Application of metabolomics to the evaluation of renal function in kidney transplantation

CIRILLO ARIANNA

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy in Biomedical and Pharmaceutical Sciences

CENTER FOR INTERDISCIPLINARY RESEARCH ON MEDICINES (CIRM)
CLINICAL METABOLOMICS GROUP (CliMe)
UNIVERSITY OF LIEGE (ULiège)

PROMOTERS: DR. PASCAL DE TULLIO AND PROF. FRANCOIS JOURET

Academic year 2023-2024

Application of metabolomics to the evaluation of renal function in kidney transplantation

Kidney transplantation (KTx) represents nowadays the best treatment for patients with end-stage kidney disease. Allograft dysfunction post-transplantation currently represents the leading cause of altered kidney function and rejection. For this reason, the long-term success of a transplantation depends on the graft quality at transplantation time and of a continuous follow-up of kidney transplant recipient. In this work we propose the application of a metabolomics approach to explore several challenges linked to KTx: (1) the accurate assessment of graft quality before-transplantation; (2) a consistent and precise monitoring of patients during post-transplantation period. To achieve these goals two different cohorts were analyzed in parallel. The first cohort consists of urine samples collected at 3-and 12-months post-transplantation with patient's stratification based on mGFR value. Combination of NMR and MS-based metabolomics approaches were used to predict and monitor kidney function decline. For the second cohort, samples of perfusate collected during transplantation were used for the analysis. NMR metabolomics was employed in this cohort with the aims to define a panel of biomarkers predicting the occurrence of kidney dysfunction in post-graft period. The primary goals of this thesis are to create tools helping clinicians in patient management and follow-up within their clinical practice.

Application de la métabolomique à l'évaluation de la fonction rénale en transplantation rénale

La greffe rénale est le traitement de choix pour les patients atteints de graves problèmes rénaux. Cependant les dysfonctions post-transplantation du greffon sont la principale cause de complications rénales et de rejet chez les patients transplantés. Afin d'améliorer la durabilité et l'efficacité des transplantations, l'évaluation de la qualité du greffon et la surveillance continue sont essentielles, bien que l'évaluation actuelle soit encore imprécise. Ce projet vise à deux objectifs principaux : (1) une évaluation précise de la qualité du greffon avant-transplantation ; (2) un suivi précis de l'état des patients post-transplantation. Pour atteindre ces objectifs, deux cohortes de patients ont été examinées. La première cohorte comprend des échantillons d'urine prélevés à 3 et 12 mois post-transplantation, classés en fonction de la valeur mGFR. La combinaison des approches métabolomiques par RMN et MS été utilisé pour prédire et monitorer la diminution de la fonction rénale. La deuxième cohorte comprend des échantillons de perfusate collectés lors de la transplantation et analysés par approche métabolomique par RMN pour identifier de nouveaux biomarqueurs précoces de la dysfonction rénale. L'objectif globale de ce travail était donc la mise en place de nouveaux outils afin d'aider les cliniciens dans le suivi et le management des patients.

Acknowledgements — Remerciements

Ce travail représente l'aboutissement de mon parcours en tant qu'étudiante en doctorat, et rien de tout cela n'aurait été possible sans la présence de tous mes collègues, amis et famille. Au cours de ces 5 années, j'ai eu l'opportunité d'apprendre, de grandir, de me mettre au défi et de surmonter des obstacles petits et grands. Tout cela a été rendu possible grâce à votre soutien continu et votre assistance, ainsi qu'aux mots gentils et encourageants, que ce soit en français, en anglais ou même en italien, qui m'ont permis de me sentir "chez moi" même lorsque tous mes repères familiers étaient loin.

Tout d'abord, je tiens à exprimer ma gratitude au Dr Pascal de Tullio, qui m'a fait découvrir la métabolomique lors de mon expérience Erasmus et m'a confié cette incroyable aventure. Votre guidage en tant que "Chef" et votre passion pour la métabolomique m'ont inspirée tout au long du parcours à viser l'excellence et à persévérer. Merci pour votre positivité, même face à l'adversité, et pour les précieuses leçons que j'ai apprises dans cet environnement enrichissant. Merci de m'avoir permis de grandir en tant que scientifique et personne au cours de cette incroyable aventure, loin de ma zone de confort.

Je tiens à remercier chaleureusement le Prof. François Jouret pour votre guidance, votre formation et votre mentorat tout au long de mes études doctorales. Merci pour les innombrables réunions, sessions de brain-storming et idées de projets auxquels j'ai eu l'honneur de participer et de contribuer. Vos paroles encourageantes et votre soutien indéfectible tout au long de ce parcours ont joué un rôle essentiel dans ma formation en tant que scientifique et individu aujourd'hui.

Je suis reconnaissante au Dr. Pierre Delanaye de m'avoir aidée pendant ma thèse en me permettent de rester à jour avec la bibliographie et les nouveaux articles liés à la métabolomique dans les maladies rénales, et pour vos suggestions et explications concernant les aspects cliniques.

Un merci spécial au Prof. Etienne Cavalier pour nos discussions passionnantes sur la métabolomique et ses applications dans les contextes cliniques.

Je voudrais remercier la disponibilité du Prof. Marianne Fillet pour la lecture de mon manuscrit et pour ses commentaires et suggestions pertinents et enrichissants concernant la métabolomique basée sur la spectrométrie de masse.

Merci au Dr. Olivier Detry pour les réunions et les explications détaillées sur les aspects de la perfusion et de la chirurgie, ainsi que pour votre aide à améliorer la forme de ce manuscrit.

Je suis reconnaissante au Prof. Emmanuelle Vidal-Petiot pour votre disponibilité et les réunions qui ont fourni des informations précieuses sur la cohorte d'échantillons et pour m'avoir aidée à comprendre comment l'analyser correctement.

Je souhaite remercier au Dr. Jean Marie Collet d'avoir pris le temps de lire mon manuscrit et pour les discussions et les suggestions qui ont contribué à son amélioration.

Un merci spécial à toutes les personnes qui ont contribué aux différents aspects de cette thèse : Yann Guitton, Justine Massais, Guillaume Resimont, Morgan Vandermulen, Tiago Pinto Coelho et Pauline Erpicum. Sans votre contribution, ce travail n'aurait pas été possible.

Je tiens à exprimer ma gratitude à tous mes collègues du LCP pour leur humour, leur positivité, leurs sourires, leur gentillesse et les "Buongiorno !", ainsi que pour les repas "boulets frites". Ensemble, vous avez créé un environnement de travail incroyable qui, au fil du temps, est devenu un environnement familial, ce qui rendait agréable de venir travailler même les jours les plus gris. Ces 5 années n'auraient pas été les mêmes sans chacun d'entre vous.

Un grand merci à Manon et Léa d'être les meilleurs membres du "metabolomics band". Merci pour les rires, les sessions de remue-méninges sur la métabolomique et les moments incroyables que nous avons partagés lors des congrès. Les sessions de karaoké sur "I can buy myself floweeeeers", la façon "belge" de descendre les escaliers, les soirées de danse inoubliables lors des congrès et les retours à l'hôtel avec 2000 arrêts de "pierre-papier-ciseaux" occuperont à jamais une place spéciale dans ma mémoire.

Un grazie sincero a Cecilia e Fabiana per il loro sostegno reciproco nei giorni grigi e per essere state la mia famiglia italiana qui in Belgio. Un ringraziamento speciale a tutti gli amici universitari, a Bruna, la mia amica di sempre, e a tutte le persone che mi hanno sostenuto durante questo viaggio, la cui presenza e supporto non sono mai stati dati per scontati. Grazie per avermi sempre aiutato, consigliato e confortato quando sentivo nostalgia di casa e la malinconia mi assaliva, e per non avermi mai fatto sentire sola.

Un grazie infinito ai miei genitori e a mia sorella. Il vostro amore supera la barriera dei chilometri, e il vostro sostegno mi ha permesso di affrontare tutte le mie paure, i miei dubbi e le mie incertezze. Se sono la donna che sono oggi, è solo grazie a voi e al vostro amore incondizionato.

À Noesy (oui, le nom vient directement de la technique de RMN utilisée dans cette thèse), dont les « fêtes » en rentrant à la maison, le bonheur et l'hyperactivité ont réussi à me faire apprécier même les jours de pluie et les jours les plus gris à travers nos promenades.

À Matthieu,

Qui est maintenant mon collègue, mon ami, ma famille et mon tout. Merci d'être entré dans ma vie si doucement et de l'avoir transformée en un tourbillon de bonheur, de sourires et de bonnes vibrations. C'est grâce à ta positivité et ta gentillesse que je m'efforce de me dépasser et de devenir une meilleure personne. Merci de toujours croire en moi, même dans les moments où je ne croyais pas en moi-même. Les cinq dernières années n'auraient pas été les mêmes sans toi et ton amour. Je t'aime tellement.

« Soyez comme l'arbre, changez vos feuilles, mais jamais vos racines.
Vous pouvez changer vos opinions, mais jamais vos principes. »
Victor Hugo

List of abbreviation

Area under the curve	AUC
Body mass index	BMI
Body surface area	BSA
Capillary electrophoresis	CE
Carr-Purcell-Meiboom-Gill	CPMG
Chemical ionization	CI
Chromium 51-ethylenediamine tetraacetic acid	^{51}Cr-EDTA
Chronic kidney disease-epidemiology collaboration	CKD-EPI
Chronic kidney disease	CKD
Cold ischemia time	CIT
Correlation spectroscopy	COSY
Creatinine clearance	ClCr
Data dependent acquisition	DDA
Data Independent acquisition	DIA
Deceased donor	DD
Delayed graft function	DGF
Discriminant analysis	DA
Donation after brain death	DBD
Donation after circulatory death	DCD
Electron Impact	EI
Electrospray ionization	ESI
End-stage renal disease	ESRD
Estimated glomerular filtration rate	eGFR
Extended donor criteria	ECD
Fourier Transform	FT
Fourier Transform-Ion Cyclotron Resonance	FT-ICR
Free induction decay	FID
Gas chromatography mass spectrometry	GC-MS
Global natural product social molecular networking	GNPS
Glomerular filtration rate	GFR
Heteronuclear single quantum correlation	HSQC
High mass resolution	HRMS
High performance liquid chromatography	HPLC
Higher-energy C-trap Dissociation	HCD
Human metabolome database	HMDB
Hydrophilic Interaction Liquid Chromatography	HILIC
Hypothermic machine perfusion	HMP
Institute Georges Lopez-1	IGL-1
Ion chromatography	IC
Ion Trap	IT
Kidney Disease: Improving global outcome	KDIGO

Kidney transplantation recipient	KTR
Kidney donor profile index	KDPI
Kidney donor risk index	KDRI
Kidney transplantation	KTx
Liquid chromatography mass spectrometry	LC-MS
Living donor	LD
Mass spectrometry	MS
Measured glomerular filtration rate	mGFR
Metabolic informative content	MIC
Metabolomics standard initiative	MSI
Modification of diet in renal disease	MDRD
Nuclear Overhauser Effect Spectroscopy	NOESY
Nuclear magnetic resonance	NMR
One dimensional nuclear magnetic resonance	1D NMR
Orthogonal Partial Least square	OPLS
Orthogonal signal correction	OSC
Partial Least square	PLS
Principal component Analysis	PCA
Probabilistic quotient normalization	PQN
Quadrupole	Q
Receiver operator characteristic curve	ROC
Relaxation delay	RD
Resolving power	RP
Retention time	RT
Reverse Phase Liquid Chromatography	RPLC
Standard donor criteria	SCD
Supercritical fluid chromatography	SFC
Tandem mass spectrometry analysis	MS/MS
Technetium 99m diethylenetriamine pentaacetic acid	99mTc-DTPA
Thin layer chromatography	TLC
Time mixing	Tm
Two-dimensional nuclear magnetic resonance	2D NMR
(3-trimethylsilyl)-2,2,3,3-tetradeuteropropionic acid	TMSP
Ultra-high performance liquid chromatography	UHPLC
Variable importance of projection	VIP
Warm ischemia time	WIT
Workflow4metabolomics	W4M

Table of content

Chapter 1

Introduction

1 Introduction

1.1 Background-Metabolomics

1.1.1 Omics sciences

“Omics” defines the group of fields analyzing large-scale data describing a specific biological system at a particular level. The core concept on which all “omics” are founded is that a specific complex system can be better understood if considered in his whole environment and as strictly interconnected to this. In the last decades the improvement reached in computer science and instrumental technologies have allowed the generation of “big” data sets that represents the subject of study of omics technologies. Specifically, omics science embraces the wholeness of tools, approach and technique able to assess and quantify relationship between a set of molecules and a biological, biochemical, physiological or pathological process. Due to the high versatility of this domain, the application area of omics science is broadly various by including toxicology, drug development, biomarker discovery, food and environmental science and personalized medicine.

Depending on the type of molecules on which the “omics” is focused we can distinguish four major fields: genomics, transcriptomics, proteomics and metabolomics.

The term "genomics" emerged in the late 1980s to describe the study of an organism's entire set of genes, or its genome. The Human Genome Project, launched in 1990, aimed to sequence the entire human genome and was completed in 2003. This marked a significant milestone in omics science, providing a foundation to understanding genetic variations and their implications[1]. Transcriptomics involves the study of all RNA molecules produced by a cell or organism, collectively known as the transcriptome. In 1990, this field gained momentum with the development of microarray technology and later, RNA sequencing (RNA-seq) [2]. Transcriptomics enabled researchers to simultaneously measure the expression levels of thousands of genes and allowed insights into gene regulation, cellular responses, and disease mechanisms. Proteomics focuses on the large-scale study of proteins and their functions within a biological system. In late 1990s the development of mass spectrometry techniques facilitated the identification and quantification of proteins in complex mixtures [3]. Proteomics research unveiled details about protein interactions, post-translational modifications, and their roles in health and disease. Metabolomics involves the study of small molecules, or metabolites, that are produced as a result of cellular processes. This field provides insights into the metabolic state of an organism and how it responds to environmental changes, nutrition, and disease.

Advancement in analytical techniques, such as nuclear magnetic resonance and mass spectrometry in early 2000s, played a key role in advancing metabolomics research[4].

The complementarity of these sciences is evident and can give a holistic picture of a biological system. From the top to the bottom of this cascade ***(Figure 1)*** a multitude of possible combinations and modifications can happen by modifying the end-product represented by the phenotype. Particularly, when we look at the omics cascade, metabolomics represents the downstream output of the genome, transcriptome and proteome and directly reflect the functional status of an organism[5,6]. For instance, metabolomics bridges the gap between molecular components of an organism and the observable traits that define its phenotype.

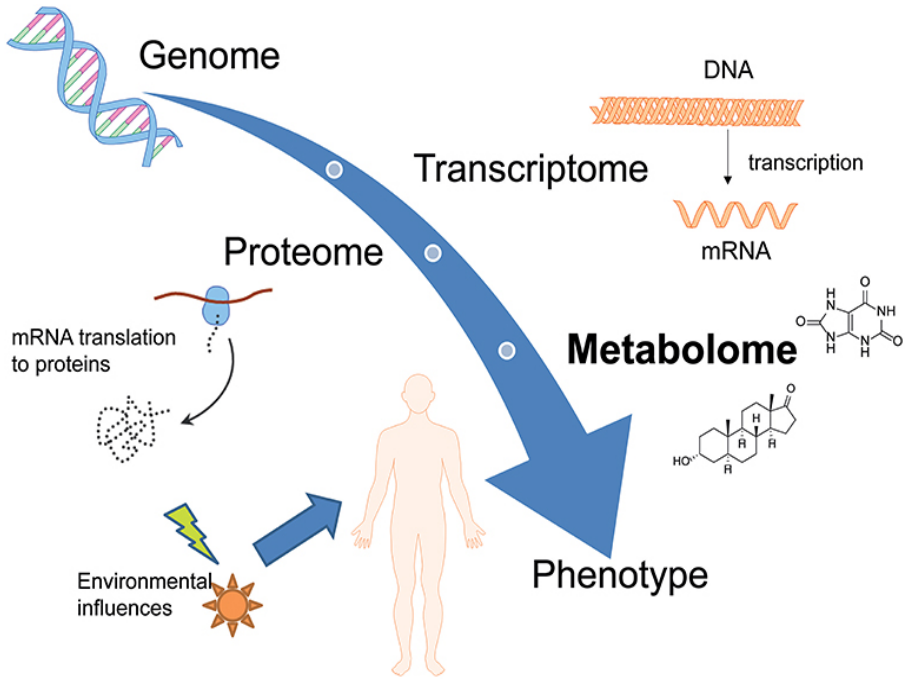


Figure 1 *Omics" cascade where Metabolomics represents the downstream output of the genome but also the upstream input from the environment (Steuer, Andrea E et al. "Metabolomic Strategies in Biomarker Research-New Approach for Indirect Identification of Drug Consumption and Sample Manipulation in Clinical and Forensic Toxicology?" DOI: : 10.3389/fchem.2019.00319.)*

1.1.2 Generalities and history

Metabolomics is a field of scientific study that involves the comprehensive analysis and profiling of small molecules, known as metabolites, within a biological system. These metabolites are the end products of various cellular processes, reflecting the biochemical activity and interactions occurring within an organism.

Metabolomics aims to provide a holistic understanding of the metabolic status and dynamics of a biological system. This approach allows a systematic profiling of metabolites and their

changes issue by environmental factor, diet, lifestyle, pathologies or drug assumption. Metabolites are represented by small molecules (80-1500 Da) produced as a result of various biochemical reactions and processes within living organisms. These molecules play the roles of intermediates and end products of metabolic pathways, which are intricate networks of chemical reactions that occur within cells.

Metabolites covers diverse range of compounds, including sugars, amino acids, fatty acids, vitamins, hormones, neurotransmitters, and many others. These molecules serve as building blocks to constructing cellular structures, as well as the energy suppliers for cellular functions; furthermore, they act as signaling molecules that orchestrate a diverse array of biological processes. Analyzing the alterations of metabolites within tissues, cells, or biofluids provides distinct insights into the physiological or pathological condition, furnishing invaluable information. The idea of follows changes in metabolites and consequently in metabolomes is not as new as we can imagine as we can see from the ***Figure 2***. The first use of metabolomics can be spot back to the ancient Chinese culture (2000–1500 BC) when traditional Chinese doctors evaluated glucose level in urine of diabetic patients using ants. In 300BC, both Egypt and Greece practiced the traditional method of diagnosing human diseases by assessing the taste of urine [7]. However, it was only with the development of modern analytical techniques, such as chromatography and spectroscopy, in the 20th century that the systematic study of metabolites gained importance. These advancements allowed scientists to identify and quantify specific metabolites in biological samples, laying the groundwork for the field of metabolomics [8, 9]. The term "metabolomics" itself emerged in late 1990s as technologies like nuclear magnetic resonance (NMR) and mass spectrometry (MS) became more sophisticated and accessible [10]; these technologies enabled researchers to analyze complex mixtures of metabolites in a comprehensive and systematic manner. Since then, the interest to metabolomics exploded with a particular interest to its domain of application potential going from agriculture to medicine and other related areas in the biological sciences. The important place occupied by metabolomics on biomedical research fields, allowed, in 2007, the publication of the first human metabolome database (HMDB) today known as the most comprehensive database of measurable metabolites in human biofluids[11]. Currently, the database records a total of 220.945 metabolites, a notable increase from the 2.180 metabolites documented in 2007. This remarkable growth serves as proof of the growing user community and the increasing amount of research being conducted in the field of metabolomics. From 2000 to 2016, numerous pathways analysis tools have been developed (including KEGG, MELTIN, iPath, MetPA,

BioCyc, etc.) to aid in the visualization of metabolite profiles within biochemical network diagrams and the identification of metabolic patterns associated with specific biological states.

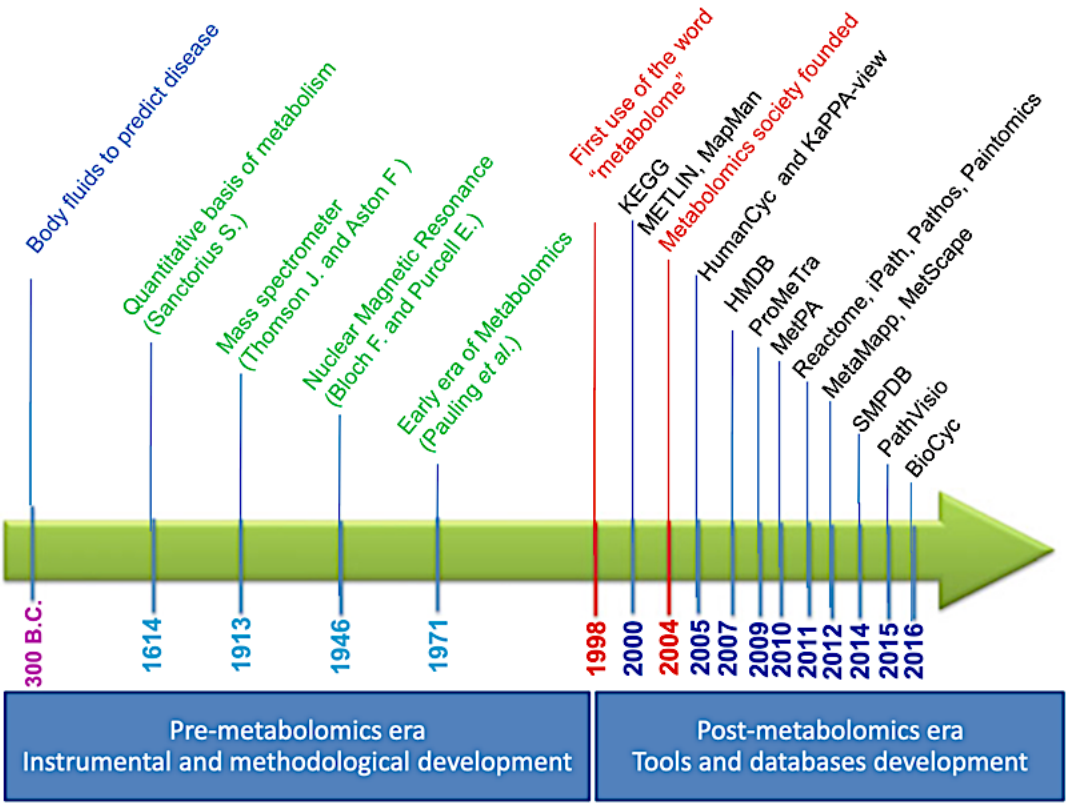


Figure 2 *Timeline of metabolomics science starting from the "pre-metabolomics era" with instrumental and technologies development until the "post-metabolomics era" with databases and tools generation. The pre metabolomics era goes from Santorio Sanctorius, founding father of metabolic balance studies (in 1614) to pioneers of mass spectrometry and nuclear magnetic resonance techniques (respectively in 1913 and 1946) until the first study on urine sample aiming to investigate biological variability conducted by Pauling and his team (1971). Few years later, we enter in the post-metabolomics era marked by the inaugural use of the term "metabolomics." Remarkable events have signed this era by passing from the foundation of "Metabolomics Society" in 2004 until the publication of the first Human Metabolome Database in 2007 and the following advances in tools and databases nowadays available and in continuous development. (Kusonmano, Kanthida et al. "Informatics for Metabolomics" DOI: 10.1007/978-981-10-1503-8_5).*

1.1.3 Modern metabolomics and trends

In the post-metabolomics era, substantial progress in analytical platforms and bioinformatics tools has led to a notable rise in the utilization and publication of metabolomics across a wide spectrum of disciplines. As shown in ***Figure 3a*** the number of articles reported in the context of metabolomics has seen a considerable grow since the first time the term "metabolomics" was coined. Specifically, the main research subjects are represented by biomedicine, agriculture, toxicology, and environment (***Figure 3b)***.

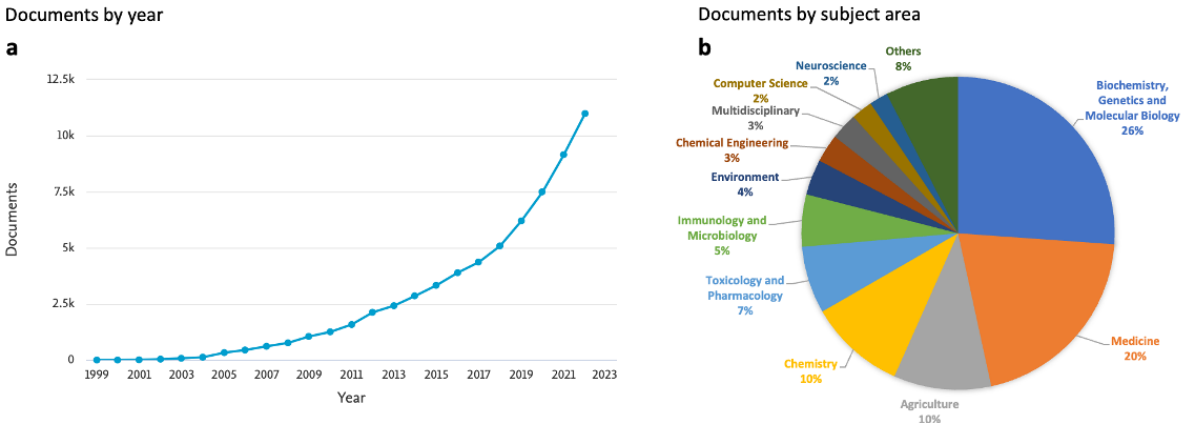


Figure 3 (a) *Number of publications per year listed in Scopus when searching [metabolomics];* ***(b)*** *percentages of documents by subject area obtained when searching [metabolomics] in Scopus.*

In the context of medical science, metabolomics has already been employed to identify biomarkers, enabling both disease diagnosis and prognosis, while also enhancing our mechanistic comprehension of metabolic pathways linked to pathological conditions [12,13]. In agriculture area, the use of this approach allowed the study of agricultural products and their relation to nutritional content, effect of pesticide and herbicide and elucidation of metabolites responsive for aroma and flavors [14].

In toxicology, the study of how toxins and pollutants disrupt normal metabolic pathways and impact health has become an indispensable tool [15]. In environmental field, a deeper understanding of how organisms and ecosystems respond to changes in their surroundings is nowadays primordial, and in this context, metabolomics see its application in pollution monitoring and climate change effect[16]. Metabolomics' interdisciplinarity and its capacity for comprehensive analysis contribute to its growing significance within the contemporary scientific landscape. Therefore, as technology and methodologies continue to evolve, metabolomics remains a crucial instrument for investigating the dynamic metabolome of diverse biological systems.

1.1.4 Metabolomics strategies

The extensive range of applications and diversified goals in metabolomics research, coupled with the high number and varied spectrum of detectable metabolites, has led to the classification of this field into two major distinct approaches: untargeted metabolomics and targeted metabolomics. The two approaches as shown in ***Figure 4*** are here explained:

- **Targeted approach:** is also called **"biology driven"** [17] and is focused on the quantification of a specific set of pre-defined metabolites. These metabolites are often

chosen based on prior knowledge of their involvement in specific pathways, diseases, or biological processes [18,19]. Because it relies on pre-existing knowledge, this method does not allow the identification of novel molecules that might have a significant role in characterizing a biological process. This strategy proves especially valuable when the research question is well-defined and demands accurate quantification of specific compounds. This method is selected to set the baseline metabolites level in an organism or to determine a threshold that can differentiate between "healthy" and "perturbed" states.

- **Untargeted approach:** has as goal to measure the complete set of molecules found in a sample. Through a single analysis, is possible to access the quantities of all molecules present within an individual, enabling the creation of a distinct metabolic profile. Despite the comprehensive character of untargeted metabolomics, the number of substances that can be evaluated in a single analysis is limited by the broad spectrum of physiochemical characteristics within the metabolome[20,21]. In comparison to targeted approach, this profiling technique is especially adapted when the aim is to delineate a metabolic profile associated with a specific condition of the studied individual or for the discovery of new biomarkers. Furthermore, by comparing metabolomics profiles between "healthy" and "perturbed" individuals, is possible to obtain predictive model for the subsequent classification of unknown samples.

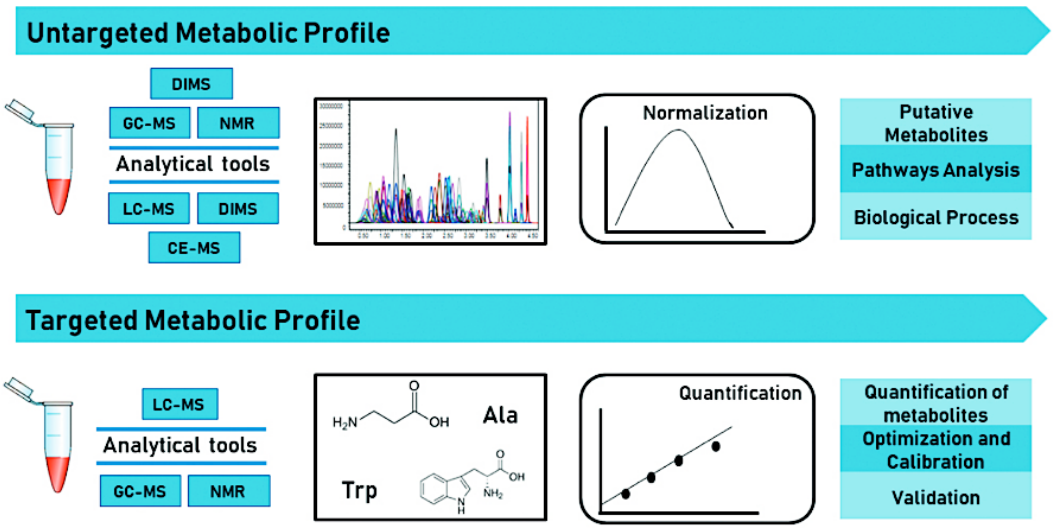


Figure 4 *Scheme of approaches used in metabolomics divided in untargeted and targeted with analytical techniques used and different goals (Adaptedfrom: De San-Martin, Breno Sena et al. "Metabolomics as a potential tool for the diagnosis of growth hormone deficiency (GHD): a review." DOI: 10.20945/2359-3997000000300)*

In summary, both approaches aim to identify variations within the metabolomes of a biological system; however, targeted metabolomics focus on one or a limited number of defined

metabolites, while untargeted metabolomics includes a wide and novel array of metabolites for analysis [22,23]. Because targeted and untargeted strategies have different scopes, a range of analytical platforms can be utilized. Alongside Fourier Transform Infrared Spectrometry (FT-IR), and Raman spectroscopy, the most frequently employed methods include Nuclear Magnetic Resonance (NMR) and mass spectrometry (MS). For targeted approaches, the MS technique (coupled to gas or liquid chromatography) is often favored due to its sensitivity, serving as an excellent method to quantify specific metabolite sets. Conversely, NMR is well-suited for an untargeted approach as it does not require prior knowledge about sample composition; furthermore, its quantitative nature makes it suitable for targeted analyses of the most concentrated metabolites.

In recent years, another emerging approach has gained prominence, combining the strengths of both targeted and untargeted techniques: the **semi-targeted approach**. This analytical strategy enables the merge of several aspects from both targeted and untargeted methods, addressing the limitations inherent in these approaches. The fundamentals of the semi-targeted method involve the quantification of a broad panel of selected metabolites without a priori hypotheses. While only a group of pre-selected metabolites is annotated, this approach allows for the reanalysis of data to identify global metabolic changes that were not part of the original focus[24,25] . An illustrative example of the potential of this approach is evident in a study within the field of cancer research, which took advantage of both hypothesis-based targeted verification and discovery through untargeted data acquisition. This study, by enabling the precise quantification of 110 cancer-related metabolites alongside untargeted profiling of thousands of features, highlighted the adaptability and all-encompassing nature of the semi-targeted approach[26].

1.1.5 Analytical platforms used in untargeted metabolomics

As previously shown, the metabolome is a complex mixture of molecules belonging to most disparate class of metabolites, each with unique structures and possessing distinct chemical properties. Indeed, from a physiochemical standpoint, metabolites display various characteristics in term of pKa, solubility, polarity, volatility, charge, reactivity and stability. Adding to these chemical and physiochemical considerations, their concentrations in biological systems vary significantly across several orders of magnitude.

Given this intricate complexity, there is currently no single analytical platform capable of comprehensively analyzing the entire metabolome. Starting from the post-metabolomics era, and benefiting from significant advancements in technological techniques, two predominant platform currently find application in metabolomics: nuclear magnetic resonance (NMR) and mass spectrometry (MS). These two methods are both complementary and supplementary to one another ***(see table 1)*** [27].

NMR	MS
Non-destructive for sample	*Destructive for sample*
Robust instrumentation	*Relatively instable*
High reproducibility	*Moderate reproducibility*
Small sample preparation	*More complex sample preparation*
No need of chromatography separation	*Requirement for chromatography*
No need of chemical derivatization	*GC need chemical derivatization*
Predictable spectra	*Spectra difficult to predict*
Allow structure determination	*Allow partial structural determination*
Inherently quantitaive	*Not inherently quantitative (need for reference standards)*
Early automated workflow	*Difficult to automate workflow*
Low cost per sample (56-98$)*	*More expensive cost per sample (115-190$)**
Poor to moderate sensitivity (µM)	**Excellent sensitivity (nM)**
Modest in metabolite coverage	**Extensive metabolite coverage**
Few software resources	**Many software resources**

Table 1 *Comparison of analytical pro and cons pf NMR and MS-based metabolomics (adapted from https://hdl.handle.net/2268/301865)*
**price for a targeted analysis available on The Metabolomics Innovative Center (TMIC : https://www.tmicwishartnode.ca/)*

Mass spectrometry is a sensitive technique widely used in metabolomics method able to detect feature at the till the nano- or picomolar range [28]. Because of the complex mixture of metabolites present in metabolomics samples, MS is generally coupled to a separation system such as liquid chromatography (LC-MS), gas chromatography (GC-MS), supercritical fluid chromatography (SFC-MS) and capillary electrophoresis (CE-MS). The most advanced LC-MS platform allows to measure tens of thousands of distinct features within a single experiment giving access to elemental formula with high mass resolution (HRMS) and matching metabolites to huge MS databases with tandem mass spectrometry analysis (MS/MS) [29]. Indeed, while a considerable number of features can be detected through LC-MS metabolomics, these features do not always lead to the identification of specific metabolites. This is due to various factors, including the challenges posed by overlapping peaks, co-elution of compounds, and the incomplete coverage of existing metabolome databases. Nevertheless, advances in the field of annotation process help today the discovery and the knowledge of the "dark metabolome" [30,31].

