Exploratory NMR Metabonomics Study of Second Trimester Maternal Urine and Blood Plasma. J Proteome Res 10: 3732–3742.
24. Watanabe H, Miyamoto Y, Honda D, Tanaka H, Wu Q, et al. (2013) p-Cresyl sulfate causes renal tubular cell damage by inducing oxidative stress by activation of NADPH oxidase. Kidney Int.
25. Bahado-Sing RO, Akolekar R, Mandal R, Dong E, Xia J, et al. (2012) Metabolomics and first-trimester prediction of early-onset preeclampsia. J Matern-Fetal Neo M.
26. Bahado-Singh RO, Akolekar R, Mandal R, Dong E, Xia J, et al. (2012) First-trimester metabolomic detection of late-onset preeclampsia. American journal of obstetrics and gynecology 208: e10–e11.
27. Bolin M, Akerud P, Hansson A, Akerud H (2011) Histidine-rich glycoprotein as an early biomarker of preeclampsia. American journal of hypertension 24: 496–501.
28. Demirci O, Tugrul AS, Dolgun N, Sozen H, Eren S (2011) Serum lipids level assessed in early pregnancy and risk of pre-eclampsia. J Obstet Gynaecol Res 37: 1427–1432.

Paper II

First Trimester Urine and Serum Metabolomics for Prediction of Preeclampsia and Gestational Hypertension: A Prospective Screening Study

Marie Austdal[†,‡], Line H. Tangerås[‡,§], Ragnhild B. Skråstad[ǁ,¶], Kjell Å Salvesen[¶,#], Rigmor Austgulen[§], Ann-Charlotte Iversen[§,*], Tone F. Bathen[†,*]

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

[‡]St. Olavs Hospital, Trondheim University Hospital, 7006 Trondheim, Norway

[§]Centre of Molecular Inflammation Research, and Department of Cancer Research and Molecular Medicine, NTNU, 7491 Trondheim, Norway

[ǁ]Department of Laboratory Medicine Children's and Women's Health, Faculty of Medicine, NTNU, 7491 Trondheim, Norway

[¶]National Center for Fetal Medicine, Department of Obstetrics and Gynecology, St. Olavs Hospital, Trondheim University Hospital, 7030 Trondheim, Norway

[#]Department of Obstetrics and Gynecology, Clinical Sciences, Lund University, 221 00 Lund, Sweden

ABSTRACT

Hypertensive disorders of pregnancy including preeclampsia are major contributors to maternal morbidity. Current prediction methods in first trimester pregnancy have low sensitivity and new predictive biomarkers are warranted. The goal of this study was to evaluate the potential of metabolomics to predict preeclampsia and gestational hypertension from urine and serum samples in early pregnancy, and to elucidate the metabolic changes related to the diseases. Nuclear magnetic resonance spectra were acquired on samples from 599 women at medium to high risk of preeclampsia, and principal component analysis, partial least squares discriminant analysis and variable selection were applied to obtain metabolic profiles of the diseases. Urinary metabolomic profiles predicted preeclampsia and gestational hypertension at 51.3% and 40% sensitivity respectively, at 90% specificity, with hippurate being the most important metabolite. Serum metabolomic profiles predicted preeclampsia at 15% sensitivity and gestational hypertension at 33% sensitivity at 90% specificity with increased lipid levels and an atherogenic lipid profile as most important. Combining maternal characteristics with the urinary hippurate/creatinine level improved the prediction rates of preeclampsia in a logistic regression model. These results show a potential clinical importance for metabolomic analysis of urine samples to predict preeclampsia.

Introduction

Hypertensive disorders of pregnancy including preeclampsia and gestational hypertension are major causes of maternal morbidity and mortality, and affect up to 10% of pregnant women.[1-3] Early identification of women at high risk of preeclampsia might enable potential prophylactic treatment to reduce or avoid the onset of symptoms.[4, 5] Late onset preeclampsia (occurring after 34 weeks of pregnancy) is more common, and has lower detection rate.[6] Predictive models for late onset preeclampsia have employed a combination of maternal

characteristics, biochemical and biophysical markers at 11+0-13+6 weeks of gestation, to predict the syndrome at 30-60% sensitivity.[5-8] New and improved predictive biomarkers are warranted. Gestational hypertension is often included in the disorder spectrum of preeclampsia, particularly if other symptoms are present. Clinical findings in gestational hypertension are often intermediate between normal pregnancy and preeclampsia.[1] In general, placental, renal or hepatic involvement are not present in gestational hypertension, and outcomes are better for mother and baby.[1]

Metabolomics represents a "top-down" view of the metabolism, which more closely characterises the phenotype of the organism than genomic and proteomic applications. Metabolomics is the detection and semi-quantitation of low molecular weight metabolites present in cells, tissues or body fluids, using high throughput analysis platforms such as Nuclear Magnetic Resonance (NMR) spectroscopy or Mass Spectrometry (MS).[9, 10] Recent interest has mounted in the metabolomics approach to predict and characterise preeclampsia. Early and late preeclampsia has been predicted using serum from weeks 11+0-13+6 of pregnancy in combination with maternal markers,[11, 12] and markers of preeclampsia have been found in urine and serum in the second trimester using metabolomics.[13-15] A phase IIa trial for MS based metabolomics prediction of preeclampsia has been launched.[16] To date, no studies have attempted to predict hypertensive disorders in pregnancy by NMR analysis of urine and serum from early pregnancy in a complete cohort of women.

The aim of this study was to evaluate whether metabolic profiles of urine and serum collected from women at gestational week 11+0-13+6 could predict preeclampsia and/or gestational hypertension. Secondly, we aimed to elucidate the metabolic changes that may accompany the early stages of these hypertensive disorders of pregnancy.

Materials and methods

Study population

The study population has been described in detail previously.[17, 18] Briefly, pregnant women who were nulliparous or had preeclampsia or gestational hypertension in a previous pregnancy were invited to attend an examination at 11+0-13+6 weeks of gestation (crown-rump length 45-84mm). At the study visit participants were interviewed about their health and pregnancy. All participants were weighed, and body mass index (BMI) was calculated in kg/m^2. The women were asked to fast for one hour before their visit. Venous blood was drawn into non-heparinised tubes and centrifuged at 1800G for 10 minutes. A serum sample (0.8mL) was separated and stored at -80°C, thawed once and an aliquot of 120µL was stored at -80°C. Spot urine samples were collected at the study visit and aliquots (1.8mL) were stored at -80°C. Blood pressure was measured with a CAS 740 MAX NIBP automated device (CAS Medical systems Inc, CT, USA).[19] The mean arterial pressure (MAP) from the arm with the highest MAP was used. Participants were examined with transabdominal ultrasound with a Siemens ACUSON Antares™ machine (Siemens Medical Solutions Inc, CA, USA), and the uterine artery pulsatility index (UtAPI) was measured.[20] The UtAPI was measured three times on each side, and the average of three measurements on each side was used. The average of the two sides was used in calculations. All scans were carried out by specialized trained midwifes who were certified by the Fetal Medicine Foundation (http://www.fetalmedicine.com). Data on pregnancy outcomes were collected from hospital records. Preeclampsia was defined as systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg in combination with proteinuria ≥ 0.3g per 24 hours measured twice within 4-6 hours, occurring after gestational week 20.[21] Gestational hypertension was defined as for preeclampsia, but occurring without proteinuria. All women gave written informed consent at study entry. The study was approved by the Regional Committee for Medical Research Ethics in mid-Norway, entries REK 2010/102 and 2013/386.

NMR metabolic analyses

Laboratory analyses were performed blinded to pregnancy outcome after all women had delivered their babies. Urine samples were thawed on ice and centrifuged at 6000RPM (Sorvall RMC 14; DuPont) for five minutes. The supernatant (540µL) was mixed with a bacteriostatic buffer (60µL) (pH 7.4, 1.5mM KH_2PO_4 in D_2O, 0.1% Trimethyl-Silyl Propionate (TSP), 2mM NaN_3) (Receipt from Bruker Biospin AG, Reinstetten, Germany) and transferred to 5mm NMR tubes (Bruker Biospin). Serum samples were thawed on ice. Serum (100µL) was mixed with a bacteriostatic buffer (100 µL) (pH 7.4 0.075mM Na_2HPO_4, 5µM NaN_3, 5µM TSP) (Bruker Biospin) and transferred to 3mm NMR tubes.

NMR analysis was performed at the MR Core Facility at the Norwegian University of Science and Technology (NTNU), Trondheim, Norway using a Bruker Avance III Ultrashielded Plus 600MHz spectrometer (Bruker Biospin GmbH, Rheinstetten, Germany) equipped with a 5mm QCI Cryoprobe with integrated, cooled preamplifiers for ^1H, ^2H and ^{13}C. Experiments were fully automated using the SampleJet^{TM} in combination with Icon-NMR on TopSpin 3.1 software (Bruker Biospin). The NMR analyses were performed blinded to pregnancy outcomes. NMR spectroscopy acquisition parameters are described in Supplementary Table S1. One-dimensional Standard Nuclear Overhauser Effect spectroscopy (NOESY) (noesygppr1d; Bruker) spectra were acquired on the urine samples for quantitative detection of small molecular weight metabolites. On the serum samples, Carr-Purcell-Meiboom-Gill (CPMG) (cpmgpr1d; Bruker) spectra were acquired. In CPMG spectra, signals from macromolecules were filtered out for better detection of small molecular weight metabolites. For additional aid in metabolite identification, two-dimensional spectra were acquired (Supplementary Table S1).

Spectra were automatically Fourier transformed, phased and baseline corrected in TopSpin with a line broadening of 0.3Hz. Spectra were imported to Matlab r2013b (The Mathworks Inc, MA, USA). Urine spectral regions containing metabolites of interest (δ0.5 to 9.0ppm) were extracted and the residual H_2O and urea signals removed. Peaks in the urine spectra were aligned by the *i*Coshift algorithm[22] with 245 manually chosen intervals, using the spectrum with the highest correlation as the reference.[23] The spectra were normalised using Probabilistic Quotient Normalization to account for dilution.[24] Finally the data was pareto scaled and mean centered.[25] Serum spectral regions containing metabolites of interest (δ 0.1 to 4.2ppm) were extracted and aligned by the left Alanine doublet peak at 1.48ppm. The spectra were normalised to equal area and mean centred.

Statistical analyses

Statistical analyses were done in Matlab and in Statistical Package for the Social Sciences (SPSS) (version 20.0; SPSS Inc, IL, USA). Baseline characteristics of the study population were tested for normality using the Kolmogorov-Smirnov test.[26] Non-normal data was reported as median (25th-75th percentile), normal data as mean (standard deviation), and categorical data as number (percentage) using either the Kruskal-Wallis test, analysis of variance (ANOVA), or Fishers exact test. The metabolic data was explored for clusters and outliers using PCA as described below. The predictive potential of the metabolic profiles for preeclampsia, gestational hypertension, and both combined was evaluated using PLS-DA. The predictive values of urine and serum NMR spectra were evaluated by sensitivity at 10% false positive rate and at the best cut-off on the Receiver Operator Characteristic (ROC) curve.

PCA is a powerful method of data extraction that finds combinations of variables, called principal components (PCs), describing the main variation in large data. These are visualised in scores and loading plots, making it possible to visualise high dimensional data using only a few dimensions. PLS-DA models the relationship between the spectra and class information

using multivariate regression methods, and is used to establish prediction models. The metabolites responsible for separation between classes are given by the latent variables (LVs).[27] Similar to PCA, the resulting model can be visualised in scores and loading plots. Variable Importance in Projection (VIP) is a method of assessing which variables are most important to the prediction. Variables with VIP scores ≤ 1 can be considered irrelevant to the prediction and excluded.[28] Competitive Adaptive Reweighted Sampling (CARS) is a variable selection method which iteratively selects variable subsets which perform best in cross validated classification.[29]

Multivariate models were constructed using PLS Toolbox 7.3.1 (Eigenvector Research, WA, USA). Preeclampsia, gestational hypertension and both combined were predicted within the entire cohort; e.g., for prediction of preeclampsia the gestational hypertension remained in the control group. The input to the classification models was either full, preprocessed spectra, or sets of variables selected by different variable selection algorithms (VIP or CARS). The classification models were evaluated using a double cross validation procedure.[30] A set of samples (20%) was set aside for independent validation of the model (the outer loop). The remaining samples were split into an inner calibration and test set used for determining the optimal number of PLS components. The inner and outer loops were both repeated 20 times. The data was split so that the ratio of case to control samples was the same in validation, calibration and test sets. Mean sensitivity, specificity and classification accuracy were calculated from the validation set. The PLS-DA classification results were validated using permutation testing, with p-values ≤ 0.05 considered significant.[30] One hundred permutations were built for each predictive model.

Urine metabolite ratios found to be important to the multivariate predictions based on high VIP scores and/or CARS selection (hippurate, lactate, dimethylamine, and 4-deoxythreonic acid to creatinine ratios) were used as input to logistic regression models in combination with

maternal characteristics. The maternal variables found to give the best prediction of preeclampsia in the same cohort by Skråstad et al[17] were combined with selected metabolite ratios, in order to improve the prediction rates for preeclampsia in the cohort. Preeclampsia was predicted using MAP and age as variables in combination with metabolite ratios. Women with maternal age more than 35 or less than 20 were categorised as high risk.[31] The results were compared to prediction rates obtained using UtAPI, MAP and age. Logistic regression analyses were done in SPSS.

Results

A total of 640 women (585 nulliparous and 55 parous women) attended the study visit between 11+0 and 13+6 weeks gestation. A flow chart describing the women included in the analysis is shown in Supplementary figure S1. After exclusions for conditions appearing at or after the study visit as described in[17], and technical reasons (failed acquisitions or missing samples), 599 women remained in total with 587 urine samples and 591 serum samples. One excluded urine sample was from a woman who later developed gestational hypertension.

Characteristics of the study participants for each outcome group are shown in Table 1. Twenty-six women (4.3%) developed preeclampsia and 21 women (3.5%) developed gestational hypertension. One woman had early onset preeclampsia (delivered < 34 weeks gestation). BMI was higher for women with gestational hypertension, and gestational age at birth and birth weights were lower in preeclampsia. MAP measured at 11+0 and 13+6 weeks gestation was higher in women who later developed preeclampsia or gestational hypertension, but below the definition of chronic hypertension.

Identification of 54 urine metabolites and 30 serum metabolites was achieved (Supplementary tables S2 and S3). The exploratory PCA (Figure 1) of the urine and serum NMR spectra showed a characteristic clustering of urine samples from women who later developed preeclampsia or gestational hypertension, indicating a difference in urinary

metabolic profiles between healthy and later hypertensive pregnancies. No apparent clustering was seen in serum samples.
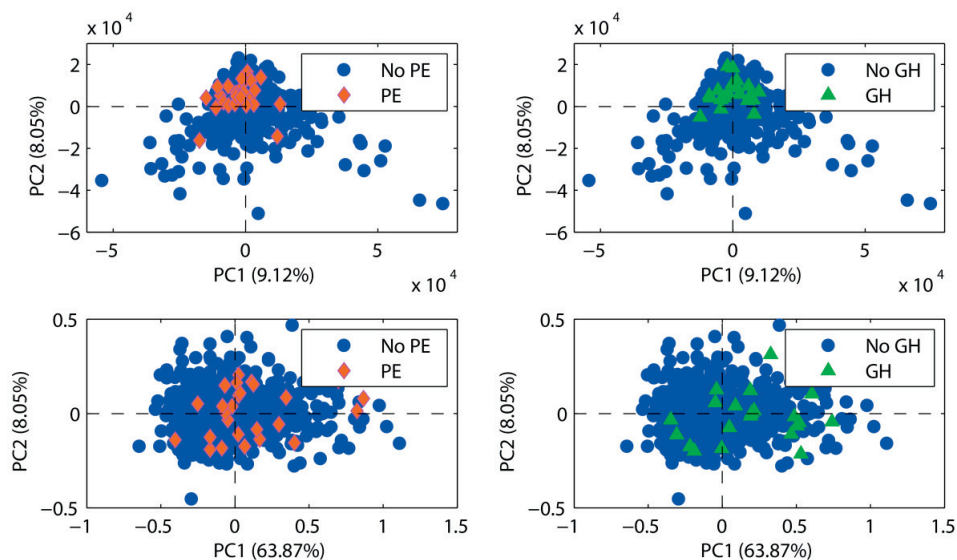
Figure 1. Score plots from PCA of urine and serum spectra from women who later developed preeclampsia or gestational hypertension. PCA score plots of the first and second principal components (PCs) with the explained variance (%), performed on urine samples (top) and serum samples (bottom) taken at 11+0 and 13+6 weeks gestation. Urine samples gave a clustering of women later developing preeclampsia (PE) (red diamonds) or gestational hypertension (GH) (green triangles), while serum samples gave no apparent clustering.