The primary limitation of NMR technique when compared to MS is its relatively low sensitivity (sub mM), which makes it less suitable to analyzing metabolites that are present in low

concentrations, anyway recent advancements, have addressed this issue by enhancing the sensitivity of NMR through hardware improvements in NMR systems [32]. Additionally, NMR's sensitivity is not influenced by ionization or separation methods of the metabolites, allowing for a broader coverage of the metabolome in a single analysis compared to MS; furthermore, minimal or absent sample preparation as well as nondestructive aspect are part of the advantages of NMR. In longitudinal studies, both reproducibility and robustness are important factors, while accurate quantification and structural elucidation play a crucial role in ensuring the quantitative and qualitative aspects [33]. Together, these elements form the core strengths of NMR.

By examining the strengths and weaknesses of each platform individually, it becomes evident that the challenge of achieving comprehensive coverage of the metabolome using a single technique is a complex issue[34] ***(Figure 5)***.

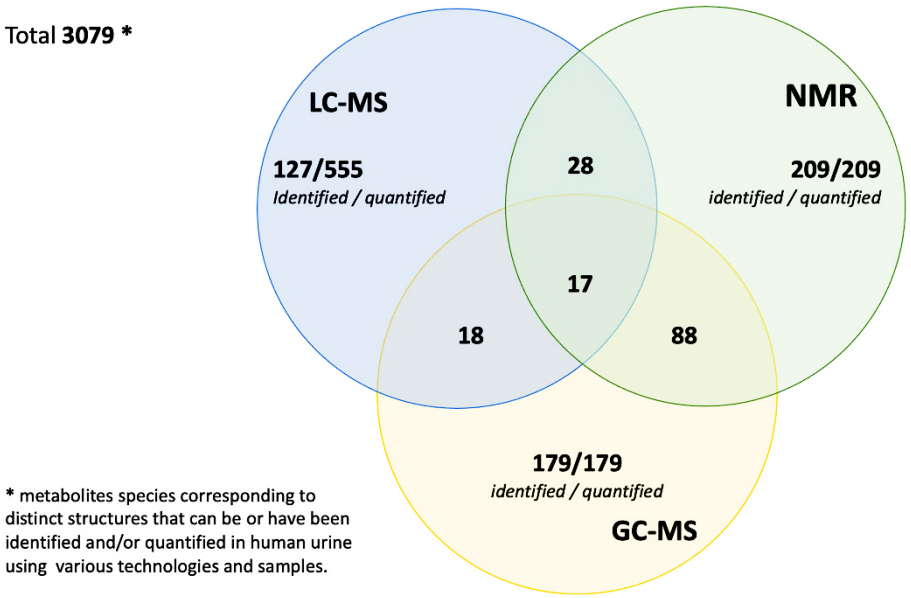


Figure 5 *The Venn diagram illustrates metabolite coverage for NMR, GC-MS, and LC-MS platforms in urine samples using a targeted approach (LC-MS quantification and identification were performed using the Biocrates AbsoluteIDQ p180 Kit, capable of quantifying a maximum of 188 metabolites). A total of 3079 metabolite structures (corresponding to known and highly probable metabolites) can be measured in urine metabolome by using NMR ,GC-MS, DFI/LC-MS/MS ,ICP-MS and HPLC assays (with UV or FD detection). In this specific reported study, NMR spectroscopy, successfully identified and quantified 209 compounds, GC-MS identified and quantified 179 compounds , while DFI/LC-MS/MS identified 555 metabolite structures and quantified 127 compounds. NMR, GC-MS, and LC-MS collectively identified a common set of 17 metabolites, including 15 amino acids, creatinine, and hexose/glucose. Nearly all non-volatile metabolites (87) identified by GC-MS analyses were also detected by NMR. Exceptions included oxalic acid, phosphate, and uric acid, each of which was identified by GC-MS but not by NMR . As we can see, NMR is able to measure ~7% (209/3079) of the human urine metabolome; GC-MS is able to measure ~6% (179/3079); DFI/LC-MS/MS is able to measure ~18% (555/3079) of the urine metabolome. When combined, the 3 analytical techniques can cover >26% of the known and probable or putative urinary metabolome (>824/3079). (Bouatra, Souhaila et al. "The human urine metabolome." DOI: 10.1371/journal.pone.0073076).*

Recognizing the high complementarity between NMR and MS platforms in detecting different chemical classes of metabolites, a solution has emerged in recent years: the adoption of a multiplatform approach.

The general concept of multiplatform approach is still not well described in literature but since 2008 an increasing number of publications are reported by highlighting the interest of this novel technique ***(Figure 6).*** While the utilization of the multiplatform approach is on the rise, several limitations can impede its implementation. These include factors such as cost, lower throughput, a shortage of expertise, and the complexity of data analysis. Nevertheless, these challenges can be overcome through preventive measures such as meticulous experimental design, selection of appropriate analytical methods or optimizing sample preparation protocols [35,36]. By addressing these challenges, researchers can improve the synergistic benefits of a multiplatform approach and achieve more comprehensive insights into complex biological systems.

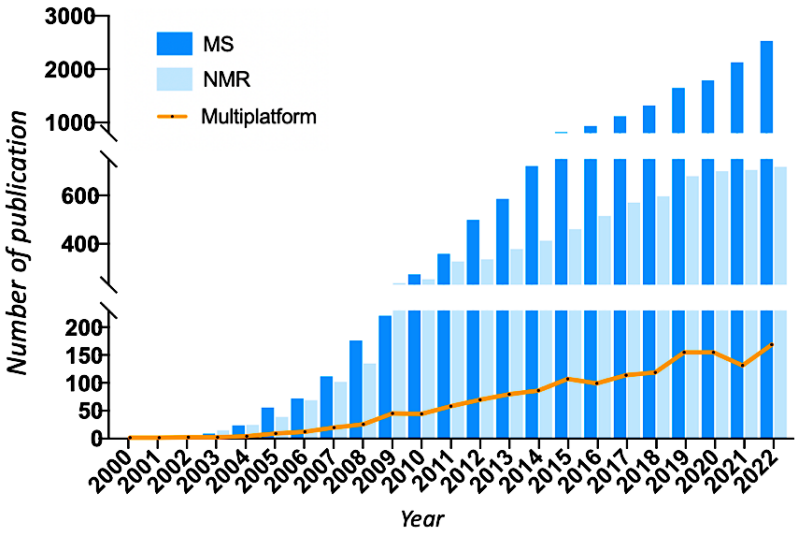


Figure 6 *Number of publications per year find by searching on Scopus [MS-metabolomics], [NMR-metabolomics] and [multiplatform metabolomics]*

1.2 NMR-based metabolomics

1.2.1 Principles of Nuclear Magnetic Resonance [37,38,39]

Nuclear magnetic resonance spectroscopy, known as NMR spectroscopy, is a spectroscopic technique used to analyze the local magnetic field around the atomic nucleus and it is based on the homonymous physical phenomena. The phenomena were described and measured by Isidor Rabi in 1938 that received for that the Nobel Prize [40], and in 1952 the technique was developed by Edward Mills Purcell and Felix Bloch that won Nobel Prize in Physics for their discoveries[41].

NMR spectroscopy exploits a fundamental intrinsic property of atomic nuclei, referred to as "**spin**." This nuclear spin property is intrinsic to subatomic particles like protons, neutrons, and electrons and is defined by quantum number **I**. Notably, certain atomic nuclei have a total spin of zero, rendering them unobservable in NMR. Only atomic nuclei with $I \neq 0$ are detectable, and specifically, this means that nuclei possessing an odd number of protons or neutrons, or both, are observable through NMR. Examples include ^{1}H, ^{13}C, ^{15}N, ^{19}F and ^{31}P, which all exhibit a spin of ½. Among these observable nuclei, ^{1}H is the most abundant in nature, making it a prominent focus of investigation in metabolomic research.

1.2.1.1 Energetic states

All atomic nuclei have a charged structure, and in certain nuclei, this charge undergoes rotation around the nuclear axis. This rotational motion of the charges within nuclei generates a localized magnetic field along the axis and its magnitude is quantified by the nuclear magnetic moment (μ). In the absence of an external magnetic field, the nuclear magnetic moment, or dipole, exhibits random orientations, resulting in no net magnetization. When an external magnetic field (denoted as **B0**) is applied, the nuclear magnetic moments align themselves in accordance with this field. The number of possible orientations that a nucleus can assume in an external magnetic field depends on its quantum spin number (I) and is determined by the formula 2I+1. For nuclei with a spin of 1/2, such as ^{1}H , two orientations are allowed based on energy levels: parallel (α-state, m = +1/2) and antiparallel (β-state, m = -1/2) to the external magnetic field. In this scenario, the parallel orientation (α-state) is energetically favored ***(Figure 7).*** This leads to a vector sum of all nuclear magnetic moments, resulting in a non-zero value directed along the B0 field. This sum, known as nuclear magnetization (M), is directly

proportional to the number of nuclei within the sample. Consequently, this magnetization is the subject of modification and measurement of NMR techniques.

In essence, the application of an external magnetic field causes the alignment of nuclear magnetic moments, generating nuclear magnetization, which forms the basis for the measurements performed in NMR experiments.

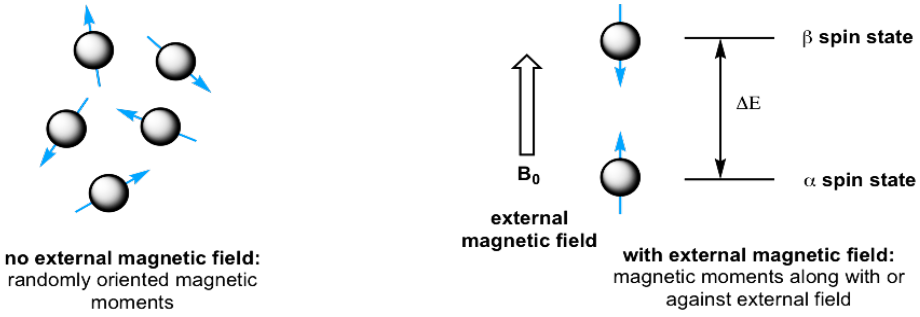


Figure 7 *Orientations of magnetic moments of protons without and with external magnetic field (https://kpu.pressbooks.pub/organicchemistry/chapter/6-5-the-nmr-theory/)*

1.2.1.2 Resonance

When an external magnetic field (B0) is applied to a sample containing nuclei with non-zero nuclear spin, the nuclear magnetic moments align themselves with the direction of the field. However, due to the intrinsic angular momentum (spin) of the nuclei, these magnetic moments do not remain perfectly aligned. Instead, they start to precess or rotate around the direction of the magnetic field at a given angle generating the precession movement ***(Figure 8).*** Its precession frequency (**ω**) which is defined as the number of precessions by second, depends upon the strength of B0 and is calculated according to the Larmor equation: ω= γB0 (where **γ** represent the gyromagnetic ratio and is a constant characteristic of each nucleus).

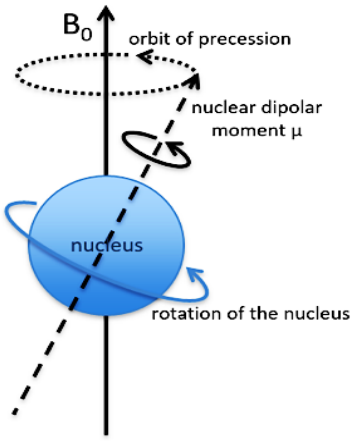


Figure 8 *Representation of the rotation of magnetic nuclei and its magnetic field in addition to the precession movement (http://brussels-scientific.com/?p=6273)*

At that moment both B0 and M are orientated in parallel in z axis. When an extra magnetic field B1(or radio frequency **RF**) is introduced perpendicular to B0 and at a precise energy matching the precession frequency of specific atomic nuclei, it becomes possible to switch M from its initial position to the y axis. This change is achieved through the absorption of energy by the atomic nuclei, prompting their transition from lower energy states to higher energy states ***(Figure 9).***

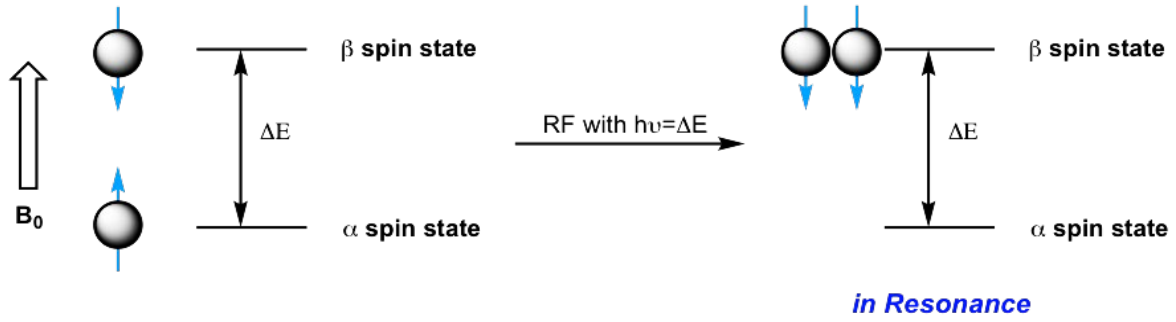


Figure 9 *Transition from the α to β state in the resonance process involves an initial magnetic field (B0) and an energetic difference (ΔE) between the α and β states.(https://kpu.pressbooks.pub/organicchemistry/chapter/6-5-the-nmr-theory/)*

Stopping RF allows the nuclei to release the excess energy acquired during resonance by reverting to a stable state. Concurrently magnetization M returns to its equilibrium alignment to B0 through a precession movement around the main magnetic field named relaxation **(*Figure 10).***

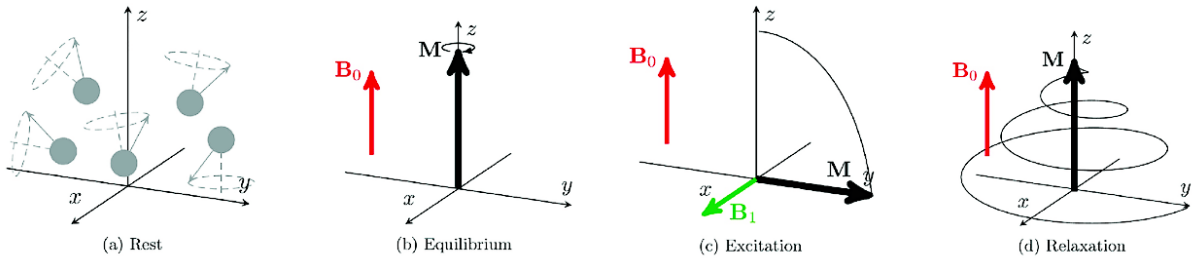


Figure 10 *Demonstration of the NMR experiment: (a) In the absence of a magnetic field, spins exhibit random orientations. Upon applying a B0 field along the z-axis, spins undergo precession around this axis, leading to (b) the establishment of equilibrium magnetization aligned along the z-axis. The equilibrium magnetization is manipulated toward the transverse xy-plane by the influence of an RF-pulse (B1) applied at the resonance frequency (c). Upon cessation of the RF-excitation, the magnetization gradually returns to its equilibrium state (d) with a precession frequency determined by the magnetic properties of the isochromatic being studied. Notably, the magnetization shift induced by the RF-pulse occurs at a rate approximately two or three orders of magnitude faster than the relaxation process.*

1.2.1.3 Signal detection

The NMR signal acquired after RF pulse is turned off and during the return of the magnetization to its equilibrium state. The way the magnetization decreases over the time is depended on two different relaxation processes: longitudinal relaxation (**T1**) and transverse relaxation (**T2**). The longitudinal relaxation T1 corresponds to the return of M in the alignment with B0 when RF pulse is stopped. Particularly, T1 corresponds to the time interval required for the longitudinal magnetization to recover 63% of its initial value. This value is influenced by molecular motions and chemical interactions and differs on different types of nuclei and their chemical environments. The transverse relaxation T2 refers to the decay of the transverse magnetization component perpendicular to the main magnetic field B0.

Specifically, T2 represents the time required from the transverse magnetization to lose 63% of its value after the RF stops. The value of T2 is always shorter or equal to T1 and depends on the interactions between neighbor spins. As the RF stops, the nuclei start to precession movement in way to return magnetization to equilibrium; this oscillating movement emits a radio frequency signal known as Free Induction Decay (**FID**). The FID contains information about the frequencies of the nuclei in precession and their interactions with their surroundings. To obtain a spectrum, a Fourier Transform (**FT**) need to be applied; this method allows to separate the several frequencies and determine their respective intensities forming a spectrum by converting the signal from a time domain (FID) to a frequency domain (NMR spectrum) ***(Figure 11).***

Figure 11 *Simplified diagrams to illustrate NMR experiment (https://kpu.pressbooks.pub/organicchemistry/chapter/6-5-the-nmr-theory/)*

1.2.2 Interpretation of ^{1}H-NMR spectra

The following sections elucidate the key parameters that defines a spectrum acquired through ^{1}H-NMR represented by chemical shift, multiplicity and intensity and peak area. These features play a crucial role in interpreting the spectral data ***(Figure 12).***

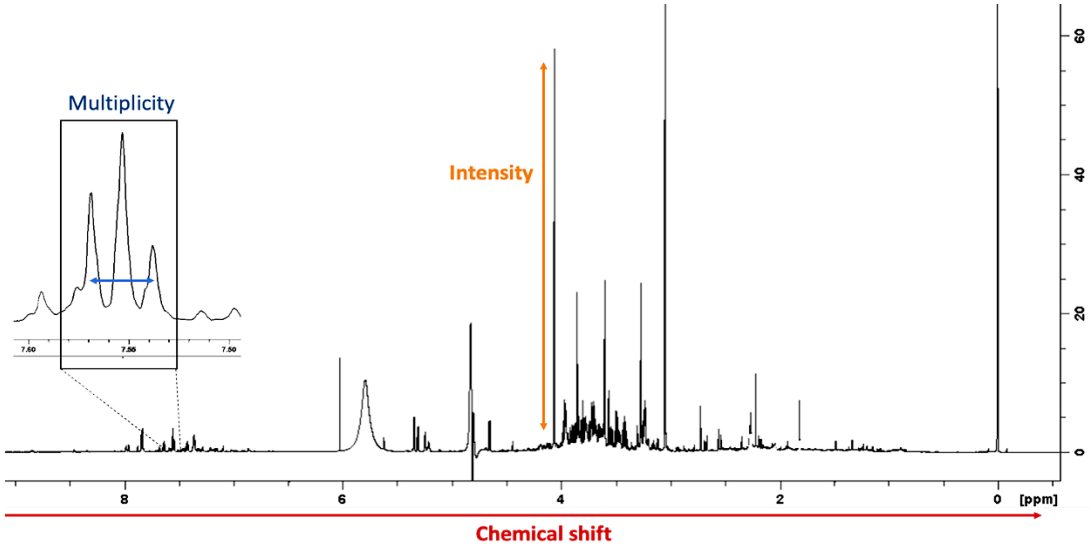


Figure 12 *Example of ^{1}H-NMR spectra of human urine sample; in red we can see the chemical shift scale for the signals; in orange the corresponding intensity; in blue the multiplicity related to each proton.*

1.2.2.1 Chemical shift

The chemical shift (δ) refers to the displacement of NMR resonance frequencies of atomic nuclei caused by their surrounding chemical environment. In particular, when looking at an ^{1}H-NMR spectrum, the position of a defined signal compound on the x-axis represents its chemical shift. This position is expressed in parts per million (ppm) by frequency and measured as follow:

$$\delta = \left(\frac{\nu - vref}{vref}\right) \times 10^{6}$$

where ν corresponds to the resonance frequency of the signal observed and ν_{ref} to the resonance frequency of the reference signal. In case of biofluids, like in this work, reference compound is represented by TMSP (3-(trimethylsilyl)-2,2,3,3-tetradeuteropropionic acid). The chemical shift represents indeed the resonance frequency of a proton relative to a reference; in a molecule

the chemical shift of each protonated group depends on its surrounding and in particular from electron density and electronegativity ***(Figure 13).***

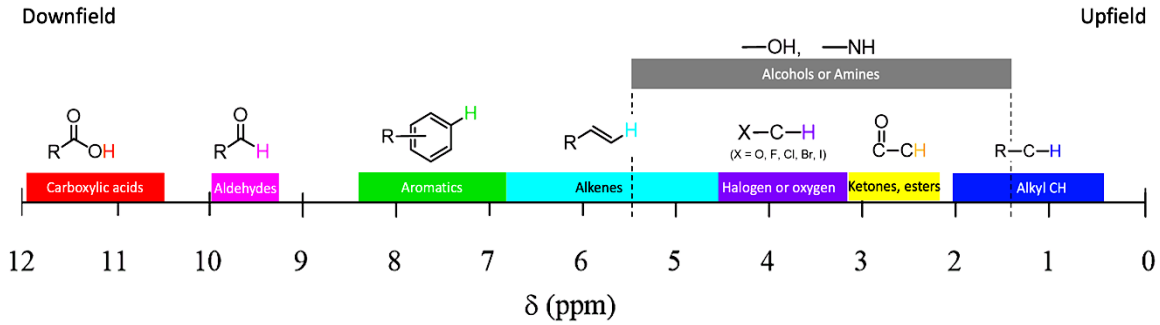


Figure 13 *Chemical shifts of protons engaged in different chemical groups.*

-Electrondensity

In a molecule the distribution of electrons around the nuclei determines the electron density within the atom. This distribution around nucleus affects its chemical shift through shielding and deshielding effects. The shielding effect happens when the electron cloud "shield" the nucleus from the full strength of the external magnetic field B0; this results in a lower observed resonance frequency (downfield shift) in the NMR spectrum. In opposite way, the deshielding effect takes place when nucleus experiences a stronger effective magnetic field, resulting in a higher observed resonance frequency (upfield shift) in the NMR spectrum. In general, a high electron density gives a lower chemical shift and vice versa.

- Electronegativity

Electronegativity is a measure of an atom ability to pull electrons towards itself in a covalent bond. Atoms with higher electronegativity have a stronger attraction on shared electrons, creating partial charges and influencing the electron distribution around neighboring atoms. This higher electron density enhances the shielding effect, reducing the effective magnetic field experienced by those nuclei. Consequently, these nuclei exhibit lower chemical shift values in the NMR spectrum.

1.2.2.2 Multiplicity or coupling

The signal originating from a proton can manifest as a collection of multiple peaks. This phenomenon is tied to the concept of multiplicity or coupling, which provides information about the count of adjacent hydrogen or adjacent carbon atoms. The multiplicity of a proton's resonance signal, represented by a peak, is determined by considering the neighboring proton(s) based on the following rule: n+1 multiplet ***(Figure 14).***

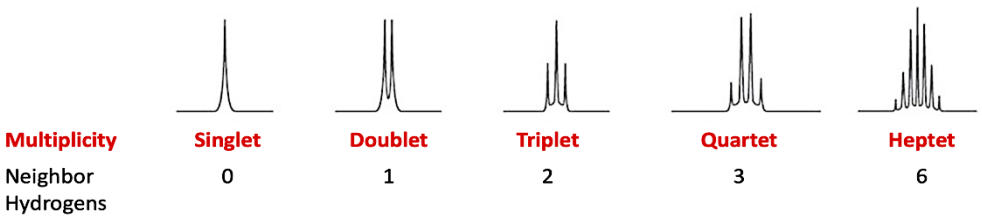


Figure 14 *Multiplicity of signals showing the number of corresponding neighbors' hydrogens.*

1.2.2.3 Intensity and area

If the peak intensity of two equivalent protons can give information concerning the relative concentration, it is more the peak area that will be used for quantification. Indeed, in NMR spectrum, peak area is directly related to both the molar concentration of the sample and the number of nuclei contributing to the signal independently from its multiplicity. A weak signal is observed in samples with lower concentrations, whereas an increase in concentration, by a factor of 2 or 3, leads to a corresponding increase in peak area ***(Figure 15).***

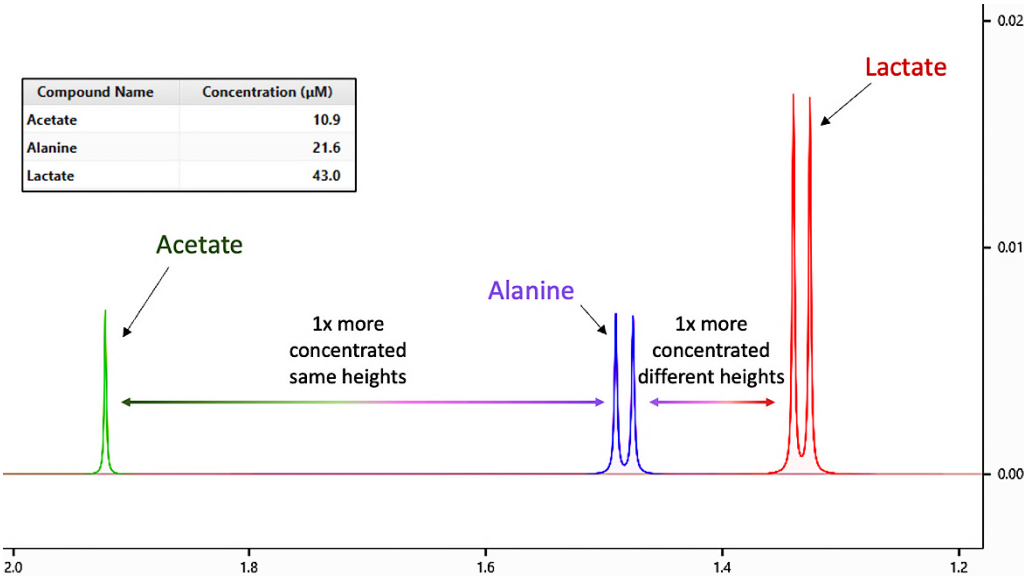


Figure 15 *The simulated spectrum shows the methyl signals from three distinct metabolites at different concentrations.*

1.2.3 Instrumental platform

The instrumental platform of NMR spectroscopy is composed of three major components: a superconducting magnet, a probe coil and the console. The superconducting magnet is kept under 9.5K with liquid helium generating the static magnetic field. The probe contains the RF coil and is situated in the bottom of the machine; the samples to be analyzed are placed in NMR tube and sent to the probe through an air flux from the top of the machine (bore). An adequate RF is generated by a RF transmitter and conducted to the sample through the probe; the console is on charge of receive the FID that are finally Fourier Transformed to produce spectra [42,43] . In the following ***Figure 16*** we can spot the different components:

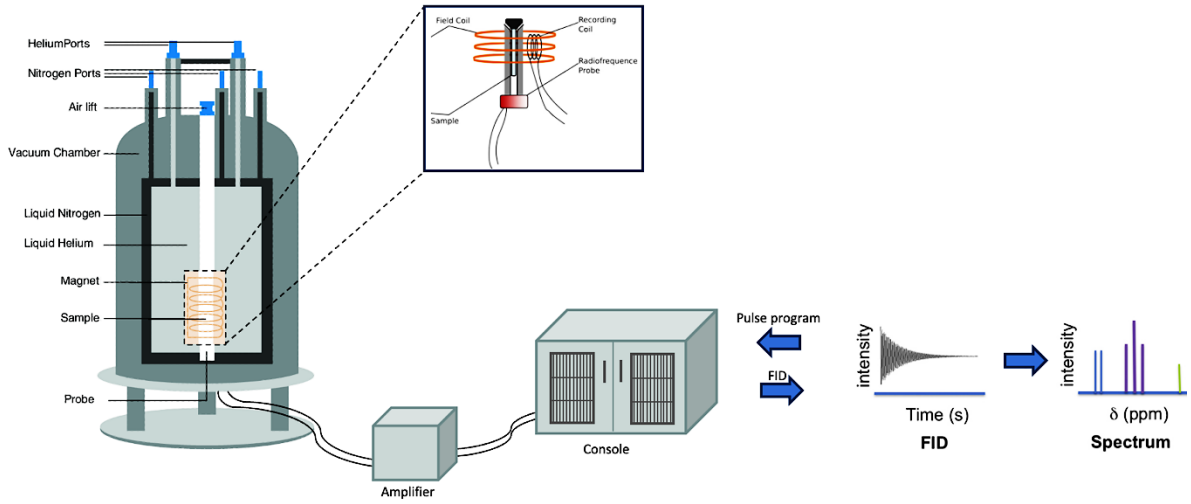


Figure 16 *Schematic representation of the instrumental platform for NMR spectroscopy (adapted from Antcliffe, David, and Anthony C Gordon. "Metabonomics and intensive care." DOI: 10.1186/s13054-016-1222-8)*

1.3 MS-based metabolomics

Mass spectrometry-based metabolomics is a highly utilized methodology in the field of metabolomics research. Among the various merits of this approach (as highlighted in the provided ***Table 1***), one of its key advantages contributing to its prominence lies in its remarkable detection sensitivity, capable of reaching measurements as low as the pico- or attomole scale. A metabolomics analysis conducted via the MS platform is typically composed of three important steps: at first, the sample is injected into a **separation system**, such as chromatography; following this, the metabolites enter the mass spectrometry system where they undergo an **ionization** procedure; subsequently, the ionized metabolites undergo analysis using a **mass spectrometry analyzer** and finally **detected** [44]. All these components will be described in detail in the following paragraphs ***(Figure 17).***

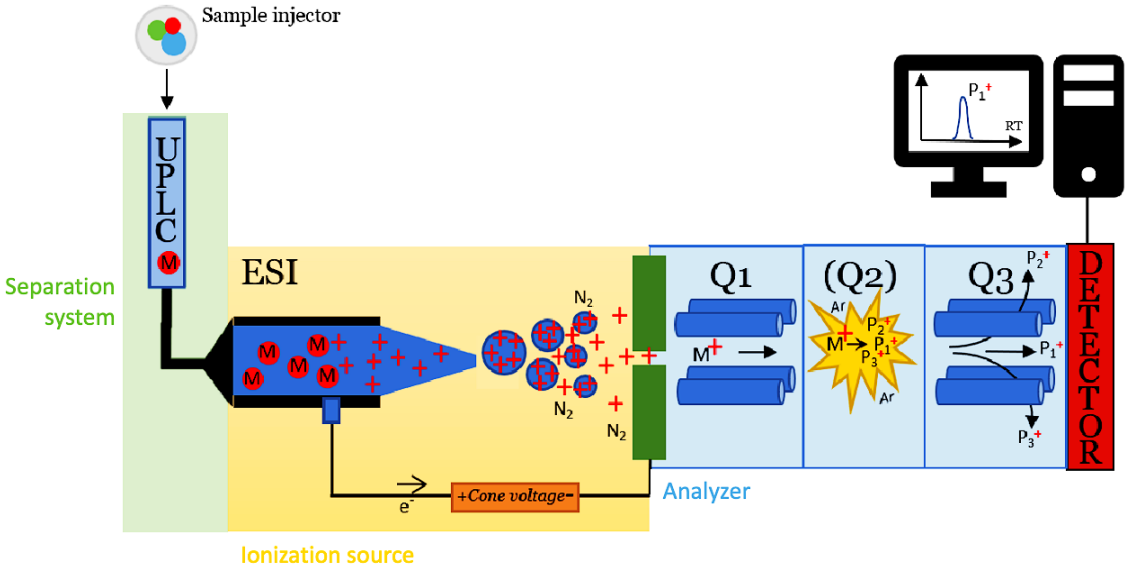


Figure 17 MS instrumentation for a Quadrupole system with all its different components (adapted from *Zimdahl, Anna. "Pharmacogenetic studies of thiopurine methyltransferase genotype-phenotype concordance and effect of methotrexate on thiopurine metabolism." DOI:10.3384/diss.diva-163614*).

1.3.1 Separation system

The predominant approaches to separating metabolites involve liquid chromatography (LC) and gas chromatography (GC), which are widely employed. Additionally, alternative methods to separation include supercritical fluid chromatography (SFC), capillary electrophoresis (CE), and ion chromatography (IC). The choice of separation technique depends on the nature of the metabolites being analyzed, the complexity of the sample, and the desired level of separation and resolution[45] as shown in ***Table 2***. In this work, liquid chromatography will be the technique mostly detailed.

Technique	Analysis time	Advantages	Disadvantages
GC	Minutes to hours	Allows high peak capacity small-molecule analyses; chromatography is highly reproducible	Is limited to volatile compounds; high temperature can induce decomposition
SFC	Minutes	Makes fast separations; is orthogonal to reverse-phase LC	Works best for nonpolar molecules
LC	Minutes to hours	Is applicable to many molecule types; allows high peak capacity small-molecule analyses	Has limited throughput because long separation times provide the best measurements
CE	Minutes	Has low sample consumption	Lacks orthogonality with DTIMS

Table 2 *Separation systems coupled with MS platforms with corresponding advantages and disadvantages (Zheng, Xueyun et al. "Coupling Front-End Separations, Ion Mobility Spectrometry, and Mass Spectrometry For Enhanced Multidimensional Biological and Environmental Analyses." DOI: 10.1146/annurev-anchem-061516-045212)*

1.3.1.1 Liquid Chromatography

The name chromatography takes its origins from the Greek words "Khroma" that means "Colors" and "Graphein" that mean "Writing" in way to define the technique able to separate the colored pigments of plants. Invented in 1903 by a Russian Italian botanist Mikhail Semyonovich Tsvet, liquid chromatography represents today the separative technique of reference [46,47].

Liquid Chromatography operates on the principle of separating molecules while in a liquid state, and this separation is characterized by the measurement of retention time (RT), which denotes the time taken by a molecule to traverse the chromatographic column.

Molecules separation is achieved through injection of a mixture of molecules is carried by the mobile phase into a column containing a stationary phase; this stationary phase is made up of functionalized particles that interact differently based on the different physio-chemical properties of molecules. Based on the type of molecules being studied, a different type of mobile and stationary phases can be used. In MS-metabolomics the two commonly used are

Reverse Phase Liquid Chromatography (RPLC or reverse phase) and Hydrophilic Interaction Liquid Chromatography (HILIC) [48] .

Reverse Phase Liquid Chromatography (RPLC or reverse phase) is mainly used to analyzing semi-polar and non-polar molecules. It is composed of a non-polar stationary phase (silica containing C18 and C8 particles) in which flows a mobile phase that varies between H_2O and organic solvent (such as acetonitrile and methanol) in way to generate a polarity gradient. Additives such as Formic acid or Ammonium formate can be added to mobile phase to adjust the pH and favorize molecules ionization [29]. This type of chromatography allows the rapid elution of higher polarity compounds that could be less retained by the stationary phase; at the opposite the less polar molecules could be strangely retained by the stationary phase and eluted at the end.

Hydrophilic Interaction Liquid Chromatography (HILIC) is mainly dedicated to the analysis of polar compounds. The silica of stationary phase is characterized by the presence of polar groups (such as alcohol, amino, amide or others) that may allow the retention of polar molecules through the mobile phase which composition varies between organic solvent and H2O. In metabolomics these methods see an increasing application for the analysis of lipids and polar compounds [49].

The efficiency in separation of LC methods is not only depending on the type of mobile and stationary phase but also from other factors linked to length of LC column and the particle size in the stationary phase [50]. Based on these column features two major methods of LC can be reported: **high performance liquid chromatography (HPLC)** and **ultra-high performance liquid chromatography (UHPLC)** that differ on the size of particles present in column (5µm and 1.8µm respectively). Precisely, the elution time and, consequently, the analysis time are influenced by the particle size and column length. A reduced internal diameter of the column and smaller particle size enhance the efficiency and resolution of LC by restricting the longitudinal diffusion of analyte particles. Additionally, this configuration leads to a shorter elution time, offering two significant advantages of UHPLC.

1.3.2 Mass spectrometry

Mass spectrometry (MS) is a sophisticated technology extensively employed in the branch of metabolomics. Beginning with the revelation of the electron's properties and its mass-to-charge ratio (*m/z*) measurement by J.J. Thomson in 1897, and subsequently the construction of the first mass spectrometer in 1912 by the same scientist, the domain of MS has undergone remarkable advancements and continues to be a central area of dedicated research. This technique measures the *m/z* in way to determine the monoisotopic mass of a given ion. Through the ion' detection information regarding its quantity in the samples can be obtained. In fact, the MS acquires a mass spectrum consisting of two dimensions: the x axis is the *m/z* range, and the ordinate is the intensity of the measured *m/z* ***(Figure 18).*** The diverse types of available spectrometers necessitate a selection based on the objectives and purposes of the intended investigation. From a broad perspective, a mass spectrometer can be envisioned as a composite of three components: an ion source (for sample ionization), an analyzer (to measure m/z values), and a detector (responsible for ion detection and generating the mass spectrum). These components are elucidated further in the ensuing sections.

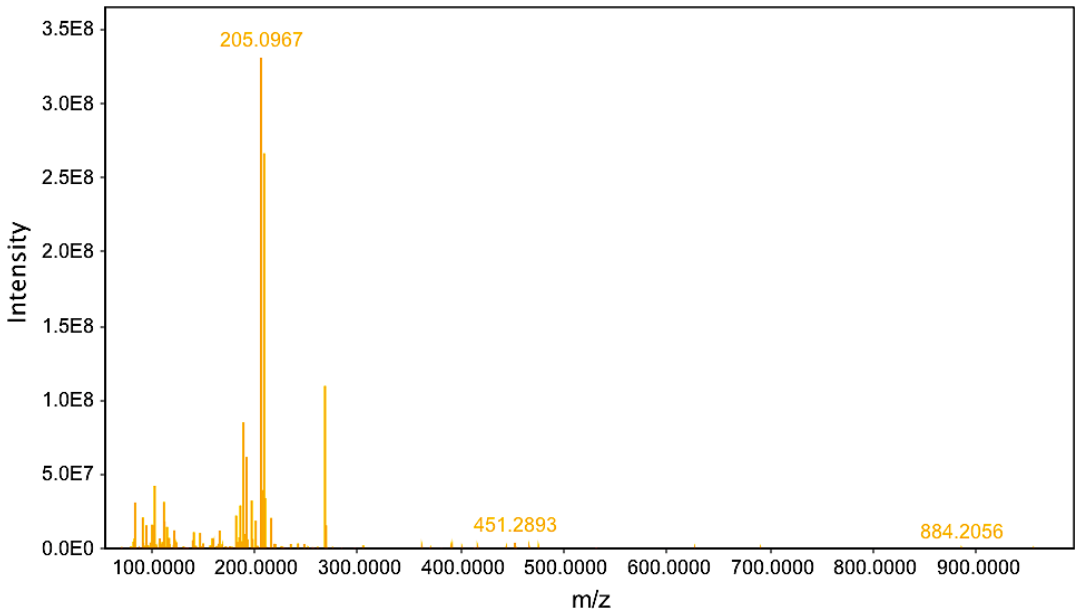


Figure 18 *Mass spectrometry urine spectrum showing in the x axis the m/z ration and on the y axis the intensity.*

1.3.2.1 Ionization sources

Different types of ionization sources are available in MS-based metabolomics, and they are often classified into hard and soft ionization methods. This classification depends on the amount

of internal energy received by the ions during their production [51]. Specifically, hard ionization techniques like Electron Impact (EI) and Chemical Ionization (CI) result in substantial fragmentation of the analyte molecules. In contrast to this technique, soft ionization such as Electrospray ionization (ESI) operates with low energy in way to ionize the molecules without dissociating them. This technique is well-suited to preserving the structural integrity of molecules, making them suitable to identifying molecular masses and determining molecular formulas. Several other techniques exist for such modes [52], but in this manuscript only one mode will be exploited: the electrospray ionization.

1.3.2.2 Electrospray ionization

Electrospray ionization is an ionization method described by Fenn in 1974 and is one of the most widely used method to ionizing liquid-phase samples [53]. The underlying principle of ESI centers on applying a high-voltage electric potential between the infusion capillary needle and the MS inlet. Specifically, the sample emerging from chromatography undergoes nebulization and is shaped into a Taylor cone, aided by elevated temperatures and a coaxial flow of gas. Within this mechanism, the droplet size diminishes due to solvent evaporation, resulting in increased electrical density on the droplet surface. Through this progression, a crucial threshold known as the "Rayleigh stability limit" is attained, at which juncture, electrostatic repulsion leads to a Coulombic explosion. This explosion, coupled with oxidation-reduction reactions, results in the release of ions in the gaseous state ***(Figure 19)*** [54].

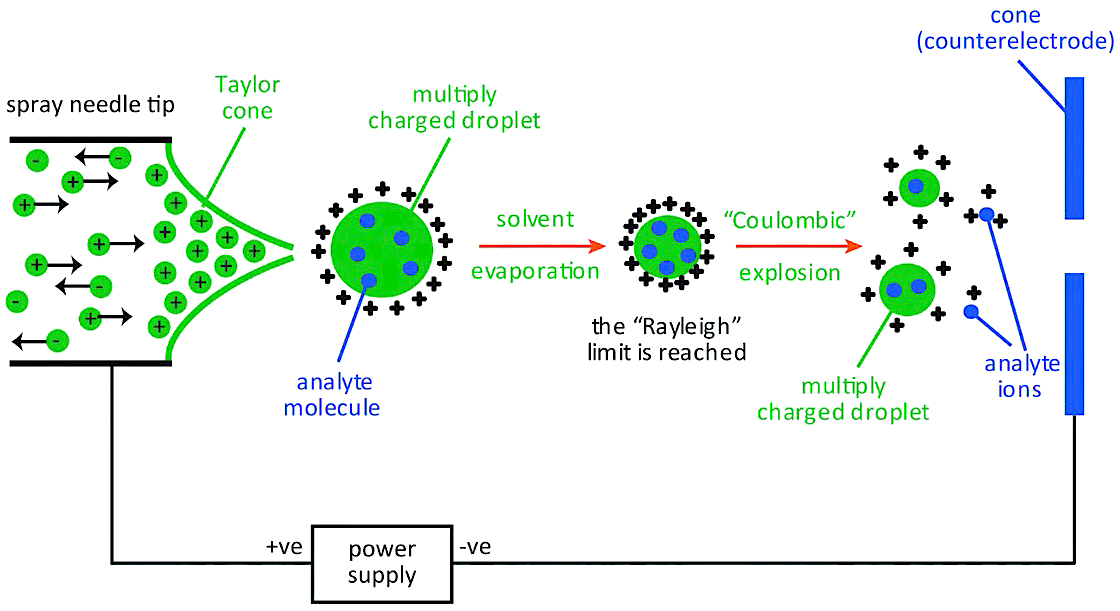


Figure 19 *Representation of how ESI source works (https://www.chm.bris.ac.uk/ms/esi-ionisation.xhtml)*

The ionization can be performed in positive mode **(ESI+)** and negative mode (**ESI-**). In positive mode the ionization of molecules with base properties is enhanced as they are able to capture positive charged proton (e.g. -NH_2 , -CO ,-CN etc.). The ESI- favors the ionization of molecules with acid properties that are vulnerable to deprotonation (e.g. -OH, -COOH etc.). ESI stands as a soft ionization technique facilitating the generation of molecular ions with minimal fragmentation by achieving high sensitivity and being easily coupled to commonly employed separation techniques like liquid chromatography (HPLC-MS and UPLC-MS)[55]. In the field of metabolomics, ESI is widely used because of its tendency to minimize ion fragmentation, which aids in the identification of molecules. However, it introduces complexities in result analysis due to redundancies in datasets. Therefore, the ability to analyze liquid samples encompassing polar and semi-polar, nonvolatile, and thermally unstable metabolites positions this approach as a cornerstone within LC-MS-based metabolomics [56,57].

1.3.2.3 Mass analyzer

Following the ionization process, the resultant ions are separated based on their m/z values through the mass analyzer. This component is responsible to performing m/z measurements through the separation of ions according to their respective m/z ratios. Generally, several attributes are employed to evaluate the performance of a mass analyzer and to guide the selection process for metabolomics analysis ***(Table 3)*** :

- **The scan rate** denotes the speed at which the mass analyzer conducts m/z ratio measurements across a specified range of masses. It is quantified as the number of mass units analyzed per second or per millisecond.
- **Mass accuracy**, often referred to as the error in the m/z value, reflects the precision with which the analyzer measures the m/z value of an ion in comparison to its theoretical value. This metric is expressed in parts per million (ppm) and is calculated using the formula: mass accuracy = [(mexperimental – mtheoretical)/ mtheoretical]x106
- **Resolving power (RP)** is associated with the capacity of a mass analyzer to differentiate between two overlapping m/z peaks (where the intensity valley between them reaches 10% of their intensities). It is computed using the formula RP = M / ΔM, where M represents the measured m/z value and ΔM signifies the disparity between the m/z ratios of the overlapping peaks.

- In terms of resolving power, analyzers can be categorized into two primary types: "low resolution" analyzers [58], which encompass the Quadrupole (Q), Ion Trap (IT), and Linear Time-of-Flight (ToF); and "high resolution" analyzers, which include the Reflectron Time-of-Flight, Fourier Transform Orbitrap, Magnetic Field and/or Electric Field Sectors, and Fourier Transform-Ion Cyclotron Resonance (FT-ICR) mass spectrometers. Within the scope of this thesis, our attention will be directed towards the Fourier Transform Orbitrap.

Instrument type	Resolving power	Mass accuracy[a]
Ion Trap	$\sim 10^3$	300–1,000 ppm
QTOF	$\sim 10^4$	5–40 ppm
FT-Orbitrap	$\sim 10^5$	<5–10 ppm
FT-ICR	$\sim 10^5$	<l–10 ppm

Table 3 *Table showing resolving power and mass accuracy for each type of mass analyzer (Ntai, Ioanna, and Neil L. Kelleher. "Approaches for Natural Product Detection and Structural Elucidation Using Mass Spectrometry with High Mass Accuracy." DOI: 10.1017/CBO9780511996634.010)*

1.3.2.4 Quadrupole

The quadrupole, a low-resolution mass analyzer, was invented in 1953 by Wolfgang Paul and Helmut Steinwedel [59], marking it one of the earliest analyzers to be employed. Its configuration consists of four parallel metallic rods organized in a square or rectangular pattern. Each pair of opposing rods is connected to the same voltage. An alternating combination of Radio frequency/Direct current (RF/DC) voltage is applied between the two pairs of rods. As an ion enters the quadrupole, an attractive force is generated on it by the rod whose charge opposes its ionic charge. With a periodic voltage applied, alternated attraction and repulsion forces in the x and y directions over time (stable trajectory), permits the ion to traverse the analyzer without colliding with the rods and ultimately reaching the detector [60]. Under specific electric conditions, only ions possessing defined m/z ratios can traverse the analyzer by making the quadrupole works as a mass filter ***(Figure 20).***

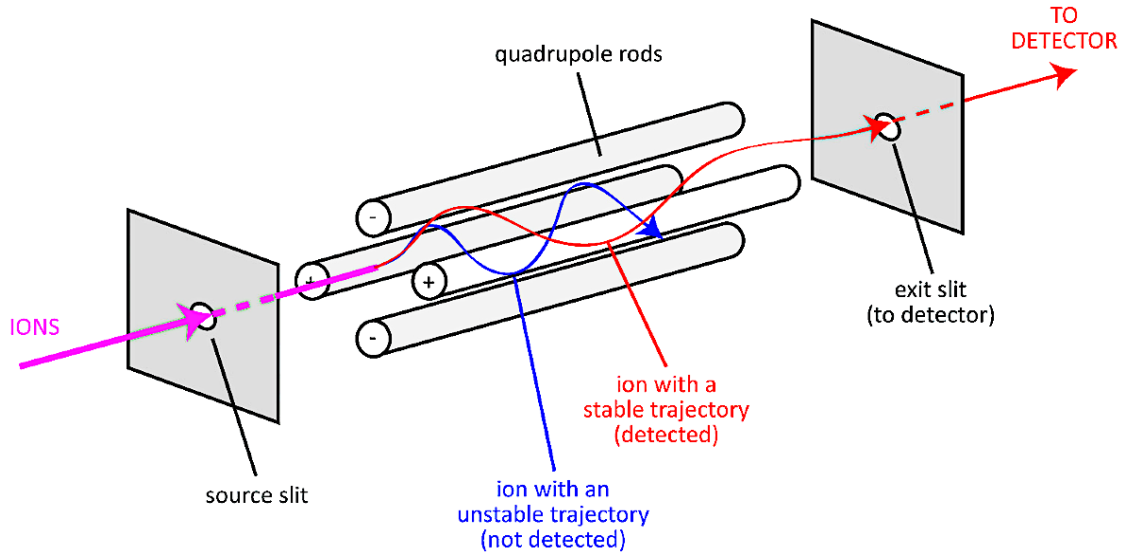


Figure 20 *Quadrupole instrument with its components and its way to work (https://www.chm.bris.ac.uk/ms/quadrupole.xhtml)*

1.3.2.5 Fourier Transform Orbitrap

The Orbitrap, a frequently employed high-resolution mass spectrometry instrument for metabolomics profiling experiments, has evolved since its initial appearance in 1923 within a physical review[61]. Numerous technological improvements have been made over the years, culminating in the 1999 version developed by Alexander Alexeyevich Makarov, which represents the base of the current commercially available[62] version.

This device is constituted by an ion trap composed by two electrodes: an external barrel-shape electrode and a coaxial electrode. Ions are trapped within the inner region of the analyzer through an electrostatic field produced by these two electrodes. The movement of ions within the electrodes involves a combined motion, including both a rotational component around the central electrode and an axial movement along the same electrode. Each ion possessing a specific m/z ratio consequently experiences a distinctive velocity (or time of flight), leading to a corresponding rotational frequency. This frequency gets recorded by a detector and subsequently amplified; then, the recorded signals undergo a Fourier Transform, enabling the calculation of the precise m/z ratio for each identified ion ***(Figure 21)*** [63,64,65].

Enhanced precision of m/z is achieved when the duration of ion trapping is extended. Furthermore, the resolution of the Orbitrap is based on the scan rate; as the scan rate rises, both the resolution and the precision of m/z measurement decrease.

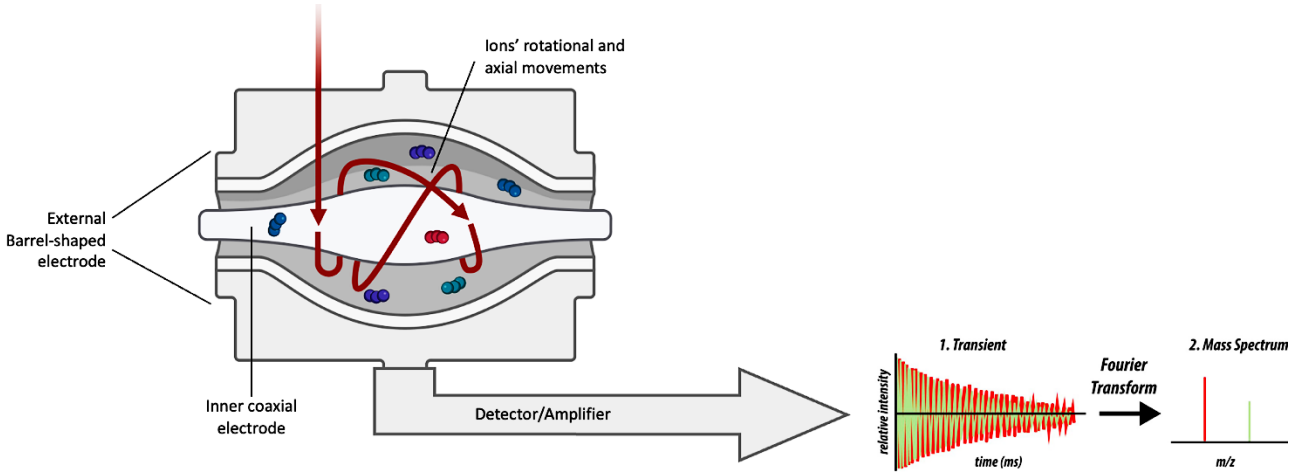


Figure 21 *Orbitrap mass analyzer: from sample injection until mass spectrum acquisition (adapted from Savaryn, John P et al. "A researcher's guide to mass spectrometry-based proteomics." DOI: 10.1002/pmic.201600113)*

1.3.2.6 Hybrid mass spectrometers for Tandem mass spectrometry

Tandem mass spectrometry (MS/MS) is a dual stage MS technique that target the fragmentation of an ion in order to identify its structure. It involves the use of two sequential mass spectrometry experiments, where the output of the first mass spectrometry stage (called "parent ion or "precursor") serves as the input for the second stage. This precursor traverses a "Collision Cell," a site where its fragmentation occurs. This process results in the acquisition of an MS/MS spectrum, which can be employed for determining structural characteristics or conducting spectral data searches. Various mechanisms exist for MS/MS fragmentation; however, our focus in this study centers on Higher-energy C-trap Dissociation (HCD). HCD operates by introducing the precursor into a collision cell containing a high-pressure inert gas (e.g., Argon, Nitrogen). Once in the collision cell, the ion experiences defined RF and collision energy, inducing its motion; this movement initiates collisions between the precursor ion and gas molecules, augmenting the ion's internal energy and leading to fragmentation in accordance with the applied collision energy [66,67]. In the case of instruments like the Q ExactiveTM Plus, two available acquisition mode are available: the first mode is represented by the simple MS acquisition (full scan), and the second is in tandem mode. After being passed through ESI source, ions pass through the quadrupole, that works as a prefilter based on the selected m/z ratios; at this point, pre-filtered ions pass in the C-trap. In full scan mode the ions pass directly

in Orbitrap analyzer for m/z detection and spectra acquisition. In case of MS/MS mode, ions are transferred from C-trap into HCD cell filled with Nitrogen gas where fragmentation happens; once the ions have been fragmented, they accumulate within the C-trap before being directed to the orbitrap for enhanced sensitivity of detection ***(Figure 22).*** .

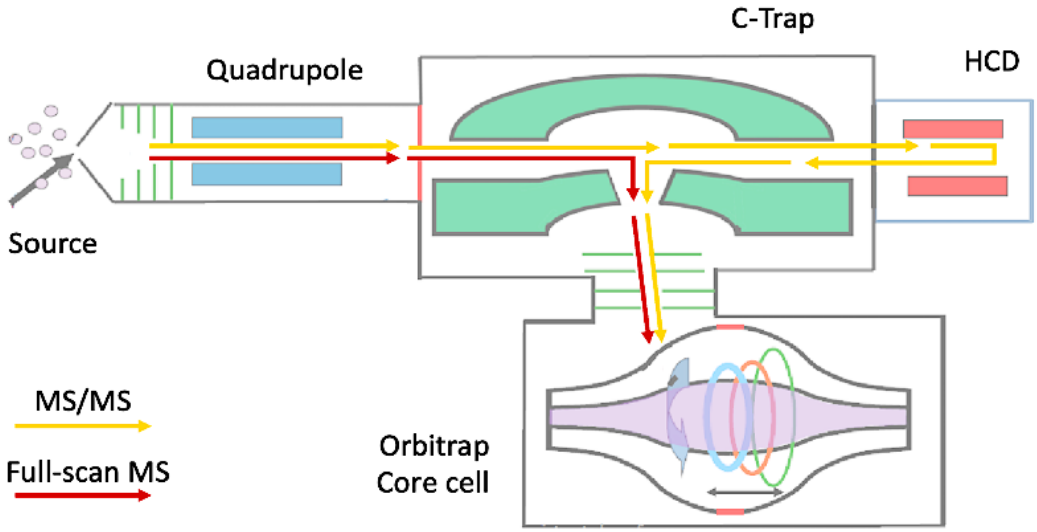


Figure 22 *Tandem Mass spectrometer with two different acquisition modes showed: MS/MS mode and full-scan MS (adapted from Kaufmann, Anton. "Analytical performance of the various acquisition modes in Orbitrap MS and MS/MS." • DOI: 10.1002/jms.4195)*

1.4 Untargeted metabolomics workflow

The metabolomics workflow involves a series of steps designed to comprehensively analyze the small molecules present in a biological sample. It aims to provide insights into the metabolic pathways, biomarkers, and overall biochemical composition. The typical untargeted metabolomics workflow aims to explore all possible metabolites linked to a defined condition in way to improve the knowledge about biochemical processes and pathways and by positioning itself as a comprehensive exploratory tool. The several generic and essential steps of a metabolomics workflow need to be accurately defined before the analysis and may differ according to several aspects, such as biofluid choice or instrumental technique ***(Figure 23).*** Several metabolomics-based studies have been published in the last decades each characterized by its distinctive original workflow[68,69,70]. In the forthcoming sections, an in-depth exploration of the multiple stages constituting the untargeted metabolomics workflow will be presented. This exploration will encompass both NMR and MS platforms.

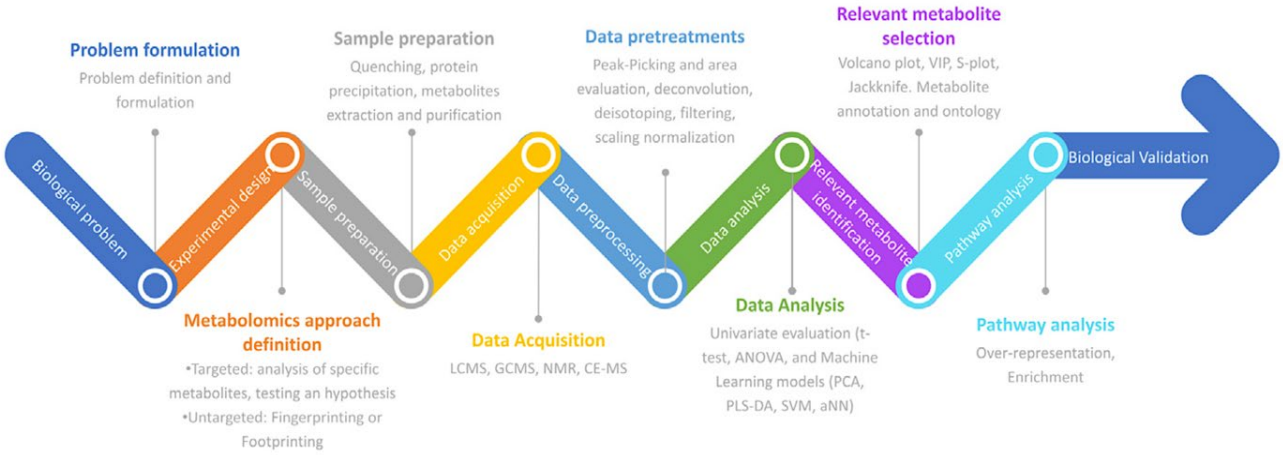


Figure 23 *General workflow for metabolomics studies (Lee AY, Troisi J, Symes SJK. "Experimental design in metabolomics." DOI: 10.1016/B978-0-323-85062-9.00002-7)*

1.4.1 Biological problem and experimental design

A well-structured experimental design is crucial to obtaining reliable and meaningful results in a metabolomics study. A thoughtful experimental design includes several key aspects that help minimize bias and ensure that the study objectives are met. The process starts by deciding on the focus and purpose of the experiment. Based on the focus of the study, the design of the metabolomics experiment can begin to take shape [71]. As first point the scientist must pose himself these questions: What aspect of the metabolome do you intend to investigate? Are you

looking for biomarkers, elucidating pathways, or exploring responses to specific treatments? Once this point is elucidated, the subsequent essential phase involves determining the appropriate sample size and type. It is crucial to meticulously select the biological specimen (Tissue? Urine? Plasma?) that can best answer the biological question. Conducting a power analysis to determine the most suitable sample size according to research's goal is another intricate aspect to address, aiming to achieve the expected effect size. This aspect is closely linked with considering the type of study design, which entails deciding among options like cross-sectional, longitudinal, case-control, or other designs, based on the inherent nature of the research question. In the case of a case-control study, ensuring an adequate number of samples for each group becomes imperative to attain a meaningful biological response. Another crucial aspect to carefully manage is the selection of participants, aiming to minimize variations in inter-individual factors (such as age, sex, BMI, lifestyle, etc.) and intra-individual factors (such as diet, circadian cycle), which may introduce confounding factors. In way to overcome these biases, the most homogenous population is needed, and in this sense, a well-defined inclusion/exclusion criteria list can be edit helping researchers or clinicians in patients' recruitment. Based on these considerations, correct sample size should be based on statistical power analysis. Particularly in metabolomics studies characterized by significant variability influencing metabolites and in which multiple outcomes are tested, ensuring an adequate and minimum sample size is mandatory. To achieve this, integrating power analysis into the experimental design of metabolomics studies becomes essential to enhance robustness of study findings by allowing detection of meaningful effects with a good level of confidence.

1.4.2 Sample collection and preparation

Inadequate sample collection, storage and pretreatment represent some of the many sources of variation and crucial limits in metabolomics studies. All these pre-analytical factors play a key role in the repeatability and robustness of metabolomics studies and need to be well controlled [72 ,73]. Nutritional status and circadian cycle are two of the major impact factors on variability of metabolites; indeed, when collecting plasma or serum, the collection of these specimens in a fasting or non-fasting condition can have a high impact in concentration of glucose and other amino acids. For example, fasting condition is mandatory to investigate plasmatic difference in population with diverse dietary habits [74]; the concomitant impact of sleeping and feeding was also reported as strongly affecting the metabolome in another interesting study [75].

In the context of urine, representing the biofluid of choice of this thesis, sampling time can be done over 24 hours or by timed collection (first- or second-day urine). Several factors differentiate the two-sampling time such as the feasibility of collection and metabolites concentration. The easiest and less-invasive approach is represented by timed sampling such as the second-day urine, like done in our work. This approach allowed us to better monitor the evolution of specific biomarkers that may be less concentrated in 24h urine and additionally to "standardize" patient urine collection in hospital during medical examination. The choice of one or another sampling time strongly affects the level of metabolites; indeed a study conducted on healthy volunteers shows how urinary content strongly depends on time collection due to feeding status, physical activity or other factors[76]. Concerning sample preservation the addition of antimicrobial additives (such as sodium azide or boric acid) to prevent bacterial overgrowth is commonly used in urine protocols [77,78] . In terms of sample storage, no changes in metabolites were observed after 6 months at −80 °C, in opposite to freeze–thaw cycles that exhibit to significantly change urine sample composition [79,80]. In case of urine specimen, dilution effect represents a major problem due to difference in metabolites concentration. For these reasons, several normalization strategies can be applied, and they can be divided in ***pre-*** or ***post-acquisition normalization*** (this last procedure will be better discussed in section 1.4.5). In case of ***pre-acquisition normalization***, samples can be diluted to a common concentration before analysis based on specific gravity or osmolarity values specific to each sample. The use of one normalization or another is strictly dependent from sample cohort.

1.4.3 Data acquisition

1.4.3.1 One dimensional ^{1}H-NMR metabolomics

Depending on the type of biofluids and the goals of the study, different acquisition sequences can be adopted in ^{1}H-NMR analysis. Indeed, the choice of a sequence which matches the specific composition of a samples is a fundamental step in way to enhance the metabolite coverage and spectral resolution. The major challenges for metabolite profiling in biofluids are represented by the presence of water, and in some cases by interference from massive amounts of proteins.

By revisiting the principles of ^{1}H-NMR, we can recognize the primary advantage of this method, which lies in the abundant presence of ^{1}H atoms within nearly all recognized metabolites facilitating their straightforward identification in biofluids. However, these same ^{1}H atoms are also present in water molecules, which serve as standard medium to suspending biological metabolites. Consequently, this main challenge arises when examining urine or filtered plasma/serum samples, as the dominating resonance of water can overshadow the NMR signal. The most effective approach to overcome the influence of water is through the use of **1-D-NOESY** (Nuclear Overhauser Effect Spectroscopy, also known as ***NOESY-presat***). This sequence yields NMR spectra characterized by remarkable reproducibility and a substantial decrease in water signal, all while maintaining the integrity of the other spectral peaks. This methodology involves a sequence of three consecutive pulses, where the final two pulses serve as the NOESY filter. These two pulses are separated by a brief mixing time, allowing a more uniform sample excitation and enhanced water suppression [27,81]. Consequently, this pulse sequence currently represents the reference choice for NMR-based metabolomics studies of biofluids.

When dealing with intact serum or plasma samples, a major challenge in metabolite profiling arises from the substantial interference posed by a substantial protein content (6-8 g/dL). Without the application of an appropriate sequence, the NMR spectra will be overwhelmed by the prevalence of wide resonance peaks coming from these molecules, resulting in insufficient resolution of the sharper signals originating from other metabolites. In such scenarios, employing a sequence that filters protein signals while amplifying the sharper peaks originating from smaller molecules, becomes essential. The Carr-Purcell-Meiboom-Gill (***CPMG)*** experiment is based on the distinction in relaxation time (T2) between large molecules like

proteins, lipoproteins, and lipids, and smaller molecules. By being coupled a water presaturation pulse to suppress the water signal, this filter effectively enhances resolution by reducing signals from quickly relaxing species. Consequently, it improves the resolution of smaller molecules. [27,82] ***(Figure 24).***

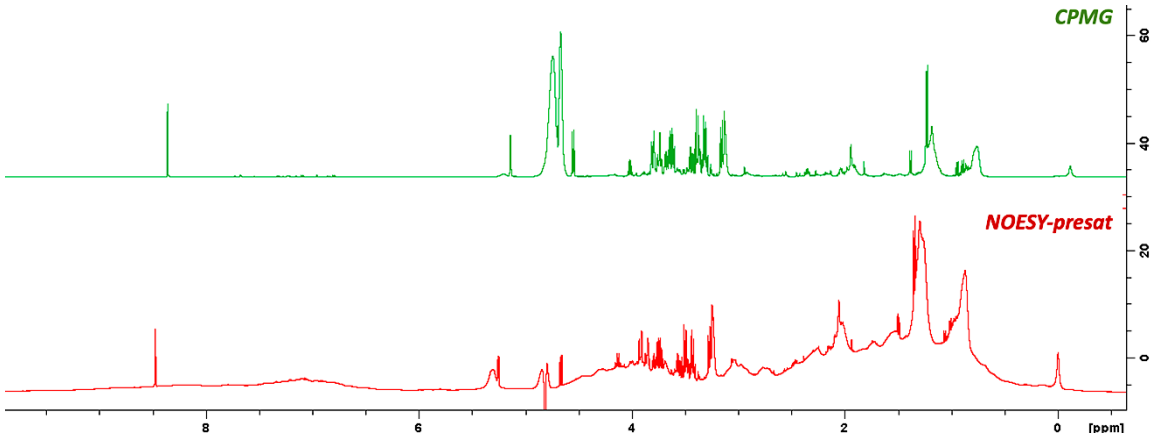


Figure 24 *Spectra of human plasma analyzed through CPMG and NOESY-presat sequences. As we can see, the green spectra (upper) represent the plasma sample on which the CPMG filter allows to obtain a better resolution of smaller molecules by filtering protein signals; in the red spectra (down) NOESY-presat sequence is applied to the same sample by showing the interference of wide peaks from protein content which generate loss of signals of all other small metabolites present in the sample.*

1.4.3.2 Two-dimensional NMR metabolomics

Two-dimensional nuclear magnetic resonance (2D NMR) is increasingly being investigated as a tool for metabolomics due to its superior resolution compared to one-dimensional NMR. Indeed, the major advantage of 2D NMR spectra is represented by the peaks distribution along two orthogonal dimensions (typically 1H–1H or 1H-13C) instead of an unique one, effectively reducing peak overlap while providing essential insights into atomic connectivity. The most commonly employed 2D pulse sequences include Correlation Spectroscopy (***COSY***) and Heteronuclear Single Quantum Correlation (***HSQC***), each focusing on distinct aspects of molecular structure analysis. The COSY sequence primarily examines homonuclear H correlations, uncovering through-bond interactions to delve into the structure of molecules and reveal their specific interconnections. In contrast, HSQC is an experiment designed to visualize connections between H spins and their directly linked C spins[83].