Metabolomic biomarkers in urine

Urine metabolic profiles at 11+0 and 13+6 weeks gestation predicted preeclampsia, gestational hypertension and both combined (Table 2, Supplementary Table S4). At 10% false positive rates using metabolomics analyses, preeclampsia could be predicted at 51% sensitivity from first trimester urine samples, gestational hypertension with 40% sensitivity, and both combined at 37% sensitivity. The loading plots from PLS-DA pinpoint the metabolites that are different between the modelled groups (Figure 2, Table 2). Women that

later developed preeclampsia had increased urinary levels of creatinine, glycine, 4-deoxythreonic acid, α-hydroxyisobutyrate, histidine and dimethylamine and decreased hippurate, lactate and proline betaine. In the first trimester of women developing gestational hypertension there was also a decrease of citrate, but overall a similar profile to preeclampsia.
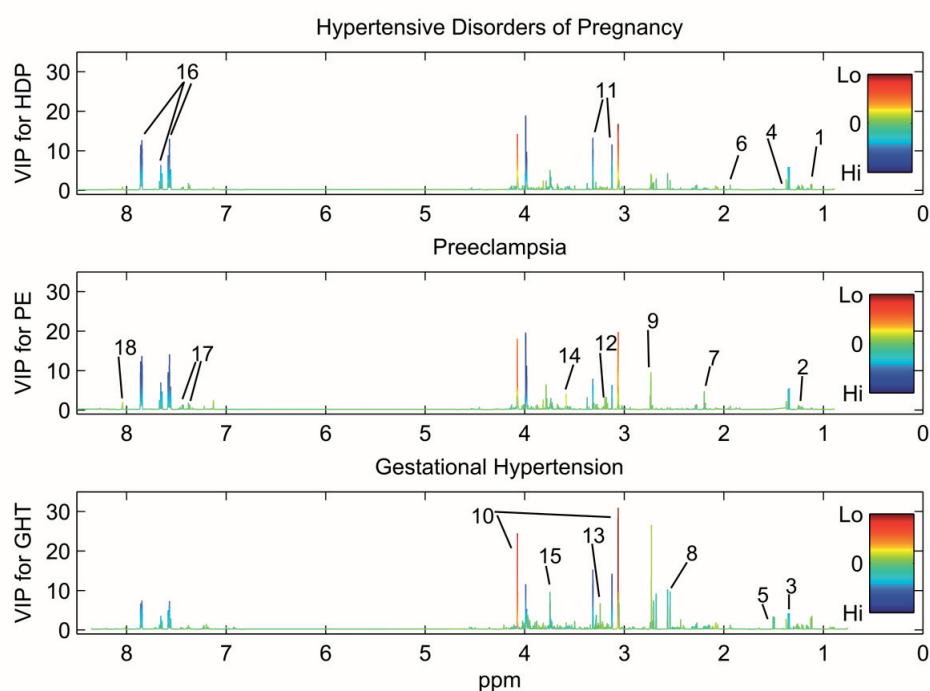
Figure 2. Urine NMR variables involved in predicting preeclampsia (PE) and gestational hypertension (GH) using PLS-DA. The variable importance in projection (VIP) scores for each variable (part per million, ppm) are shown on the vertical axis, with higher VIP scores meaning increasing importance in the predictive model. The variables are coloured by the loadings from the corresponding PLSDA model. Red means increasing levels of metabolite in the indicated condition and blue means decreasing levels. Metabolites by number: 1, 4-Deoxyerythronic acid; 2, 4-Deoxythreonic acid; 3, Lactate and Threonine; 4, α-Hydroxyisobutyrate; 5, Alanine; 6, Acetate and Lysine; 7, Glutamine; 8, Citrate 9,

Dimethylamine; 10, Creatinine; 11, Proline Betaine; 12, Carnitine (tentative); 13, Betaine; 14, Glycine 15, Ascorbic acid; 16, Hippurate; 17, Phenylacetylglutamine; 18, Histidine.

Metabolomic biomarkers in serum

Serum metabolic profiles in first trimester predicted preeclampsia, gestational hypertension and both combined (Table 3, Supplementary Table S5). At 10% false positive rates, 15%, 33% and 30% of preeclampsia, gestational hypertension and both combined could be predicted. The loading plots from PLS-DA pinpoint the metabolites that are different between the modelled groups (Figure 3, Table 3). Mainly, increased lipid levels were evident in both hypertensive groups, and the increased signals originated primarily from triglycerides. Decreased levels of phosphatidylcholines, with signals originating from lipids in HDL, glucose, lactate and alanine were also found important for first trimester prediction of the hypertensive pregnancy disorders.
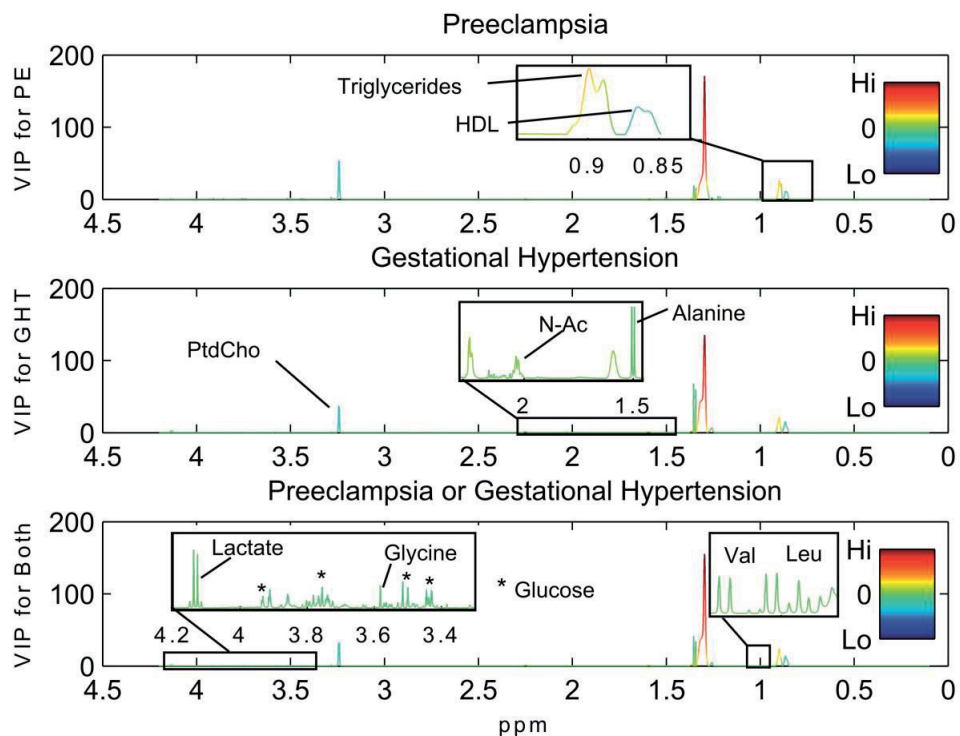
Figure 3. Serum variables involved in first trimester prediction of preeclampsia, gestational hypertension, and both. The variable importance in projection (VIP) scores for each variable are shown in the vertical axis. The variables are coloured by the loadings as described in Figure 3. Abbreviations: HDL, High density lipoprotein; Leu, Leucine; N-Ac, N-acetylated carbohydrate side chains of glycoproteins; PtdCho, Phosphatidylcholine; Py, Pyruvate; Val, Valine.

Prediction of preeclampsia based on a combination of metabolomic and clinical biomarkers

The best logistic regression models for prediction of preeclampsia are shown in Table 4. Urinary hippurate:creatinine combined with MAP and maternal age gave better prediction rates (AUC 0.778) than UtAPI combined with MAP and maternal age (AUC 0.738) (Figure 4).
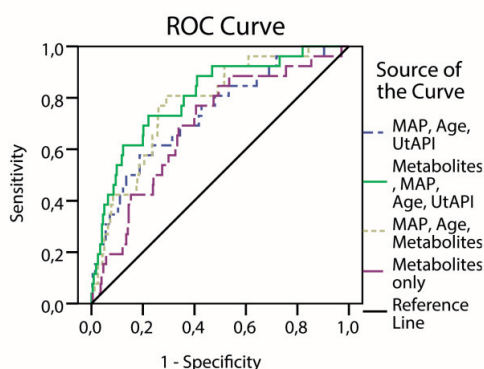
Figure 4: ROC comparison of logistic regression analyses. Prediction of preeclampsia using logistic regression, with risk of preeclampsia as dependent variable and maternal age and MAP in combination with urinary metabolites (Hippurate and Creatinine) or UtAPI as independent variables. Abbreviations: MAP: Mean Arterial Pressure. UtAPI: Uterine Artery pulsatility index.

Discussion

Main findings

Metabolic profiles in urine and serum samples from pregnant women at 11-13 weeks gestation identified a significant difference between women who developed preeclampsia or gestational hypertension, and women with normotensive pregnancies. Both urine and serum metabolic profiles could predict preeclampsia and gestational hypertension, with urine profiles giving the best prediction. Increased urinary creatinine, decreased urinary hippurate, and increased levels of serum lipids were the most important differences between women who developed preeclampsia or gestational hypertension, and normotensive women.

Strengths and limitations

The major strength of our study was the prospective design with complete follow-up of almost 600 women with medium to high risk of preeclampsia. The multivariate statistical analyses were performed up to standards of the field, with rigid cross validation.[32] Weaknesses of our study include the limited numbers of cases, which is difficult to overcome given the low incidence of the disorders. An advantage with multivariate analyses as used in metabolomics approaches is that the covariance between metabolites is taken into account in the modelling. However, in order to more conveniently translate metabolic findings to the clinic, a selection of metabolites could be made which convey most of the information contained in the source. Combining sets of metabolites with maternal characteristics may improve the prediction rates. Previous publications within the metabolomics field have used case-control studies, or samples obtained at a later gestational age, to demonstrate a change in metabolic profile before onset of preeclampsia.[11, 12, 15, 33, 34] This study aimed to predict two hypertensive disorders of pregnancy in a cohort of women and this reflects the predictive power of the metabolomics approach more accurately than in a case-control design. This difference in design may also explain why our results have lower prediction rates than previous metabolic studies.[1]

13

Interpretation

This is the first metabolomics study on first trimester urine samples for prediction of preeclampsia and gestational hypertension. Preeclampsia has been successfully predicted with second trimester urine in a case-control study,[13] and their results partially align with ours for instance by increased 4-deoxythreonic acid, possibly related to ketone body production.[13] Non-metabolomics methods have been used to predict preeclampsia from urine samples, with focus on urinary albumin to creatinine ratios as a measure of kidney function,[35] and on creatinine levels in urine.[36] Increased urinary creatinine was found to be predictive of preeclampsia in a study as early as 8-10 weeks gestation.[36] Our corresponding findings of increased creatinine in maternal urine may be an effect of the increased BMI and MAP of the women who developed preeclampsia and gestational hypertension,[36] or possibly a marker of early renal involvement. Studies have reported increased glomerular filtration rates in pre-hypertensive subjects.[37] Changes in hippurate excretion preceding preeclampsia are novel to this study. Decreased urinary hippurate has been shown to correlate with increased blood pressure,[38] and may be related to diet or to blood pressure related changes in the gut microflora, where this metabolite is produced.[38, 39] The predictive metabolic profile for preeclampsia also included increased glycine and 4-deoxythreonic acid, as well as decreased lactate and creatine. 4-Deoxythreonic acid is a degradation product of 3-hydroxybutyrate,[40] which is increased in maternal serum in our study. Both preeclampsia and gestational hypertension was associated with an increase in urinary dimethylamine at 11-13 weeks getation. This metabolite may have dietary origins,[41] but is also derived from asymmetric dimethylarginine, a biomarker of increased cardiovascular risk.[41, 42].

The prediction of preeclampsia and gestational hypertension using metabolic profiling of urine performed with similar sensitivity as previous approaches using maternal biophysical and biochemical markers on the same cohort[17, 18], and in other studies.[5, 8] The serum metabolic profiling did however not predict the syndromes equally well. This may reflect the lower

sensitivity of NMR spectroscopy to detect small molecular weight metabolites in serum, due to viscosity, lipid signal overlap and lower concentrations of NMR-visible metabolites.

Combining urinary hippurate:creatinine ratios with maternal MAP and age increased the prediction rates of preeclampsia compared to using the uterine artery Doppler measurement with MAP and age. Replacing the UtAPI measurement, which requires skilled ultrasound technicians and time, with an easily accessible urinary marker, would be an advantage especially in low resource areas. However, clinical application of the metabolic profiling prediction method will require confirmation in cohorts from other populations, where the metabolites identified in this study form the basis of prediction. Currently, early identification of women at risk for developing hypertensive disorders of pregnancy would enable closer follow-up of these women.[43] Extensive research is examining potential prophylactic treatment, especially for preterm or severe preeclampsia.[4, 44, 45] In order for these treatments to work, it is important to identify the women at risk for developing disease at an early timepoint. This research shows that there is potentially predictive information contained in a simple urine sample.

First-trimester plasma or serum for prediction of preeclampsia by metabolomics methods have been reported previously.[11, 12, 15, 33] A model with maternal characteristics combined with four serum metabolites predicted late onset PE.[11] These results pointed to disturbed lipid metabolism. MS-based metabolomics studies have found increased serum levels of metabolites in carnitine, fatty acid and lipid classes reliably predicting preeclampsia.[15, 33] However, no previous studies have used metabolomics to predict gestational hypertension in early pregnancy. Changes in lipid metabolism evident early in pregnancies of women who develop hypertensive disorders of pregnancy are known.[46-48] Abnormal lipid metabolism may play a role in the aetiology of the disease.[48] Elevated lipid and low-density lipoprotein levels in maternal serum may induce endothelial dysfunction secondary to oxidative stress.[46] The

decrease in phosphatidylcholine levels related to both later gestational hypertension and preeclampsia may indicate altered choline metabolism.[49] Choline is an essential nutrient which functions in phospholipid metabolism, and choline levels may influence inflammation and angiogenesis.[49-51] Decrease in phosphatidylcholine serum levels have also been found in individuals with increased cardiovascular risk factors.[52] The decrease in maternal serum pyruvate concurrent with the increase in 3-hydroxybutyrate in women who later developed preeclampsia may indicate an early shift in metabolism from glycolysis to ketosis for energy production,[53] and this was reflected also in the urine metabolic profiles.

Conclusions

Metabolic profiling of urine and serum in early pregnancy revealed specific, significant changes in the metabolism of women who later developed preeclampsia or gestational hypertension. Preeclampsia and gestational hypertension could be successfully predicted in early pregnancy using urine and serum metabolic profiles. The consistent changes in the urinary metabolome represent an attractive avenue for clinical prediction, since spot urine is easily accessible. Combining a panel of urine metabolites with maternal characteristics may improve the accuracy of prediction of late onset preeclampsia.

ASSOCIATED CONTENT

Supporting information includes a flow chart of study participants (Figure S.1), NMR spectroscopy parameters (Table S.1), lists of metabolites identified in urine (Table S.2) and serum (Table S.3), and additional PLS-DA model parameters for urine (Table S.4) and serum (Table S.5) prediction of the hypertensive pregnancy disorders. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Corresponding authors. Tone F. Bathen, telephone +47 73551355, tone.f.bathen@ntnu.no; Ann-Charlotte Iversen, telephone +47 72573305, ann-charlotte.iversen@ntnu.no.

Author Contributions

The manuscript was written by contributions from all authors. All authors have given approval to the final version of the manuscript.