By leveraging the primary strengths of these experiments, 2D NMR techniques were employed in this thesis to enhance signal resolution and improve metabolite elucidation. However, despite these advantages, several limitations characterize 2D NMR. The time-consuming nature of

these experiments can potentially impact the stability of analyzed biofluids over time. Furthermore, the lengthy acquisition time makes the analysis of large sample cohorts impractical. Another significant drawback is the current lack of tools capable of adequately preprocessing 2D NMR data. Extracting meaningful information from raw data remains a challenging task, and routine utilization of these techniques is still distant [13,84].

1.4.3.3 MS metabolomics

Different strategies for data acquisition have been introduced in metabolomics to better capture the composition of metabolites within a biological sample. The choice of acquisition method, like other parameters, depends on the research objectives, the complexity of the sample or the sensitivity required. These methods allow researchers to gain insights into the types and quantities of metabolite and are represented by three major modes: full-scan, data-dependent acquisition (DDA), and data-independent acquisition (DIA).

Full scan mode allows, through the processes explained in *paragraph 1.3.2.4*, the acquisition of an MS spectra composed by the entire set of m/z ratio and metabolites abundance. It provides an overview of the entire mass range and a detailed chromatographic peak shape in MS1 data[85,86].

The principle of **data dependent acquisition (DDA)** involves an initial high-resolution MS analysis where the complete range of m/z ratio is acquired. In the subsequent step, the most intense signals (ion exceeding a predefined threshold) are selected, fragmented, and the subjected to MS/MS analysis. The MS/MS data generated in DDA holds valuable information for metabolite identification by comparing it with a spectral library[86,87].

In **data independent acquisition (DIA)** in contrast to DDA mode, a larger precursor ion mass width to fragment is selected to obtain MS/MS data. Through this mode a higher quantity of MS2 spectra is acquired with the major advantage of detection of metabolic features with lower MS abundance. This aspect determines also its primary handicap related to increased spectra complexity and higher difficulties in features identification since the link between precursors and their fragment ions are dissociate (fragment ions may result from multiple precursor ions present in the same m/z range).

By taking advantage of the extensive coverage of full scan mode and the qualitative insights from DDA allowing an easier features identification, modern untargeted metabolomics studies, including this work, combine these approaches. As first, each sample undergoes LC-full scan MS analysis to gather precise m/z ratio and relative abundance data for all metabolic features. Following data preprocessing and the selection of interesting features, DDA analysis is performed to acquire fragmentation spectra, thus verifying the chemical identity of the chosen metabolites [88] ***(Figure 25).*** In the context of this thesis, a full-scan analysis was done on all samples and DDA mode was only performed on pool sample for features identification purposes.

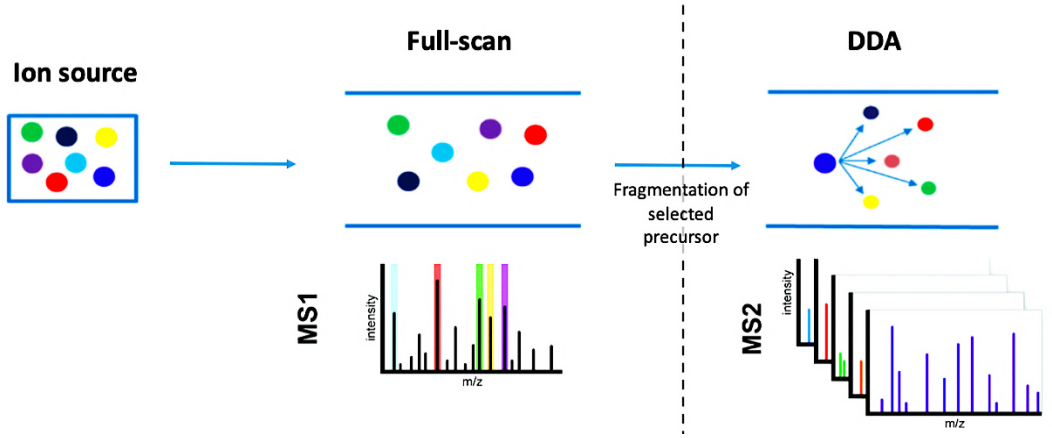


Figure 25 *Differences between Full-scan and DDA acquisition modes. In full-scan all the m/z ration detected are reported; in DDA mode only selected ions are fragmented for generate singles MS spectra. (adapted from: Villalta, Peter W., and Silvia Balbo "The Future of DNA Adductomic Analysis" DOI: 10.3390/ijms18091870)*

1.4.4 Data preprocessing

Analysis through NMR and LC-MS platforms generate a big amount of data with high level of complexity. In way to pass from spectral data to interpretable results, a fundamental step is need which can ensure the quality of raw data and limit possible biases. This step is represented by the preprocessing. Preprocessing aims to correct for technical variations, remove noise, and standardize the data, ultimately ensuring that meaningful patterns and differences in metabolite levels are accurately captured. Effective preprocessing enhances the reliability and interpretability of metabolomics data, enabling meaningful comparisons, pattern recognition, and biomarker discovery. Throughout the preprocessing phase, there is a transition in terminology from "feature" to "peak" and ultimately to "metabolite." Specifically, the term ***"feature"*** serves as a broad descriptor for a measurable characteristic derived from analytical data, such as data points or variables in the dataset. More specific terms include ***"metabolite",*** which refers to a specific chemical compound identified in the sample, and ***"peak",*** which represents a distinct signal in spectra indicating the presence of a particular compound (observed at specific m/z values or chemical shifts). This progression in terminology reflects the refining of the analytical information from broad characteristics to specific chemical entities during the preprocessing of metabolomics data which enable the switch from spectra information to interpretable biological results. The specific preprocessing steps adopted can vary depending on the nature of the data, the used analytical techniques, and the research objectives. In this thesis, the principal preprocessing steps of NMR and MS based metabolomics will be detailed.

1.4.4.1 NMR preprocessing

NMR-preprocessing is composed of several steps which allow to pass from the FID or raw data to a table containing all the features used for the statistical analysis and data interpretation. In the last decades several commercial software and web-based platforms have been develop each one including all the essential steps for a correct data preprocessing. The commercial tools nowadays available are composed by MestReNova© and TopSpin© which allows an easy and friendly solution for NMR preprocessing. In the case of TopSpin© only simple spectra manipulations are allowed such as baseline and phase corrections; instead, MestReNova© software allows a complete manipulation of spectra by proposing different steps that must be well chosen and handled by the user in way to obtain a robust workflow. Free web-based

available tools and R packages are composed by WorkflowforMetabolomics (W4M) NMRProcFlow and PepsNMR. Concerning W4M [89] which includes several steps of PepsNMR packages, is a tool working on Galaxy platform which allows data treatment starting from the FID until the statistical analysis. This represent a powerful tool whose only disadvantage is represented by its non-friendly interface. NMRProcFlow [90] is an open-source software providing a complete set of tools for processing and visualization of NMR data, with and interactive interface based on a friendly spectra visualization. As R package, PepsNMR [91] developed by M. Martin and B. Govaerts from UCLouvain in collaboration with J. Leenders and P. de Tullio from ULiege allows robustness and flexibility among a complete series of pre-processing functions. All the cited tools are operator dependent thanks to the variety of parameters that can be set based on the specific biospecimen or cohort analyzed. In the current of this work TopSpin© for baseline and phase correctios and MestReNova© for alignment and bucketing were used for preprocessing of NMR spectra ***(Figure 26).***

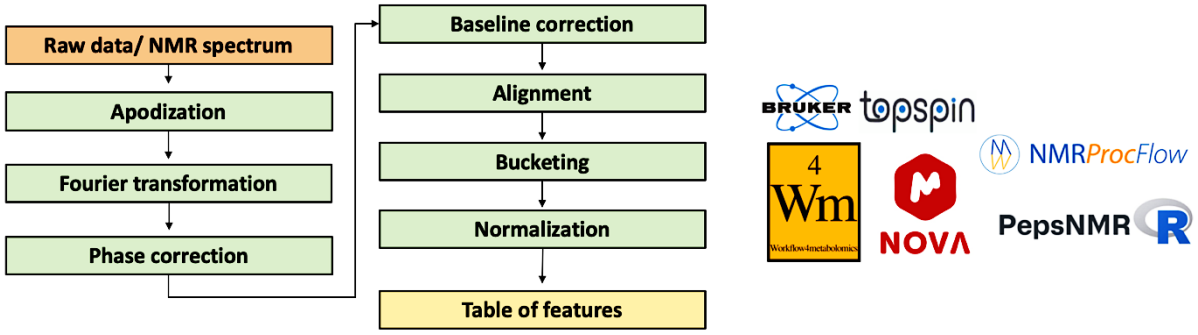


Figure 26 *General workflow for NMR data preprocessing reporting the main step from data importation until generation of features' table; several preprocessing tools and platforms are reported.*

Apodization

Apodization refers to a signal processing technique used to modify the appearance of a spectral line in an NMR spectrum. This step involves multiplying the FID signal by a mathematical function before performing a Fourier Transform to convert the signal into a frequency-domain spectrum. For instance, considering the function e^(-πLB), where LB represents line broadening, when LB is increased, the signal-to-noise ratio (SNR) also increases. In the case of NOESY-presat, a value of 0.3 is typically employed for LB [92,91].

Fourier Transform

This procedure enables the conversion of the intricate signal into a comprehensible spectrum. Each signal undergoes transformation into peaks characterized by a specific height, position, and width, determined by the proton's aptitude, frequency, and relaxation time (T1). The resulting spectrum exhibits peaks with a chemical shift, commencing from 0 ppm and progressing towards the left.

Phase correction

The purpose of phase correction is to eliminate unwanted phase shifts that can result from instrumental imperfections, magnetic field inhomogeneities, or other factors. The goal is to align the baseline of the spectrum and eliminate unwanted distortions such as baseline tilting or peak distortion. Two type of phase correction can be applied. Correction to order 0 is applied to the whole spectrum in the same way and aims to account to any phase shift independently of the signal frequency. Correction to order 1 is a correction type proportional to the frequency of the signal.

Baseline correction

Baseline distortion can appear in NMR spectra due to pulse imperfections and non-linearity of electronic detection process. The process of baseline corrections consists in correcting the baseline to zero or a consistent reference level, allowing for better visualization and analysis of the actual NMR peaks.

Alignment

Alignment refers to the process of adjusting and matching the positions of peaks or signals in multiple NMR spectra. I isa crucial preprocessing step in cases where multiple spectra need to be compared though bucketing process as happens in metabolomics studies. Several parameters can influence the misalignment of signals in different samples such as instrumental factor or sample variability. The major cause of misalignment is generated by pH variation, that especially in urine samples, represents a great source of errors. Different algorithms are available for alignment process each with its advantages and limitation [93,94]; the choice of one compared to another strictly depends from the obtained results.

Bucketing

Bucketing (or binning) is a step that allows to reduce the complexity of the data and facilitate downstream analysis by generating a table composed by all the features present in the spectrum. It consists in cut the spectrum into discrete regions or "buckets" all having the same width (usually between 0.05 and 0.02 ppm). Bucketing simplifies the data by reducing the number of variables (individual data points) while maintaining the relevant information about the metabolite profiles by making them more amenable to statistical analysis, visualization, and pattern recognition. Moreover, in analysis of biospecimen such as urine, it allows to partially overcome problems due to peaks alignment by englobing all the signals placed in a defined spectra width to a single bucket. Even if this technique has several advantages like the variable reduction, this segmentation generates sometimes the separation of a unique signal in two different buckets dividing the chemical information between several bins. As previously showed, the application of this method must be done carefully since the "quality" of the bins are highly dependents from the alignment step.

Spectral deconvolution

As we have previously discussed, a straightforward method to quantifying metabolites in ^{1}H-NMR spectra involves integrating peaks following a bucketing process. In an ideal scenario, each bucket would correspond to a metabolite, and its integrated value would directly relate to its concentration in the sample. However, NMR spectra often display overlapping peaks, particularly in complex samples, leading to integrated values that may not accurately reflect actual metabolite concentrations. In response to this challenge, spectral deconvolution has been developed to address complex overlapping signals or peaks in NMR spectra. While creating a spectrum from a peak list, known as convolution, is a clear-cut task, the reverse process of determining a peak list from a given spectrum, known as deconvolution, proves to be more intricate[95]. In essence, spectral deconvolution is a computational technique that endeavors to reconstruct the experimental NMR spectrum by using a linear combination of reference compounds present in a library[96]. Various algorithms, available in commercialized tools like Chenomx NMR Suite [97], B.I. QUANTN[98] and FoodScreener[99] as well as in freely accessible tools like Batman [100], rDolphin[101], ASICS[102], and AQuA[103] have been proposed for spectral deconvolution. These tools vary in their applications (biofluids or food and beverage), technical requirements (e.g., 600 MHz-only, Bruker-only such as Bruker IVDR), and the extent of their automated quantification capabilities, often preceding a preprocessing step[33]. However, these

methods even offering a promising semi-targeted approach to address identification and quantification challenges in H-NMR spectra, they do not provide the capability to identify entirely novel and unknown features. Nonetheless, they contribute to advancing the field by facilitating more accurate and comprehensive analyses of complex NMR spectra.

1.4.4.2 MS preprocessing

MS preprocessing can be done using several available tools and software reported in literature[104]. Each one of them allows to start from MS spectra until to arrive the identification of features and statistical analysis. Of these tools MZmine3 [105], Metaboanalyst [106], and XCMS [107] are the most employed by the user thank to the friendly interface and the robustness of workflow. In this work we used Workflow4metabolomics platform which afford different tools for data preprocessing. All the steps are based on R software implemented in the platform as XCMS algorithm for peak piking, peak grouping, and RT correction, or the CAMERA algorithm for features annotation ***(Figure 27).***

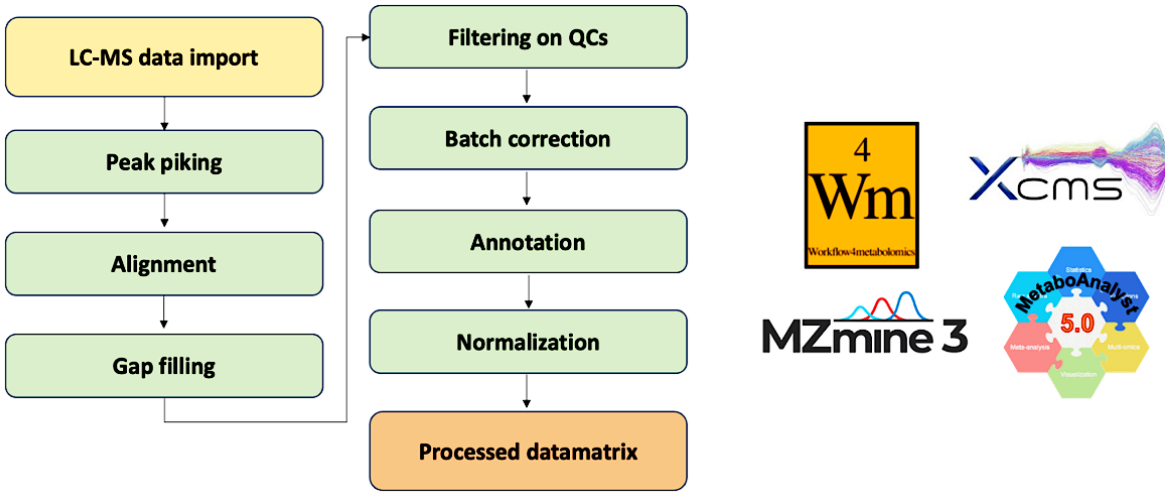


Figure 27*. Workflow for MS data preprocessing; several steps compose this process and can made using several tools of which some are here reported.*

Peak-peaking

The primary objective of peak picking is to identify and characterize each measured ion within a sample, subsequently assigning it to a specific feature represented by a combination of m/z (mass-to-charge ratio) and retention time (RT). Peak picking algorithms aims to captures and deconvolutes peaks found in the extracted ion chromatograms by taking into consideration the inherent structures of baselines and noise, allowing the accurate identification of peaks [108 , 109].

Alignment

Shift in RT during chromatographic separation can lead to the elution of a same metabolites at different retention time between samples. This shift can be caused by different reasons such as changes in the mobile phase or variations in temperature or pressure. In way to face this problem alignment algorithm can be used aiming to group detected peaks across the samples with respect to a *m/z* and a RT window [104, 109]. These grouped peaks are then assigned to a feature in the final data matrix.

Gap filling

After peak peaking process our data matrix may contain several missing values (gap). These missing values can occur due to various reasons such as technical limitations, instrument variations, or the presence of low-abundance metabolites that fall below the detection threshold. Gap filling is a computational technique that aims to estimate or predict these missing values to create a completer and more informative dataset for downstream analysis. Several algorithms are described in literature and can be employed; however, care must be taken in choosing appropriate imputation methods, as poorly handled gap filling can introduce bias or artifacts into the data [109].

QCs-based features filter

Another supplementary step that can be done in the MS-preprocessing workflow is represented by the feature filtration. Several algorithms permit to filter the features based on their presence in the QCs. In particular, this step allows to select the fraction of QCs in which the feature must be present for retain it in the data matrix list. The general assumption of this step is that the QCs are well representative of features presents in the samples.

Batch correction

When planning the analysis of large cohort, it can happen to divide the samples in different groups that will be analyzed at different time by using the same instrumental parameters. These groups known as “batch” can generate a common problem known as “batch effect”. Batch effects can arise from different factor such as instrument performance, reagents, or other experimental conditions; the presence of these variations can overshadow true biological differences in the data and lead to inaccurate or misleading results. In way to overcome this

effect, a batch correction is needed. This correction can be done through simple normalization or scaling processes used to bring the data from different batches onto a comparable scale[110] or through different algorithms such as Combat [111].

1.4.5 Normalization

Normalization is a data preprocessing technique, common to both platforms, used to standardize the intensity levels of signals across different spectra. The goal of normalization is to remove systematic variations in signal intensity that can arise due to factors such as sample dilution, instrument settings, or experimental conditions. Normalization process involves scaling the intensity values of the data points within a spectrum by a certain factor. The choice of normalization method depends on the characteristics of the data and the research objectives. Some common methods include total by sum normalization, Probabilistic Quotient Normalization (PQN), or normalization by an internal standard. **Total by sum normalization** represent the most standard method used for serum and plasma samples and scales the spectra in way that all of them have the same concentration. **PQN normalization** involves dividing each variable in a spectrum by the intensity of the corresponding variable in a reference spectrum; subsequently, the total spectrum is divided by the median of these quotients. This method can handle higher variation in features values and is particularly adapted to urine samples. **Normalization by internal standard** allows the sample normalization using the intensity or the integral area of this standard. The normalization step has a great importance on urine samples, in which the metabolite content can change due to the dilution factor. In health condition, urine samples can be normalized by using creatinine as internal reference even if an increasing number of studies link creatinine to variation to age, weight exercise or gender. This aspect will be better explored in chapter 3.

1.4.6 Scaling

Once the preprocessing of raw data is done, as end-product of our preprocessing workflow a table as the one showed in ***Figure 28*** is obtained. This table known as ***data matrix*** contains in rows the samples name or code and in the columns all the features. Each column represents a metabolites or feature identified in our spectra and each row reports features concentration or intensities relative to each sample. This data matrix will be used for the subsequent statistical analysis. Simply examining this data table allows us to identify metabolites or features that may exhibit significantly higher intensity compared to the others. Indeed, in metabolomics

spectra, it is not uncommon to observe peaks of certain metabolite be significantly more prominent than others. In such cases, groups of metabolites with higher concentrations might overshadow the rest, potentially skewing the results of statistical analyses. This dominance of a few concentrated metabolites can mask the presence of less abundant features, rendering their significance and importance largely unrecognized. In such case scaling, metabolites intensity in a data matrix is mandatory. This step allows to transform data through the application of different mathematical operation in way that at the end of the process all the variables have the same weight and "level of importance". Different scaling methods are today used in metabolomics including pareto scaling, autoscaling, centering or range scaling [112,108]. The choice of scaling method depends on the characteristics of the data, the specific requirements of the analysis, and whether preserving certain features of the data is essential. Furthermore, the use of each one of these methods can impact how the data is represented, how features are weighted, and the outcome of subsequent analyses.

sample nam	0.859992	0.879992	0.899992	0.919992	0.939992	0.959992	0.979992	0.999992	1.05999	1.07999	1.09999	1.11999	1.13999	1.15999	1.17999	1.19999	1.21999
3_3	-0.0328186	-0.0090765	-0.0283616	-0.0571212	0.0408077	-0.0194064	0.02174205	-0.0499463	-0.1671938	-0.1845717	-0.270134	-0.4472061	-0.2001705	-0.0201003	-0.2149678	-0.0393223	0.01149942
6_3	0.09264411	0.35356572	0.06758252	0.12165422	0.48018533	0.57872619	0.5627722	0.43056891	0.54119886	0.21290644	0.65653306	0.86854613	0.24413672	-0.2775997	0.33631144	-1.2229656	-1.0362468
4_3	-0.1508802	-0.2536082	0.00916708	0.04509599	-0.0386168	-0.1041848	-0.0940307	-0.2414685	-0.2660219	-0.4060395	-0.3942576	-0.0376378	-0.120351	-1.5305708	-0.2102155	-0.0390429	0.04962644
20_3	-0.1050553	-0.1774058	-0.1042701	-0.1598548	-0.1247401	-0.1980276	-0.2409141	-0.3479231	-0.2038742	-0.1646054	-0.189192	0.09370985	-0.1605297	-0.3646899	-1.5696496	-1.4787418	-0.7182459
42_3	0.09996126	0.04823292	-0.0575929	0.04126427	0.15452895	0.31822181	0.27648173	0.35939387	0.58853572	0.02787308	0.28296068	0.00387269	0.10926628	-0.5229861	0.09286532	-0.1520081	-0.8715663
41_3	-0.1266936	0.28612525	0.0406984	-0.1371723	0.27511029	-0.0134742	-0.3539441	-0.5138629	-0.6489846	-0.5705526	-0.5406802	-0.6952729	-0.4775873	0.20481106	0.01902244	-0.4978183	-0.3913744
63_3	0.11249668	0.23595582	0.03502379	0.14857945	0.1591447	0.32517368	0.45387907	0.49327093	0.35555811	0.42836671	0.16249965	-0.5039471	0.13478664	0.2552981	0.43742109	0.08018874	0.30833527
61_3	0.04104594	0.19014392	0.04075813	0.09433015	0.03305172	0.05080371	0.12643801	0.19706543	0.2142255	0.28909628	0.01532454	-0.5237646	-0.150995	0.10333167	0.20333045	0.18324877	0.15249559
48_3	0.02782696	0.12668713	-0.0031342	0.03545745	0.0041516	-0.174507	-0.654659	0.11075934	0.16777471	0.21002717	0.31250068	0.16838014	0.21700106	0.428308	0.76536087	1.14482596	1.04901209
67_3	-0.1275171	-0.310036	-0.1100158	-0.1285194	-0.3924103	-0.4084549	-0.2767279	-0.2227038	-0.510573	-0.3090323	-0.4614963	-0.5598962	-0.3859882	-0.0617136	-0.264936	-0.0653118	-0.0445956
59_3	0.24148366	0.52431037	0.23163614	0.46848324	0.41055825	0.62556806	1.38874314	0.91463299	0.8051322	1.00387605	0.97926161	0.97312285	0.75337737	0.68935721	1.09297354	0.75571412	0.91886172
69_3	0.02335696	0.08580941	-0.0064811	0.06750545	-2.16E-05	0.12757913	0.28299066	0.18204197	0.10891886	0.13311437	0.06852872	-0.0808607	0.06086848	0.23275076	0.15773892	0.24564762	0.25294972
79_3	0.10107396	0.12662933	0.14281547	0.26145679	-0.0625392	0.17661402	0.46284743	0.41361306	0.50029255	0.65459262	0.89271695	1.12670263	0.82343507	0.55336369	0.52618977	0.44998257	0.52018309
83_3	0.14906121	0.25852145	0.07508735	0.19376815	0.11077354	0.1142811	0.16916635	0.29034538	0.31848474	0.40431218	0.48220243	0.33610943	-0.0471223	0.47722585	0.21422351	-1.8158399	-1.1983438
3_12	0.03281859	0.00907649	0.02836162	0.05712115	-0.0408077	0.01940639	-0.021742	0.0499463	0.16719384	0.18457171	0.27013402	0.44720614	0.20017054	0.02010029	0.21496776	0.03932229	-0.0114994
4_12	0.15088019	0.25360822	-0.0091671	-0.045096	0.03861685	0.10418478	0.09403075	0.24146853	0.26602189	0.40603951	0.39425763	0.03763777	0.12035099	1.53057081	0.2102155	0.03904291	-0.0496264
6_12	-0.0926441	-0.3535657	-0.0675825	-0.1216542	-0.4801853	-0.5787262	-0.5627722	-0.4305689	-0.5411989	-0.2129064	-0.6565331	-0.8685461	-0.2441367	0.27759965	-0.3363114	1.22296564	1.03624681
20_12	0.10505534	0.17740576	0.10427012	0.15985481	0.1247401	0.19802757	0.24091414	0.34792315	0.20387421	0.16460543	0.18919195	-0.0937099	0.16052967	0.36468994	1.56964964	1.4787418	0.7182459
41_12	0.12669356	-0.2861252	-0.0406984	0.13717226	-0.2751103	0.01347424	0.35394412	0.51386286	0.64898463	0.5705526	0.54068024	0.69527288	0.47758727	-0.2048111	-0.0190224	0.49781828	0.39137443
42_12	-0.0999613	-0.0482329	0.05759294	-0.0412643	-0.154529	-0.3182218	-0.2764817	-0.3593939	-0.5885357	-0.0278731	-0.2829607	-0.0038727	-0.1092663	0.52298613	-0.0928653	0.1520081	0.87156625
48_12	-0.027827	-0.1266871	0.00313418	-0.0354575	-0.0041516	0.17450705	0.65465896	-0.1107593	-0.1677747	-0.2100272	-0.3125007	-0.1683801	-0.2170011	-0.428308	-0.7653609	-1.144826	-1.0490121
50_12	-0.2414837	-0.5243104	-0.2316361	-0.4684832	-0.4105583	-0.6255681	-1.3887431	-0.914633	-0.8051322	-1.0038761	-0.9702616	-0.9731229	-0.7533774	-0.6903572	-1.0030735	-0.7557141	-0.9188617

Figure 28 *Data matrix table representing in row (first row colored in yellow) the list of samples and in columns the "buckets" and their corresponding value for each sample (in green)*

1.4.7 Chemometrics and Statistical analysis

In metabolomics-based studies, the substantial volume of data resulting from its vast quantity, intricacy, and diversity poses a substantial challenge. Indeed, once the spectra acquired hundreds of variables describe each sample, complicating the interpretation and comprehension of biological impacts. For these reasons, in metabolomic studies, in which the quantity of generated data is much higher than the number of samples, a set of tools able to reduce data dimensionality is mandatory. Within this framework, over the past decades, chemometrics tools have been developed to address this challenge by becoming more comprehensive and powerful with the growing complexity of study designs. These methods range from straightforward univariate statistical models to chemometrics and sophisticated machine learning techniques by having as aims overview of samples, clustering samples, and biological pathways investigation A wide list of tools and platforms are available for metabolomics chemometrics and statistical analysis[113]. Of all the tool and platforms available, SIMCA® represents a complete tool for statistical analysis widely used and user-friendly; Metaboanalyst[114], Biostatflow[115], and W4M are complete web-based platforms performing both univariate statistics and chemometric analysis. This wide family of tools can perform statistical analysis by using two distinguished approaches: unsupervised approach and supervised approach whose goals are respectively to build models giving a samples overview and generate classification and predictive models.

1.4.7.1 Unsupervised analysis

As the main goal of unsupervised analysis is to explore the data without "a priori", the tool that better represent this philosophy is constituted by **Principal component Analysis (PCA).** PCA is an unsupervised tool that allows data mining and outlier detection by transforming a high-dimensional dataset into a new set of uncorrelated variables called principal components (PC). These principal components are ranked in order of the amount of variance they capture from the original data. The first principal component captures the maximum variance, the second principal component captures the second most, and so on. Once applied, original data can be projected onto the principal components, effectively reducing its dimensionality while retaining much of the data variability ***(Figure29a)***. The representation of this variability is showcased in a *score plot*, while the influence of each variable is projected in a *loading plot* ***(Figure29b)***. As aforementioned, this approach serves not only to identify outliers, but also to spotlight

instrument or preprocessing discrepancies, and even to discern sample clusters devoid of prior assumptions.

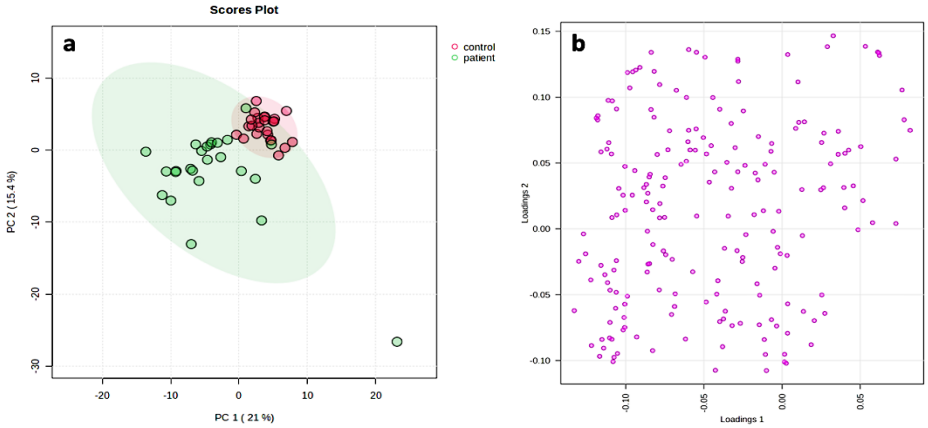


Figure 29 *Unsupervised analysis with PCA score plot and loading plot**; (a)** PCA score plot showing an overview of our samples dispersion; **(b)** loading plot showing the dispersion of our features across the space according to the sample dispersion. . All the models were generated through a test dataset on Metaboanalyst.*

As seen for PCA models, outliers' detection is a fundamental step for a correct data analysis and interpretation. Specifically, the term ***outlier*** refers to a data point that significantly deviates from the overall pattern or distribution of the data set.; they can arise due to various reasons, including technical errors, contamination, instrumental issues, or biological variations. In way to improve the accuracy of the outlier detection, the PCA model can be combined with other methods, such as different distance such as: Hotelling's T^2 for detection of "strong outliers" and DModX for "moderate outliers". **Hotelling's T^2** is a multivariate extension of the univariate t-test; it quantifies how far a data point (sample) deviates from the mean of a group of samples in a multivariate space. It is used to identify outliers or samples that significantly differ from the central cluster by defining the normality area corresponding to 95% confidence.[116,117,118,119].

DModX is another metrics used to ameliorate the detection of" moderate outliers". The "X" in DModX indicates the reference group, which can be a group of samples with similar characteristics, or a model built from the data. DModX quantifies the number of standard deviations by which a data point deviates from the center of the reference group. A high DModX value indicates that the data point is distant from the reference group and could be considered an outlier [117,118,119] .

1.4.7.2 Supervised analysis

Supervised analysis has as principal aim to maximize the separation between groups of predefined samples by using an "a priori" knowledge. This model allows to find a correlation

between the X matrix and the Y response where Y can be a quantitative variable (age, BMI etc.) or qualitative (healthy, ill etc.).

Partial Least Square (PLS) represents one of the most known supervised model and widely used in metabolomics studies [120]. PLS works by creating new variables, called latent variables or components, which are linear combinations of the original predictor variables. These new components are constructed in a way that maximizes their covariance with the response variable (in opposite to PCA in which the maximization is done on the variance). The outcome of this model is represented by a *score plot* ***(Figure 30a)*** in which the cluster between the two groups is maximize, and a *loading plot* in which the variables contributing to the discrimination are shown. In addition to loading plot, *variable importance in projection (VIP) plot* can be generated by showing the significant features explaining the discrimination between groups ***(Figure 30b)***. Two types of PLS models can be distinguished: a PLS discriminant analysis (PLS-DA), and a PLS regression model. PLS-DA is a well-known tool which has as goal to build a model that separates different classes based on the predictor variables (example: control vs treated). In PLS regression, the goal is to build a predictive model for a continuous response variable (example: correlate metabolome with the increasing Body mass index (BMI)). The second highly applied model is represented by the **Orthogonal projection to latent structured discriminant analysis (OPLS-DA).** This tool combines the power of Orthogonal signal correction (OSC) to the PLS discriminant model. Specifically, OSC allows to remove systematic variation of X not correlated to Y in way to delate the inter-subject variability and to describe maximum separation based on class[121].

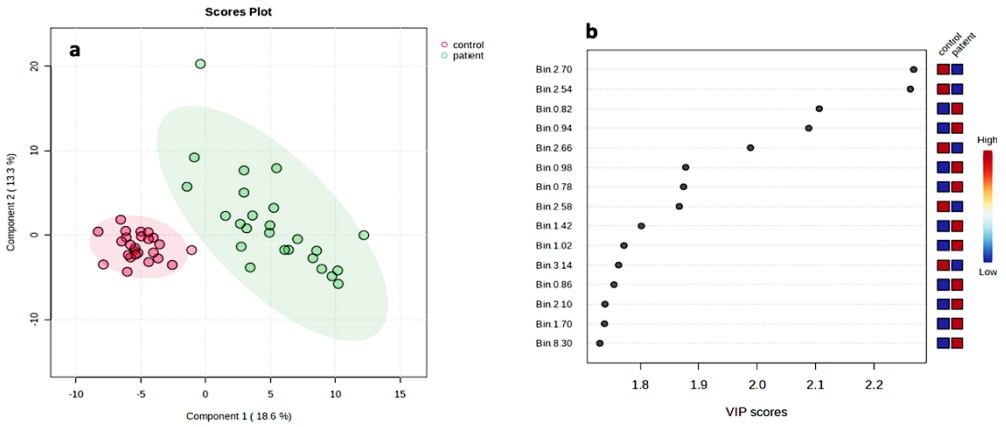


Figure 30 *PLS-DA score plot and VIP score; (a) PLS-DA score plot illustrating the separation between control and patient groups; (b) VIP scores listing all the important variables which allows the separation in PLS-DA score plot. All the models were generated through a test dataset on Metaboanalyst.*

Despite the useful aspect of supervised models, they need to be used with carefully since they can lead to overfitting and unreliable results. In way to avoid this scenario, for each of our supervised model, different metrics can be used to evaluating their performance and validity.

All the presented tools are part of machine learning techniques which are methods or algorithms that enable computers to classify, extract meaningful information from complex datasets, identify patterns, and make predictions. Indeed, if in traditional programming, developers write explicit instructions for a computer to follow, machine learning allows computers to learn from data and improve their performance over time. In addition to previously illustrated algorithm, other tools exist such as Random Forest (RF), support vector machine (SVM) and logistic regression (as part of supervised learning algorithm) or K-means and hierarchical clustering (as unsupervised learning techniques).

1.4.7.3 Validation and performance models

The main goal of metabolomics is to identify a list of putative and reliable biomarkers that are linked to a defined biological status. To obtain these features, supervised models can be employed; however, relying solely on this model is insufficient. As previously discussed, supervised models are prone to the overfitting effect, leading to a misinterpretation of results. Indeed, overfitting refers to the process by which a model exhibits excellent performance on training data but fails to effectively generalize to a new dataset. To overcome this problem, validation algorithms have been developed in way to look at the validity of our models. The most common used validation models in metabolomics are represented by cross-validation methods and permutation test.

Cross-validation method is based on two parameters that assess the model's performance: R^2 (x and Y) and Q^2Y. The R^2 value represents the explained variance proportion of the matrix of x and y variables, while Q^2Y represents the "goodness of the fit" better known as the predictive quality of the model[122,123] . The closer these values are to 1, the more the model separation is good. The theory of this method is to split the total samples into two groups, named training and testing groups. The training group is used to generate a predictive model whose performance will be evaluated on the testing group. This process is repeated as many times as possible to ensure that each sample has been included in the testing group at least once. The objective of the permutation test is to validate that the initially generated model is superior to all other models created by permuting class labels and randomly assigning them to different

individuals[123,124]. The performance of this model is evaluated through a threshold named p-value that needs to be inferior to 0.05 ***(Figure 31a).*** Once validity of the classification models is proven, their performance needs to be evaluated. The two most common performance algorithms are represented by: receiver operator characteristic (ROC) curve and confusion matrix.

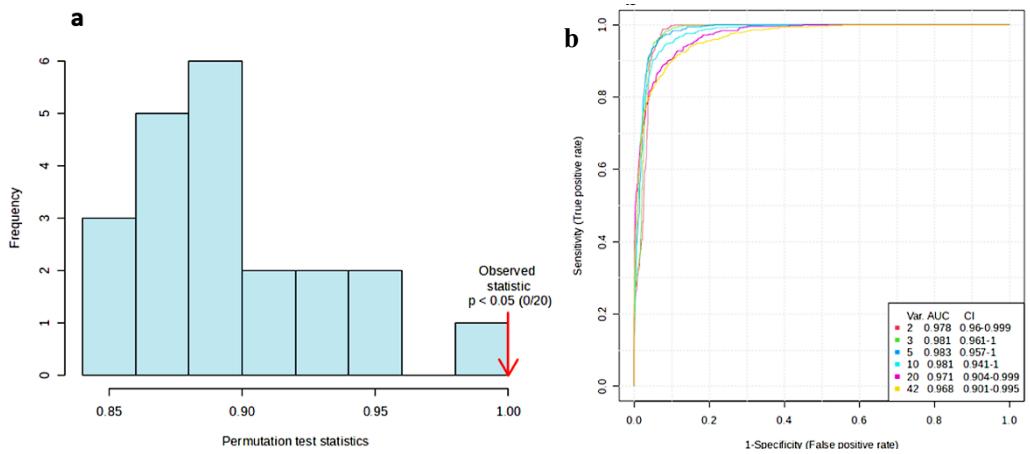


Figure 31 *Permutation test and ROC curve for a defined dataset;* ***(a)*** *Permutation test indicating the validation of the model since its p-value <0.05;* ***(b)*** *multivariate ROC curves showing different AUC based on the number of selected variables. All the models were generated through a test dataset on Metaboanalyst.*

The multivariate ROC curve serves as a metric indicating the efficacy of a model in distinguishing between two distinct classes or categories; it differs from the classical ROC in terms of the number of variables employed for its construction (1 variable for classical one, 2 or more variables for multivariate). In a multivariate ROC curve graph the rate of true positive in function of the false-positive rate is shown. This curve is expressed in terms of sensitivity (percentage of true positive correctly classified) and specificity (percentage of true negative correctly classified). The area under the curve (AUC), often used to summarize this model, is an index for the quality of the discrimination model ***(Figure 31b)***. Specifically, the closer this value is to 1, the better the separation; a value equal to or less than 0.5 indicates a lack of separation between the two classes.

Confusion matrix[125] is a performance measurement for machine learning classification problem where output can be two or more classes. This model is presented as a table with 4 different combinations of predicted and actual values by allowing to understand how the classification model is "confused" when it makes predictions. To evaluate its performance and to better understand and analyze the models, we can look at several metrics such as: accuracy, recall and precision. **Accuracy** simply measures how often the classifier makes the correct prediction. It is represented by the ratio between the number of correct predictions and the total number of predictions. The major limitation of accuracy can be underlined in imbalanced data;

in this case, if the model predicts that each sample belongs to the majority class label, the accuracy will be high, but the model will inaccurate. This is why, in unbalanced studies, it is important to have other metrics to which refers. **Precision** is a measure of correctness that is achieved in true prediction; particularly, it is defined as the ratio of the total number of correctly classified positive classes divided by the total number of predicted positive classes. **Recall** is a measure of actual observations which are correctly predicted and is defined as the ratio of correct predictions (true positives) divided by the total number of positive classes ***(Figure 32).***

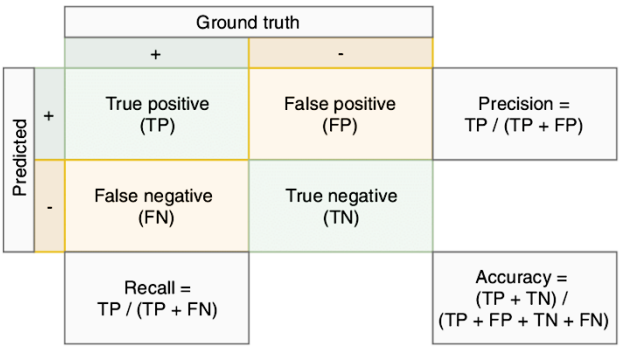


Figure 32 *Scheme of confusion matrix to simplify its understanding; precision, accuracy, and recall value are reported with their respective formula (TP= true positive; FP= false positive; FN= false negative; TN=. True negative)*

1.4.7.4 Univariate statistical analysis

Univariate analysis in metabolomics involves assessing individual variables (such as metabolite concentrations or peak intensities) one at a time in way to extracting meaningful insights from complex datasets though statistical models . Univariate data analysis is fundamental to uncover variables that might be overlooked in multivariate models. Indeed, while multivariate analysis considers all metabolites and their relationships simultaneously, it may fail to identify metabolites that have significant effects when examined individually. Identifying specific metabolites that exhibit significant changes under different experimental conditions or disease states is essential for biomarker discovery. These aspects underscore the importance of integrating univariate analysis into metabolomics experimental designs[126,127].

Student's* t-*tests can be used in order to discover statistical significance in univariate datasets consisting of two sample comparisons. Several variations to this test exist in way to meet the assumptions on which the test is based (e.g., normal distribution of data).

The Mann-Whitney U test, also known as the Mann-Whitney-Wilcoxon test, is a non-parametric statistical test used to assess whether two independent groups differ significantly in

their distribution of a continuous variable. It is particularly useful when the data are not normally distributed or when the sample sizes are small.

1.4.7.5 Multiplatform approach

Multiplatform approach has already been documented in metabolomics literature. This approach has been defined as the integration of data blocks from different analytical platforms into a single comprehensive model. The aim of this unique block is to increase the resultant predictive performance of the model by combining the strength of each analytical platform and by increasing the metabolomic coverage. Different strategies have been proposed for the combination of the several analytical techniques: low-level, mid-level and high-level data fusion[128] ***(Figure 33).***

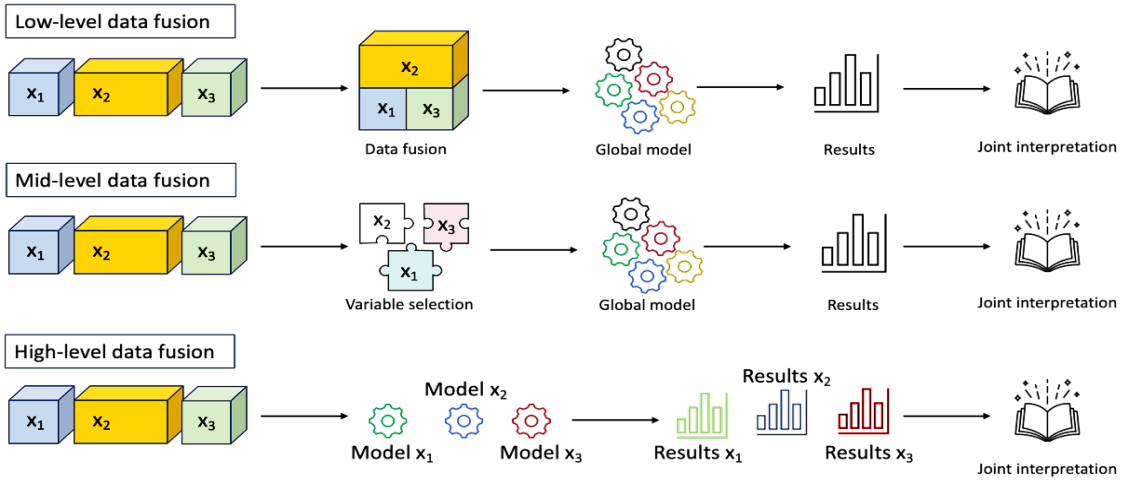


Figure 33 *Multiplatform approach with the 3 major strategies explained; low-level data fusion allows a direct combination of dataset by generating a global model and related results; in mid-level data fusion variable selection for each dataset is done prior the global model and results generation; in high level-data fusion separate models and results for each* dataset are generated before to combine it for a joint interpretation. (adapted from: *Boccard, J. & Rudaz, S. « Harnessing the complexity of metabolomic data with chemometrics: Metabolomic data analysis with chemometrics. » DOI: 10.1002/cem.2567*)

Low-level data fusion is the easiest approach, obtained by the horizontal concatenation of data matrices coming from the different analytical platforms. The great pitfall of this method is related to the great difference between intensity ranges of features and the high number of variables coming from different platforms[129]. Mid-level data fusion presents itself as a solution to the big amount of data that limits biomarkers discovery in low-level approach. Through a previous feature selection, operated through chemometrics tool (such as PCA or PLS), mid-level approach proposes the concatenation of reduced dataset by which it is easier to extrapolate important information[130]. Another proposed approach, that represents also the one applied in

the current work, is represented by the high-level data fusion. This approach is based on the generation of separate discriminant models, each one coming from a different platform, whose results are finally "fused". By applying an high level data fusion we can expect an increase of prediction performance and a better classification model [131].

The adoption of the multi-platform approach has seen an increase in its application, within metabolomics and, more broadly, in omics studies. By integrating data from various sources, this approach not only expands metabolomic coverage but also offers a significant advantage in enhancing the robustness of biomarker discovery by validating results across different platforms. It serves as a valuable strategy for overcoming challenges in identifying biomarkers and understanding metabolic pathways, often challenging aspects of the metabolomics workflow. While the integration of information from multiple sources remains a potent strategy to gain deeper insights into complex biological systems, the use of multiple analytical platforms introduces challenges. Differences in scale, resolution, and variability between platforms can be challenging, necessitating specialized expertise to effectively handle and interpret data from each distinct platform. Despite these challenges, the multi-platform approach remains a valuable and powerful method in metabolomics and omics research.

1.4.8 Biomarkers identification

In the last decades, with the increase in the application of metabolomics to different field, an important discussion about identification step is born by ending with the constitution of Metabolomics Standards Initiative (MSI) [132]. The MSI has as major objective to uniform and give guidelines to metabolomics community in way to help scientists in the difficult process of metabolite identification. This initiative allowed the redaction of a list of 4 levels of known metabolite identification ***(see table 4)*** to which all the scientists should refer in way to better standardize the process, harmonize procedures and obtain inter-laboratory reproducibility.

1.4.8.1 NMR identification

After univariate statistics and chemometric analysis, a list of significant features to discriminate the biological factor of interest is available. In way to go further in the analysis and have a real insight of the biological process that occurs, placing a name on each significant bucket or feature is imperative. Features identification in ^{1}H-NMR spectra can be a challenging step that needs a multi-step process crossing experimental data, databases, predictive and complementary analytical techniques to confidently assign chemical identities to the observed peaks. In this context, the researcher adopts the role of a detective, employing all available clues derived from the spectra to uncover the "culprit". In the initial step of metabolite identification, employing chemical shift lists specific to each biofluid from the literature enables the identification of all potential matching features within the desired chemical shift zone[34,133,134]. Once a number of candidate metabolites have been identified for a particular spectral signal, the subsequent step involves referring to 2D NMR spectra. Indeed, ^{1}H-NMR spectra, as discussed in *section 1.4.3.2,* can be challenging to interpret due to significant signal overlap, making it difficult to definitively attribute a signal to a single metabolite with absolute certainty. 2D NMR spectra play a crucial role in resolving the uncertainties encountered in 1D peak assignments[135]. The key advantage of this technique is its ability to provide a distinct and unique signature for each molecule, which needs, however, to be present in a sufficiently high concentration compared to other metabolites within the complex mixture. An online and complete database is available to compare both 1D and 2D NMR spectra for a specific molecule, along with its concentration profiles under both normal and pathological conditions. HMDB stands out as the most comprehensive database for metabolites present in various human biofluids, enabling to research molecules based on their name as well as on their chemical shift by giving back a matching score. Another tool allowing metabolite identification,

and already discussed in the context of spectral deconvolution, is represented by Chenomx. This software enables the assignment and the identification of signals in 1D spectra using an internal library and by considering both chemical shift and sample's pH for accurate metabolite identification.

1.4.8.2 MS identification

As seen for NMR identification, after statistical analysis, a list of selected features, significant for discrimination of biological response, is generated. The identification process in MS, also known as annotation, is a challenging step in MS-based metabolomics studies. In fact, if in ^{1}H-NMR spectra, identification is difficult due to significant signal overlap, in MS metabolomics the scenario is even worst. Indeed, with the advancement in instrumental technique, the sensitivity of machines has greatly increased by allowing, in untargeted studies and though a single run, to detect the global coverage of metabolome. Despite the great number of detected features, the main problem is represented by the impairment between the experimental output composed by a long list of features and signals and the minimal portion of annotated metabolites. As illustration of this disproportional ratio, in a study where a single run detected more than 25,000 features, the number of unique metabolites identified was lower than 1,000 [136].

Indeed, in MS analysis a single metabolite can be represented by multiple features due to presence of adducts of different fragments. This complexity is due to the fact that majority of features with their accurate mass and fragmentation do not match to the available metabolomics databases by making them often be listed as "unknown metabolite". The annotation in LC-MS is generally based on several measures referring to the feature including accurate mass (AM), retention time (RT) and MS/MS (fragmentation spectra). All these criteria were included in the guidelines proposed by metabolomics community in way to obtain a general workflow for identification [137,132] ***(see table 4***) .

Confidence Level	Description	Minimum Data Requirements
Level 0	Unambigous 3D structure: Isolated, pure compound, including full stereochemistry	Following natural product guidelines, determination of 3D structure
Level 1	Confident 2D structure: uses reference standard match or full 2D structure elucidation	At least two orthogonal techniques defining 2D structure confidently, such as MS/MS and RT or CCS
Level 2	Probable structure: matched to literature data or databases by diagnostic evidence	At least two orthogonal pieces of information, including evidence that excludes all other candidates
Level 3	Possible structure or class: Most likely structure, isomers possible, substance class or substructure match	One or several candidates possible, requires at least one piece of information supporting the proposed candidate
Level 4	Unkown feature of insterest:	Presence in sample

Table 4 *Confidence intervals of metabolites annotation as discussed by the Compound identification work group at Metabolomics Society annual meeting in 2017 (from: Blaženović, Ivana et al. "Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics." DOI: 10.3390/metabo8020031)*

These suggestions are divided on different level of annotation based on the information found on the features. **Level 0** is the strongest level of annotation and includes stereochemistry discrimination**;** at **level 1** the detected feature is identified by comparing its mass spectrometry data (at least 2 orthogonal parameters between AM, RT, or MS/MS) to an authentic chemical standard. **Level 2** annotation involves matching the mass spectrometry data with a class-specific standard; **level 3** is based on confirmation by one parameter (e.g., AM); **level 4** signifies that the detected feature cannot be assigned to a known compound, chemical class, or group and is often referred to as "unknowns" [138]. It is important to underlying that mass fragmentation spectra (MS/MS spectra) obtained through DDA mode of acquisition, are of fundamental importance to add information to the molecular structure[139]. To achieve one of the annotation levels described, various tools are documented in the literature. CAMERA for example is a tool also implemented in W4M that allows to search for common relationship between peaks facilitating the annotation step [140]. Another way to facilitate this step is based on the creation of molecular networking. Molecular networking relies on the calculation of spectral similarities between the detected features by comparing the mass spectra of these features. Features with high spectral similarity are connected to form a network in which each node represents a feature (metabolite), and edges between nodes indicate spectral similarity. GNPS (Global Natural Products Social Molecular Networking) is the most known tool used to generate molecular networking by providing an overview of the relationships between detected features even if their main application is on natural product field[141]. In addition to these, homemade databases represent in both MS an NMR, fundamental tools allowing the identification step. All these tools and their utilization can assist scientists in the process of annotating metabolites. However, the future of metabolite annotation will primarily rely on databases and repositories that must

be continually expanded through the collective contribution of data by the metabolomics community.

1.4.9 Metabolic pathway analysis

Over the last decades, the metabolomics workflow has seen a great enhancement in all the following steps by starting from more accurate sample preparation protocol, passing through more sensitive instrumental platforms until powerful statistical analysis. The milestone of these workflow, in line with the goal of metabolomics, is represented by the investigation of all the metabolites that affect a biological process. In this context metabolic pathway analysis is key corner of metabolomics study and represents a powerful tool to link changes in metabolic compounds to biological pathways. This process can be achieved using different available database containing a wide range of information concerning gene, metabolites and related biochemical information. One of the most famous libraries is represented by Kyoto encyclopedia of genes and genomes (KEGG) whose databases (GENE, PATHWAY and LIGAND) are daily updated [142]. Other online platforms provide insight into biochemical events by highlighting the pathways most affected through the inclusion of an identified list of metabolites. These platforms are represented by Metaboanalyst[143] , Cytoscape[144] and MetExplore [145] ***(Figure 34).***

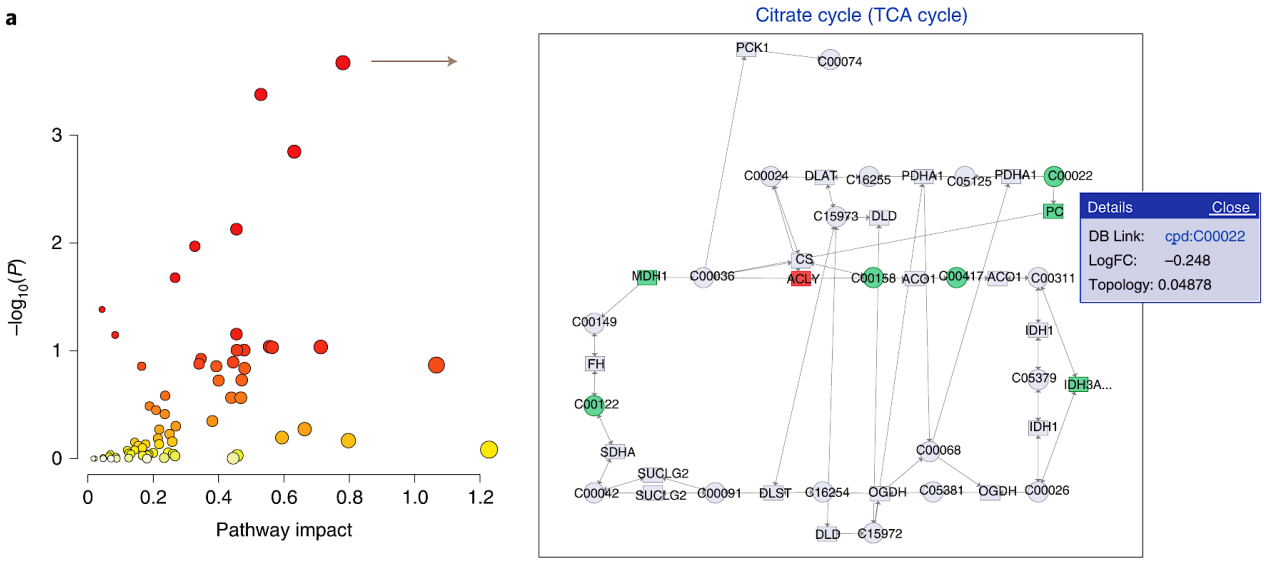


Figure 34 *Pathways analysis for an example dataset on Metaboanalyst; on the left plot, all the biochemical pathways implicated (represented by the circles with shading colors from yellow to red) are showed; on the right panel a single metabolic pathways is showed specifically by highlighting the metabolites (in red) implicated in the used dataset*

1.5 Metabolomics and personalized medicine

Personalized medicine, also known as precision medicine, is an innovative approach to medical treatment and healthcare that adapts medical decisions, practices, and treatments to individual characteristics of each patient. By considering each individual as "unique", with variations in their genetics, environment, lifestyle, and biology, this approach has as principal aim to provide an adjusted and customized healthcare for each person. In this scenario, metabolomics places itself as a fundamental instrument for this approach. Indeed, metabolomics by being output of genomics and input of environmental stimuli represents the most representative mirror of the individual at a precise moment. Over the past few decades, metabolomics has gained prominence within the biomedical field, with its most significant contributions, as emphasized in this thesis, in the identification of novel therapeutic targets and biomarkers[146–148]. These advancements have the potential to greatly assist clinicians in early disease prediction and improved patient care thereby empowering the role of personalized medicine and patient well-being[149].

Defined as the stethoscope of twenty-one century[150], metabolomics have a great number of potential application that in the future could became the new must-have tool for clinicians in either illness prediction, diagnostic, or prognostic. In oncology field, for example, in which the prevention and the diagnosis are fundamental for patient survival, metabolomics signatures have be found for ovarian cancer [151], lung cancer [152] and endometrial cancer [153]. Other biomedical fields such as neurology, rheumatology and endocrinology benefit of the powerful of this approach even if the transposition process from laboratory to clinic represents a long way. Moreover, several metabolomics discoveries have seen the light in clinical application such as AminoIndex® Cancer Screening (AICS®) which by using LC/MS-based measurement of amino acids in plasma enabling early and non-invasive detection of cancer [154]; another application is represented by Theoreo Srl which through GC-MS platform allows the early noninvasive diagnosis of fetal malformations and endometrial carcinoma[155]. In addition to cancer field, another interesting application takes place in nephrology domain. *AXINON*® GFR (NMR) test[156] combines the use of novel biomarkers quantified through NMR-based metabolomics approach with traditional renal markers (creatinine and cystatin C) to obtain accurate GFR results compared to standard eGFR equations[157,158].

In summary, metabolomics plays a crucial role in the field of precision medicine by providing valuable insights into an individual's unique metabolic profile. Moreover, as research in metabolomics continues to advance, and as more evidence of its clinical utility emerges, it is likely that metabolomics will play an increasingly significant role in clinical diagnostics and personalized medicine in the future.

1.6 Chronic kidney disease (CKD) and kidney transplantation (KTx)

1.6.1 Background – Kidney physiopathology

The kidneys are a pair of bean-shaped organs situated on opposite sides of the spine in the retroperitoneal space along the posterior abdominal wall. Each kidney measures approximately 10 to 12 centimeters in length, roughly the size of a closed fist, and is encased by a fibrous renal capsule that offers structural support to the soft internal tissue. On the inner side of each kidney, there is a central depression referred to as the renal hilum or pedicle where various structures such as renal blood vessels, nerves, lymphatic vessels, and the ureter enter and exit the kidney. The kidney is divided into two primary components: the cortex and the medulla. The cortex is the outer portion situated just under the renal capsule and comprising renal corpuscles, as well as the proximal and distal convoluted tubules. The renal medulla is the inner part of the kidney and is also composed by a network of capillaries. Within the kidney's inner region are pyramid-shaped units called Malpighian pyramids, which are oriented with their bases facing toward the cortex and their apexes directed toward the hilum[159,160 161] ***(Figure 35).***

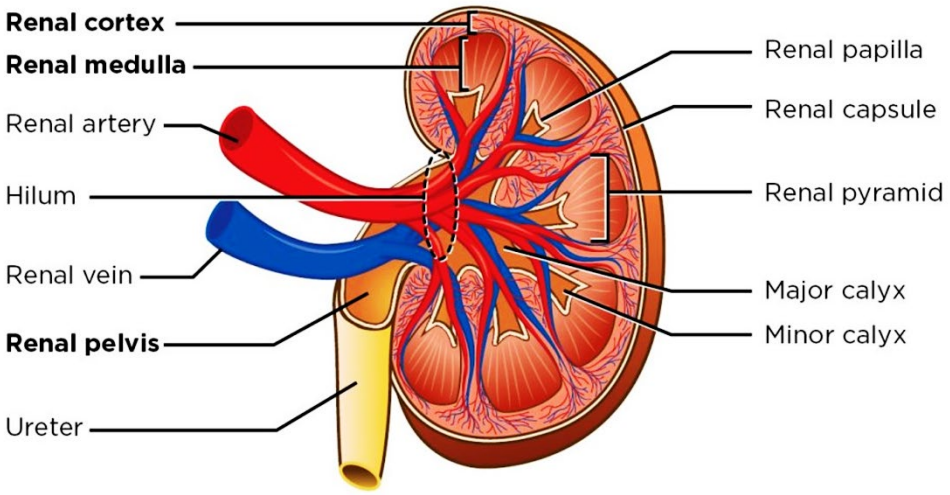


Figure 35 *Representation of kidney anatomy with all its most important components (from: https://www.nursingtimes.net/clinical-archive/renal/renal-system-1-the-anatomy-and-physiology-of-the-kidneys-23-01-2023/)*

Scattered throughout the cortex and the medulla there are nephrons which represent the functional unit of kidney. Around 2 millions of nephrons per kidney are present in an adult individual. Each nephron consists of a renal corpuscle (glomerulus and Bowman's capsule) and a renal tubule (proximal convoluted tubule, loop of Henle, distal convoluted tubule, and collecting duct). Throughout the nephron's length, substances are selectively reabsorbed from

the filtrate into the bloodstream, and waste products and excess substances are secreted into the filtrate. This intricate process allows the kidneys to regulate electrolyte balance, fluid volume, and the excretion of metabolic waste products, ultimately producing urine. The nephron's efficient filtration and reabsorption processes are essential to maintain overall bodily homeostasis[159,160]. As part of nephron corpuscle, the glomerulus is another fundamental component of kidney which is on charge of filtration function. Glomerulus is composed by a tuft of capillaries loop and surrounded by Bowman's capsule ***(Figure 36)***. The glomerular filtration process is a passive mechanism responsible for the formation of plasma ultrafiltrate called "primitive urine". Specifically, blood coming from the bloodstream through the renal artery enters in the glomerular tuft where the high blood pressure and fenestrated endothelial capillaries act as filtration barrier. Blood cells and molecules greater than 68kDa are retained in contrast to water, sodium, urea and glucose which are filtrated and then reabsorbed. In this process, the volume of primitive urine formed each minute by all the glomeruli of both kidneys is defined as glomerular filtration rate (GFR) and represent a fundamental parameter for kidney function measurement.

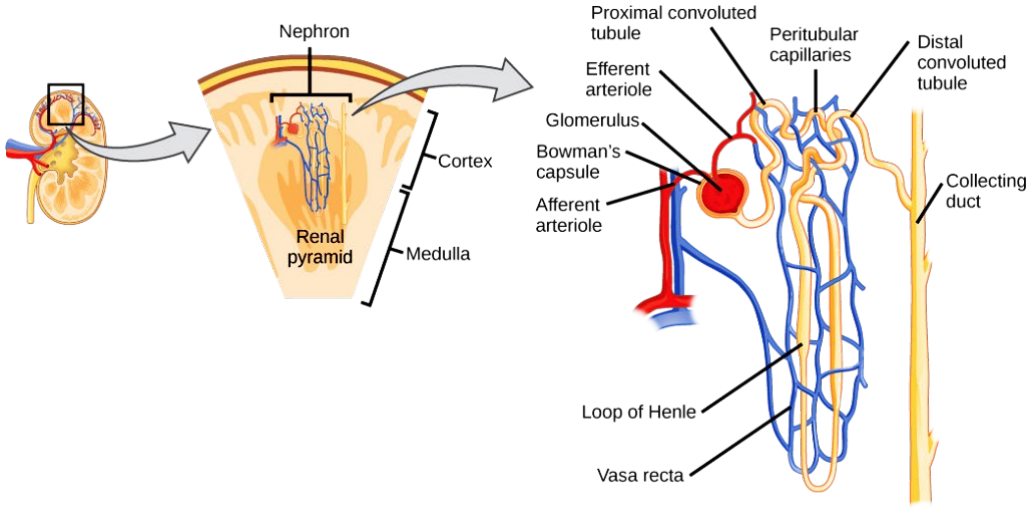


Figure 36 *Image of nephron structure and its location in the kidney (from: https://courses.lumenlearning.com/suny-dutchess-ap1/chapter/nephrons-structure/)*

In biological system, kidneys have several functions that can be resumed in these points:

- **Excretion of metabolic waste products and urine production:** the primary role of kidney is to filter blood and remove waste products produced during the body's metabolic processes helping prevent the buildup of toxins in the body. The process of filtration, which results in the production of urine, not only clean body from waste products but also plays a crucial role in preserving the body acid-base equilibrium and in the regulation of the body fluid and electrolyte balance.

- **Erythropoiesis regulation:** kidneys produce and release erythropoietin, a hormone that stimulates the bone marrow to produce red blood cells (erythrocytes). This process helps maintain adequate oxygen-carrying capacity in the blood.

-Hormone production: in addition to erythropoietin, the kidney converts inactive vitamin D (calcidiol) into its active form, calcitriol which is involved in calcium metabolism and bone health, and prostaglandins, which are involved in various physiological processes.

-Blood pressure regulation: this process is done by controlling the volume of blood in the circulatory system and by releasing the enzyme renin, which helps regulate blood pressure. The renin-angiotensin-aldosterone system is a complex hormonal system that influences blood pressure.

The central role of kidney in maintaining the body in a state of homeostasis, is fundamental in way to keep subject in healthy status; when the kidneys functions become compromised or abnormal, it can lead to a range of health issues, including chronic kidney disease (CKD).