ABBREVIATIONS

α-HIB, α-hydroxyidobutyrate; BMI, Body Mass Index; CARS, Competitive Adaptive Reweighted Sampling; Cre, Creatine; Crn, Creatinine;DMA, Dimethylamine; GA, gestational age; GH, Gestational hypertension;Gly, Glycine; Hist, Histidine; Hipp, Hippurate; IQR, Interquartile Range; Lac, Lactate; LV, Latent Variable; NMR, Nuclear Magnetic Resonance; MAP, Mean Arterial Pressure; MS, Mass spectrometry; PAG, Phenylaceylglutamine; PC, Principal Component; PCA, Principal Component Analysis; PE, preeclampsia; PLS-DA, Partial Least Squares Discriminant Analysis; ProlB, Proline Betaine; ROC, Receiver Operator Characteristic; Thr, Threonine; UtAPI, Uterine Artery Pulsatility Index; VIP, Variable Importance in Projection

REFERENCES

1.   Myatt, L.; Redman, C. W.; Staff, A. C.; Hansson, S.; Wilson, M. L.; Laivuori, H.; Poston, L.; Roberts, J. M., Strategy for standardization of preeclampsia research study design. *Hypertension* **2014,** 63, (6), 1293-301.
2.   Sibai, B. M., Diagnosis and management of gestational hypertension and preeclampsia. *Obstet gynecol* **2003,** 102, (1), 181-92.
3.   Roberts, J. M.; August, P. A.; Bakris, G.; Barton, J. R.; Bernstein, I. M.; Druzin, M.; Gaiser, R. R.; Granger, J. R.; Jeyabalan, A.; Johnson, D. D.; Karumanchi, S. A.; Lindheimer, M.; Owens, M. Y.; Saade, G. R.; Sibai, B. M.; Spong, C. Y.; Tsigas, E.; Joseph, G. F.; O'Reilly, N.; Politzer, A.; Son, S.; Ngaiza, K., Hypertens Pregnancy. *Obstet gynecol* **2013,** 122, (5).
4.   Roberge, S.; Villa, P.; Nicolaides, K.; Giguere, Y.; Vainio, M.; Bakthi, A.; Ebrashy, A.; Bujold, E., Early administration of low-dose aspirin for the prevention of preterm and term preeclampsia: a systematic review and meta-analysis. *Fetal diagn ther* **2012,** 31, (3), 141-6.
5.   Akolekar, R.; Syngelaki, A.; Poon, L.; Wright, D.; Nicolaides, K. H., Competing risks model in early screening for preeclampsia by biophysical and biochemical markers. *Fetal diagn ther* **2013,** 33, (1), 8-15.
6.   Poon, L. C.; Kametas, N.; Maiz, N.; Akolekar, R.; Nicolaides, K. H., First-trimester prediction of hypertensive disorders in pregnancy. *Hypertension* **2009,** 53, (5), 812-818.
7.   Scazzocchio, E.; Figueras, F.; Crispi, F.; Meler, E.; Masoller, N.; Mula, R.; Gratacos, E., Performance of a first-trimester screening of preeclampsia in a routine care low-risk setting. *Am J Obstet Gynecol* **2013,** 208, (3), 203.e1-203.e10.
8.   Akolekar, R.; Syngelaki, A.; Sarquis, R.; Zvanca, M.; Nicolaides, K. H., Prediction of early, intermediate and late pre-eclampsia from maternal factors, biophysical and biochemical markers at 11–13 weeks. *Prenat Diagn* **2011,** 31, 66-74.
9.   Fanos, V.; Atzori, L.; Makarenko, K.; Melis, G. B.; Ferrazzi, E., Metabolomics application in maternal-fetal medicine. *BioMed Red Int* **2013,** 2013, 720514.

10.   Lindon, J. C.; Nicholson, J. K.; Holmes, E.; Everett, J. R., Metabonomics: Metabolic processes studied by NMR spectroscopy of biofluids. *Concept Magnetic Res* **2000,** 12, (5), 289-320.

11.   Bahado-Singh, R. O.; Akolekar, R.; Mandal, R.; Dong, E.; Xia, J.; Kruger, M.; Wishart, D. S.; Nicolaides, K., First-trimester metabolomic detection of late-onset preeclampsia. *Am J Obstet Gynecol* **2013,** 208, (1), e10-e11.

12.   Bahado-Singh, R. O.; Akolekar, R.; Mandal, R.; Dong, E.; Xia, J.; Kruger, M.; Wishart, D. S.; Nicolaides, K. H., Metabolomics and first-trimester prediction of early-onset preeclampsia. *J Matern Fetal Neonat Med* **2012**.

13.   Diaz, S. O.; Barros, A. S.; Goodfellow, B. J.; Duarte, I. F.; Galhano, E.; Pita, C.; Almeida Mdo, C.; Carreira, I. M.; Gil, A. M., Second trimester maternal urine for the diagnosis of trisomy 21 and prediction of poor pregnancy outcomes. *J Proteome Res* **2013,** 12, (6), 2946-57.

14.   Pinto, J.; Maciel, E.; Melo, T. S.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. d. C.; Carreira, I. M.; Gil, A. M., Maternal plasma phospholipids are altered in trisomy 21 cases and prior to preeclampsia and preterm outcomes. *Rapid Commun. Mass Spectrom* **2014,** 28, 1635-1638.

15.   Kenny, L. C.; Broadhurst, D. I.; Dunn, W. B.; Brown, M.; North, R. A.; McCowan, L.; Roberts, C.; Cooper, G. J. S.; Kell, D. B.; Baker, N. B., Robust early pregnancy prediction of later preeclampsia using metabolomic biomarkers. *Hypertension* **2010,** 56, 741-749.

16.   Navaratnam, K.; Alfirevic, Z.; Baker, P. N.; Gluud, C.; Gruttner, B.; Kublickiene, K.; Zeeman, G.; Kenny, L. C., A multi-centre phase IIa clinical study of predictive testing for preeclampsia: improved pregnancy outcomes via early detection (IMPROvED). *BMC pregnancy and childbirth* **2013,** 13, 226.

17.   Skråstad, R. B.; Hov, G. G.; Blaas, H.-G. K.; Romundstad, P. R.; Salvesen, K. Å., A prospective study of screening for hypertensive disorders of pregnancy at 11 to 13 weeks in a Scandinavian population. *Acta Obstet Gynecol Scand* **2014,** 93, (12), 1238-47.

18.   Skrastad, R.; Hov, G.; Blaas, H. G.; Romundstad, P.; Salvesen, K., Risk assessment for preeclampsia in nulliparous women at 11-13 weeks gestational age: prospective evaluation of two algorithms. *Br J Obstet Gynaecol* **2014**.

19.   Mancia, G.; Fagard, R.; Narkiewicz, K.; Redon, J.; Zanchetti, A.; Böhm, M.; Christiaens, T.; Cifkova, R.; De Backer, G.; Dominiczak, A.; Galderisi, M.; Grobbee, D. E.; Jaarsma, T.; Kirchhof, P.; Kjeldsen, S. E.; Laurent, S.; Manolis, A. J.; Nilsson, P. M.; Ruilope, L. M.; Schmieder, R. E.; Sirnes, P. A.; Sleight, P.; Viigimaa, M.; Waeber, B.; Zannad, F., 2013 ESH/ESC Guidelines for the management of arterial hypertension. *Blood Pressure* **2013,** 22, (4), 193-278.

20.   Khalil, A.; Nicolaides, K. H., How to record uterine artery Doppler in the first trimester. *Ultrasound Obstet Gynecol* **2013,** 42, (4), 478-9.

21.   Diagnosis and management of preeclampsia and eclampsia. *Int J Obstet Gynaecol* **2002,** 77, (1), 67-75.

22.   Savorani, F.; Tomasi, G.; Engelsen, S., icoshift: a versatile tool for the rapid alignment of 1D NMR spectra. *J Magn Res* **2010,** 443, (1), 190-202.

23.   Giskeodegard, G. F.; Bloemberg, T. G.; Postma, G.; Sitter, B.; Tessem, M. B.; Gribbestad, I. S.; Bathen, T. F.; Buydens, L. M., Alignment of high resolution magic angle spinning magnetic resonance spectra using warping methods. *Anal Chim Acta* **2010,** 683, (1), 1-11.

24.   Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H., Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem* **2006,** 78, (13), 4281-90.

25. van den Berg, R.; Hoefsloot, H.; Westerhuis, J.; Smilde, A.; van der Werf, M., Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **2006,** 7, (1), 142.

26. Massey, F. J., The Kolmogorov-Smirnov test for goodness of fit. *J Amer Statist Assoc* **1951,** 46, (253), 68--78.

27. Wold, S.; Sjöström, M.; Eriksson, L., PLS-regression: a basic tool of chemometrics. *Chemometr Intell Lab* **2001,** 58, (2), 109-130.

28. Chong, I.-G.; Jun, C.-H., Performance of some variable selection methods when multicollinearity is present. *Chemometr Intell Lab* **2005,** 78, (1–2), 103-112.

29. Li, H.; Liang, Y.; Xu, Q.; Cao, D., Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Anal Chim Acta* **2009,** 648, (1), 77-84.

30. Westerhuis, J.; Hoefsloot, H. J.; Smit, S.; Vis, D.; Smilde, A.; Velzen, E. J.; Duijnhoven, J. M.; Dorsten, F., Assessment of PLSDA cross validation. *Metabolomics* **2008,** 4, (1), 81-89.

31. Saftlas, A. F.; Olson, D. R.; Franks, A. L.; Atrash, H. K.; Pokras, R., Epidemiology of preeclampsia and eclampsia in the United States, 1979-1986. *Am J Obstet Gynecol* **1990,** 163, (2), 460-5.

32. Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S., Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics* **2013,** 9, (2), 280-299.

33. Odibo, A. O.; Goetzinger, K. R.; Odibo, L.; Cahill, A. G.; Macones, G. A.; Nelson, D. M.; Dietzen, D. J., First-trimester prediction of preeclampsia using metabolomic biomarkers: a discovery phase study. *Prenat Diagn* **2011,** 31, 990-994.

34. Kuc, S.; Koster, M. P.; Pennings, J. L.; Hankemeier, T.; Berger, R.; Harms, A. C.; Dane, A. D.; Schielen, P. C.; Visser, G. H.; Vreeken, R. J., Metabolomics profiling for identification of novel potential markers in early prediction of preeclampsia. *PLoS One* **2014,** 9, (5), e98540.

35. Baweja, S.; Kent, A.; Masterson, R.; Roberts, S.; McMahon, L. P., Prediction of pre-eclampsia in early pregnancy by estimating the spot urinary albumin: creatinine ratio using high-performance liquid chromatography. *Br J Obstet Gynaecol* **2011,** 118, (9), 1126-1132.

36. Kuromoto, K.; Watanabe, M.; Adachi, K.; Ohashi, K.; Iwatani, Y., Increases in urinary creatinine and blood pressure during early pregnancy in pre-eclampsia. *Ann Clin Biochem* **2010,** 47, (4), 336-42.

37. Palatini, P., Glomerular hyperfiltration: a marker of early renal damage in pre-diabetes and pre-hypertension. *Nephrol Dial Transplant* **2012**.

38. Holmes, E.; Loo, R. L.; Stamler, J.; Bictash, M.; Yap, I. K.; Chan, Q.; Ebbels, T.; De Iorio, M.; Brown, I. J.; Veselkov, K. A.; Daviglus, M. L.; Kesteloot, H.; Ueshima, H.; Zhao, L.; Nicholson, J. K.; Elliott, P., Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* **2008,** 453, (7193), 396-400.

39. Nicholson, J. K.; Holmes, E.; Kinross, J.; Burcelin, R.; Gibson, G.; Jia, W.; Pettersson, S., Host-Gut Microbiota Metabolic Interactions. *Science* **2012,** 336, (6086), 1262-1267.

40. Appiah-Amponsah, E.; Shanaiah, N.; Nagana Gowda, G. A.; Owusu-Sarfo, K.; Ye, T.; Raftery, D., Identification of 4-deoxythreonic acid present in human urine using HPLC and NMR techniques. *J Pharm Biomed Anal* **2009,** 50, (5), 878-885.

41. Tsikas, D.; Thum, T.; Becker, T.; Pham, V. V.; Chobanyan, K.; Mitschke, A.; Beckmann, B.; Gutzki, F. M.; Bauersachs, J.; Stichtenoth, D. O., Accurate quantification of dimethylamine (DMA) in human urine by gas chromatography-mass spectrometry as pentafluorobenzamide derivative: evaluation of the relationship between DMA and its

precursor asymmetric dimethylarginine (ADMA) in health and disease. *J Chromatogr B* **2007,** 851, (1-2), 229-39.

42.   Boger, R. H., Association of asymmetric dimethylarginine and endothelial dysfunction. *Clin chem lab med* **2003,** 41, (11), 1467-72.

43.   Poon, L. C.; Nicolaides, K. H., Early prediction of preeclampsia. *Obstet Gynecol Int* **2014,** 2014, 297397.

44.   Roberge, S.; Giguere, Y.; Villa, P.; Nicolaides, K.; Vainio, M.; Forest, J. C.; von Dadelszen, P.; Vaiman, D.; Tapp, S.; Bujold, E., Early administration of low-dose aspirin for the prevention of severe and mild preeclampsia: a systematic review and meta-analysis. *Am J perinatol* **2012,** 29, (7), 551-6.

45.   Villa, P. M.; Kajantie, E.; Raikkonen, K.; Pesonen, A. K.; Hamalainen, E.; Vainio, M.; Taipale, P.; Laivuori, H.; Predo, S. g., Aspirin in the prevention of pre-eclampsia in high-risk women: a randomised placebo-controlled PREDO Trial and a meta-analysis of randomised trials. *Br J Obstet Gynaecol* **2013,** 120, (1), 64-74.

46.   Enquobahrie, D. A.; Williams, M. A.; Butler, C. L.; Frederick, I. O.; Miller, R. S.; Luthy, D. A., Maternal plasma lipid concentrations in early pregnancy and risk of preeclampsia. *Am J Hypertens* **2004,** 17, (7), 574-581.

47.   Demirci, O.; Tugrul, A. S.; Dolgun, N.; Sozen, H.; Eren, S., Serum lipids level assessed in early pregnancy and risk of pre-eclampsia. J*ournal Obstet Gynaecol Res* **2011,** 37, (10), 1427-32.

48.   Demir, B.; Demir, S.; Atamer, Y.; Guven, S.; Atamer, A.; Kocyigit, Y.; Hekimoglu, A.; Toprak, G., Serum levels of lipids, lipoproteins and paraoxonase activity in pre-eclampsia. *J Int Med Res* **2011,** 39, (1473-2300 (Electronic)).

49.   Jiang, X.; Bar, H. Y.; Yan, J.; Jones, S.; Brannon, P. M.; West, A. A.; Perry, C. A.; Ganti, A.; Pressman, E.; Devapatla, S.; Vermeylen, F.; Wells, M. T.; Caudill, M. A., A higher maternal choline intake among third-trimester pregnant women lowers placental and circulating concentrations of the antiangiogenic factor fms-like tyrosine kinase-1 (sFLT1). *FASEB J* **2013,** 27, (3), 1245-53.

50.   Mehta, A. K.; Singh, B. P.; Arora, N.; Gaur, S. N., Choline attenuates immune inflammation and suppresses oxidative stress in patients with asthma. *Immunobiology* **2010,** 215, (7), 527-34.

51.   Mehedint, M. G.; Craciunescu, C. N.; Zeisel, S. H., Maternal dietary choline deficiency alters angiogenesis in fetal mouse hippocampus. *Proc Natl Acad Sci U S A* **2010,** 107, (29), 12834-9.

52.   Bye, A.; Vettukattil, R.; Aspenes, S. T.; Giskeødegård, G. F.; Gribbestad, I. S.; Wisløff, U.; Bathen, T. F., Serum Levels of Choline-Containing Compounds Are Associated with Aerobic Fitness Level: The HUNT-Study. *PLoS One* **2012,** 7, (7), e42330.

53.   Laffel, L., Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes. *Diabetes Metab Res Rev* **1999,** 15, (6), 412-26.

TABLES

Table 1. Characteristics of the Study Participants

Characteristic	Preeclampsia	Gestational Hypertension	Normotensive pregnancies	P-value[a]
Number of women included	26	21	552	
Age in years, median (IQR)	26 (7)	28 (6)	28 (5)	NS
BMI in kg/m^2, median (IQR)	24.8 (5.6)	27.1 (7.6)	23.5 (4.9)	<0.01
Smoking, n (%)	4 (15.3)	2 (9.5)	63 (11.4)	NS
MAP at enrolment, median (IQR)	87.0 (11.1)	92.1 (8.1)	82.7 (8.8)	<0.001
UtAPI at enrolment, median (IQR)	1.75 (0.70)	1.49 (0.39)	1.46 (0.52)	<0.05
GA at enrolment, days, median (IQR)	92 (4)	92 (5)	90 (6)	NS
GA at delivery, days, median (IQR)	266 (22)	281 (9)	282 (13)	<0.001
Birth weight, g, median (IQR)	3243 (705)	3460 (1096)	3475 (600)	<0.01

Abbreviations: BMI, Body Mass Index; GA, Gestational age; IQR, Interquartile range; MAP, Mean arterial pressure; UtAPI, Uterine Artery Pulsatility Index
[a] P-values calculated by nonparametric Kruskal-Wallis test for continuous variables or Fishers exact test for categorical variables.