1.7 Chronic kidney disease (CKD)

1.7.1 Definition, causes and symptoms

Chronic kidney disease (CKD) is a progressive with no cure disease with high morbidity and mortality that affects 10-15% of global adult population [162]. CKD is defined by the "Kidney Disease: Improving Global Outcomes (KDIGO) clinical practice guideline" (KDIGO) as abnormalities of renal structure or functions due to a loss of functional nephrons. An adult patient is defined as affected by CKD when in a period of three months or longer shows a glomerular filtration rate (GFR) lower than 60 ml/min/1.73 m^2 or as presenting markers of kidney damage (for example albuminuria). Based on the value of GFR and albuminuria, 18 stages of the pathology can be distinguished [163] ***(Table 5):***

	GFR descriptors and range	Range (mL/min/1·73m²)	Persistent albuminuria categories, descriptors and ACR range		
			Normal to mildly increased (<30mg/g)	Moderately increased (30–300mg/g)	Severely increased (>300 mg/g)
G1	Normal or high	≥90	1 if CKD	1	2*
G2	Mildly decreased	60–89	1 if CKD	1	2*
G3a	Mildly to moderately decreased	45–59	1	2	3†
G3b	Moderately to severely decreased	30–44	2	3	3†
G4	Severely decreased	15–29	3*	3*	≥4†
G5	Kidney failure	<15	≥4†	≥4†	≥4†

Table 5 *Stage of CKD based on GFR and albuminuria values.*

The affected population is heterogeneous, and the major causes vary globally including preexisting pathologies, lifestyle factor (such as tobacco use), environmental exposure or family history. The risk factors linked to CKD progression can be distinguished in non-modifiable risk factors[164] such as: older age, male gender, a non-Caucasian ethnicity (such as African Americans, Afro-Caribbean individuals, Hispanics, and Asians); modifiable risk factors including systemic hypertension, diabetes, proteinuria, and metabolic determinants [165]. The pathophysiology of this disease, particularly in the early stages of CKD injury, involves a complex and diversified process that may remain "invisible" for an extended period, underscoring the primary challenge associated with CKD: the challenging detection of its initial phases.

Indeed, no symptoms appear in the early stages of CKD making its diagnosis difficult to achieve. As CKD advances, the declining kidney function leads to the accumulation of

substances known as uremic toxins. While the precise biochemical and physiological effects of these molecules remain partially understood, within the context of CKD, they contribute to processes such as inflammation, neurological disturbances, gastrointestinal issues, and other mechanisms that give rise to a diverse array of symptoms[166,163]. The presence of these non-specific symptoms often leads to patients being diagnosed in the advanced stages of CKD. Once they reach end-stage renal disease (ESRD), dialysis or transplantation becomes the only viable option.

1.8 Kidney transplantation

For individuals facing ESRD, kidney transplantation is often considered the best treatment option, thanks to substantial advancements made over the past half-century. Indeed, since the first successful kidney transplant performed by Dr. Joseph Murray in 1954, there have been remarkable advancements in surgical techniques, organ preservation methods, immunosuppressive medications, and organ matching systems[167]. These improvements have significantly increased the success rate of kidney transplants and offers a better quality of life and longer life expectancy compared to patients on chronic dialysis [168,169].

Moreover, even if kidney transplantation represents the best treatment for patients in end-stage of CKD, it is important to note that this option is not suitable for every ESRD patient, and the decision depends on various factors including the patient's overall health, age, and availability of suitable donors ***(see Figure 37).*** Additionally, there is a need for a lifelong commitment to post-transplant care and immunosuppressive medications to ensure the success of the transplant.

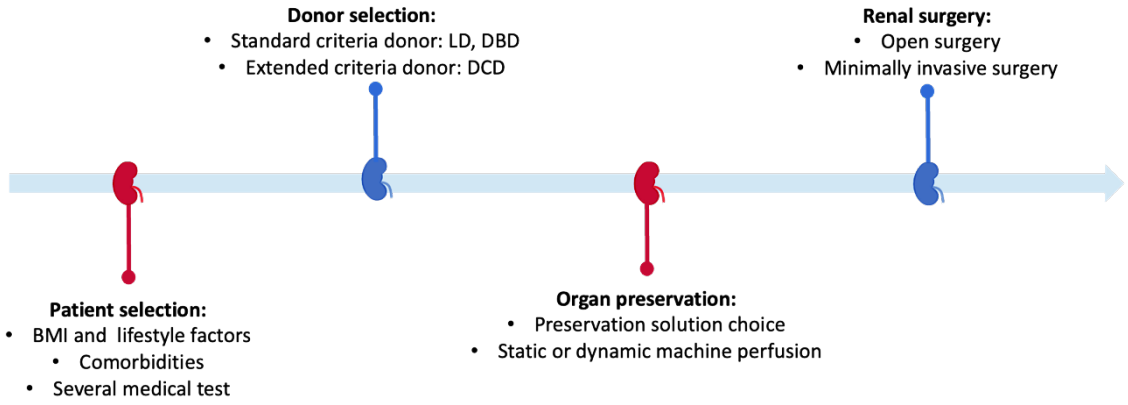


Figure 37 *Timeline in kidney transplantation composed by several step starting from patient selection until kidney surgery; (LD= living donor; DBD= donation after brain death; DCD= donation after circulatory death)*

1.8.1 Patients selection

The selection of patients for kidney transplantation is a meticulous procedure that encompasses the assessment of a wide range of medical, psychological, and logistical factors to maximize the likelihood of successful outcomes. It is important to recognize that many individuals with ESRD often experience multiple coexisting medical conditions that could potentially impact the success of the transplant. From a medical standpoint, these comorbidities can significantly influence the long-term health of both the recipient and the transplanted organ. Consequently, additional criteria beyond age and body mass index (BMI) have been incorporated into the evaluation process. These supplementary criteria encompass a thorough examination of the

patient's medical history, as well as specific tests to assess cardiovascular, cerebrovascular, pulmonary, and gastrointestinal health [170,171].

Indeed, every potential kidney transplant candidate must undergo a battery of tests and examinations to provide a comprehensive evaluation This evaluation serves to determine the patient's suitability for transplantation and identify any potential risks or complications that may arise during or after the procedure. Furthermore, specific criteria and evaluation process may vary slightly from one transplant center to another. The goal is to identify suitable candidates who will benefit from kidney transplantation while minimizing the risks associated with the procedure[172].

1.8.2 Donor selection

Two distinct types of donors are possible in kidney transplantation: living donors or deceased donors. Living donors (LD) are an excellent source of kidney grafts for transplantation offering the best graft and recipient survivals. Several advantages contribute to an enhanced graft survival in this type of donor such as improved compatibility which result in a better match between the donor and recipient, and shorter cold ischemia time that helps to preserve the organ's function .In opposition to these, a significant drawback associated with living donation is the increased complexity of surgical interventions, which is closely tied to another major concern represented by the risks faced by the donor. However, by taking all that into account, in the context of this thesis, our attention was focused on deceased donors. This type of donors can be for their part divided in two classes: *donation after brain death (DBD) and donation after circulatory deaths (DCD)*. Donation after brain death is possible when death occurs after primary brain injury or disease, resulting in the irreversible cessation of brain function. In DBD, cardiac circulation and respiration are maintained by medical measures. Dead circulatory donors are represented by individuals in which death occurs due to permanent cardiac arrest before the organ procurement was done. Based on this classification, two types of donor criteria have been established: standard donor criteria (SDC) and extended donor criteria (EDC). Standard donor criteria include donors under the age of 50 years with no history of hypertension and normal values of creatinine whose death occurs under DBD condition. In opposite to this description, extended donor criteria (EDC) include donors over the age of 60 with history of high blood pressure and/or high creatinine values whose death occurred in DCD conditions. While SDC donors are generally preferred, the increasing gap between demand and supply for kidney transplantation in the last decades enhanced the use of ECD donors. Indeed,

according to its definition, ECD kidneys exhibit lower graft survival rates when compared with those from SCDs, displaying a relative risk of graft failure exceeding of 1.7 [173].

Nevertheless, during the early years of the 21st century, several studies highlighted the necessity for improved decision-making guidelines that provide a more precise definition of graft quality not solely based on the classification of kidneys as SDC or EDC [174,175]. A novel tool has been introduced to assess the risk of graft failure by combining various donor and recipient variables. This tool is termed the Kidney Donor Risk Index (KDRI) and is primarily used to assess the risk of post-transplant kidney graft failure in deceased donor kidneys, specifically in an average adult recipient. KDRI provides an estimate of the relative risk of graft failure for a particular deceased donor compared to a reference donor. Notably, this formula has demonstrated increased predictive accuracy for allograft survival when compared to the ECD classification criteria [176]. Following the calculation of KDRI, another important metric, known as the Kidney Donor Profile Index (KDPI), can be derived. KDPI combines a range of donor-related factors, including clinical parameters and demographics, to condense the overall quality of deceased donor kidneys into a single numerical score relative to other available kidneys. Lower KDPI scores are indicative of kidneys with an extended long-term estimated function, while higher KDPI scores suggest kidneys with a shorter mid-term estimated function [177,178].

1.8.3 Organ preservation

Organ preservation involves the careful management and maintenance of the donor kidney from the time it is recovered from the donor until it is transplanted into the recipient. During this process, once the heart stop beating, kidney undergoes a process known as ischemia. Two different ischemic process can be distinguished during kidney transplantation: cold ischemia time (CIT) and warm ischemia time (WIT). Cold ischemia time refers to the duration during which the kidney is stored at 4°C; during this process, the utilization of a cold storage solutions mandatory to help slow down the metabolic processes in kidney. Warm ischemia time (WIT) represents the span of time in which organ is without a blood supply and experiences warm conditions. Distinguishing between DBD and DCD is feasible by considering the presence of one or two WIT. In the case of DCD donors, following circulatory arrest and before the onset of CIT, an initial WIT (referred to as WIT-1) takes place; subsequently, as the cold preservation solution is flushed into the organ, CIT begins. The process then proceeds until reaching the

second WIT (WIT-2), which is measured from the moment the organ is removed from cold storage to allograft reperfusion. This interval is commonly referred to as anastomosis time.
In contrast, in DBD donors, where blood circulation is still active after death, the kidney is directly flushed with a cold storage solution, initiating CIT without any preceding. The prolonged duration of these processes has been demonstrated to affect negatively the graft outcome (as we will see more in details in the following paragraphs) [179,180,181,182]. Another critical point of organ preservation is represented by the cold preservation solution flushed in the organ during CIT. Preservation solutions are specifically formulated to rapidly cooling the organ while preserving and protecting the organ from cellular damage. These solutions contain a combination of various components, including impermeant, colloids, buffers, electrolytes, and other substances which primary goal is to maintain cellular viability during the storage of the kidney at a temperature of 4 °C [183]. Among all the possible preservation solutions used in kidney transplantation practice (such as UW, HTK etc..), a special focus is directed in this work to IGL-1. This preservation solution has in fact demonstrated that its low potassium concentration in extracellular compartments allows a reduction of vasospasm occurrence linked to potassium-induced smooth muscle cell depolarization[184] that translates in clinical practice in equal or superior graft outcome compared to classical storage solutions[185,186,187]. In addition to preservative solutions, two preservation methods can be distinguished for kidney graft: static and dynamic. The static approach involves flushing the kidney with a preservation solution and subsequently storing it in a static state at 4 °C. This procedure helps extend the viability of the kidney during the storage period by reducing metabolic activity which minimizes the risk of cellular damage and deterioration occurring when a non-constant oxygen and nutrients supplies are available. Dynamics preservation methods requires dynamic movements of their fluids to improve preservation state. Three different methods are nowadays used and are represented by: hypothermic machine perfusion (HMP), normothermic machine perfusion and venous oxygen per sufflation. HMP technique involves perfusing the kidney with a cold preservation solution at temperatures below normal body temperature by allowing a continuous supply of kidney tissue in oxygen. The main advantage of this technique compared to static storage is represented by an improved graft quality post-transplantation. Indeed, several studies have shown to improve recovery of DCD kidneys and to reduce primary kidney dysfunction post transplantation [188,189,190].

1.9 Post-transplantation follow-up

While kidney transplantation represents the beast treatment for patients in ESRD, there can be various problems or complications that may arise in the post-transplantation period. This phase is critically important, necessitating mandatory patient follow-up to improve graft outcomes. Indeed, if survival rate and quality of life are increased and enhanced compared to dialyzed patients, survival of KTx recipients remains lower than in individuals without ESRD. The primary source of challenges in transplanted patients is related to instances of graft rejection, which can occur shortly after the transplant or even years later. Allograft rejection can be classified in hyperacute, acute, and chronic rejection based on histopathology and immunological characteristics and multiple factors can influence it after KTx[191,192] . If hyperacute and acute rejections can happen between first hours and three months post-transplantation, chronic rejection can occur even years post-KTx making patients follow-up fundamental for management of kidney recipient patients (***Figure 38)*** .

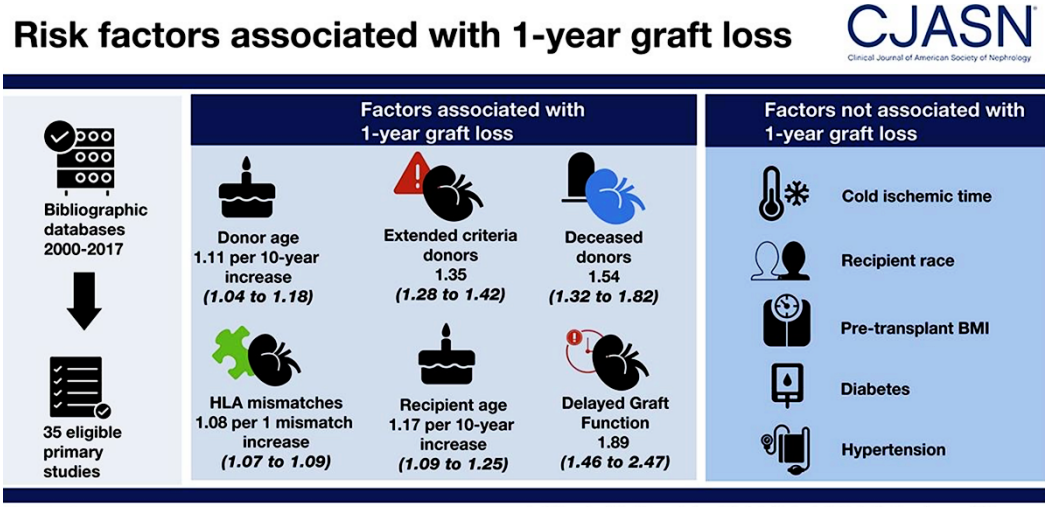


Figure 38 *Risk factors linked with graft loss including extended criteria donor, kidney dysfunction post transplantation and age (Foroutan, Farid et al. "Risk Factors for 1-Year Graft Loss After Kidney Transplantation: Systematic Review and Meta-Analysis." DOI: 10.2215/CJN.05560519)*

Between all factors related to graft-loss a major classification can be done by distinguishing them based on the pre-peri-or post-operative factors [193]. Pre-operative factors include characteristics of both the donor and recipient, including age, gender, race, and immunological compatibility. Peri-operative factors involve parameters such as the duration of cold and warm ischemia times during the surgical procedure. Additionally, post-operative considerations include the occurrence of delayed graft function (DGF) and the management of immunosuppressive treatment. Among all these factors, occurrence of delayed graft function (DGF) will represent the focus of the next paragraph. Independently from the risk of kidney

rejection, the constant and continuous patient follow-up remains a key point at least in the first-year post-graft. During this period patients are required to present themselves to clinical visits in way to evaluate health status, the onset of complications, the correct dosage of immunosuppressors or most important the monitoring of kidney function. These clinical evaluations, that are common to clinical test effectuated during follow-up of CKD patients will be in depth illustrated in the following sections.

1.9.1 Delayed graft function

Delayed graft function defines a condition characterized by the inability of the transplanted kidney to function immediately after transplantation, requiring the initiation of dialysis within the first week following the transplant. The clinical definition of this varies among transplant centers, and within the context of this work, it is defined as the failure of serum creatinine decreasing at least of 10% daily on 3 consecutive days during the first 7 days post-Tx, considering the lowest serum creatinine value observed in the 3 months post-Tx. DGF represents a clinical complication that currently impacts approximately 25-30% of transplant recipients [194]. From the 1990s to the first decade of the 2020s, a notable increase in incidence of DGF occurrence was observed with the percentage rising from 14.7% to 23% in the United States [195,196]. This phenomenon was closely linked to the utilization of EDC and a higher number DCD. As in the case of rejection, DGF also have factors that can be classified based on donor, recipient and peri-operative characteristics ***(see Table 6)***.

Donor Related	**Recipient Related**	**Perioperative**
Deceased donor	Pretransplant dialysis	Hemodynamic instability
Expanded criteria donor	Previous kidney transplant(s)	Calcineurin inhibitors
Donation after cardiac death	HLA mismatch	Nephrotoxic antibiotics
Longer WIT and/or CIT	ABO incompatibility	Nephrotoxic analgesics
Organ quality	Comorbidities	Surgery
Donor age	Higher body mass index	
Acute Kidney Injury	African American ethnicity	
Higher body mass index		
Shipping distance		
African American ethnicity		

Table 6 *Factors related to the development of delayed graft function dysfunction divided in donor related, recipient related and perioperative (Salazar Meira, F et al. "Factors Associated With Delayed Graft Function and Their Influence on Outcomes of Kidney Transplantation." DOI: 10.1016/j.transproceed.2016.06.007).*

As previously seen, donor type represents a fundamental factor for the outcome of graft and is also strictly related to the occurrence of DGF condition. Generally, deceased donors have demonstrated a higher risk of DGF compared to LD [197,198]. This increased incidence can be attributed to the fact that LDs typically have shorter CIT and are healthy individuals, often with only minor comorbidities; additionally, LDs typically exhibit normal renal function, unlike deceased donors, whose medical history is often unknown or incomplete. Another factor that can influence this higher rate of DGF in deceased donor, and particularly for DCD, is the higher period of WIT to which they are subjected. Several studies have shown an increase of DGF risks in WIT longer than 45 minutes because of the possible hypoxia process produced during this ischemic event [182,181,199]. The same hypoxic mechanistic process is probably involved in CIT which duration has to be less than 24h to be accepted, even if a decrease in its length is suggested as reported in a study demonstrating how each additional hour increases the risk of graft failure [200]. Organ preservation conditions are also essential to reduce the risk of DGF incidence. In this context, the choice of machine perfusion rather than static preservation seems to prevent from dysfunction event[201], similarly to preservation solution selection. Indeed, several studies have demonstrate reduced DGF when using IGL-1 solution [185,187] which is also the preservation solution employed in the samples analyzed in this work. In addition to donor related factors some other factors related to recipient can generate a higher probability to develop DGF. Indeed, transplanted patients are often under chronic dialysis which can generate hemodynamic instability and hypotension problems that may aggravate with higher risk of DGF[202]. All these factors can play an important role in triggering kidney dysfunctions. Based on these considerations, several precautions and preventive decisions can be taken prior, during and after transplantation to overcome it. However, the lack of precise tool able to predict DGF and the absence of a precise knowledge about the biochemical events involved in this pathological phenomenon open the door to the use of innovative and original approaches.

1.9.2 Graft function monitoring

Patient follow-up after kidney transplantation is a crucial aspect of post-transplant care which involves a comprehensive and coordinated approach to ensure the well-being of the transplant recipient and the success of the transplantation. As key components of patient follow-up in kidney transplantation, graft function monitoring represents a fundamental point allowing early detection of any issues or complications. Glomerular filtration is one of the key point roles

between the many functions carried out by the kidney and monitoring of GFR represents the best tool for assess evolution in kidney function.

Since 1937 when the GFR was measured for the first time, substantial progresses have been done. GFR represents the rate at which a glomerulus filters plasma in way to produce "primitive urine". Since real GFR cannot be measured, the monitoring of its rate is done by measuring the clearance or serum levels of exogenous or endogenous compounds. In clinical practice different techniques are nowadays used and they can be distinguished in measured GFR (mGFR) or estimated GFR (eGFR), as a measurement or an estimation is conducted for the rate.

1.9.2.1 Measured GFR (mGFR) [203,204,205,206]

When in 1937 Homer Smith, the pioneer of GFR measurement, measured for the first time mGFR, he used the concept of urinary clearance to evaluate the excretion rate of inulin in urine for the first time[207]. The notion of clearance describes the volume of plasma from which a marker (not subject to secretion and reabsorption) is removed through excretion per unit of time; specifically, two different clearance concepts can be used for GFR measurement: (1) ***plasma clearance*** which defines elimination of the marker from plasma without knowledge of its route of elimination; (2) ***urinary clearance*** refers to the elimination of the marker from plasma by urinary excretion. During the last decades, in addition to inulin which imposed itself as gold-standard, other exogenous biomarkers have been used to measure GFR through plasmatic or urinary clearance such as iohexol, iothalamate, technetium 99m diethylenetriamine pentaacetic acid (^{99m}Tc-DTPA) and chromium 51-ethylenediamine tetraacetic acid (^{51}Cr-EDTA). All these markers are, as defined by Homer Smith, ideal exogenous compounds since they do not bound plasma proteins and are freely filtered by the glomeruli while having a molecular weight lower than 20.000Da. If exogenous markers need to be introduced in the body through intravenous injection or bolus administration, the use of endogenous compounds allows an easy measurement of marker clearance without need of reaching a steady state plasma level. As part of endogenous markers creatinine, urea and cystatin C are the most common used in clinics. Creatinine is a derivate of phosphocreatine metabolism in muscle but can also derives from dietary meat intake; its level in blood is related to muscle mass and is strongly dependent from subject's factors age, sex, race or drug assumption. Creatinine clearance measurement is the most commonly available assay in laboratories and can be done though 24h urine collection and a single measurement of steady-state serum creatinine. In ***Table 7*** strengths and limitations of mGFR approaches and markers are highlighted.

Nevertheless the highly accuracy of mGFR to evaluate kidney function, its invasive aspect associated with high cost, need of specialized personal and equipment and time-consuming procedure make it be partially replaced in clinical practice by estimated GFR (eGFR).

Approach	Strengths	Limitations
Methods		
Urinary clearance		
Bladder catheter and continuous intravenous infusion of markers	Gold standard method	Invasive
Spontaneous bladder emptying	Patient comfort Less invasive	Possibility of incomplete bladder emptying Low flow rates in people with low level of GFR
Bolus administration of marker	Shorter duration	Rapidly declining plasma levels at high levels of GFR Longer equilibration time in extracellular volume expansion
24-h urinary collection		Cumbersome Prone to error
Plasma clearance	No urine collection required Potential for increased precision	Less accurate
Nuclear imaging	No urine collection or repeated blood samples required Relatively short duration	Less accurate
Markers		
Inulin	Gold standard No side effects	Expensive Difficult to dissolve and maintain into solution Limited supply
Creatinine	Endogenous marker, no need for administration Assay available in all clinical laboratories	Secretion can vary among and within individuals
Iothalamate	Inexpensive Long half-life for ^{125}I as tracer, suitable for central laboratory for research	Probable tubular secretion, leading to overestimation of GFR Requirement for storage, administration, and disposal of ^{125}I as tracer Potential thyroid uptake of ^{125}I Use of nonradioactive iothalamate requires expensive assay Cannot be used in patients with allergies to iodine
Iohexol	Not radioactive Inexpensive Sensitive assay allows for low dose	Possible tubular reabsorption or protein binding, leading to underestimation of GFR Use of low doses requires expensive assay Cannot be used in patients with allergies to iodine Nephrotoxicity and risk for allergic reaction at high doses
EDTA	Widely available in Europe chelated to ^{59}Cr as tracert	Probable tubular reabsorption, leading to underestimation of GFR Requirement for storage administration, and disposal of radioactive substances when ^{51}Cr is used as tracer
DTPA	Widely available in the United States chelated to ^{99m}TC as tracer New sensitive and easy to use assay for gadolinium	Requirement for storage, administration, and disposal of radioactive substances when ^{99m} Tc used as tracer Short half-life of 99m not suitable for central laboratory Requires standardization for ^{99m}Tc Dissociation and protein binding of ^{99m}Tc, leading to underestimation of GFR Concern for NSF when gadolinium is used as tracer

Table 7 *strengths and limitations of GFR measurement methods and filtration markers (Levey, A S, and L A Inker. "Assessment of Glomerular Filtration Rate in Health and Disease: A State of the Art Review." DOI: 10.1002/cpt.729)*

1.9.2.2 Estimated GFR (eGFR) [203,204,205,206]

Estimation of GFR is a convenient and non-invasive method commonly used in routine clinical practice to diagnose and manage kidney disease. The basics of eGFR are represented by the generation of a regression equation which aim to estimate the mGFR by using the level of endogenous metabolite in serum (such as creatinine) and several other independent factors such as age, sex, BMI etc... As also seen in the case of measurement of GFR, creatinine has a central role in monitoring kidney activity and represents although the predefined marked for estimation of GFR. The main advantage of eGFR is linked to the absence of clearance calculation since it

is evaluated through levels of creatinine in serum. The value of serum creatinine used in combination with all other factors in the equation offers a valuable tool to monitor kidney function through a cost-effectiveness and non-invasive method.

Starting from the second half of 90's, in way to obtain the most precise formula, several equations have been developed all having their limitations linked to the high dependency of creatinine from muscle mass and diet; indeed, even though assay standardization for creatinine and adjustments on sex, race, age, BMI the equations remain imprecise. The first equation that was used to estimate GFR, introduced in 1976, is the Cockcroft-Gault equation; this formula estimates creatinine clearance (CrCl) based on age, sex, body weight and serum creatinine measured through Jaffe reaction a colorimetric assay made with picrate and serum under alkaline conditions. Several limitations are linked to (1) the use of weight factor by systematically underestimating CrCl in obese, (2) the measurement of serum creatinine on Jaffe reaction that cannot be standardized; (3) not include the race underestimating CrCl African and Asian races. The second equation is the MDRD which estimate GFR by using serum creatinine, body surface area, age, sex and race. Even by being more accurate compared to Cockcroft-Gault equation, its main limitations are linked to (1) the race factor which only includes white, African, and Asian (2) systematically underestimates mGFR for eGFR levels >60 mL/min/1.73 m2. The CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) is the newest of all the mathematical equation developed and is considered more accurate and precise than earlier GFR estimation formulas. This equation formed in 2009 and based on serum creatinine as the previous formulas, take into account all the parameters used by the MDRD equation by including correction factors for sex and race (Black or non-Black). The CKD-EPI formula offers, compared to all previous equations, a higher accuracy across the range of GFR with an increased precision in GRF value higher than 60 mL/min/1.73. A summary of equation for GFR estimation in adults based on creatinine with reference methods and limitations in reported in ***Table 8***.

Age	Marker	Reference Method	Standardized Assay	Derivation Study Characteristics	Equation	Comment
Creatinine ($eGFR_{cr}$)						
Adult	Cockcroft-Gault (1976)[a]	mCL_{cr}	No	249 men; 0% Black participants (presumed)	(140 − age × weight)/(72 × Scr) × 0.85 if female	Underestimates mCL_{cr} in older age, obesity, and edematous states
Adult	MDRD Study (2006)[b]	Urinary iothalamate	Yes	983 men/645 women; mGFR 40 mL/min/1.73 m^2; age 50.6 y; 12% Black participants	$175 \times Scr^{-1.154} \times age^{-0.203} \times 0.745$ if female × 1.212 if Black	Underestimates mGFR in high-normal GFR values
Adult	CKD-EPI $eGFR_{cr}$ (2009)[c]	Urinary iothalamate, other mGFR	Yes	4,648 men/3,606 women; mGFR 68 mL/min/1.73 m^2; age 47 y; 30% Black participants	$141 \times \min(Scr/\kappa, 1)^{\alpha} \times \max(Scr/\kappa, 1)^{-1.209} \times 0.993^{age} \times 1.018$ if female × 1.159 if Black | $\alpha = -0.329$ (female); −0.411 (male); $\kappa = 0.7$ (female); $\kappa = 0.9$ (male)	Unbiased across range of GFR; recommended in adults

Table 8 *Equations Estimating GFR from creatinine endogenous Filtration with large representation of north American (Levey, A S, and L A Inker. "Assessment of Glomerular Filtration Rate in Health and Disease: A State of the Art Review." DOI: 10.1053j.ajkd.2021.04.016)*

The CKD-EPI equation, recommended for GFR estimation since 2012 according to international guidelines, has sparked debates in recent years due to the inclusion of a race coefficient, perceived as a potential source of discrimination. This inclusion of a "race" correction factor in medical algorithms has come under increasing scrutiny in the medical field, as it is regarded as a social construct rather than a biological one[201]. For these reasons, in 2021 a new equation has been proposed without the race factor. The recently introduced European Kidney Function Consortium (EKFC) equation[208] is a race-neutral formula derived from the median value of SCr ($EKFC_{SCr}$). Adjustments to the Q value are applied to standardize creatinine levels before incorporating them into the equation, and these adjustments can be tailored for various populations based on factors such as age and gender. An intriguing aspect of this equation lies in its adaptability to biomarkers beyond Creatinine, as explored in another notable study conducted on a large Swedish cohort[209]. This alternative equation, termed $EKFC_{CysC}$, substitutes rescaled serum creatinine with Cystatin C as selected biomarker of interest. Cystatin C, an endogenous biomarker widely employed in clinics for estimating GFR, holds a key advantage in being unaffected by muscle mass, thus showing reduced dependence on age and sex factors. Despite the encouraging results obtained on different cohorts, the performance of these two equations is not increased compared to CKI-EPI equation but the bias is the same for both black and non-black populations[200,210]. In addition to the previous equations, the combination of Cystatin C and Creatinine in the EKFC equation demonstrated an improved performance in estimating GFR. All the mentioned equations are reported in ***Table 9***. [211,212]

Equation	Age (years)	eGFR equation	
$EKFC_{SCr}$	2-40	SCr/Q < 1.0 SCr/Q ≥1.0	$107.3 \times (SCr/Q)^{-0.322}$ $107.3 \times (SCr/Q)^{-1.132}$
	>40	SCr/Q < 1.0 SCr/Q ≥1.0	$107.3 \times (SCr/Q)^{-0.322} \times 0.990^{(AGE-40)}$ $107.3 \times (SCr/Q)^{-1.132} \times 0.990^{(AGE-40)}$
$EKFC_{ScysC}$	18-40	ScysC/Q < 1.0 ScysC/Q ≥1.0	$107.3 \times (SCr/Q)^{-0.322}$ $107.3 \times (SCr/Q)^{-1.132}$
	>40	ScysC/Q < 1.0 ScysC/Q ≥1.0	$107.3 \times (SCr/Q)^{-0.322} \times 0.990^{(AGE-40)}$ $107.3 \times (SCr/Q)^{-1.132} \times 0.990^{(AGE-40)}$
$EKFC_{SCr+ ScysC}$	≤18		Median of $EKFC_{SCr}$ and $EKFC_{Cys}$

Table 9 *New equations estimating GFR introduced by European Kidney Function Consortium in 2021. SCr= serum creatinine; SCysC=serum cystatin C*

1.10 Metabolomics and kidney transplantation: state of art

As highlighted in the preceding sections, the process of kidney transplantation, encompassing both its pre-transplant and post-graft phases, demands meticulous monitoring to enhance the success of the graft and the overall well-being of the patient. In this context, metabolomics through a personalized medicine approach places itself as an important tool to evaluate, monitor and diagnose kidney diseases. Over the last decades, the application of metabolomics has led to the identification of significant biomarkers associated with biochemical pathways impacted by physio-pathological events. This link between disease and metabolome has shed light on how pathologies are distinguished by their unique metabolomic signatures, which encapsulate the combined effects of both external and internal cellular processes. The evolving status of patients during pathological condition makes metabolomics play a crucial role in detecting and explaining the dynamic system composed by metabolites.

Thanks to its potential to provide insight into the disease mechanisms, metabolomics has been extensively used to explore and investigate CKD pathology. Indeed, the first publications diagnosing CKD using a metabolomic approach have been published in the last decade [213,214], but since then a lot of other studies have been published. Based on the complexity of this pathology, metabolomics has been used to investigate the most disparate aspect of CKD by going from identification of diagnostic biomarkers to the discrimination of panel of features linked to pathology' stages until the generation of model estimating eGFR. Concerning diagnosis of kidney disease several publication reported the identification of specific metabolites and pathways as responsive for CKD progression such as uremic toxins[209], plasma phospholipids[209,], or several amino acids (valine, isoleucine, and tryptophan)[210] . Another important focus of interest for the metabolomics community in the field of CKD is represented by the improvement of GFR estimation. In fact, this side of the pathology is still nowadays challenging in clinical practice and several studies reported interesting results in this field. D-threitol, myo-inositol, 4-deoxyerithronic acid and galacturonic acid were reported to be strictly correlated to GFR evolution in serum samples of CKD patients[211]. In kidney transplant recipients, serum valine and myo-inositol were found to improve GFR estimation though NMR technique [151].

As best treatment for patients with ESRD, kidney transplantation represents another axis of interest for metabolomics community. Indeed, in KTx context patients' evolution through days,

weeks, months, and years' post-graft is dynamically changing and the application of metabolomics could represent a real help in clinical practice. Indeed, since the pre-transplantation event metabolomics can be employed in kidney monitoring in way to evaluate graft quality and eventually predict the graft outcome by adding interesting and useful data to clinical measurements already available. During the post-graft period, monitoring patient progress is essential for clinicians to enhance patient outcomes, address post-transplant dysfunction, or manage rejection, ultimately improving the patient's overall status.

By taking advantage of the dynamic aspect of metabolomics, several studies have been conducted in the context of graft quality prediction and post kidney transplantation events (as shown in ***Table 10***).