Table 2. Urine Metabolite Multivariate Models Predicting Preeclampsia, Gestational Hypertension, or Either.[a]

Hypertensive group	Accuracy (%)	Specificity (%)	Sensitivity (%)	Sensitivity at 10% FPR (%)	P-value[b]	Indicated metabolites[c]
Preeclampsia (n=26)						
Full urine spectra	61.4	65.3	57.5	11.3	<0.01	↑Crn, Gly, α-HIB, Hist, DMA ↓ Hipp, Lac/Thr, ProlB
VIP ≥ 1 variables	68.2	60.1	76.3	23.8	<0.01	↑Crn, Gly, α-HIB, Hist, DMA ↓ Hipp, Lac/Thr
CARS variables	70.8	74.2	67.5	51.3	<0.01	↑ Gly, 4-DEA, DMA ↑ Hipp, Lac, Cre, ProlB
Gestational hypertension (n=20)						
Full urine spectra	59.1	68.2	50.0	11.7	<0.01	↑ Crn, α-HIB, DMA ↓ Hipp, Lac/Thr, ProlB, Citrate
VIP ≥ 1 variables	63.7	65.7	61.7	16.7	0.01	↑ Crn, α-HIB, DMA ↓ Hipp, Lac/Thr, ProlB, Citrate
CARS variables	63.8	89.3	38.3	40.0	0.04	↑ DMA ↓ PAG, Ala
Preeclampsia or gestational hypertension (n=46)						
Full urine spectra	61.5	56.1	66.8	14.4	<0.01	↑ Crn, α-HIB DMA, ↓ Hipp, Lac/Thr, ProlB
VIP ≥ 1 variables	64.0	56.2	71.9	20.0	<0.01	↑ Crn, α-HIB, DMA ↓ Hipp, Lac/Thr, ProlB

CARS variables	66.4 75.9	56.9	36.9	<0.01	↑ α-HIB, DMA ↓ Hipp, PAG, Lys, Ala, noise variables

Abbreviations: 4-DEA, 4-Deoxythreonic acid; α-HIB, α-Hydroxyisobutyrate; Ala, Alanine; Cre, Creatine; Crn, Creatinine; CARS, Competitive Adapted Reweighted Sampling; DMA, Dimethylamine; Gly, Glycine; Hipp, Hippurate; Hist, Histidine; Lac, Lactate; Leu, Leucine; PAG, Phenylacetylglutamine; ProlB, Proline Betaine; Thr, Threonine; VIP, Variable Importance in Projection.

[a] VIP or CARS variable selection was performed, and results were evaluated using accuracy, specificity and sensitivity from double cross validation.

[b] Model validity was estimated by 100 permutation tests.

[c] The metabolites are listed as increased (↑) or decreased (↓) in the hypertensive disease group compared to the normotensive pregnancy group.

Table 3. Serum Metabolite Multivariate Models Predicting Preeclampsia, Gestational Hypertension, or Either[a].

Hypertensive group	Accuracy (%)	Specificity (%)	Sensitivity (%)	Sensitivity at 10% FPR (%)	P-value[b]	Indicated metabolites[c]
Preeclampsia (n=26)						
Full serum spectra	59.4	73.8	45.0	20.0	>0.05	NS
VIP ≥ 1 variables	58.3	70.3	46.3	26.3	>0.05	NS
CARS variables	64.6	65.4	63.8	15.0	0.05	↑ Signals from triglycerides, 3-HB ↓ Pyruvate, PtdCho, Lac
Gestational Hypertension (n=21)						
Full serum spectra	59.1	74.8	43.3	25.0	>0.05	NS
VIP ≥ 1 variables	58.1	75.0	41.3	22.5	>0.05	NS
CARS variables	66.1	55.0	76.9	33.3	0.02	↑ Signals from triglycerides, ↓ Variables corresponding to HDL, Lac, N-Ac, PtdCho, Glc
Preeclampsia or gestational hypertension (n=47)						
Full serum spectra	62.6	70.8	54.4	24.4	0.01	↑ Lipid signals, signals from triglycerides ↓ Signals from HDL, Glc, Val, Leu, Lac, Ala, PtdCho
VIP ≥ 1 variables	63.0	70.4	55.6	27.5	<0.00	↑ Lipid signals, signals from triglycerides ↓ Signals from HDL,

						Glc, Leu, Val, Ala, Lac, PtdCho
CARS variables	64.5	69.1	60.0	30.0	0.02	↑ Variables corresponding to triglycerides ↓ Lac, PtdCho

Abbreviations: 3-HB, 3-Hydroxybutyrate; Ala, Alanine; CARS, Competitive adaptive reweighted sampling; FPR, False positive rate; GH, Gestational hypertension; Glc, Glucose; HDL, high density lipoprotein; Lac, Lactate; Leu, Leucine; N-Ac, N-acetyl Glycoproteins;NS, Not significant PE, Preeclampsia; PtdCho, Phosphatidylcholine; Thr, Threonine; Val, Valine; VIP, Variable importance in projection.

[a] VIP or CARS variable selection was performed, and results were evaluated using accuracy, specificity and sensitivity from double cross validation.

[b] Model validity was estimated by 100 permutation tests.

[c] The metabolites are listed as increased (↑) or decreased (↓) in the hypertensive group compared to the normotensive group.

Table 4. Prediction of Preeclampsia Based on Urinary Metabolites and Maternal Characteristics in Logistic Regression

Variable	AUC (95% CI)	Sensitivity (%)[a]	PPV	NPV	P-value[b]
Urine Metabolites[c] only	0.694 (0.595-0.793)	0.192	0.082	0.960	0.004
MAP, age[d], UtAPI	0.738 (0.637.0.839)	0.346	0.138	0.967	<0.001
Metabolites, MAP, age[d]	0.778 (0.695-0.862)	0.423	0.164	0.971	<0.001
Metabolites, MAP, age[d], UtAPI	0.807 (0.721-0.893)	0.538	0.200	0.977	<0.001

Abbreviations: AUC, area under the receiver operator characteristic curve; CI, confidence interval; MAP, mean arteral Pressure; NPV, negative predictive value; PPV, positive predictive value; UtAPI; uterine artery pulsatility index.

[a] Sensitivity is given at 10% false discovery rate.

[b] Omnibus chi-square significance level of the model.

[c] Metabolites were chosen based on highest VIP scores. The final metabolites were hippurate/creatinine ratio in urine.

[d] Women with maternal age <20 or >35 were categorised as high risk.

TOC GRAPHIC
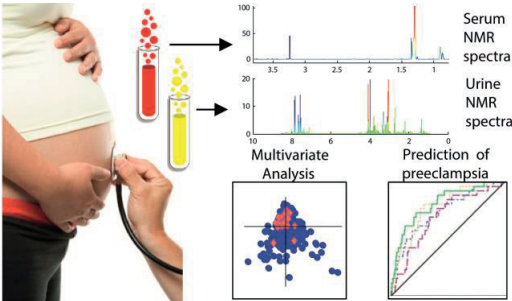
Supplementary information to manuscript titled:

First Trimester Urine and Serum Metabolomics to Predict Preeclampsia and Gestational Hypertension: A Prospective Screening Study

Marie Austdal[†*], Line H. Tangerås[‡§], Ragnhild B. Skråstad[∥¶], Kjell Å Salvesen[¶#], Rigmor Austgulen[§], Ann-Charlotte Iversen[§*], Tone F. Bathen[†*]

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

[‡]St. Olavs Hospital, Trondheim University Hospital, 7006 Trondheim, Norway

[§]Centre of Molecular Inflammation Research, and Department of Cancer Research and Molecular Medicine, NTNU, 7491 Trondheim, Norway

[∥]Department of Laboratory Medicine Children's and Women's Health, Faculty of Medicine, NTNU, 7491 Trondheim, Norway

[¶]National Center for Fetal Medicine, Department of Obstetrics and Gynecology, St. Olavs Hospital, Trondheim University Hospital, 7030 Trondheim, Norway

[#]Department of Obstetrics and Gynecology, Clinical Sciences, Lund University, 221 00 Lund, Sweden

*Corresponding authors. Tone F. Bathen, telephone +47 73551355, tone.f.bathen@ntnu.no; Ann-Charlotte Iversen, telephone +47 72573305, ann-charlotte.iversen@ntnu.no.

```
Attended the 11+0 – 13+6 study visit:
640 pregnancies
```

```
38 were excluded due to conditions apparent at
the first study visit (twins, missed abortions,
CRL > 84 mm, used ASA or low-molecular
weight heparin, para > 0 with no previous PE
or GH)
```

```
Included in the study:
602 pregnancies
```

```
13 were excluded from the analysis due to
conditions appearing after the study visit:
-   2 cases of termination of pregnancy
-   1 case of spontaneous abortion
-   4 cases could not be venipunctured
-   6 cases missing urine samples
```

```
7 were excluded for technical reasons
-   6 failed urine spectrum acquisitions
-   4 failed serum spectrum acquisitions
```

```
Included in the prediction analysis:
-   587 urine samples
-   591 serum samples
```

Figure S.1 Flow chart describing participants included in the study.

Abbreviations: ASA, Acetyl Salisylic acid; CRL, Crown rump length; GH, gestational hypertension; PE, preeclampsia.

Table S.1 NMR spectroscopy parameters

Biofluid	Spectrum shorthand	Pulse sequence	Temp (K)	FID size	Scans	Spectral width (ppm)
Urine	NOESY	noesygppr1d	300	65536	32	20,5682
	JRES	jresgpprqf	300	8192x40	2	16.6602x0.1302
	HSQC	hsqcedetgpsisp.2	300	2048x256	32	16.0194x165.6500
	COSY	cosygpprqf	300	4096x512	8	16.0194x16.0194
	HSQC-TOCSY	hsqcdietgpsisp.2	300	2048x256	32	12.0146x165.6580
	HMBC	hmbcetgpl3ndpr	300	4096x256	128	16.0194x209.9990
Serum	NOESY	noesygppr1d	310	98304	32	29.8927
	CPMG	cpmgpr1d	310	65536	64	20.0243
	JRES	jresgpprqf	310	8192x40	1	16.6602x0.1302

Abbreviations: COSY, Correlated spectroscopy; CPMG, Carr-Purcell-Meiboom-Gill pulse sequence; HMBC, Heteronuclear multiple bond correlation spectroscopy; HSQC, Heteronuclear single quantum coherence spectroscopy; JRES, J-resolved spectroscopy; NOESY, Nuclear overhauser effect spectroscopy; TOCSY, Total correlation spectroscopy.

Table S.2 Metabolites Identified in Urine NMR spectra

Metabolite	^{1}H NMR shift, multiplicity[a], (^{13}C shift)[b]
1-Methylhistidine	3.140m (30.9), 3.730m (36.5), 3.930s (56.0), 7.210t(122.0), 8.220s(140.0), (130)
1-Methylnicotinamide	4.469s(51.4) 8.886d(146.4) 8.957d(150.0)
3-Aminoisobutyrate	1.200t(17.8) 2.037m(46.1)
3-Hydroxyisobutyrate	1.075d(16.5), 2.490m(47.6), 67.6m, (186.9)
3-Hydroxyisovalerate	1.274s(31.0), 2.370s(52.3), (72.5)
3-Methylhistidine	3.270m(27.6), 3.320m(27.6), 3.770m(35.0)
4-Deoxyerythronic acid	1.11d(18.26), 4.08d(78.67), 4.1m(71.54)
4-Deoxythreonic acid	1.236d(21.2), 3.85m(79.1), 4.06m(71.7), (176.8)
4-Hydroxyphenylacetate	3.450s(46.8), 6.850d(118.2), 7.16d(133.0)
Acetate	1.93s(26.1)
Acetoacetate	2.287(32.7), 3.343s(55.5)
Alanine	1.489d(19.1), 3.809q(53.3)
Allantoin	5.390m(65.9)
Ascorbic acid	3.770s(65.3), 4.520d(81.1), (178)
Betaine	3.270 (55.5), 3.890 (69.0)
Carnitine	2.450(45.7), 3.202(56.4), 3.430(73.0)
Choline	3.230(57.1), 3.520(70.1)
Cis-aconitate	3.130(46.2), 5.74(127)
Citrate	2.550d(47.7), 2.700d(47.7)
Creatine	3.040s(39.6), 3.940(56.4)
Creatinine	3.052s(32.7), 4.070s(59.0), (172) (191)
Dimethylamine	2.730s(37.3)
Dimethylglycine	2.933s(46.4)
Formate	8.500s(168.0)
Fumarate	6.528s
Glucose	5.210(94.7)
Glucuronic acid	3.520(74.5), 3.520(78.4), 5.250d(94.7), 3.560(74.0)
Glycine	3.580s(44.2)
Guanidoacetate	3.806(47.9)
Hippurate	3.970d(46.6), 7.550t(131.0), 7.640t(134.0), 7.830t(129.0)
Histidine	3.220(30.1), 3.282(30.1), 4.016(57.0), 7.167s(120.2), 8.060s(138.0) (133)
Hypoxanthine	8.177s(144.0), 8.200s(148.0)
Lactate	1.337d(22.4), 4.127q(71.5), (185)
Lysine	1.449(24.1), 1.524(24.1), 1.740(29.5), 1.910(33.0), 3.030(42.3), 3.790(57.0)
Methylamine	2.617s(27.8)
Methylmalonate	3.158(44.3)
Myo-inositol	3.250(77.1), 3.300(76.9)
p-Cresol sulfate	2.346s(23.1), (133.0), (139.0)
Phenylacetylglutamine	2.27m (34.0) 3.676(45.5), 4.010 (57.3), 7.352t(131.7), 7.362t(130.0), 7.428t(131.6)
Proline betaine	3.110s(48.6), 3.310s(54.8), 2.29m(69),2.17m (79), 2.51m(28.1)
Propylene Glycol	1.150d(20.4), 3.470d(69.1), 3.550d, 3.900m(70.6)
Pyruvate	2.382s(29.3)
Scyllo-inositol	3.366s(76.6)

Succinate	2.363s(36.8)
Tartaric acid	4.350S(76.7)
Taurine	3.280(50.3), 3.440(38.4)
Threonine	1.337d(22.4), 4.273q(68.8), (182.4)
Trigonelline	8.835t(148.0), 9.121s(148.4)
Trimethylamine-N-Oxide	3.280(62.5)
Tyrosine	6.900(118.0), 7.190(133.7)
Valeric acid	0.880m(15.9), 0.930m(15.9), 1.560(28.0), 2.200(40.0)
Valine	0.995d(19.5), 1.050d(20.9), 2.276(34.4), (63.3)
α-Hydroxyisobutyric acid	1.365s(29.5), (76.5), (187.1)
α-Ketoglutarate	3.01(57.3)

[a] Multiplicities annotated as s, singlet; d, doublet; t, triplet; q; quartet; m, multiplet

[b] ^{13}C chemical shifts directly coupled to the ^{1}H in Heteronuclear Single Quantum Coherence spectra.

Table S.3 Metabolites Identified in Serum NMR Spectra

Metabolite	^{1}H Shifts[a] and multiplicities
3-Hydroxybutyrate	1.219d
Acetate	1.935s
Acetoacetate	2.300s
Acetone	2.248s
Alanine	1.499d
Asparagine	2.871d, 2.951d
Citrate	2.554d, 2.701d
Creatine	3.059s, 3.948s
Creatinine	3.062s, 4.072s
Formate	8.478s
Glucose	3.269t, 3.403-3.576m, 3.716-3.939m
Glutamine	2.152m, 2.475m, 3.772m
Glycerol	3.677m
Glycine	3.58s
Histidine	7.808s, 7.081s, 3.14m
Isoleucine	1.028d
Lactate	1.347d, 4.131q
Leucine	0.983d, 0.963d, 1.734m
Lysine	1.916m
Methanol	3.38s
Methionine	2.662t
N-Acetylated groups	2.059s
Phenylalanine	7.449t, 7.393t, 7.352d
Proline	2.37m, 3.354m
Propylene Glycol	1.139d
PtdCho	3.238b
Pyruvate	2.39s
Threonine	3.593d, 1.347d, 4.271q
Tyrosine	7.217d, 6.920d
Valine	1.06d, 1.009d, 3.63d

[a] Multiplicities annotated as s, singlet; d, doublet; t, triplet; q; quartet; m, multiplet; b, broad signal

Table S.4 Additional prediction model parameters for urine spectra

Patients	Variable selection method	No. variables	No. latent variables in prediction model	p
PE (n=26)	None	27084	1	<0.01
	VIP ≥ 1	3747	1	<0.01
	CARS	54	1	<0.01
GH (n=20)	None	27084	1	<0.01
	VIP ≥ 1	4653	1	0.01
	CARS	8	1	0.04
PE+GH (n=46)	None	27084	1	<0.01
	VIP ≥ 1	3789	1	<0.01
	CARS	45	4	<0.01

Abbreviations: CARS, Competitive adaptive reweighted sampling; GH, Gestational Hypertension; PE, preeclampsia; VIP, Variable importance in projection.

Table S.5 Additional prediction model parameters for serum spectra

Patients	Variable selection method	No. variables	No. latent variables in prediction model	p
PE (n=26)	None	13420	4	>0.05
	VIP ≥ 1	1081	7	>0.05
	CARS	105	3	0.05
GH (n=20)	None	13420	1	>0.05
	VIP ≥ 1	1259	7	>0.05
	CARS	87	4	0.02
PE+GH (n=46)	None	13420	4	0.01
	VIP ≥ 1	1328	4	<0.00
	CARS	51	6	0.02

Abbreviations: CARS, Competitive adaptive reweighted sampling; GH, Gestational Hypertension; PE, preeclampsia; VIP, Variable importance in projection.

Paper III

Metabolic profiles of placenta in preeclampsia using HR-MAS MRS metabolomics

Marie Austdal[a,b], Liv Cecilie Vestrheim Thomsen[c,d,e], Line Haugstad Tangerås[b,e], Bente Skei[e], Seema Mathew[c], Line Bjørge[c,d], Rigmor Austgulen[e], Tone Frost Bathen[a], Ann-Charlotte Iversen[e]

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

[b] St. Olavs Hospital, Trondheim University Hospital, 7006 Trondheim, Norway

[c] Department of Gynecology and Obstetrics, Haukeland University Hospital, 5021 Bergen, Norway

[d] Department of Clinical Science, University of Bergen, 5021 Bergen, Norway

[e] Centre of Molecular Inflammation Research, and Department of Cancer Research and Molecular Medicine, NTNU, 7491 Trondheim, Norway

Author email addresses: marie.austdal@ntnu.no, liv.vestrheim@k2.uib.no , line.tangeras@ntnu.no, bente.skei@ntnu.no, Seema Mathew, line.bjorge@k2.uib.no , rigmor.austgulen@ntnu.no, tone.f.bathen@ntnu.no, ann-charlotte.iversen@ntnu.no

Corresponding author: Ann-Charlotte Iversen, ann-charlotte.iversen@ntnu.no, +47 93283877, Centre of Molecular Inflammation Research, and Department of Cancer Research and Molecular Medicine, NTNU, 7491 Trondheim, Norway

23 **Abstract**
24 **Introduction:** Preeclampsia is a heterogeneous gestational disease characterized by maternal
25 hypertension and proteinuria, affecting 2-7% of pregnancies. The disorder is initiated by
26 insufficient placental development, but studies characterizing the placental disease
27 components are lacking. **Methods:** Our aim was to phenotype the preeclamptic placenta using
28 High-resolution magic angle spinning nuclear magnetic resonance spectroscopy (HR-MAS
29 MRS). Placental samples collected after delivery from women with preeclampsia (n=19) and
30 normotensive controls (n=15) were analyzed for metabolic biomarkers including amino acids,
31 osmolytes, and components of energy and phospholipid metabolism. The metabolic
32 biomarkers were correlated to clinical characteristics and maternal serum inflammatory
33 biomarkers. **Results:** Principal component analysis showed inherent differences in placental
34 metabolic profiles between preeclamptic and normotensive pregnancies. Significant
35 differences in metabolic profiles were found between placentas from severe and non-severe
36 preeclampsia, but not between pregnancies with and without fetal growth restriction. The
37 placental metabolites correlated with the placental stress marker sFlt-1, suggesting a variation
38 in placental stress signaling between different placental phenotypes. **Discussion:** HR-MAS
39 MRS is a sensitive method for defining the placental disease component of preeclampsia,
40 identifying several altered metabolic pathways. Placental HR-MAS MRS analysis may
41 provide improved insights into which processes are disturbed in placental subtypes of
42 preeclampsia, and introduce a new long-required tool for more detailed and sensitive placental
43 phenotyping of this heterogeneous disease.

44 **Keywords:** metabolism; metabolomics; NMR; preeclampsia; placenta; profiling.

45 **Abbreviations:** BMI, body mass index; CS, cesarean section; FGR, fetal growth restriction;
46 GA, gestational age; GLM, general linear model; HDL, high-density lipoprotein, HR-MAS
47 MRS; high-resolution magic angle spinning magnetic resonance spectroscopy; hs-CRP, high
48 sensitivity C-reactive protein; MSEA; metabolite set enrichment analysis; PCA, principal
49 component analysis; PE; preeclampsia; PLS-DA, partial least squares discriminant analysis;
50 sFlt-1, soluble fms-like tyrosine kinase receptor 1

51 **Highlights**:

52 • The largest metabolic profiling of placentas to date reveals differentially expressed
53   metabolic pathways in preeclamptic and normotensive placenta
54 • The preeclampsia specific metabolite changes include taurine and nutrient depletion,
55   and enrichment of cell membrane breakdown products.
56 • HR-MAS MRS is a novel sensitive method for placental tissue profiling with limited
57   need for sample preparation

58 Word count: 2999 (excl footnotes)

59

Introduction

60 **Introduction**

61 Preeclampsia is a gestational disease that originates in the placenta and affects 2-7% of
62 pregnancies [1]. Preeclampsia may be an excessive maternal inflammatory response to
63 insufficient placentation or to pregnancy itself [2, 3]. The current hypothesis in regard to its
64 development states that the uterine spiral arteries develop insufficiently during placentation,
65 causing placental ischaemia and abnormal inflammation as the pregnancy develops [1, 4]. The
66 oxidatively stressed placenta releases increasing amounts of inflammatory and angiogenic
67 factors to the maternal circulation, eventually causing the clinical manifestations of
68 preeclampsia; endothelial dysfunction, intravascular inflammation and activation of the
69 hemostatic system evidenced by proteinuria and hypertension [1, 4].

70 Preeclampsia is a heterogeneous disease, but clinical markers in the placenta identifying
71 disease subgroups are scarce [5]. Common subtypes of preeclampsia are today defined by
72 severity of maternal features, time of diagnosis, and presence of fetal growth restriction
73 (FGR) [6, 7]. These subgroups are based on end stage maternal or fetal factors and placental
74 histology findings, and maternal serum markers overlap between groups [2]. Similar end stage
75 presentations of preeclampsia may stem from different pathologic placental processes [8].
76 Studying the altered metabolism in the preeclamptic placenta and how it is affected with
77 disease severity and when combined with FGR may give insight into which processes are
78 shared and specific for these disorders. Placental phenotyping by metabolic expression may
79 provide better insights into the initial cause of the disease and which processes are disturbed
80 during and after its development, and help identify potential targets for treatment in the future.

81 Metabolite expression is the final level of regulation over gene and protein expression, and
82 can be measured directly in tissue samples using high resolution magic angle spinning
83 magnetic resonance spectroscopy (HR-MAS MRS) [9]. HR-MAS MRS has been used
84 successfully in investigating molecular subtypes of breast cancer for improved treatment and
85 outcome stratification [10]. To our knowledge, HR-MAS MRS of intact placental tissue has
86 not previously been performed on normal or preeclamptic pregnancies. Additionally,
87 correlating known placental stress and inflammatory markers in maternal serum with the
88 placental phenotype may improve understanding of the "missing link" between the placental
89 and maternal manifestations of preeclampsia.

90 We hypothesize that placental metabolic profiling provides a sensitive method for detailed
91 identification of the placental component of preeclampsia. This may provide information
92 about underlying pathogenic processes reflecting the initial stages of the syndrome. We
93 further aim to investigate whether the placental metabolites correlate to maternal serum
94 measurements of placental stress and inflammatory markers.

95

Materials & Methods

Study participants, serum samples and placental tissue biopsies

The study was approved by the Norwegian Regional Committee for Medical and Health Research Ethics (REC 2012/1040) and informed consent was obtained from all participants. Complete clinical characteristics of the pregnancies and placentas studied are available in the Supplementary information, Tables S.1 and S.2 [11]. Women with singleton pregnancies delivering by cesarean section (CS) were recruited at Haukeland University Hospital (Bergen, Norway) from 2009 to 2013. Preeclampsia was defined as persistent hypertension (systolic/diastolic blood pressure 140/90 mmHg) plus proteinuria ($\geq$0.3g/24h or $\geq$1+ by dipstick developing after 20 weeks gestation [6]. Superimposed preeclampsia was defined as pre-existing hypertension plus proteinuria developing after 20 weeks of gestation [6] and was included in the preeclampsia group. Preeclampsia was subclassified as severe if diagnosed with one or more of the severe features (Supplementary Table S.3), or else designated as non-severe [7]. FGR was diagnosed by serial ultrasound measurements showing reduced intrauterine growth. In absence of serial ultrasound measurements, neonates were defined with FGR if their birth weights were <5th percentile for gestational age (GA) according to Norwegian fetal weight reference curves [12].

Women with pregnancies complicated by preeclampsia were included as cases. Healthy pregnant women with no previous history of pregnancies with preeclampsia or FGR were included as normotensive controls. Women with gestational hypertension, HELLP syndrome, or pregnancies with fetal chromosomal or congenital abnormalities were not included in the study. The preeclamptic women underwent CS due to the main disease, while normotensive women underwent CS for unrelated reasons (i.e. breech position, previous CS or maternal request). Only pregnancies without labor activity were included. Information about the pregnancy and a familial history of preeclampsia were collected from medical journals and interviews.

Maternal venous blood was collected prior to CS, left to clot for $\leq$30 minutes, centrifuged at 1800G for 10 minutes, and serum aliquots (1mL) stored at -80°C until analysis. A tangential section (100mg) from the maternal central side of the placenta was collected after delivery, placed in a cryotube and frozen either in liquid nitrogen or directly at -80°C within 101 ± 49 minutes (mean ± SD) after delivery by CS.

Maternal serum analyses

Soluble fms-like tyrosine kinase receptor 1 (sFlt-1) was measured in duplicate using a quantitative sandwich ELISA according to the manufacturer's instructions (#DVR100B, R&D Systems, Abingdon, UK). High sensitivity C-reactive protein (hsCRP) (turbidimetric assay, Modular P analyzer, Roche, Burgess Hill, UK), total cholesterol, high-density lipoprotein (HDL), triglyceride and creatinine (enzymatic colorimetric assays, Modular P analyzer) were measured by accredited methods at the Department of Clinical Chemistry, St. Olavs Hospital, Trondheim, Norway.

HR-MAS MRS placental analyses

The placental biopsies were analyzed in random order, blinded to pregnancy outcome. Samples were prepared on a metal plate bathed in liquid nitrogen in order to minimize the effect of tissue degradation. Biopsies (7.5 ± 1.4 mg) were cut to fit 30μL disposable inserts (Bruker Biospin Corp, USA) filled with 3μL D_2O containing 25mM formate for shimming. Spin-echo spectra were acquired on a Bruker Avance DRX600 spectrometer with a $^{1}H/^{13}C$

141 MAS probe with gradient (Bruker Biospin GmbH, Germany) using the following parameters:
142 5KHz spin rate, 5°C probe temperature, cpmgpr1d pulse sequence (Bruker Biospin) with
143 78ms total echo time, spectral width of 20 ppm and 256 scans. Two samples were additionally
144 analyzed by ^{13}C-^{1}H spectroscopy (HSQC, HMBC) for aid in metabolite identification.

Data analysis
146 Spectra were Fourier transformed into 65.5k points following 0.3 Hz line broadening, and
147 automatically phased and baseline corrected. The left alanine peak was set to 1.478ppm
148 (Topspin 3.1, Bruker Biospin). The spectra were peak aligned using iCoshift [13], and
149 normalized to unit area to account for sample size differences. Metabolites were identified by
150 comparing chemical shift values to spectral databases [14] and correlation of metabolite peaks
151 using Statistical Total Correlation Spectroscopy [15]. Semi-quantitative metabolite levels
152 were measured by integrating the spectral regions of identified metabolites (Matlab r2013b,
153 The Mathworks Inc., Natick, MA, USA).

154 Multivariate analysis was performed in PLS Toolbox v. 7.3.1 (Eigenvector Research Inc.,
155 WA, USA), and with MetaboAnalyst [16]. Quantified metabolite levels were autoscaled
156 before multivariate modeling. The metabolic profiles were evaluated by principal component
157 analysis (PCA) for initial data exploration [17]. Partial least squares discriminant analysis
158 (PLS-DA) was used to discriminate metabolic features between subgroups of preeclampsia
159 (presence of severe maternal features or FGR), and between normotensive and preeclamptic
160 placentas. Results were evaluated by five-fold cross-validation which was repeated twenty
161 times. Classification accuracy, sensitivity, specificity and number of latent variables were
162 reported for cross-validated results. To evaluate the validity of the regression and
163 classification results, 1000 permutation tests were performed with models considered valid at
164 $p \leq 0.05$ [18].

165 Univariate statistical analyses were performed in SPSS v. 20 (SPSS, Chicago, IL). Clinical
166 characteristics were compared between disease subgroups using the one-way ANOVA with
167 Tukey post-hoc test for groupwise normally distributed data (GA, birth weight) or the Kruskal
168 Wallis test with post-hoc pairwise Mann-Whitney U tests for nonparametric data (maternal
169 age, blood pressure, body mass index (BMI), placental weight, parity, metabolite levels,
170 serum measurements), or Fishers exact test for categorical variables. Metabolite levels were
171 compared using the Mann-Whitney U test. False discovery rate correction for multiple testing
172 (Benjamini-Hochberg) was applied [19]. To adjust for confounding effects of GA, linear
173 regression models with log-transformed metabolites as dependent variable and preeclampsia
174 and GA as independent variables were generated. Interaction terms between preeclampsia and
175 covariates were included in the models if significant, otherwise excluded. Standardized
176 residuals were assessed with the Kolmogorov Smirnov test. Quantitative metabolite set
177 enrichment analysis (MSEA) was performed for inferring the metabolic pathways associated
178 with disease [20].

179

180

181

Results

Clinical characteristics

The clinical characteristics of the study participants are described in Table 1. A total of 34 pregnant women were included in the study (preeclampsia; n=19 and normotensive; n=15). Birth weight and GA were lower in women with preeclampsia than normotensive pregnancies, and preeclamptic women were more likely to be primiparous. Maternal serum levels of creatinine, uric acid and sFlt-1 were significantly higher before CS in women with preeclampsia compared to normotensive (Table 2).

HR-MAS analyses

In total 25 metabolites were identified in the HR-MAS NMR analysis of placental biopsies (Table 2, Supplementary Table S.4). Median spectra from the normotensive and preeclamptic placental biopsies are shown in Figure 1. The area under the metabolite peak corresponds to the concentration in the tissue. Sample processing times were not significantly different between the preeclamptic and normotensive groups.

Placental metabolic profiles in preeclampsia

PCA was performed for an overview of the metabolic profiles (Fig. 2). The score plots show the between-samples variation and are colored by ellipses of 95% confidence intervals for preeclampsia subgroups defined by presence or absence of preeclampsia (Fig. 2A), severe or non-severe preeclampsia (Fig. 2B) or preeclampsia with or without FGR (Fig. 2C). Subdivision by early- and late onset disease was not explored because only two of the preeclampsia cases presented after 34 of weeks gestation. The women with superimposed preeclampsia appeared in the overlapping region of the score plot between preeclampsia and normal pregnancy (Fig. 2A). Following initial PCA, one sample (severe preeclampsia with FGR) was excluded from further analysis due to high levels of lipids, which obscured the baseline and interfered with the quantification of metabolites. One sample (non-severe preeclampsia) was removed because it contained very low levels of all metabolites.