References	Year	Species	nº sample	Sample type	Analytical platform
Kidney transplantation					
Serkova et al. [218]	2005	Rat	6	Kidney tissue, blood	NMR
Stenlund et al. [219]	2009	Human	19	Urine	NMR
Suhre et al. [220]	2016	Human	241	Urine, kidney tissue	GC-MS, LC-MS
Blydr-Hansen et al. [221]	2017	Human	396	Urine	LC-MS
Gagnebin et al. [222]	2020	Human	66	Plasma	LC-MS
Iwamoto et al. [223]	2022	human	50	Plasma, urine saliva	CE-MS
Colas et al. [224]	2022	Human	56	Urine	LC-MS
Kidney preservation					
Nath et al. [225]	2014	Human, pig	22	Perfusate	NMR
Guy et al. [226]	2015	Human	26	Perfusate	NMR
Wang et al.	2017	Human	36	Perfusate	NMR
Faucher et al. [228]	2022	Human	38	Perfusate	LC-MS
Liu et al. [229]	2023	Human	232	Perfusate	LC-MS

Table 10 *Studies concerning kidney transplantation and kidney preservation with metabolomics approach reported in literature.*

All the studies have been conducted by using different analytical platforms such as NMR and LC-MS and by using urine, plasma, or serum as starting materials. A particular case is reported for studies concerning organ preservation and kidney quality prediction in which the biospecimen of choice is represented by perfusate. The cohorts studied in the context of pre- and post-transplantation events have highlighted a variety of metabolites linked to the degradation of kidney function and/or kidney dysfunctions by reporting hippuric acid, TMAO, and other uremic toxin as strongly related to disease status. Indeed a study on KTR identified a dynamic model able to estimate the recovery process of kidney-transplanted patients by using

NMR on urine sample[219]; characterization of plasma, urine and saliva of the same kind of cohort allowed the identification of 3-indoxyl sulfate as predictor of acute rejection[223]. Suhre et al. [220] in a study combining mRNA signature with metabolites identified though GC-MS 3-sialyllactose, xanthosine and quinolinate as marker of kidney allograft reject. Other studies have explored the pre-transplantation event though the analysis of perfusate samples for predict DGF outcome. Interesting is the study of Wang et al. [227] which reported 37 metabolites (i.e. alpha-glucose taurine, citrate, betaine) being related to kidney dysfunction condition; alpha-ketoglutarate, a propane derivate and two phosphatidylcholines were found be related to higher risk of graft failure and have an significant interaction with donation after cardiac death [229] .

Despite the diversity of works reported in literature, a concrete lack of studies with direct applications in clinical practice is missing even if tangible efforts have been done. Indeed, multiple research studies have utilized the metabolomics approach to investigate the kidney transplantation process, in its early and later phases. These analyses share the common goal of advancing our understanding of the mechanistic processes associated with kidney dysfunction and seeking ways to mitigate adverse outcomes following KTx. However, a crucial element is still missing in the effort to translate these studies from laboratory to clinics. Indeed, there is a need to develop new diagnostic tools that could help clinicians to routinely manage kidney transplant recipients. More particularly, the detection of biomarkers predicting the risk of post-transplantation renal function decline could be a useful predictive diagnosis tool in KTx practice.

Chapter 2

Aim of the thesis

2 Aims of the thesis

Kidney transplantation (KTx) is considered as the optimal treatment for individuals with end-stage renal disease. Because of the increasing gap between demand and supply for KTx, deceased donors, including those from brain and circulatory death, have been incorporated alongside living donors. Over recent decades, donors have been categorized based on graft survival rates into standard and extended criteria donors. However, kidneys from extended criteria donors, such as those obtained after circulatory death, are subjected to prolonged ischemic events, resulting in heightened risks of both early and long-term graft complications. These data underscore the critical significance of two key aspects related to transplantation: (1) the accurate assessment of graft quality prior to transplantation, which can help reduce rejection rates through better patient management or more refined donor selection criteria; (2) a more precise and prognostic monitoring and follow-up of transplant patients, enabling proactive and personalized management. Despite the progress made over the past few decades in enhancing graft quality assessment through risk factor indices and improving patient follow-up with more accurate GFR equations, these guidelines do not provide a complete answer to the problem of graft rejection and restoration of renal function in clinical practice. In this context, there is a great need for new tools and innovative approaches to meet challenges and metabolomics has emerged as an interesting solution. Indeed, as discussed in the introduction, metabolomics offers an innovative, holistic approach that enables a focus on the dynamic aspects of biological systems. Hence, metabolomics seems to be particularly well suited in the context of transplantation process to better evaluate the graft quality in pre-KTx as well as to predict kidney function in 1-year post-graft period.

Thus, in this thesis, we aim to investigate the two mains transplantation process challenges:

1. ***Post-transplantation axis (Metarein project).*** On this part we focused on the prediction of kidney function in post-transplantation period. The peculiarity of this axis is represented by the longitudinal collection of urine samples at 3- and 12-month intervals following transplantation, with patient stratification based on mGFR delta values. The main goal of this section is to identify a panel of biomarkers that can indicate at M3 a decline of kidney function at M12. The predictive nature of this experimental design aimed to discover a list of potential metabolites associated with kidney degeneration over time. Through this study, we expected to develop a tool that can in advance predict

kidney function decline and enhance patient monitoring and management during the post-transplantation period.

2. ***Pre-transplantation axis (Metaperfusate project).*** In this part, we focused on the biological events impacting the graft before the transplantation and during the ischemic time. This part of the work aimed to identify a metabolomic signature in perfusate sample that could be correlated to donor type and to the kidney function recovery in early post-KTx. Specifically, the set of putative biomarkers identified may be linked to the different ischemic events characterizing diverse donor types composing our cohort. The primary objective along this pre-transplantation axis is to offer valuable insights to clinicians in the pre-graft decision-making process by developing a model evaluating the impact of donor type on graft metabolomes and warning about probably at-risk patients.

Analyzing these two complementary cohorts enabled us to develop distinct experimental designs for each project, emphasizing the unique strengths of each. Firstly, Metarein cohort composed by a longitudinal cohort of KTR patients which will allow the delineation of a predictive model for kidney function decline. Secondly, Metaperfusate cohort, with its primary focus on the pre-transplantation process and its innovative use of perfusate samples, grants direct access to the donor's metabolome, facilitating correlations with kidney dysfunction in the recipient ***(Figure 39).***

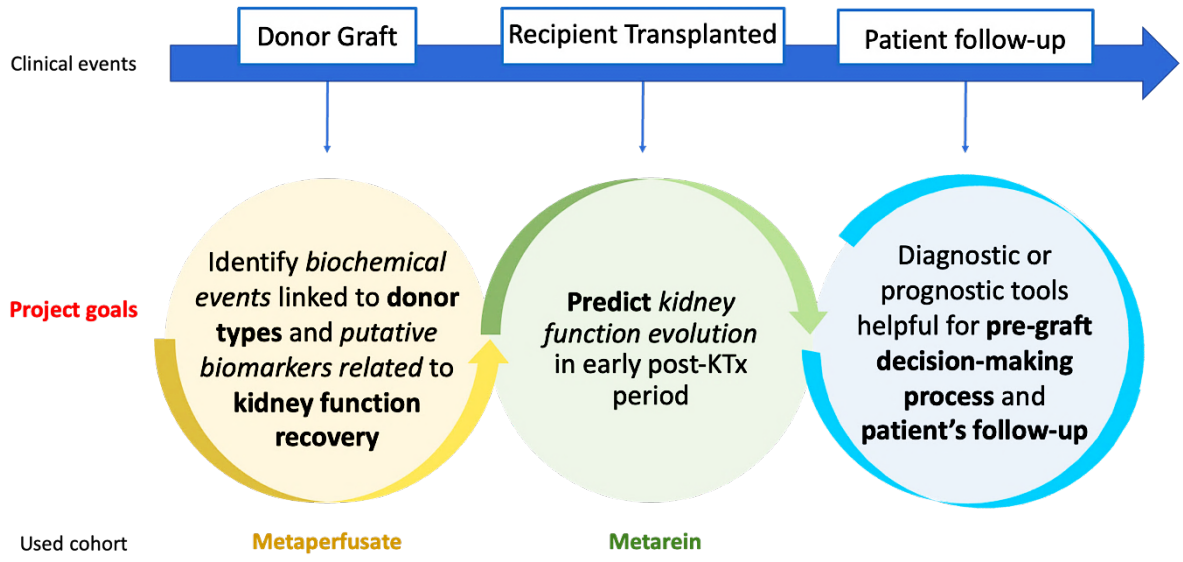


Figure 39 *Timeline of KTx with highlighted key point events; the main goals of the project are reported in relation to each phase of KTx.*

In this work both NMR- and MS-based metabolomics approaches were employed on the first axis to integrate information coming from the two platforms, thereby enhancing the quantity of biological information related to kidney function decline. On the second axis, an NMR-based metabolomics approach was applied to investigate the metabolome related to early post-KTx dysfunction. Merging the information coming from these axes will allow us to shed light on the clinical events characterizing kidney transplantation, with the objectives of: (1) identifying potential biomarkers associated with pre- and post-KTx; (2) explore impact of donor's types on the biochemical pathways; and (3) developing diagnostic instrument for patient treatment and follow-up. From initial exploratory analyses to the development of predictive models, the primary goal of this thesis is to create tools that may assist clinicians in navigating the intricate landscape of kidney transplantation events within their clinical practice.

Chapter 3

Urine in ^{1}H-NMR-based metabolomics

3 Urine in ^{1}H-NMR-based metabolomics

3.1 Background - Urine biofluid

3.1.1 Interest of urine in metabolomics

The significant role of urine in metabolomics, especially as a diagnostic biofluid, is underscored by historical references. Since Hippocrates and by passing though Byzantine era and Middle Age, analysis of urine samples have always been in the first line for medical diagnosis ***(Figure 40)*** [230,10]. Although this practice declined over time, urine has continued to be a cornerstone in modern medical diagnostics, playing a crucial role in identifying kidney, and bladder pathologies[34].

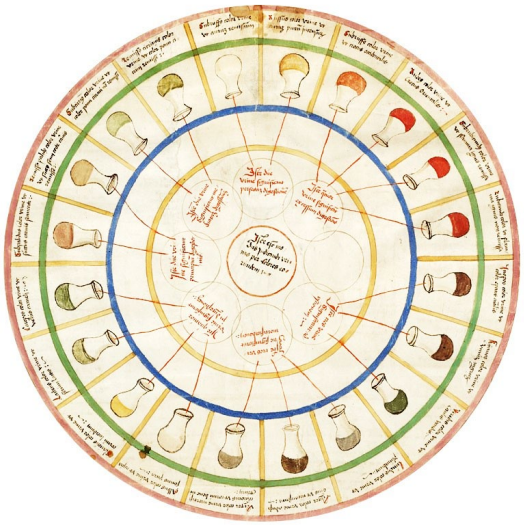


Figure 40 *Urine wheel chart describing the possible color, taste and smell and how to use it for diagnosis (Nicholson, Jeremy K, and John C Lindon. "Systems biology: Metabonomics." DOI: 10.1038/4551054a)*

Nowadays, urine has gained extensive popularity in metabolomics research due to its non-invasive sampling method, capacity for large-volume collection, and lack of interference from proteins and lipids (in non-pathological condition). As a filtrate of blood plasma, urine contains a wide range of metabolites, offering valuable insights into different metabolic pathways and their dysfunctions making it especially well-suited for the personalized medicine approach. Nevertheless, the intricate chemical composition of this biofluid presents challenges in achieving a comprehensive understanding of its complexities. Indeed, in metabolomics several analytical platforms (such as NMR, LC-MS, GC-MS etc.) have been used to enhance the metabolome coverage in urine samples. Despite all the efforts, today only a small part of the metabolites present in urine can be identified by using a single platform and in an untargeted mode. Within this complex biofluid, alongside endogenous metabolites reflecting the biological status of an individual, there is a notable abundance of exogenous metabolites resulting from diet, drug intake, and other lifestyle factors. The complexity and wide range of chemical

molecules and resulting urine spectrum is only one of the problems linked to this biofluid. Biological variability, pH variation, and normalization between urine samples represent other fundamental challenges in metabolomics studies which aim to detect urinary biomarkers linked to pathological conditions. In the next section the development and set up of urine protocol for untargeted metabolomics will be elucidated by focusing on analytical and practical aspects for urine preparation.

3.1.2 Urine spectral complexity

Urine spectra coming from ^{1}H-NMR technique are characterized by high signal complexity and very information-rich aspect. As we can see in ***Figure 41***, which represents a spectrum from a healthy individual acquired through a 600MHz spectrometer, a great quantity of signals can be distinguish. Indeed, a known study on urine metabolome shown the possibility to identify and quantify 209 unique compounds within a cohort of 22 healthy volunteers through ^{1}H-NMR technique[34] by placing this platform as the most suitable for urine analysis in terms of quantification and comprehensiveness.

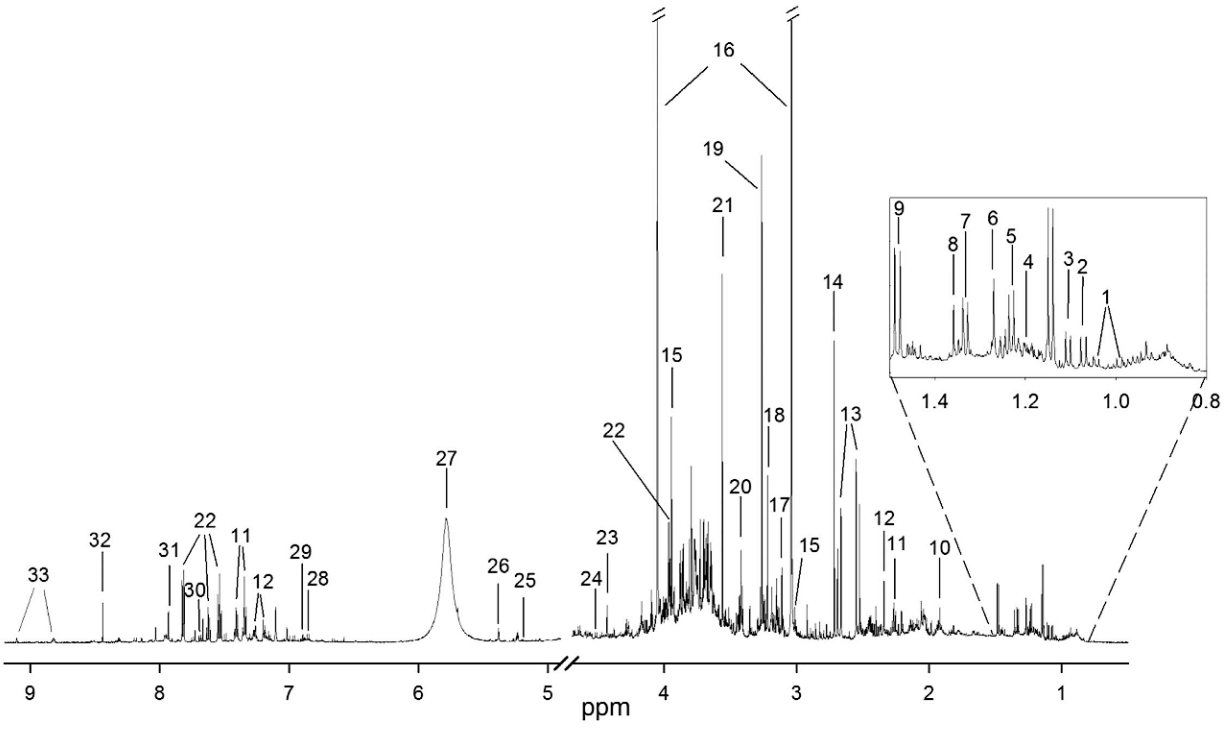


Figure 41 *The metabolites are numbered accordingly: 1: valine, 2: 3-hydroxyisobutyrate, 3: 4-deoxyerythronic acid (4-DEA), 4: 3-aminoisobutyrate, 5: 4-deoxythreonic acid (4-DTA), 6: 3-hydroxyisovalerate, 7: threonine/lactate, 8: 2/alfa-hydroxyisobutyrate, 9: alanine, 10: acetate, 11: phenylacetylglutamine, 12: p-cresol sulfate, 13: citrate, 14: dimethylamine, 15: creatine, 16: creatinine, 17: proline betaine, 18: carnitine, 19: trimethylamine-N-oxide (TMAO), 20: taurine, 21: glycine, 22: hippurate, 23: trigonelline, 24: ascorbate, 25: xylose, 26: allantoin, 27: urea*, 28: 4-hydroxyphenylacetate, 29: tyrosine, 30: 3-indoxyl sulfate, 31: histidine, 32: formate, 33: trigonelline, *urea peak was not quantified. (Giskeødegård, Guro F et al. "Diurnal rhythms in the human urine metabolome during sleep and total sleep deprivation." DOI: 10.1038/srep14843)*

Despite the high number of metabolites that can be identified through NMR platform, two main problems characterize urine biospecimen: (1) the first strictly related to NMR platform is **variation in pH** between samples which represent a well-known phenomenon in urine biofluids; (2) the second is represented by the **normalization** step which is a general challenge in urine samples independent from analytical the platform (NMR, LC-MS etc.). Both these aspects are subjects of discussion for researchers since they highly impact the results of an analysis and their biological interpretation. For these reasons, these elements will be explored and elucidated in detail in the next sections by allowing the set-up of a defined urine protocol.

3.2 Urine protocol set-up: pH variation and chemical shift

3.2.1 Introduction

Inter-sample pH variation is a common phenomenon in human urine samples highly dependent from diet, lifestyle, circadian rhythm and/or drug consumptions. In NMR spectra, when pH variations occur, some metabolites with acid and/or basic functions are subject to variation in chemical shifts of some of their signals which determine their "shift" through the x-axis of the NMR spectrum. In this case, a same metabolite will have different chemical shifts depending on the pH value of the sample. In metabolomics studies, when cohorts of urine human samples (especially in non-healthy status) are analyzed, variability between pH value is normal and needs to be taken in account. Indeed, as seen in previous chapter, one way to preprocess NMR data by reducing the complexity of spectra is based on bucketing step. When using this data-reduction technique, the x-axis of spectrum is divided in region of equal width, named bucket, each of them representing a specific feature (or metabolites) in the final data matrix. Variation in pH among samples can lead to differences in the chemical shifts of some metabolites; therefore, the spectral bucket no longer accurately represents specific feature but instead captures a different metabolite for each spectrum on the basis of pH value. To avoid these unexpected shifts, several solutions have been proposed. In this chapter we will study 2 of these approaches and compare it to the gold standard sample preparation. The three different sample preparation conditions that were tested are :

1. Use of phosphate buffer (pH= 7.4) as widely reported in literature;
2. Addition of KF solution to urine aliquot;
3. Freeze-drying urine aliquot prior analysis.

Among the three protocols evaluated, the current standard in literature for NMR metabolomics involves adding phosphate buffer to urine aliquots, aiming to stabilize and mimic the physiological pH value (pH=7.4) [78,231]. Another tested protocol, already established and utilized in our lab, focuses on freeze-drying the urine aliquot. The interest of this protocol is represented by the fact that the samples is totally "dehydrated" with the aim to suppress pH effect on chemical shift. Once freeze dreeing process is complete, samples are solubilized in a phosphate buffer at known pH= 7.4 in deuterated water. The third protocol involves the use of KF in urine aliquots. This protocol, detailed in literature[232, 233], primarily aims to remove by chelation divalent cations like Ca^{2+} and Mg^{2+} , known to form complexes with some metabolites (particularly with polyvalent carboxylates functions such as citrate, Hippurate...) and thus to

influence their chemical shifts despite pH buffering according to these divalent ion concentrations. Indeed, addition of KF to urine samples has demonstrated the ability to bind Mg^{2+} and Ca^{2+} to form water insoluble MgF_2 and CaF_2, thereby preventing these ions from binding to certain metabolites and causing variations in their chemical shift. Evaluation of the efficacy and robustness of these 3 different protocols was done in way delineate the best method able to overcome problematics related to pH variations inter-samples.

3.2.2 Materials and methods

Urine collection and study cohort

Second morning urine samples from 4 individuals (2 males and 2 females) were collected at three time points (day1, day 2, day 3). For each timepoint per individual triplicate samples were aliquoted ***(Figure 42).*** Theses aliquots were supplemented with 2% Na azide solution in way to overcome bacterial contamination and stored at -80°C until analysis.

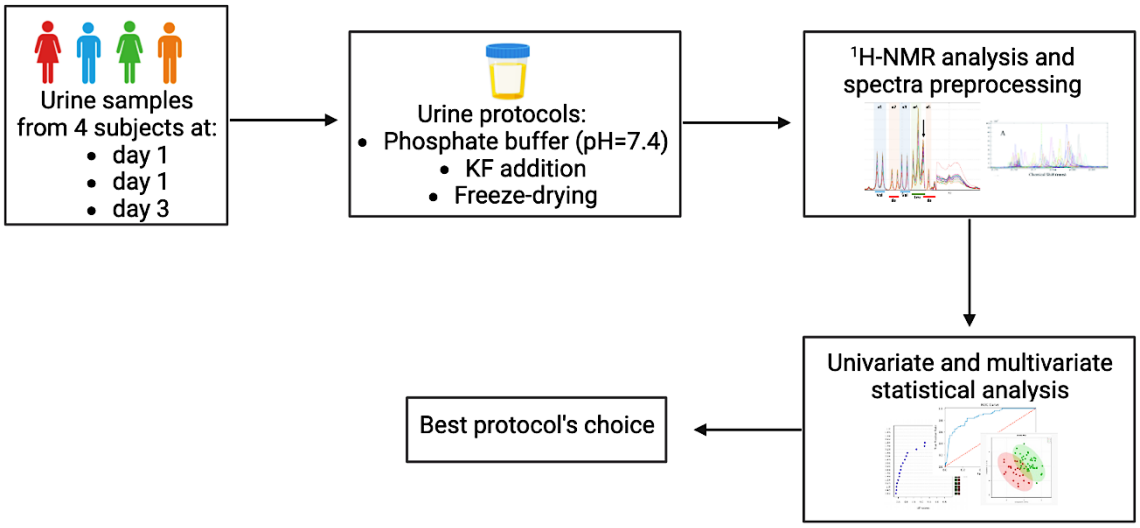


Figure 42 *Workflow used for Urine protocol set-up: pH variation and chemical shift.*

Sample preparation

Prior sample preparation pH value was measured for each sample; three different sample preparations were done:

1. 500µl of urine samples was supplemented with 300µl of phosphate buffer (pH= 7.4) in D_2O.
2. 550µl of urine samples was supplemented with 40µl of KF solution at 5M, vortexed and centrifugated (14489g, 10 min, 4°C). 500µl of supernatant were taken and

supplemented with 300µl of phosphate buffer (pH= 7.4) in D_2O, then again vortexed and centrifugated (14489g , 10 min, 4°C) (detailed protocol in Jiang e al.[232]).

3. 500µl of urine samples was freeze-dried with a 24h cycle and then solubilized in 800µl of phosphate buffer (pH= 7.4) in D_2O.

pH values were measured again for each sample and for each used protocol (see measured pH values in supplementary information Table S1). Prior ^{1}H-NMR analysis the samples were supplemented with 100µl of maleic acid 5mM used as internal standard and 10µl of TMSP (10mg/ml) as chemical shift reference to calibrate the scale of δ to 0 ppm .

NMR-metabolomics

All samples were measured at 298K on a 700MHz Bruker Avance HD spectrometer operating at 700.17MHz for proton detection. The sequence used is a 1D NOESY sequence with presaturation for urine samples. The Noesypresat experiment used a RD-90°-T1-90°-Tm- 90°-acquire sequence with a relaxation delay of 4 s, a mixing time (Tm) of 10 ms and a fixed T1 delay of 4 µs. Water suppression pulse was placed during the relaxation delay (RD). The number of transients is 64 (64K data points) and a number of 4 dummy scans is chosen. Acquisition time is fixed to 3.2769001 s. The data were processed with Bruker Topspin 4.0.8 for baseline and phase correction and the entire δ scale was set to 0 ppm using the internal standard TMSP.

NMR data preprocessing

Data preprocessing on spectra was done by using MestReNova lab (v14.1.1). Normalization on totally by sum was performed on the spectra; bucketing process was done by using spectral width equal to 0.05. No algorithms for alignment were applied in way to better evaluate the effect of our three protocols on chemical shift variation.

Multivariate analysis

SIMCA® software was used to generate PCA score plot in way to allows data-reduction and overview. In addition to PCA, the concept of Metabolic informative content (MIC) and inertia were used by allowing to calculate the intergroup and intragroup variability. Inertia calculations were performed by using a R code. The concept of the MIC is closely tied to the signal-to-noise ratio within a spectrum. Essentially, the MIC seeks to quantify the extent to which we can

differentiate the "signal" attributed to a specific group or condition of interest in our analysis from the "noise," which includes various factors like sampling variations, bacterial degradation, and freezing/thawing cycles[91,234]. In its computation, MIC employs the calculation of two metrics: *inertia between groups*, which assesses the variability among different sample classes, and *inertia*, within groups which measures the variability within samples belonging to the same class.

3.2.3 Results and discussion

No significant differences between protocols were shown when considering pH values after sample preparation (see supplementary information ***TableS1***). To evaluate the impact of each tested protocol on chemical shift variation, after spectra acquirement and bucketing process PCA score plot was generated. Thanks to this non supervised approach, it was possible to examine samples and groups dispersions over the plot. In each plot of ***Figure 43*** it is possible to highlight the 4 individuals encoded from 1 to 4. From these plots we can see how each replicate and each individual are grouped independently from their time-point. In addition, when comparing the three protocols, already visually is possible to spot the intra-group variations. Indeed, in freeze-dried protocol we can see a higher intra-group dispersion than in deuterated buffer and KF protocols.

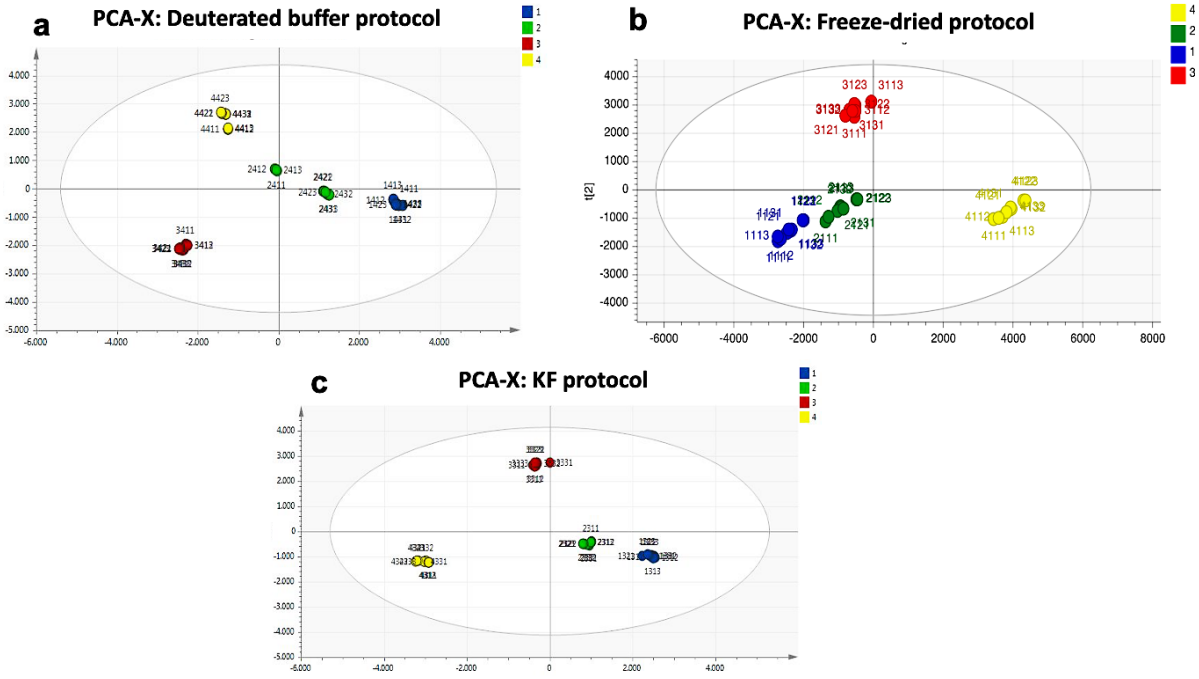


Figure 43 *PCA score plots for each protocol for urine preparation;* ***(a)*** *PCA on deuterated buffer protocol (R2=0.921;);* ***(b)*** *PCA on freeze-dried protocol (R2=0.958);* ***(c)*** *PCA on KF protocol (R2=0.957)*

To better evaluate the effect of the three different protocols on our samples another supplementary metrics was used and is represented by: Metabolomic Informative Content (MIC)[234]. The MIC is an indicator which aims to quantify the percentage of total variation of the data (total inertia) due to the variability between groups and within groups. In the case of this study, our interest will be focused on the inertia linked to variability within groups. Indeed, the lower the value of inertia within groups, the more effectively the used protocol manages to overcome the problem of chemical shift variations due to pH modifications. As highlighted in the summary ***Table 11***, the lowest value of inertia within groups is reached when using the KF protocol which shows a value equal to 4.532 compared to the other two protocols showing a much higher variability value. In opposite to sample preparation though KF which showed the lowest inertia within groups, freeze-dried protocol exhibited the highest variability. The primary reason to this result may be attributed to the phenomenon of hydrogen-deuterium exchange in creatinine when urinary sample is reconstituted in D_2O solution, leading to a decrease in its peak intensity[235]. In our study, under freeze-dried conditions, this phenomenon could alter the intensity of creatinine peaks among samples, resulting by consequences in higher variability within group.

Protocol	Inertia between groups	Inertia within groups
Deuterated Buffer	85.24	14.76
Freeze-dried	77.74	22.26
KF	95.47	4.532

Table 11 *Metabolic Inertia Content (MIC) calculated for the urine protocols tested; variability is expressed in terms of inertia between groups and inertia within groups. Inertia between groups describes the variability between samples appertaining to different groups (in these case different individual); inertia within groups represents the variabiliy of samples appertaining at the same subject(variation within replicates).*

3.2.4 Conclusions

In conclusion, the study allowed us to set-up a protocol for urine preparation. By evidence, the PCA score plot and the MIC values, highlighted the importance of a correct sample preparation with KF when facing pH variation inter samples. In particular, this protocol is especially well-suited when using bucketing method for data reduction in which features alignment is of fundamental importance for a correct biological interpretation. This protocol will be the protocol-of-choice for the urine sample preparation of Metarein project that will be shown in chapter 5.

3.3 Urine protocol set-up: normalization

3.3.1 Introduction

A common problematic related to urine analysis as biofluid is represented by the natural variability in urine's concentration. Indeed, depending on hydration and physiological status the concentration of metabolites in urine samples can widely vary between individuals and even between samples coming from the same subject. Normalization step, in this context, represents a fundamental solution in way to minimize variations resulting from individual urine outputs. In physiological condition, this problematic can be overcome by normalizing urine content on creatinine value representing an endogenous metabolites whose renal excretion in urine is expected to be constant [231, 232]. However, the concentration of creatinine can vary depending on sex, age, diet and activity in normal condition; in addition to this, in pathological condition and specifically in kidney impairment status, creatinine levels are not anymore constant in urine and its excretion could be altered [233, 234]. In metabolomics-based studies where most often pathological conditions are the main focus, several strategies have been developed and can be classified in three main categories: (1) pre-acquisition by adjusting the volume to normalize some specific parameters (i.e. osmolarity); (2) post-acquisition by using a data treatment step; (3) pre- and post- acquisition by using a combination of both previously cited methods.

The irreversible nature of pre-acquisition normalization poses a significant disadvantage, as once normalization is applied, the original raw data cannot be recovered. Given this major disadvantage and the challenging nature of the dataset, our study will focus on post-acquisition normalization by elucidating the different approaches that can be used.

In addition to creatinine normalization, whose concentration can vary in kidney impairment status, another technique reported in literature as gold-standard method used in clinics, is represented by osmolarity. Osmolarity, defined as the concentration of solute particles in a solution and expressed as osmoles of solute particles per liter of solvent, is widely used in clinical studies to assess the concentration of solutes in biofluids such as blood and urine. This normalization method allows a comprehensive evaluation of sample concentration with reported better results for normalization when compared to creatinine method [235]. Another widely used method in metabolomics studies of urine is represented by Probabilistic Quotient Normalization (PQN) algorithm. This technique is based on the calculation of a most probable dilution factor by looking at the distribution of the quotients of the amplitudes of a test spectrum by those of a reference spectrum[236]. The benefic effects of this normalization are shown in several reported studies in literature [237,238 ,239]. Despite the increasing number of metabolomics

studies on urine samples, normalization strategies are still source of debate within metabolomics community due to the lack of a unique method well-fitting to each study. Indeed, today scientists agree on the fact that the choice of the most suitable method may depend on the study objectives, characteristics of the cohort, and the metabolites of interest. Due to these considerations, various normalization methods were tested in order to determine the most suitable method for representing our datasets.

3.3.2 Materials and methods

Urine collection and study cohort

Second-morning urine from 4 individuals (2 males and 2 females) was collected. Additionally, to the starting aliquot, for each sample per individual a dilution 1:1 and 1:2 in triplicate was done. A pool sample from each starting and diluted aliquot was done ***(Figure 44).*** The aliquots were supplemented with 2% Na azide solution in way to overcome bacterial contamination and stored at -80°C until analysis.