Normotensive pregnancies were clearly separated from preeclamptic pregnancies based on the placental metabolic expression. The preeclamptic pregnancies showed a more heterogeneous metabolic expression than the more unified group of normotensive pregnancies (Figure 2A).

The preeclampsia subgroups defined by severe preeclampsia or FGR largely overlapped in metabolic expression (Fig. 2B and C). The severe preeclampsia placental profiles clustered closer to the normotensive profiles than preeclampsia without severe features (Fig. 2B). In contrast, placental metabolic profiles of preeclampsia with presence of FGR appeared more separated from normotensive profiles than for preeclampsia without FGR (Fig. 2C), although both had wide distributions. The loading plot (Fig. 2D) shows the metabolites contributing to the sample distributions; increasing aspartate, phosphocholine and glycerophosphocholine and decreasing glutamate, taurine, ascorbate and glutamine corresponded to a preeclamptic phenotype.

PLS-DA defined metabolic profiles for preeclampsia and its subgroups (Table 3). A significant difference in metabolic profile was found between preeclamptic and normotensive placenta, and for subgroups between placenta from severe and non-severe preeclampsia. There was no significant difference in the metabolic profile between preeclampsia with or without FGR. PLS regression revealed correlations between the placental metabolic profiles

6

225 and maternal serum markers of disease; triglycerides, sFlt-1, uric acid, and creatinine (Table
226 4). Serum sFlt-1 showed the highest correlation to placental metabolites (R^2=0.49, p<0.001).
227 Of special interest, maternal serum sFlt-1 was clearly correlated with increased placental
228 glycerophosphocholine levels.

229 Placentas from preeclamptic pregnancies showed enrichment of phospholipid biosynthesis
230 and depletions in bile acid biosynthesis, taurine metabolism, ammonia and urea cycles and
231 protein biosynthesis, compared to placentas from normotensive pregnancies (Supplementary
232 Table S.5).

233 Metabolites were evaluated by univariate analysis to confirm the multivariate analyses (Table
234 2). Twelve out of 25 metabolites were significantly different between the placentas from
235 preeclamptic and normotensive pregnancies after correction for multiple testing. Placental
236 levels of choline and lysine were increased in severe preeclampsia (Supplementary Table
237 S.6).

238 **Assessment of gestational age**
239 Following linear regression modeling of the metabolites differing between preeclamptic and
240 normotensive placentas, GA was found to significantly affect only the levels of ethanolamine
241 and glycine (increasing with GA). Glutamine, glutamate, taurine, valine, 3-hydroxybutyrate
242 and ascorbate remained significantly different between preeclamptic and normotensive groups
243 after correction for GA (Supplementary information, Table S.7).

244

245

246

247

248

Discussion

We have analyzed the metabolic profile of placental tissue from preeclamptic and normotensive pregnancies using HR-MAS MRS. The placentas showed a highly significant altered metabolic state in preeclampsia. Preeclampsia is a heterogeneous disease, clearly reflected by the scattering of the metabolic profiles compared to normotensive pregnancies (Figure 2). Our study demonstrated metabolic differences between severe and non-severe preeclampsia, and showed that the presence of FGR was not reflected in the placental metabolites. The metabolic placenta profile also correlated with maternal serum markers for angiogenic imbalance, kidney function and lipid levels. Finally, we have analyzed the effect of GA, a common confounder in placental studies comparing preeclamptic and normotensive pregnancies. This is the largest metabolic profiling of placental tissue from preeclamptic pregnancies published to date, and the first using HR-MAS MRS technology, with a potential to reveal the placental component of preeclampsia.

Our study confirms and expands upon previous *in vitro* studies on placental explants and cell culture [21-23]. Metabolic profiling of whole placental tissue by HR-MAS MRS is a nondestructive analysis of the placental component of preeclampsia, which allows further proteomic or genomic analysis of the same sample. Studies in cancer have found metabolic profiles to correlate to prognostic factors and survival [24]. We have shown placental metabolites correlating with the disease severity and with maternal serum factors reflecting the placental disease. The heterogeneous nature of the metabolic distribution in preeclamptic pregnancies indicates that this holistic approach may provide a sensitive phenotyping of the placental component of preeclampsia. This provides a novel tool for understanding the underlying mechanisms, and thereby subgrouping the disease based on placental involvement, not only end stage maternal and fetal features. The findings warrant further investigation in a larger cohort.

Knowledge of the metabolic pathways affected in the placenta may give clues to molecular targets for screening and treatment and have implications towards future management of preeclampsia. Placental sampling must be done after delivery, but importantly, increased understanding of underlying placental disease components will enable a more targeted search for disease markers that more accurately reflect the diversity of the placental disease. Identification of a placental phenotype correlating to maternal serum markers as shown here is an important step in this direction, as a direct serum marker of specific placental metabolic alterations may enable tailored treatment of the particular preeclampsia phenotypes.

Several metabolic pathways were shown affected in preeclampsia; most notably the taurine, glutamate and phospholipid metabolism. Taurine is an essential nutrient in fetal metabolism, as the fetus and placenta lack the enzyme for taurine synthesis [25]. Reduced activity of a placental taurine transporter has been found in preeclampsia and FGR [26, 27]. Reduced taurine in the placenta may impair syncytiotrophoblast cell renewal, reducing cell turnover, and further lead to decreased nutrient transfer to the fetus and increased release of necrotic material to the maternal circulation [27]. In our study, taurine levels were similar in placentas from preeclampsia with or without FGR, suggesting that taurine depletion is not specific for FGR and remains important also for preeclampsia without fetal growth impairment.

Glutamine and glutamate are crucial to the fetal carbon and nitrogen metabolism as precursors to protein, purine and pyrimidine synthesis [28]. Glutamate is also a precursor to glutathione, an important antioxidant, and has been shown to be lower in medium of placental explants from hypoxic normal tissue and preeclamptic tissue [21]. Consistent with this finding,

295 placental ascorbate was found to be significantly lower in placentas from preeclamptic
296 women. Intracellular ascorbate protects endothelial cells from hypoxia-reoxygenation induced
297 apoptosis [29].

298 Choline, phosphocholine, glycerophosphocholine and ethanolamine play important roles in
299 the phospholipid biosynthesis pathway. Two possible reasons for increased
300 glycerophosphocholine in preeclampsia are suggested. First, the increase may be due to
301 excessive cell death in preeclamptic placentas [30, 31]. Placental lipid signaling may be
302 disrupted by antiphospholipid antibodies in preeclampsia, causing aberrant cell death in the
303 syncytiotrophoblast and release of necrotic debris [32]. This conforms to our findings of
304 increased glycerophosphocholine especially in those pregnancies with increased maternal
305 serum sFlt-1. Phosphatidylcholine catabolism releases glycerophosphocholine and
306 arachidonic acid by the phospholipase PLA_2, possibly playing a role in increased
307 inflammation, a central process in the preeclamptic placenta. PLA_2 activity is increased in
308 preeclamptic placental tissue [33]. Second, the increase may stem from placental cell
309 membrane catabolism for regeneration of choline methyl groups due to folate deficiency [34].
310 Choline levels were here similar in preeclamptic and normotensive placentas, indicating a
311 compensatory mechanism. Another component of phospholipid biosynthesis, ethanolamine,
312 was decreased in the preeclamptic placenta. Choline and ethanolamine are competitively
313 transferred into the placenta by a common transporter [35]. Increased maternal serum choline
314 levels might inhibit placental uptake of ethanolamine [35, 36]. Ethanolamine kinase deficient
315 mice have low birth weight offspring and increased placental thrombosis and apoptosis,
316 indicating an important role of ethanolamine in placental and fetal development [37].

317 Limitations of our study include the variation in GA between preeclamptic and normotensive
318 pregnancies. Differences were accounted for using linear regression, but variation due to GA
319 cannot easily be overcome in placenta research due to the nature of the preeclampsia
320 diagnosis. Additionally, only one sample per placenta was analyzed, thus, intra-individual
321 variability was not assessed. Strengths of our study are the whole tissue profiling without need
322 for extraction and derivatization, and the sensitivity as reflected by metabolite correlation to
323 the placental derived stress factor sFlt-1 in maternal serum. All deliveries were by CS
324 precluding any labor-induced variation. Sensitive placental profiling as shown here is missing
325 from preeclampsia research.

326 Metabolomics represents the closest measure to the phenotype, which is reflected in the
327 highly significant results. Differences in GA have been assessed and appear secondary to
328 disease-induced metabolic changes. In this study we investigated aberrant pathways affected
329 by preeclampsia, and connected the changes to maternal disease severity and serum markers.
330 An interesting direction of further research will be to metabolically classify the placental
331 component of preeclampsia in larger cohorts, and identify unique factors that may benefit
332 from separate treatment options. We present the HR-MAS MRS method as an excellent tool
333 for placental and pregnancy research possibly leading to future improved screening,
334 prediction and follow-up of pregnant women at risk for preeclampsia.

References

[1] Chaiworapongsa T, Chaemsaithong P, Yeo L and Romero R. Pre-eclampsia part 1: current understanding of its pathophysiology. Nature reviews Nephrology. 2014;10(8):466-80.

[2] Roberts JM and Escudero C. The placenta in preeclampsia. Pregnancy Hypertens. 2012;2(2):72-83.

[3] Redman CW and Sargent IL. Immunology of pre-eclampsia. American journal of reproductive immunology. 2010;63(6):534-43.

[4] Roberts JM and Hubel CA. The two stage model of preeclampsia: variations on the theme. Placenta. 2009;30 Suppl A:S32-7.

[5] Roberts DJ and Post MD. The placenta in pre-eclampsia and intrauterine growth restriction. Journal of clinical pathology. 2008;61(12):1254-60.

[6] Staff A, Andersgaard AB, Henriksen T, Langesæter E, Magnusson E, Michelsen TM, Thomsen LC and Øian P. Hypertensive svangerskapskomplikasjoner og eklampsi. 2014. Oslo: Legeforeningen.

[7] Sibai B, Dekker G and Kupferminc M. Pre-eclampsia. Lancet. 2005;365(9461):785-99.

[8] Roberts JM and Cooper DW. Pathogenesis and genetics of preeclampsia. Lancet. 2001;357:53-6.

[9] Dettmer K, Aronov PA and Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev. 2007;26(1):51-78.

[10] Cao M, Lamichhane S, Lundgren S, Bofin A, Fjosne H, Giskeodegard G and Bathen T. Metabolic characterization of triple negative breast cancer. BMC Cancer. 2014;14(1):941.

[11] Nelson DM and Burton GJ. A technical note to improve the reporting of studies of the human placenta. Placenta. 2011;32(2):195-6.

[12] Johnsen SL, Rasmussen S, Wilsgaard T, Sollien R and Kiserud T. Longitudinal reference ranges for estimated fetal weight. Acta obstetricia et gynecologica Scandinavica. 2006;85(3):286-97.

[13] Savorani F, Tomasi G and Engelsen S. icoshift: a versatile tool for the rapid alignment of 1D NMR spectra. J Magn Res. 2010;443(1):190-202.

[14] Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, Djoumbou Y, Mandal R, Aziat F, Dong E, Bouatra S, Sinelnikov I, Arndt D, Xia J, Liu P, Yallou F, Bjorndahl T, Perez-Pineiro R, Eisner R, Allen F, Neveu V, Greiner R and Scalbert A. HMDB 3.0--The Human Metabolome Database in 2013. Nucleic Acids Res. 2013;41(Database issue):D801-7.

[15] Sands CJ, Coen M, Maher AD, Ebbels TM, Holmes E, Lindon JC and Nicholson JK. Statistical total correlation spectroscopy editing of 1H NMR spectra of biofluids: application to drug metabolite profile identification and enhanced information recovery. Anal Chem. 2009;81(15):6458-66.

[16] Xia J, Mandal R, Sinelnikov IV, Broadhurst D and Wishart DS. MetaboAnalyst 2.0—a comprehensive server for metabolomic data analysis. Nucleic Acids Res. 2012.

[17] Wold S, Esbensen K and Geladi P. Principal component analysis. Chemometrics and Intelligent Laboratory Systems. 1987;2(1–3):37-52.

[18] Westerhuis J, Hoefsloot HJ, Smit S, Vis D, Smilde A, Velzen EJ, Duijnhoven JM and Dorsten F. Assessment of PLSDA cross validation. Metabolomics. 2008;4(1):81-9.

[19] Benjamini Y and Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Annals of Statistics. 2001;29(4):1165-88.

381 [20] Xia J and Wishart DS. MSEA: a web-based tool to identify biologically meaningful
382 patterns in quantitative metabolomic data. Nucleic Acids Res. 2010;38(Web Server
383 issue):W71-7.
384 [21] Dunn WB, Brown M, Worton SA, Crocker IP, Broadhurst D, Horgan R, Kenny LC,
385 Baker PN, Kell DB and Heazell AE. Changes in the metabolic footprint of placental
386 explant-conditioned culture medium identifies metabolic disturbances related to
387 hypoxia and pre-eclampsia. Placenta. 2009;30(11):974-80.
388 [22] Heazell AE, Brown M, Dunn WB, Worton SA, Crocker IP, Baker PN and Kell DB.
389 Analysis of the metabolic footprint and tissue metabolome of placental villous explants
390 cultured at different oxygen tensions reveals novel redox biomarkers. Placenta.
391 2008;29(8):691-8.
392 [23] Dunn WB, Brown M, Worton S, Davies K, Jones R, Kell DB and Heazell AP. The
393 metabolome of human placental tissue: investigation of first trimester tissue and
394 changes related to preeclampsia in late pregnancy. Metabolomics. 2011;8(4):579-97.
395 [24] Giskeodegard GF, Lundgren S, Sitter B, Fjosne HE, Postma G, Buydens LM,
396 Gribbestad IS and Bathen TF. Lactate and glycine-potential MR biomarkers of prognosis
397 in estrogen receptor-positive breast cancers. NMR in biomedicine. 2012;25(11):1271-9.
398 [25] Gaull G, Sturman JA and Raiha NC. Development of mammalian sulfur metabolism:
399 absence of cystathionase in human fetal tissues. Pediatr Res. 1972;6(6):538-47.
400 [26] Cetin I, Corbetta C, Sereni LP, Marconi AM, Bozzetti P, Pardi G and Battaglia FC.
401 Umbilical amino acid concentrations in normal and growth-retarded fetuses sampled in
402 utero by cordocentesis. Am J Obstet Gynecol. 1990;162(1):253-61.
403 [27] Desforges M, Ditchfield A, Hirst C, Pegorie C, Martyn-Smith K, Sibley C and
404 Greenwood S. Reduced Placental Taurine Transporter (TauT) Activity in Pregnancies
405 Complicated by Pre-eclampsia and Maternal Obesity. 2013; (Ed.) El Idrissi A and
406 L'Amoreaux WJ. pp. 81-91. Springer New York.
407 [28] Wu X, Xie C, Zhang Y, Fan Z, Yin Y and Blachier F. Glutamate–glutamine cycle and
408 exchange in the placenta–fetus unit during late pregnancy. Amino acids. 2015;47(1):45-
409 53.
410 [29] Dhar-Mascareno M, Carcamo JM and Golde DW. Hypoxia-reoxygenation-induced
411 mitochondrial damage and apoptosis in human endothelial cells are inhibited by vitamin
412 C. Free Radic Biol Med. 2005;38(10):1311-22.
413 [30] Bogin L, Papa MZ, Polak-Charcon S and Degani H. TNF-induced modulations of
414 phospholipid metabolism in human breast cancer cells. Biochim Biophys Acta.
415 1998;1392(2-3):217-32.
416 [31] Ackerstaff E, Glunde K and Bhujwalla ZM. Choline phospholipid metabolism: a
417 target in cancer cells? J Cell Biochem. 2003;90(3):525-33.
418 [32] Pantham P, Heazell A, Mullard G, Begley P, Chen Q, Brown M, Dunn W and Chamley
419 L. Antiphospholipid antibodies alter cell-death-regulating lipid metabolites in first and
420 third trimester human placentae. Placenta. 2014;35(9):A67-A8.
421 [33] Jendryczko A, Tomala J, Drozdz M and Wloch S. Placental phospholipase A2
422 activities in pre-eclampsia. Ginekol Pol. 1989;60(5):280-2.
423 [34] Mikael LG, Pancer J, Jiang X, Wu Q, Caudill M and Rozen R. Low dietary folate and
424 methylenetetrahydrofolate reductase deficiency may lead to pregnancy complications
425 through modulation of ApoAI and IFN-gamma in spleen and placenta, and through
426 reduction of methylation potential. Mol Nutr Food Res. 2013;57(4):661-70.

427   [35] Grassl SM. Ethanolamine Transport in Human Placental Brush-Border Membrane
428   Vesicles. Journal of Pharmacology and Experimental Therapeutics. 2001;298(2):695-
429   702.
430   [36] Braekke K, Ueland PM, Harsem NK, Karlsen A, Blomhoff R and Staff AC.
431   Homocysteine, Cysteine, and Related Metabolites in Maternal and Fetal Plasma in
432   Preeclampsia. Pediatr Res. 2007;62(3):319-24.
433   [37] Tian Y, Jackson P, Gunter C, Wang J, Rock CO and Jackowski S. Placental Thrombosis
434   and Spontaneous Fetal Death in Mice Deficient in Ethanolamine Kinase 2. Journal of
435   Biological Chemistry. 2006;281(38):28438-49.
436   [38] Jewison T, Su Y, Disfany FM, Liang Y, Knox C, Maciejewski A, Poelzer J, Huynh J, Zhou
437   Y, Arndt D, Djoumbou Y, Liu Y, Deng L, Guo AC, Han B, Pon A, Wilson M, Rafatnia S, Liu P
438   and Wishart DS. SMPDB 2.0: big improvements to the Small Molecule Pathway Database.
439   Nucleic Acids Res. 2014;42(Database issue):D478-84.
440
441
442

443 **Tables**

444

445 **Table 1**

446 Clinical maternal, fetal and pregnancy characteristics of the study participants (n=34).

Variable	Total PE (n=19)	PE+FGR (n=11)	PE severe (n=14)	Normotensive (n=15)
Maternal age in years, md (IQR)	28 (9)	25 (6)*,***	28 (10)	34 (5)
Gravidity, md (IQR)	1 (2)	1 (2)	1 (2)	2 (1)
Primipara, n (%)	9 (47.4)*	6 (54.5)*	7 (50.0)	2 (13.3)
Systolic BP in mmHg, md (IQR)^a	160 (29)**	160 (30)**	170 (20)	124 (20)
Diastolic BP in mmHg, md (IQR)^a	100 (15)**	104 (11)**	105 (36)**	75 (11)
BMI in kg/m^2, md (IQR)^b	24.2 (13.0)	24.2 (13.4)	23.8 (11.1)	23.6 (4.5)
CVD diagnosis^c, n (%)	3 (15.8)	1 (9.1)	3 (21.4)	0 (0)
GA at delivery in weeks, mn (SD)	31.8 (3.5)**	30.8 (2.26)**	31.7 (3.4)**	39.2 (0.41)
Birth weight in g, mn (SD)	1707 (1074)**	1212 (342)**	1621 (785)**	3501 (309)
Placental weight in g, md (IQR)	300 (290)**	275 (125)**	295 (277)**	620 (150)
FGR, n (%)^d	11 (57.9)**	11 (100)	8 (57.1)	0 (0)
Severe PE, n (%)	14 (73.7)	8 (72.7)	14 (100)	0 (0)
Superimposed PE, n (%)	2 (10.5)	1 (9.1)	2 (14.3)	0
Early onset PE (<34 weeks) n (%)	17 (89.5%)	11 (100)	13 (92.9)	0 (0%)
Gestational diabetes, n (%)	2 (10.5)	0 (0)	2 (14.2)	0 (0)

447 P values are from a one-way ANOVA with Tukey post-hoc test or the Kruskal Wallis test
448 with post-hoc pairwise Mann-Whitney U tests. Categorical values were compared using
449 Fishers exact test.
450 ^{a} Blood pressure measured at the last regular prenatal visit, 0-2 weeks before cesarean section.
451 ^{b} Maternal weight for BMI calculation was measured at the first prenatal care visit, before
452 week 12 of pregnancy. Weight information was missing for one woman with severe PE and
453 FGR, and one normotensive woman.
454 ^{c} Pregestational CVD diagnoses included: pregestational hypertension, cardiomyopathies,
455 congenital cardiac defects.
456 ^{d} Birth weight ≤5th percentile according to fetal weight reference curves [12] was used as a
457 proxy for one of the FGR diagnoses.
458 * p<0.05 compared to normotensive.
459 ** p<0.001 compared to normotensive.
460 *** p<0.05 compared to preeclampsia.
461 Abbreviations: BMI; body mass index; BP; blood pressure; CVD, cardiovascular disease;
462 GA; gestational age; FGR, fetal growth restriction; IQR; interquartile range; NA, not
463 applicable; md; median; mn, mean; n, number; PE, preeclampsia; SD, standard deviation

Table 2

465 Metabolite levels in placentas from preeclamptic and normotensive pregnancies, and maternal
466 serum measurements.

Metabolite, median (IQR)	PE (n=19)	Normotensive (n=15)	P-value	P (adj)[a]
Phospholipid biosynthesis				
Ethanolamine	6.6 (1.3)	9.80 (3.1)	<0.001	0.005*
Choline	75.5 (11.1)	74.7 (20.3)	0.864	0.891
Glycerophosphocholine	22.5 (12.1)	13.2 (14.8)	0.007	0.018*
Phosphocholine	11.3 (5.6)	9.0 (4.5)	0.056	0.097
Dihydroxyacetone	1.04 (1.25)	2.3 (1.9)	0.003	0.010*
Glycerol	25.9 (9.7)	25.9 (4.0)	0.945	0.945
Myoinositol	16.8 (3.8)	16.5 (3.6)	0.811	0.863
Ammonia recycling, Urea cycle, Bile acid biosynthesis				
Glutamine	5.4 (2.3)	7.6 (2.4)	0.004	0.013*
Aspartate	8.3 (4.4)	6.1 (2.4)	0.047	0.097
Glutamate	16.4 (4.0)	21.7 (1.3	<0.001	0.005*
Acetate	2.8 (1.1)	3.0 (0.8)	0.372	0.512
Glycine	8.5 (2.5)	11.1 (2.0)	0.001	0.005*
Alanine	9.1 (3.4)	9.4 (3.0)	0.179	0.281
Taurine	20.9 (7.5)	29.5 (4.6)	<0.001	0.005*
Protein biosynthesis				
Leucine	11.23 (3.65)	11.77 (1.9)	0.286	0.410
Isoleucine	1.8 (0.4)	2.0 (0.3)	0.056	0.097
Valine	3.2 (0.8)	3.6 (0.9)	0.006	0.017*
Threonine	3.6 (1.1)	4.3 (0.7)	0.021	0.046*
Lysine	8.3 (3.5)	10.0 (0.9)	0.006	0.017*
Glycolysis, ketone body metabolism				
Lactate	42.2 (10.7)	41.9 (7.2)	0.391	0.515
Glucose	2.5 (2.0)	3.3 (3.1)	0.111	0.183
Succinate	3.0 (1.3)	3.7 (1.3)	0.056	0.097
3-Hydroxybutyrate	2.7 (0.9)	4.4 (3.0)	0.019	0.045*
Catecholamine biosynthesis				
Ascorbate	2.2 (0.6)	2.9 (0.4)	0.001	0.005*
Glycine and serine metabolism				
Creatine	5.5 (1.9)	6.0 (1.9)	0.256	0.384
Serum markers, median (IQR)[b]				
Cholesterol [mM]	6.8 (1.9)	6.1 (2.2)	0.580	0.660
Creatinine [μM]	66.0 (20.0)	58.0 (9.0)	0.001	0.005*
Uric acid [μM]	399 (157)	279 (105)	0.002	0.008*
HDL [mM]	1.6 (0.4)	1.7 (0.5)	0.656	0.722
Triglycerides [mM]	3.7 (2.4)	2.7 (1.1)	0.421	0.515
Calcium [mM]	2.3 (0.3)	2.4 (0.2)	0.421	0.515
hsCRP [μg mL^{-1}]	4.8 (15.6)	3.3 (5.4)	0.486	0.573
sFlt-1 [ng mL^{-1}]	960 (1350)	239 (162)	<0.001	0.005*

467 Metabolite levels are in arbitrary units relative to total spectral intensity. Metabolites and
468 serum values were compared between preeclamptic and normotensive groups using the Mann-

469 Whitney U test. Metabolites grouped by metabolic pathways described in the small molecule
470 pathway database [38]. The metabolites may be involved in several pathways. Abbreviations:
471 HDL, high density lipoprotein; hsCRP, high sensitivity C-reactive protein; IQR, interquartile
472 range; sFlt-1, soluble Fms-like tyrosine kinase receptor 1; PE, preeclampsia.
473 [a] Corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate.
474 [b] One placenta-serum pair was excluded due to missing blood sample.
475 *Significantly different between preeclamptic and normotensive pregnancies.

476

477 **Table 3**

478 Results from partial least squares discriminant analysis. Placental metabolic profiles were
479 compared between groups and the discriminatory ability assessed with 5-fold cross validation

PLS-DA Model	LVs	Sens.	Spec.	Accuracy	P-value[a]	Metabolites (Rel to 1st mentioned)
Total PE (n=19) vs. normotensive (n=15)	1	0.870	0.987	0.928	<0.001	Increase: GPC, PCho, Asp Decrease: EtAm, Tau, Glu, Asc, Gly
Severe PE (n=14) vs. non-severe PE (n=5)	1	0.780	0.893	0.837	0.003	Increase: Cho, Lys, Ala, Glucose, Myo, Tau, Asp, Gln Decrease: 3-HB
PE (n=8) vs. PE+FGR (n=11)	1	0.619	0.736	0.678	0.163	Not significant

480 [a] p value from 1000 permutation tests.
481 Abbreviations: 3-HB, 3-hydroxybutyrate; Ala, alanine; Asc, ascorbate; Asp, aspartate; Cho,
482 choline; EtAm, ethanolamine; FGR, fetal growth restriction; Gln, glutamine; Glu, glutamate;
483 GPC, glycerophosphocholine; Gly, glycine; Lys, lysine; LVs, latent variables; Myo, myo-
484 inositol; Tau, taurine; PCho, phosphocholine; PE, preeclampsia; PLS-DA, partial least
485 squares discriminant analysis; rel, relative; sens, sensitivity; sFlt-1, soluble fms-like tyrosine
486 kinase receptor 1; spec, specificity; Suc, succinate; Val, valine.

487

488 **Table 4**

489 Results from partial least squares regression for correlation between metabolic placenta
490 profiles and maternal serum measurements.

PLS regression Metabolites vs Y [a]	LVs	R[2][b]	Y explained	P-value [c]	Metabolites
Serum sFlt-1	1	0.490	0.614	<0.001	Increase: GPC Decrease: Glu, Tau, Gln, Val, EtAm, Suc
Serum creatinine	1	0.069	0.343	0.032	Increase: GPC, Asp, PCho, Cre Decrease: Gly, EtAm
Serum uric acid	1	0.122	0.348	0.009	Increase: Asp, GPC Decrease: Glu, Tau, Gln, Val, EtAm, Suc
Serum	2	0.174	0.166	<0.001	Increase: Cho, Gln, Gly

15

triglycerides	Decrease: 3-HB, Cre

491 The maternal serum values are for 33 placenta/serum pairs (34 placentas - 1 excluded from
492 serum measurements due to missing blood sample).
493 [a] Y denotes the dependent variable, e.g, the serum measurement
494 [b] R^2 values give the correlation of the cross-validated predicted Y values to the real Y values.
495 [c] p value from 1000 permutation tests.
496 Abbreviations: 3-HB, 3-hydroxybutyrate; Asp, aspartate; Cho, choline; Cre, creatine EtAm,
497 ethanolamine; Gln, glutamine; Glu, glutamate; GPC, glycerophosphocholine; Gly, glycine;
498 LVs, latent variables; Tau, taurine; PCho, phosphocholine; PLS; partial least squares
499 regression; sFlt-1, soluble Fms-like tyrosine kinase receptor 1; Suc, succinate; Val, valine.

500

501 **Figures**
502
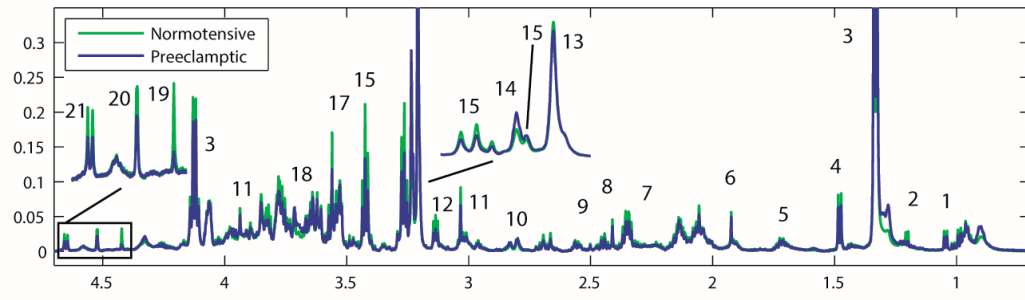
503
504 **Fig. 1.** High resolution magic angle spinning magnetic resonance spectroscopy of placental biopsies from normotensive
505 pregnancies (n=15, green) and preeclampsia (n=19, blue). Median spectra from the two groups are shown. Twenty-five
506 metabolites were identified in the spectrum. Metabolites by number: 1, valine; 2, 3-hydroxybutyric acid; 3, lactate; 4, alanine;
507 5, lysine; 6, acetate; 7, glutamate; 8, succinate; 9, glutamine; 10, aspartate; 11, creatine; 12, ethanolamine; 13, choline; 14,
508 glycerophosphocholine; 15, phosphocholine; 16, taurine; 17, glycine; 18, glycerol; 19, dihydroxyacetone; 20, ascorbate; 21,
509 glucose.
510

511

512
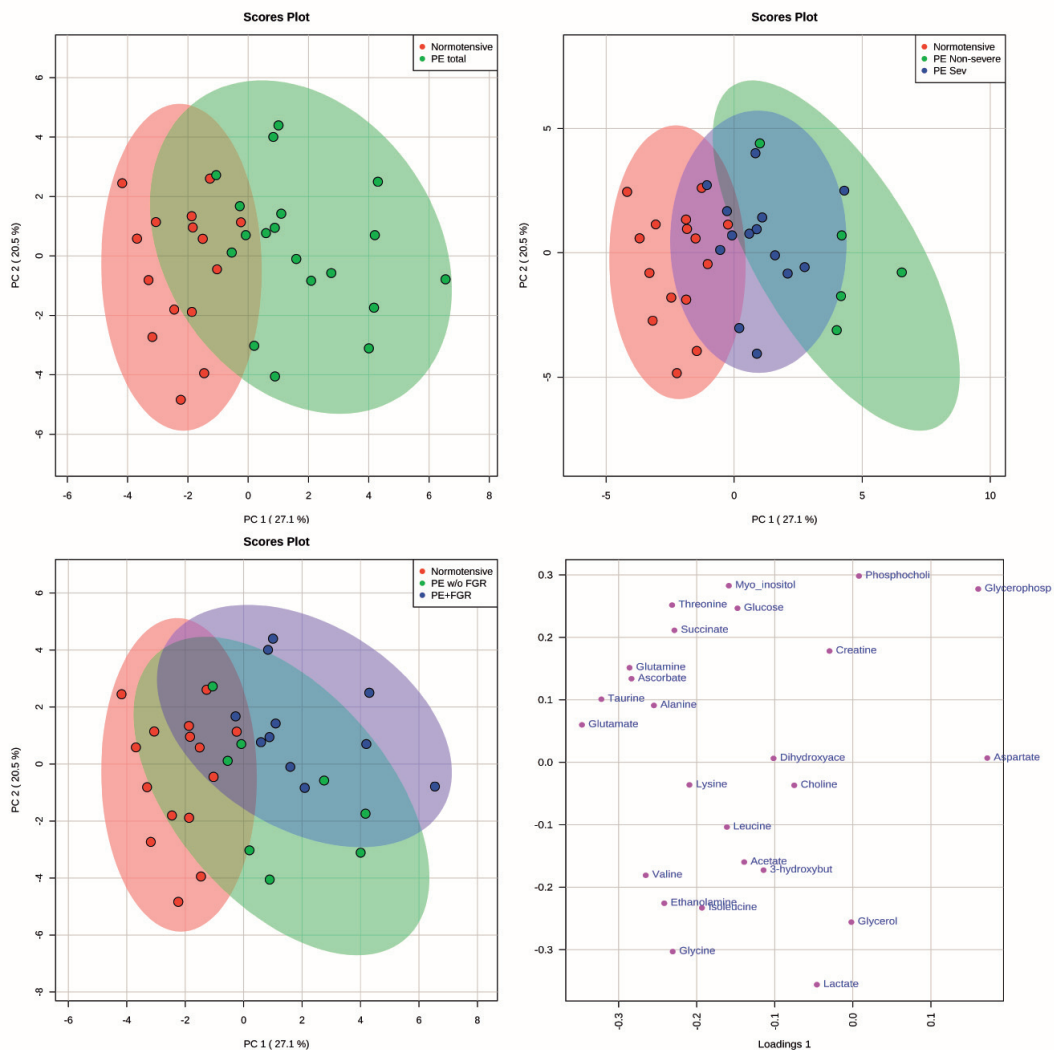
Fig. 2. Results from the unsupervised Principal Component Analysis of 34 placental metabolic profiles. Score plots show a
map of samples in the principal component space; similar samples appear close together. Loading plots show the metabolite
contribution to the variation in the score plots; e.g. samples with high score values for PC1 (to the right in the score plots)
have increased levels of metabolites to the right in the loading plot. (A) Score plot of placental samples colored by disease
group, preeclampsia or normotensive. (B) Score plot of placental samples colored by subgrouping preeclampsia by severity.
(C) Score plot of placental samples colored by subgrouping preeclampsia by the presence or absence of fetal growth
restriction. (D) Loading plot of metabolite contributions in the principal component analysis. Increasing aspartate,
phosphocholine and glycerophosphocholine and decreasing glutamate, taurine, ascorbate and glutamine correspond to the
preeclamptic phenotypes.
Abbreviations: FGR, fetal growth restriction; PC, principal component; PE, preeclampsia; Sev, severe**.**w/o, without

523

Table S.1

Parameter	Clinical Characteristics of Pregnancies for Placentas Studied (Normotensive)			
Gravidity	Median = 2	25-75% = 1	Range= 3	
Parity	Median = 1	25-75% = 1	Range =3	
Gestational age (weeks)	Average = 39.2	SD = 0.4		
Maternal age (years)	Average = 33.1	SD = 3.5	Range = 12	
Race	Black =0	White = 15	Other=0	Unknown =0
Ethnicity				Scandinavian: 15
Prenatal medications	Iron	Other (list):	None	Unknown, n = 15
Drugs	Cigarettes	Alcohol	Other (list):	Unknown, n = 15
Previous prenatal admission(s)	Yes	No	Diagnoses	Unknown, n = 15
Blood pressures <140/90 mm Hg	Yes, n = 15	No, n = 0	Unknown	
Screened for diabetes	Yes, n = 0	No, n = 0	Unknown n=15	
Antibiotics in labor	None, n = 15	Penicillin, n = 0	Other: (list) n = 0	Unknown, n = 0
Beta strep status	Positive	Negative	Unknown n = 15	
Antenatal steroids:	Yes, n = 0	No, n =15	If yes, week GA	Unknown, n = 0
Magnesium sulfate	Yes, n = 0	No, n = 15		Unknown, n = 0
Anesthesia	Epidural = 15	Narcotics	General	Unknown, n = 0
Cervical ripening agent	prostaglandin E_1	prostaglandin E_2	Mechanical	No, n = 15
Labor	No, n = 15	If yes, hours:		
Delivery mode (N= # patients)	C-section, Repeat, no labor: n =	C-section, Repeat, with labor: n =	C-section, primary, no labor: n = 2 (Unknown, no labor: n=13)*	C-section, primary, with labor: n =
Maternal Oxygen given at delivery?	Yes, n = 0	No, n = 15	Unknown = 0	
Birth weight (grams)	Average = 3501	SD = 309		
Placental weight (grams)	Average = 665	SD = 98	Unknown = 0	
Baby's sex	Female = 8	Male = 7	Unknown = 0	
Delivery to processing (mins)	Average = 93	SD = 56	Unknown = 2	

* 13 individuals had given birth previously, unspecified if vaginally or by C-section. In this study they had a C-section without labor

Table S.2

Parameter	Clinical Characteristics of Pregnancies for Placentas Studied (Preeclamptic)			
Gravidity	Median = 1	25-75% = 2	Range = 4	
Parity	Median = 1	25-75% = 1	Range = 2	
Gestational age (weeks)	Average = 31.8	SD = 3.5		
Maternal age (years)	Average = 29.8	SD = 6.3	Range = 24	
Race	Black = 0	White = 19	Other = 0	
Ethnicity				Scandinavian, n = 19
Prenatal medications	Iron	Other (list):	None	Unknown, n = 19
Drugs	Cigarettes n =	Alcohol n =	Other (list): n =	Unknown, n = 19
Previous prenatal admission(s)	Yes, n =	No, n =	Diagnoses if yes:	Unknown, n = 19
Blood pressures <140/90 mm Hg	Yes, n = 0	No, n = 19	Unknown	
Screened for diabetes	Yes, n = 2	No	Unknown, n = 17	
Antibiotics in labor	None n = 19	Penicillin, n =	Other: (list) n =	Unknown, n = 0
Beta strep status	Positive	Negative	Unknown = 19	
Antenatal steroids:	Yes, n = 0	No, n = 19	If yes, week GA	Unknown, n = 0
Magnesium sulfate	Yes, n = 0	No, n = 19	Unknown, n = 0	
Anesthesia	Epidural	Narcotics	General	Other/none
Cervical ripening agent	prostaglandin E_1	prostaglandin E_2	Mechanical	No, n = 19
Labor	No, n = 19	If yes, hours:		
Delivery mode (N= # patients)	C-section, Repeat, no labor: n =	C-section, Repeat, with labor: n = 0	C-section, primary, no labor: n = 9 (Unknown, no labor n = 10)*	C-section, primary, with labor: n =0
Maternal Oxygen given at delivery?	Yes, n = 0	No, n = 19	Unknown, n = 0	
Birth weight (grams)	Average = 1707	SD = 1074		
Placental weight (grams)	Average = 365	SD = 209	Unknown	
Baby's sex	Female = 10	Male = 9	Unknown	
Delivery to processing (mins)	Average = 106	SD = 43	Unknown = 1	

*10 individuals had given birth previously, unspecified if vaginally or by C-section. In this study they had a C-section without labor

Supplementary information

(See Supplementary tables 1 and 2 in separate documents).

Table S.3

Criteria for diagnosis of severe preeclampsia as detailed in Sibai et al [1]; one or more of the following severe maternal features of preeclampsia.

Symptom	Details
Persistent severe hypertension	≥160/110 mmHg
Severe proteinuria	≥3g protein excretion into urine/day
Pulmonary symptoms	Pulmonary edema, dyspnea, cyanosis
Seizures/eclampsia	
Oliguria	<500mL urine/day
Thrombocytopenia	Platelet count <100 000 per μL
Abnormal liver enzymes	Increased serum aspartate aminotransferase, alanine aminotransferase, lactate dehydrogenase
Hemolysis	Low serum haptoglobin (<0.2g/L)
Microangiopathic hemolytic anemia	
Epigastric or right upper quadrant pain	
Central nervous system symptoms	Altered mental status, headaches, blurred vision, blindness
HELLP syndrome	Hemolysis, elevated liver enzymes, low platelets

Table S.4

Identified metabolites in ^{1}H and ^{1}H-^{13}C nuclear magnetic resonance (NMR) spectra of placental biopsies (n=34).

Metabolite name	NMR shifts, ^{1}H (^{13}C)
3-Hydroxybutyrate	1.197d
Acetate	1.927s (26.9)
Alanine	1.478d (19.0), 3.779q (53.4)
Ascorbic acid	4.515s
Aspartic acid	2.818dd (39.5)
Choline	3.207s (56.7), 3.527t (70.4), 4.051b (58.5)
Creatine	3.029s (41.8), 3.937s
Dihydroxyacetone	4.417s
Ethanolamine	3.135t/dd(44.3), 3.82d (61.1)
Glucose	3.269t (77.14), 3.394m (72.4), 3.468m (78.8), 3.521m (74.3), 3.898m (63.5), 4.652d (97.8)
Glutamic acid	2.055m (27.4), 2.338m (36.4), 3.759t (57.0)
Glutamine	2.138m (27.4), 2.444m (35.9)
Glycerol	3.650dd (65.4), 3.780m (74.9)
Glycerophosphocholine	3.234s (56.7), 4.33m (62.0)
Glycine	3.650s (53.6)
Isoleucine	0.955t, 1.003d
Lactic acid	1.318d (23.0), 4.123q (71.3)
Leucine	0.967d
Lysine	1.718quin 1.902m, 3.021t (42.2)
Myoinositol	3.27t (77.1), 3.556dd (74.2), 3.620t (75.1)
Phosphocholine	3.222s (56.7)
Succinic acid	2.410s (36.8)
Taurine	3.263t (50.9), 3.423t (38.2)
Threonine	1.318d (23.0), 3.586d (65), 4.255m (69.4)
Valine	0.988d, 1.042d (20.42)

Multiplicity of peaks are given as follows: s, singlet; d, doublet; t, triplet; q, quartet, quin, quintet; dd, doublet of doublet; m, multiplet.

Table S.5

Metabolite Set Enrichment Analysis of quantitative metabolite data from placenta from women with preeclampsia (n=19) and normotensive pregnancies (n=15).

Pathway	Total Cmpd[a]	Hits[b]	Q Statistic[c]	Expected Q[d]	Raw p[e]	FDR[f]
Bile acid biosynthesis	49	2	39.30	3.03	1.90×10^{-7}	6.07E-06
Taurine and hypotaurine metabolism	7	1	51.60	3.03	1.72×10^{-6}	2.76×10^{-5}
Phospholipid biosynthesis	19	2	29.17	3.03	6.07×10^{-6}	6.47×10^{-5}
Ammonia recycling	18	3	21.32	3.03	2.82×10^{-5}	2.25×10^{-4}
Protein biosynthesies	19	6	15.75	3.03	1.44×10^{-4}	9.24×10^{-4}
Urea cycle	20	2	18.48	3.03	9.32×10^{-4}	0.005
Glutathione metabolism	10	1	26.99	3.03	0.002	0.007
Porphyrin metabolism	22	1	26.99	3.03	0.002	0.007
Pyrimidine metabolism	36	1	25.47	3.03	0.002	0.007
Purine metabolism	45	1	25.471	3.03	0.002	0.007
Glycine, serine and threonine metabolism	26	3	13.69	3.03	0.003	0.007
Propanoate metabolism	18	1	24.97	3.03	0.003	0.007
Glutamate metabolism	18	2	17.69	3.03	0.009	0.022
Methionine metabolism	24	2	13.50	3.03	0.010	0.024
Valine, leucine and isoleucine degradation	36	2	14.90	3.03	0.016	0.033
Lysine degradation	13	1	13.93	3.03	0.030	0.056
Biotin metabolism	4	1	13.93	3.03	0.030	0.056
Beta-alanine metabolism	13	1	11.49	3.03	0.050	0.080
Aspartate metabolism	12	1	11.49	3.03	0.050	0.080
Malate-aspartate shuttle	8	1	11.49	3.03	0.050	0.080
Citric acid cycle	23	1	9.90	3.03	0.070	0.102
Mitochondrial electron transport chain	15	1	9.90	3.03	0.070	0.102

Metabolite set enrichment analysis results (Performed on Metaboanalyst online, www.metaboanalyst.ca [2]).
[a] Total number of compounds (metabolites) in the pathway
[b] Number of measured metabolites found in pathway
[c] Q-statistic describing the correlation between compound concentration profiles and phenotype labels
[d] Expected Q statistic given no correlation between compounds and phenotype labels
[e] P value for the probability of obtaining the Q statistic
[f] P values corrected for multiple correction using false discovery rates

Table S.6

Placental metabolite levels in severe vs non-severe preeclampsia.

Metabolite, median (IQR)	Non-severe PE (n=5)	Severe PE (n=14)	P-value	P (FDR)[a]
Phospholipid biosynthesis				
Ethanolamine	6.6 (3.8)	6.7 (1.2)	0.964	1.000
Choline	67.7 (18.2)	78.1 (7.4)	0.003	0.038*
Glycerophosphocholine	21.3 (20.8)	23.1 (16.02)	1.000	1.000
Phosphocholine	8.0 (8.3)	11.4 (4.2)	0.931	1.000
Dihydroxyacetone	1.0 (0.7)	1.1 (1.4)	0.559	0.736
Glycerol	20.9 (16.9)	25.9 (7.1)	0.500	0.694
Myoinositol	13.3 (6.4)	17.0 (2.0)	0.107	0.243
Ammonia recycling, urea cycle, bile acid biosynthesis				
Glutamine	4.8 (2.0)	5.6 (2.5)	0.070	0.242
Aspartate	6.8 (12.7)	8.3 (4.6)	0.687	0.818
Glutamate	13.8 (4.3)	17.0 (3.4)	0.044	0.242
Acetate	2.6 (1.7)	2.8 (1.0)	0.444	0.653
Glycine	7.3 (2.5)	8.8 (2.1)	0.219	0.391
Alanine	6.3 (3.4)	9.7 (2.0)	0.056	0.242
Taurine	18.0 (7.3)	21.9 (6.1)	0.130	0.271
Protein biosynthesis				
Leucine	9.0 (2.5)	11.6 (3.1)	0.087	0.242
Isoleucine	1.6 (0.4)	1.8 (0.5)	0.257	0.402
Valine	3.0 (0.3)	3.5 (0.7)	0.107	0.243
Threonine	2.9 (1.0)	3.8 (0.7)	0.070	0.242
Lysine	5.9 (1.5)	9.1 (3.8)	0.003	0.038*
Glycolysis, ketone body metabolism				
Succinate	2.7 (1.9)	3.1 (1.2)	0.257	0.402
Lactate	42.2 (7.7)	41.3 (11.1)	0.823	0.935
Glucose	1.6 (1.0)	2.8 (1.6)	0.034	0.242
3-Hydroxybutyrate	3.2 (2.0)	2.7 (0.8)	0.186	0.358
Catecholamine biosynthesis				
Ascorbate	1.9 (0.9)	2.3 (0.6)	0.087	0.242
Glycine and serine metabolism				
Creatine	5.7 (5.1)	5.3 (2.2)	0.687	0.818

Metabolites were compared between groups using the Mann-Whitney U test. Metabolites grouped by metabolic pathways described in the small molecule pathway database [3]. The metabolites may be involved in several pathways. Metabolite levels are in arbitrary units relative to total spectral intensity.

Abbreviations: FDR, false discovery rate; IQR, interquartile range; PE, preeclampsia.

[a] Corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate.

* Significantly different between severe and non-severe preeclampsia after correction for multiple testing.

Table S.7

Comparison between metabolite levels in placentas from women with preeclamptic and normotensive pregnancies, with adjustments for gestational age.

Metabolite	P (PE vs Normotensive)	P (adjusted for GA)
Ethanolamine	0.005*	0.075
Glycerophosphocholine	0.018*	0.344
Dihydroxyacetone	0.011*	0.142
Glutamine	0.013*	0.029*
Glutamate	0.005*	<0.001*
Glycine	0.005*	0.070
Taurine	0.005*	0.031*
Valine	0.017*	0.019*
Lysine	0.017*	0.398
Threonine	0.046*	0.054
3-HB	0.045*	0.023*
Ascorbate	0.005*	0.045*

For preeclamptic vs normotensive groups, the Mann-Whitney U test (adjusted for multiple comparisons) is given. For the gestational age adjusted p values, linear regression models were made with log transformed metabolite levels as dependent variable and gestational age as independent variable. The metabolite is then evaluated at the average values GA=245 days. Abbreviations: 3HB, three-hydroxybutyrate; GA, gestational age; PE, preeclampsia.
* Significantly different between normotensive and preeclamptic placentas after correction for difference in gestational age.

References

[1] Sibai B, Dekker G and Kupferminc M. Pre-eclampsia. Lancet. 2005;365(9461):785-99.
[2] Xia J and Wishart DS. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. Nucleic Acids Res. 2010;38(Web Server issue):W71-7.
[3] Jewison T, Su Y, Disfany FM, Liang Y, Knox C, Maciejewski A, Poelzer J, Huynh J, Zhou Y, Arndt D, Djoumbou Y, Liu Y, Deng L, Guo AC, Han B, Pon A, Wilson M, Rafatnia S, Liu P and Wishart DS. SMPDB 2.0: big improvements to the Small Molecule Pathway Database. Nucleic Acids Res. 2014;42(Database issue):D478-84.