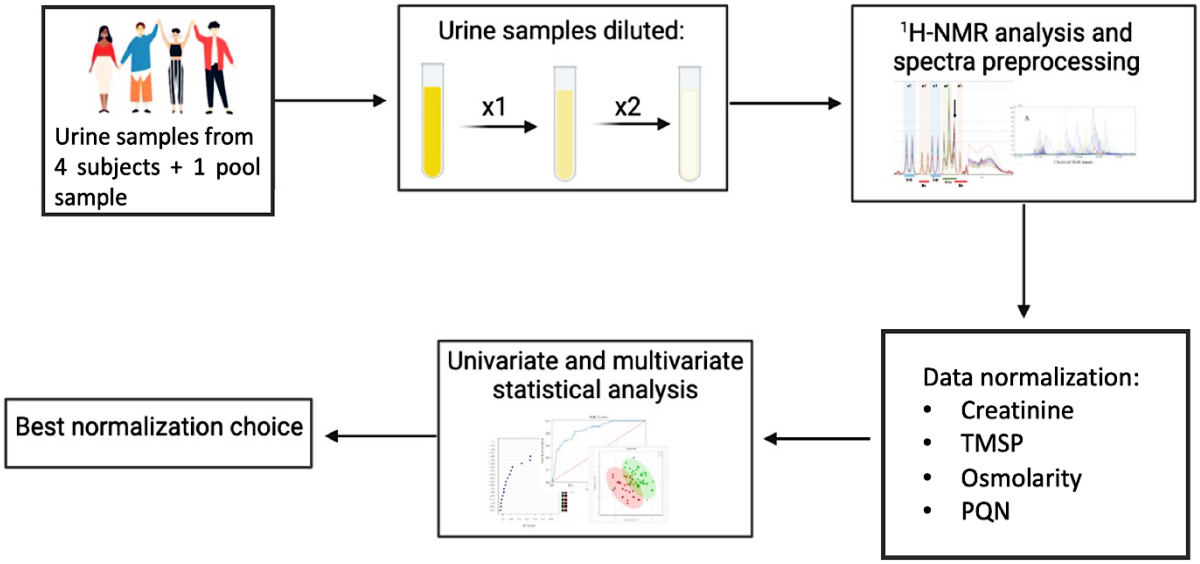


Figure 44 *Workflow for "urine protocol set-up: normalization"*

Sample preparation

For sample preparation 550µl of urine was supplemented with KF solution and prepared by following the protocol set-up in *section 3.2.* For diluted samples, to aliquots of 500µl urine a supplementation with 550µl (x1 diluted sample) and 1100µl (x2 diluted sample) of water was done; then for each diluted sample a withdraw of 550µl of samples was done prior preparation

with KF protocol. Once the samples were centrifugated and pH measured, 100µl of maleic acid 5mM used as internal standard and 10µl of TMSP (10mg/ml) as chemical shift reference were added to the aliquot. Creatinine and osmolarity measurement were conducted on sample prior NMR analysis. Creatinine measurement was done by using COBAS 6000 C501 device (Roche-Hitachi1). Osmolarity measurements were additionally done by using osmometer (Type 13, Löser Messtechnik; Berlin).

NMR-metabolomics

All samples were measured at 298K on a 700MHz Bruker Avance HD spectrometer operating at 700.17MHz for proton detection. The sequence used is a 1D NOESY sequence with presaturation for urine samples. The Noesypresat experiment used a RD-90°-T1-90°-Tm- 90°-acquire sequence with a relaxation delay of 4 s, a mixing time (Tm) of 10 ms and a fixed T1 delay of 4 µs. Water suppression pulse was placed during the relaxation delay (RD). The number of transients is 64 (64K data points) and a number of 4 dummy scans is chosen. Acquisition time is fixed to 3.2769001 s. The data were processed with Bruker Topspin 4.0.8 for baseline and phase correction and the entire δ scale was set to 0 ppm using the internal standard TMSP.

NMR data preprocessing

Data preprocessing on spectra was done by using MestReNova lab (v14.1.1). After phase and baseline correction, an alignment on global spectra was done. Peaks of maleic acid and water were deleted from the spectra and prior to bucketing step. Different normalization types were tested on resulting bucket table:

1. *Normalization on Creatinine:* creatinine value was measured for each sample prior NMR analysis. Once the raw data were processed and bucketing was done, the data matrix was normalized by using value of creatinine measured specific to each sample.
2. *Normalization on TMSP:* Peak of TMSP placed at 0 ppm on the spectra was used as internal standard for normalization. TMSP value was set to 1 for all the spectra and following a bucketing of 0.02 spectral width was performed.
3. *Normalization on Osmolarity:* the bucket table was generated by following the description of previously cited normalization on TMSP. Once data matrix was obtained each sample was normalized by its value of osmolality.

4. *Normalization on PQN:* spectra were normalized by using the algorithm probabilistic quotient normalization widely used in NMR-metabolomics.

Multivariate analysis

SIMCA® software was used to generate PCA score plot in way to allows data-reduction and overview. In addition to PCA, the concept of inertia was used by allowing to calculate the intergroup and intragroup variability. Inertia calculations were performed by using a R code.

3.3.3 Results and discussion

The main goal of our study was to identify the best normalization approach allowing to overcome problems due to the dilution effect in urine. By taking advantage of the "controlled" dilution effect generated to each urine subject we applied PCA model and inertia calculation to evaluate the reduction of dilution effect obtained through different normalization approach. Taking advantage of unsupervised model generated by PCA it was possible to spot the distribution of samples over the plot. As we can see, each point represents a sample classified on the basis of individuals (n=5) and on the basis of the dilution effect applied (native sample; diluted x1; diluted x2) .

By looking at PCA score plots of the different normalization methods we can spot that in TMSP score plot, it is more difficult to clearly cluster samples based on subject, since all the individuals are more dispersed over the plot compared to other methods. Indeed, normalize on the internal standard as in this case, allows to reduce experimental differences between samples but not to overcome variations resulting from individual urine outputs. Concerning the other normalization techniques, is difficult to visually spot which method allows the better surmount of dilution effects on our dataset ***(Figure 45).***

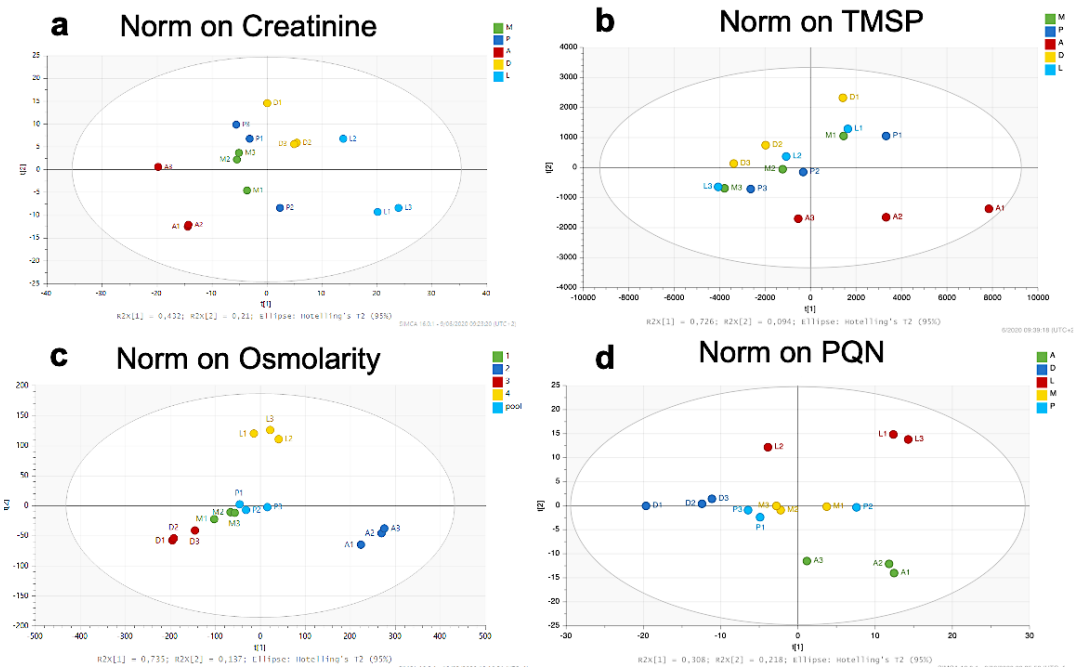


Figure 45 *PCA score plots on the cohort using different normalization techniques (letters M, A, D, L, P refers to the 4 subjects and pool samples; the following number refers respectively to 1= native sample; 2= diluted x1; 3= diluted x2);* ***(a)*** *normalization on creatinine (R^2=0.893; Q^2=0.675);****(b)*** *normalization on TMSP (R^2=0.82; Q^2=0.672);****(c)*** *normalization on osmolarity (R^2=0.905; Q^2=0.828);****(d)*** *normalization on PQN (R^2=0.716; Q^2=0.431).*

For this reason, in addition to PCA models, inertia between and within groups were measured. By comparing the value of inertia within groups, was possible to see the effect that each normalization methods have on the samples coming from the same individual with different dilution factor. Specifically, lower is the value of inertia withing groups, less is the separation between diluted samples from a same subject which means that the corresponding normalization technique works well.

In the context of this work, the normalization techniques can be classified as following based on inertia within groups and consequently on their effectiveness in reducing dilution effects ***(Table 12):***

Osmolarity > PQN > Creatinine> TMSP.

Normalization method	Inertia between groups	Inertia within groups
Creatinine	84.32	15.66
TMSP	30.32	69.68
Osmolarity	95.15	4.86
PQN	89.53	10.47

Table 12 *Metabolic Inertia Content (MIC) value for each normalization method divided in inertia between and within groups.*

3.3.4 Conclusions

By taking advantage of this study, we tested the efficacy of several normalization methods reported in literature in a sample cohort "mimicking" the dilution effects. By being aware of the limits of this study represented by the linear dilution in samples that does not represent a real case study we were still able to look at the effect of normalization on a dataset. However, through this experimental design we were able to spot how normalization step impact the dataset and how important is to correctly set this procedure. Indeed, in a metabolomics study, reducing the variability between samples due to dilution effects is important in way to increase the metabolomics effect due to pathological condition of interest. In conclusion, this study allowed us to demonstrate that the normalization technique the most adapted and the most efficient is represented by normalization on Osmolarity which will be the technique used in our "real case study" (section 3.4).

3.4 Application of urine protocol: a real case study

3.4.1 Introduction

Urine sample preparation is an important aspect in metabolomics workflow due to its source of variability and its great interest for the number of detected metabolites. As discussed in the previous sections concerning urine protocol set-up, two fundamental steps are involved in urine preparation: (1) chemical shift variation and (2) normalization. These two steps have been widely explored and at the end of each section a precise protocol have been proposed as being the better fitting. In this section the identified methods respectively for urine preparation and data pre-treatment have been applied to a "real case study". This cohort (Metarein cohort for details see chapter 4) composed of urine coming from transplanted patients and suffering from CKD has been used to test the validity and the applicability of the previously established urine protocols.

In this particular section, apart from the more effective normalization protocol discussed earlier (normalization on Osmolarity), two additional methods were evaluated: normalization based on ***creatinine levels*** and normalization based on the ***dilution factor*** [245] **.** Concerning creatinine, represents nowadays, the normalization method of choice used in clinical practice; the dilution factor is a novel approach that seems to better fit the reality of urinary dilution in patients with kidney function impairments. The cohort utilized in this real study presented a significant challenge associated with urine dilution, which proved difficult to address using the methods tested previously. Indeed, while the normalization techniques examined in Section 3.3 were well-suited and effective in addressing a "light" and linear dilution effect observed among healthy individuals, they were not as effective when applied to a cohort of individuals with CKD.

3.4.2 Materials and methods

Urine collection and study cohort

Second-morning urine from 56 individuals were collected 3- and 12- month post-transplantation. The aliquots were supplemented with 2% Na azide solution in way to overcome bacterial contamination and stored at -80°C until analysis.

Sample preparation

For sample preparation 550µl of urine was supplemented with KF solution and prepared by following the protocol set-up in section 3.2. Once the sample was centrifugated and pH

measured, 100µl of maleic acid 5mM used as internal standard and 10µl of TMSP (10mg/ml) as chemical shift reference were added to the aliquot.

NMR-metabolomics

All samples were measured at 298K on a 700MHz Bruker Avance HD spectrometer operating at 700.17MHz for proton detection. The sequence used is a 1D NOESY sequence with presaturation for urine samples. The Noesypresat experiment used a RD-90°-T1-90°-Tm- 90°-acquire sequence with a relaxation delay of 4 s, a mixing time (Tm) of 10 ms and a fixed T1 delay of 4 µs. Water suppression pulse was placed during the relaxation delay (RD). The number of transients is 64 (64K data points) and a number of 4 dummy scans is chosen. Acquisition time is fixed to 3.2769001 s. The data were processed with Bruker Topspin 4.0.8 for baseline and phase correction and the entire δ scale was set to 0 ppm using the internal standard TMSP.

NMR data preprocessing

Data preprocessing on spectra was done by using MestReNova lab (v14.1.1). After phase and baseline correction, an alignment on global spectra was done. Peaks of maleic acid and water were deleted from the spectra. Bucketing process with a spectral width of 0.02ppm was performed.

Data normalization

Normalization on osmolarity was done in way reduce dilution effects as showed in section 3.3. In addition to this method, normalization on Creatinine Equilibrium, dilution factor[245] and PQN were performed. Creatinine equilibrium has been measured in a urinary sample collected 1 hour after the injection of ^{51}CrEDTA (which is the time needed for the marker to reach an equilibrium state). The dilution factor was calculated for each sample using ^{1}H-NMR spectra. This value specifically represents the ratio between the integral of all signals originating from the metabolites in the samples and the integral of the internal standard (TMSP) added to the urine aliquot during sample preparation.

$$\text{Dilution factor} = \frac{\text{Area of all metabolites}}{\text{Area of internal standard}}$$

Multivariate analysis

SIMCA® software was used to generate PCA score plot on pareto-scaled data in way to allows data-reduction and overview. PLS-DA were used as supervised model in way to look at the discrimination between the two groups of interest. In addition to these models, PLS regression was generated to correlate the whole metabolome with continuous clinical value.

3.4.3 Results and discussion

The analysis of this real case study cohort was conducted with the objective of addressing issues in data preprocessing that could impact statistical analysis and data interpretation. Indeed, since data preprocessing step represents a "manipulation" or raw data into a structured table from which all subsequent results are derived, it is of paramount importance to handle data accurately and effectively during this phase.

During the normalization process, it is crucial to ensure that the chosen normalization factor does not introduce internal bias or non-biological distinctions within our model. To address this concern, an initial Mann-Whitney U test was conducted to assess the significance of various clinical and experimental variables between our two groups of interest. Specifically, the two groups of interest are represented by "decrease" and "stable" whose stratification was based on kidney function evolution between 2 visits (for details on stratification criteria see Chapter 4); As illustrated in ***Figure 46***, no significant differences were observed in osmolarity, creatinine Eq, and dilution factor values between the two clusters. However, in contrast to these findings, the measurement of Creatinine peak in NMR spectra was significant to discriminate between group 1 and group 2 and for this reason was not considered as normalization factor. As a result, it was not included as a normalization factor.

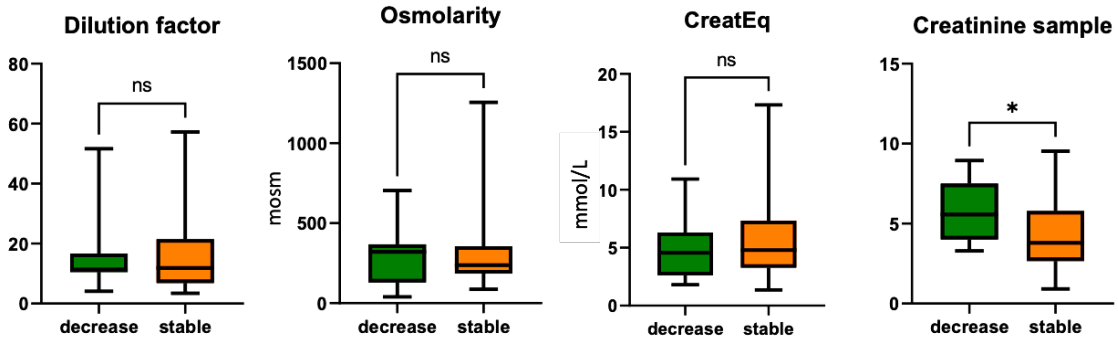
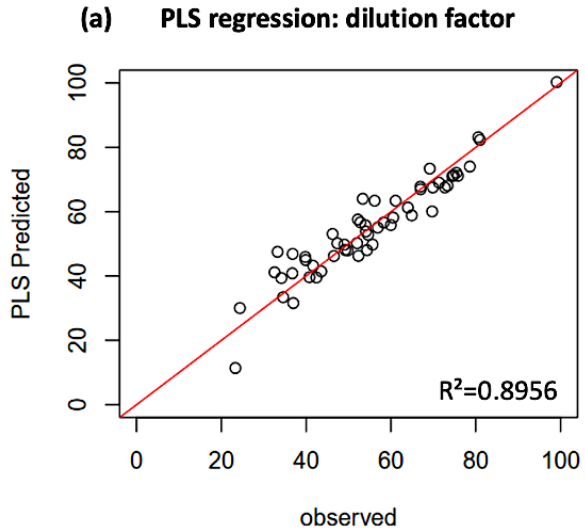


Figure 46 *Mann Whitney U test for the several experimental and clinical measure as we can see none of these tests were significant expect for Creatinine value measured through NMR (p-value <0.05)*

Once the presence of confounding factors was excluded, we proceeded to evaluate the suitability of these experimental and clinical values to normalize our dataset. Specifically, dilution factor, osmolarity and creatinine Eq in addition to PQN algorithm were employed as normalization factor.

As first, PLS model were generated to evaluate the impact of dataset normalization. PLS regression models work by evaluating the relationship between independent spectral values (metabolites) and non-spectral dependent values (such as BMI and age).

In the context of this cohort (urinary samples collected at 3 months; n=56), where sample stratification was carried out based a clinical value measuring kidney function (mGFR), a PLS regression model was generated to examine the relationship between this clinical measure and the complete metabolomic content. Particularly, when we normalized the dataset using various experimental or clinical factors, we could observe the impact of normalization strategies on linear regression ***Figure 47a.*** As indicated in ***Figure 47b,*** the highest R^2 values were achieved when employing PQN and the dilution factor as normalization methods (see FigureS1 for PLS regression plots of CreatEq, Dilution factor and Osmolarity) . Following this approach, the most representative normalization techniques for our dataset were, in order of significance, PQN, dilution factor, osmolarity, and lastly, creatinine Equilibrium.



(b)

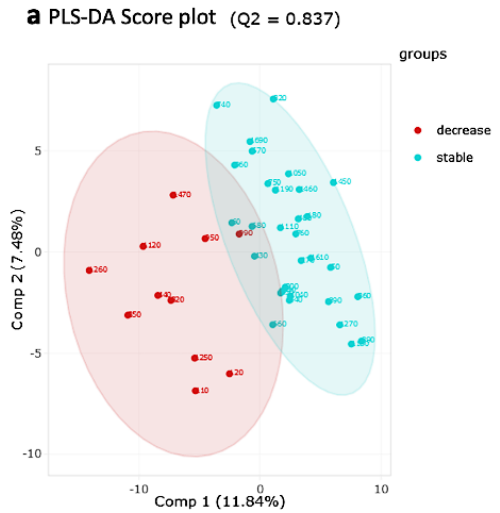
Normalization method	R^2 value
CreatEq	0.6163
Dilution factor	0.8956
Osmolarity	0.6786
PQN	0.9207

Figure 47 (a) *Example of PLS regression model;* ***(b)*** *table summarizing the value of R2 for each normalization method tested.*

By examining the outcomes of this analysis concerning normalization strategies and juxtaposing them with those attained in section 3.3, we observe a decrease in the effectiveness of osmolarity as a normalization factor. Indeed, our preliminary investigation in section 3.3, involved an ideal sample cohort wherein urinary biospecimens underwent osmolarity

measurement and NMR analysis soon after sample collection. In contrast to this, the sample cohort used in this specific real case study had undergone multiple freeze-thaw cycles before osmolarity measurement, potentially altering the osmolarity values between sample collection and analysis.

To support this hypothesis and in way to find out the best normalization strategy, a parallel analysis was conducted employing PLS-DA model ***(Figure 48a).*** In this instance, samples were stratified into two classes representing contrasting pathological conditions, and the Q^2 value was computed for each normalization technique to assess the model discrimination's power. ***Figure48b*** presents a concise summary of the Q^2 values obtained for each normalization technique, alongside the associated results of permutation tests.



b

Normalization method	Q^2 value	Permutation test
CreatEq	0.542	0.4618
Osmolarity	0.704	0.1041
Dilution factor	0.774	0.0361
PQN	0.837	0.0631

Figure 48 (a) *Example of PLS-DA model by using PQN as normalization method;* ***(b)*** *table summarizing Q2 value and p-value permutation test for each normalization method tested.*

These findings underscore the superiority of the dilution factor and PQN normalization methods over other approaches. More specifically, while the PQN model exhibits a higher Q^2 value, the permutation test's p-value for the dilution factor indicates a lower score by implying a reduced probability of model's overfitting.

3.4.4 Conclusion

This study has highlighted the critical role played by data pre-processing in the context of statistical analysis and data interpretation. It is evident that sample concentration is a significant source of variability in urine, which becomes even more crucial when dealing with kidney impairments. In our research, we aimed to enhance patient stratification by assessing the impact of various normalization techniques on the dataset, as reported in

the literature. This comparative analysis enabled us to showcase the influence of this pre-processing step on statistical analysis and, consequently, the interpretation of results. Particularly, employing the KF method for urine preparation along with normalization on dilution factor enabled us to establish the most effective urine protocol for NMR-based metabolomics studies. All the analyses conducted on this dataset serve as a starting point for the investigations detailed in Chapter 5, as the cohort employed in this study forms the core of the Metarein project. In conclusion, this work underscores the importance of selecting an appropriate pre-processing method fitting to the specific characteristics of the dataset under examination.

Chapter 4

Metarein project

4 Metarein project

4.1 Background- Metarein cohort

In collaboration with Pr. P. Delanaye and Dr. G.Resimont from CHU of Liège, and Pr. E.Vidal-Petiot and Pr. M. Flamant from Bichat-Claude-Bernard Hospital where sample collection was done, urine biospecimens from 56 KTRs were prospectively collected during their protocol visits at 3 and 12 months post-KTx (cohort is called "MetaRein") an classified on the basis of renal function decline between the two time points (***Figure 49***). Clinical and biological measurements were collected at each visit (such as protein, albumin, Na+, K+, creatinine etc.) in addition to mGFR measurement via urinary and plasmatic clearance of ^{51}CrEDTA which represent the gold standard for the renal function measurement.

All samples were transferred to CHU of Liege for NMR-based metabolomics analysis. Before this, numerous tests were conducted to establish the correct protocol for urine analysis (as detailed in Chapter 3). The goal was to identify sample preparation and normalization techniques that were better suited to this particular cohort. In addition to this, an aliquot of each urine sample was sent at "Laboratoire d'Etude des Residus et Contaminants dans les Aliments ("LABERCA"), in Nantes for MS-based metabolomics analysis.

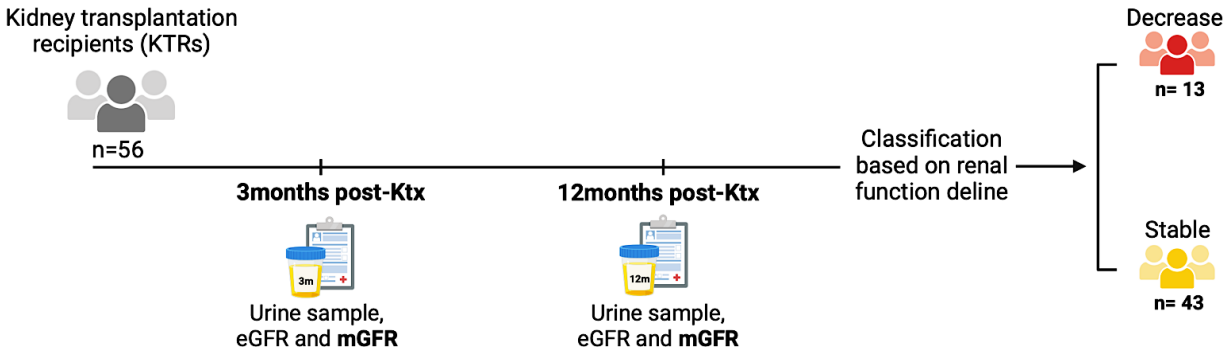


Figure 49 *Overview of the Metarein cohort: Samples were collected from 56 transplanted recipients (KTRs), for whom urinary samples and clinical data (mGFR and eGFR) were obtained at two time points (3 months and 12 months post-transplantation). Stratification of samples was based on the decline in kidney function between these two time points, classifying patients into either the "decrease" or "stable" class.*

An untargeted metabolomics approach was applied to explore the correlation between the whole metabolome and kidney function evolution.

This study, which seeks to identify a set of predictive biomarkers for the decline of kidney function in the early post-transplantation period, is currently undergoing the submission process. The cohort, the experimental workflow, and a concise summary of the results of this

study, titled " ***How metabolomics can help in the follow-up of kidney transplantation recipients: An untargeted metabolomics-based multiplatform study.***" are encapsulated in the graphical abstract presented in ***Figure 50***.

4.1.1 Metarein cohort and measured glomerular filtration rate

The major strength of this cohort is based on the measurement of GFR , which was conducted for each patient at both 3- and 12- months post-transplantation time-points. These data enhanced the phenotyping of this cohort, setting it apart from most studies in the existing literature, where only the estimation of GFR is typically available. Indeed, mGFR value now stands as the gold standard for assessing renal function in clinical practice by enabling a more accurate and precise characterization of kidney status compared to its estimated value. Consequently, the addition of mGFR to the clinical data usually available, provided a unique opportunity to establish a direct link between the metabolome of transplanted patients and the precise status of their renal function. Furthermore, the inclusion of GFR measurements for these patients at both time points significantly contributed to the diagnostic assessment of renal function. Specifically, in our study, this aspect not only represented a significant advantage in terms of patient characterization but also played a fundamental role in the development of a predictive model for the decline of renal function.

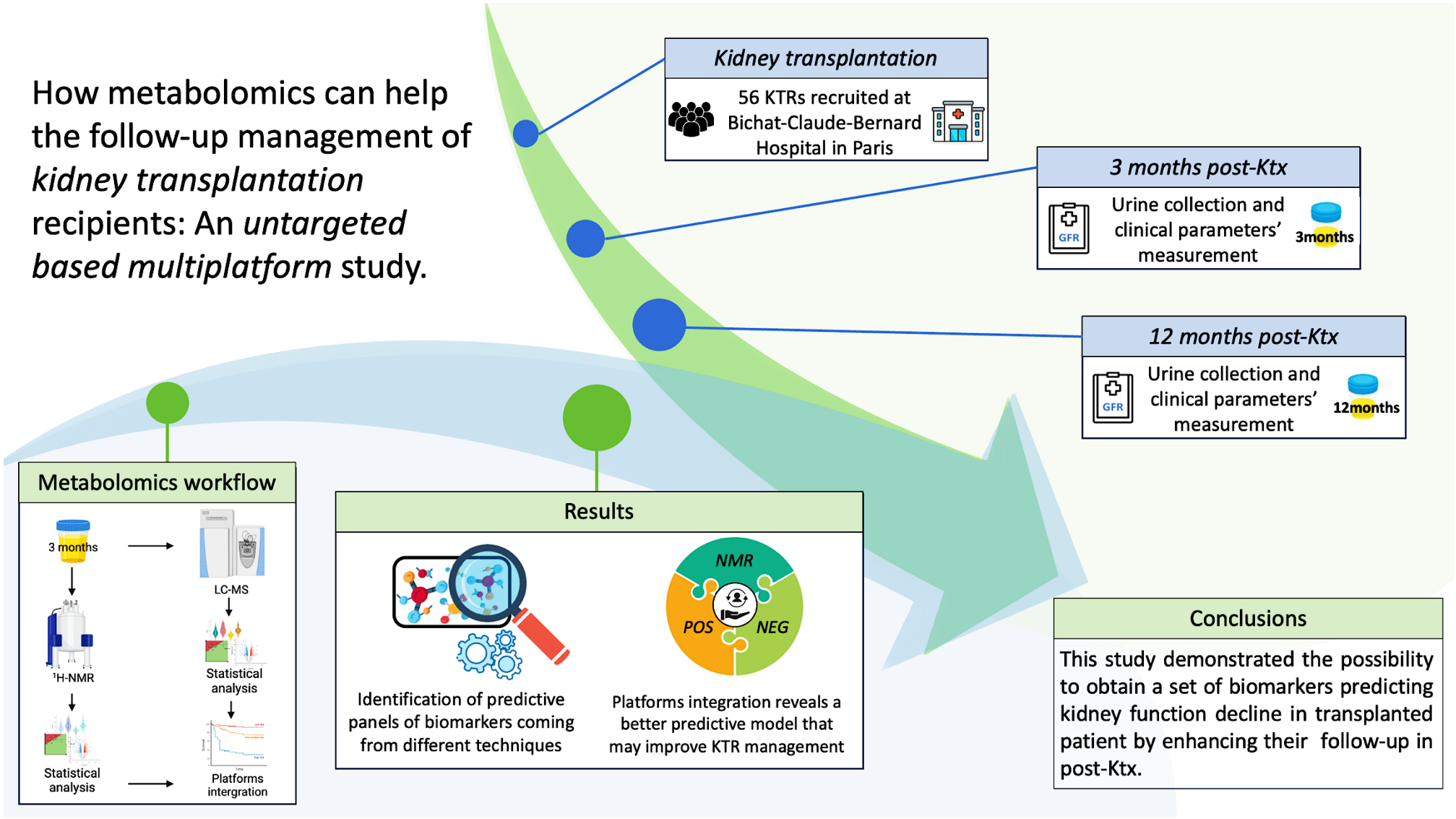


Figure 50 *Graphical abstract for "How metabolomics can help the follow-up management of kidney transplantation recipients? An untargeted based-metabolomics multiplatform study."*

4.2 How metabolomics can help in the follow-up of kidney transplantation recipients: An untargeted based multiplatform study.

Cirillo A.[1], Resimont G.[2], Massias J.[3], Jouret F.[2,4], Vidal-Petiot E.[5], Flamant M.[5], Guitton Y.[3], Delanaye P.[2], de Tullio P.[1]

1. Clinical Metabolomics Group, Center for Interdisciplinary Research on Medicines (CIRM), University of Liège, Liège, Belgium.
2. Department of Nephrology-Dialysis-Transplantation, University of Liege, CHU Sart-Tilman, Liege, Belgium
3. Laboratoire d'étude des Résidus et Contaminants dans les Aliments (LABERCA) , Oniris, INRAE, Nantes, France
4. Groupe Interdisciplinaire de Génoprotéomique Appliquée (GIGA), Cardiovascular Sciences, University of Liège, Liège, Belgium.
5. Assistance Publique-Hôpitaux de Paris, Renal Physiology Unit, Bichat Hospital Paris, France

4.2.1 Abstract

Introduction: In the context of chronic kidney disease, K represents the most favorable solution in terms of quality of life, morbidity and mortality for patients affected by ESRD. However, KTx is not devoid of risks and the follow-up of kidney graft function is crucial in the management of kidney transplantation recipients (KTRs). Currently used techniques for monitoring kidney function remains rather imprecise in the case of its estimation or time-consuming and rarely performed in case of its measurement. Because of these limits the need of new biomarkers able to precisely reflect the renal function or even predict its evolution in KTRs is a key challenge to improve patients‘ management.

Objective: This study aims to identify a new panel of biomarkers able to predict kidney function post KTx.

Methods: 56 Patients from a well-phenotyped French KTR cohort has been followed for 1-year *post* transplantation. Urinary samples have been collected at 3 and 12 month and the patients have been stratified as “decrease” or “stable” based on the decline or the stability of their kidney function after 1 year (decrease= 13; stable= 43). Untargeted NMR- and MS-based metabolomic approaches were applied to the cohort followed by integration of results coming from this dual-method.

Results: Chemometrics analyses deriving by both techniques allowed the identification of panels of biomarkers that could be linked to a GFR evolution and used as predictive markers. Merged results from the dual approach, enabled an increase of discrimination and predictive

performances delineating a metabolomic signature that may forecast kidney function decline at 12-month post-KTx though analysis of urinary sample at 3 months.

Conclusions: Analysis of urinary metabolome of KTR patients in early post- transplantation shows the possibility to predict GFR evolution at 1 year. These findings may represent an innovative and helpful tool for clinicians to enhance patient care in post-transplantation period.

4.2.2 Introduction

Chronic kidney disease is a common disease characterized by a chronic alteration and progressive decline in kidneys' function that affects 5 to 10% of the general population [206]. By definition, CKD affects patients at different degrees and when its most severe form called "end stage renal disease" (ESRD) occurs dialysis or kidney transplantation (KTx) are mandatory [241]. Nowadays, KTx represents the best treatment for patients with ESRD and is currently regarded as the most favorable patient's care in terms of quality of life, morbidity and mortality [247].

However, KTx is not devoid of risks and its main issues are linked to allograft dysfunction that led to altered kidney function and/or to graft rejection. Indeed, 30-40% of the patients lose their graft less than 10 years after KTx[248,249]. It remains clear and evident that the long-term success of a transplantation highly depends on the quality of the graft at the time of transplantation and of a continuous follow-up of the kidney transplant recipient.

In this context, the monitoring of kidney function in the kidney transplantation recipients (KTRs) is mandatory and represents a crucial issue during the follow-up phase. Measured GFR represents the gold-standard by being the best and most accurate method to determinate GFR[242,243].For practical reasons, the use of equations estimating the GFR, mainly based on blood and patient's data, is nowadays much more current in clinical practice[211]. However, the limited precision of these formulas in accurately assessing kidney function underscores the current need to improve the prediction of kidney dysfunction and enhance the management of KTR follow-up. In this sense, the discovery of novel biomarkers predicting for early post-KTx kidney function, and its decline might play a significant role in helping clinicians in patients care and follow-up.

In this context, metabolomics could represent a cornerstone. Indeed, metabolomics, as part of the omics sciences allows to have a snapshot of an organism status at a defined time point [244,245]. Metabolomics is nowadays increasingly used for detection of disease' s profile such as renal dysfunction. Several metabolomics studies have already highlighted the link between

biomarkers and CKD [246,247,224,248,249,250]. Nuclear magnetic resonance (NMR) and mass spectrometry (MS) stand as the two mains and complementary analytical platforms utilized in metabolomics studiess.[251]

We therefore planned to use the multi-platform NMR-MS metabolomic approach to explore whether the metabolome and certain metabolites could enable us to predict the evolution of renal function in transplant patients. To do this, we will be able to take advantage of a cohort and urine samples from more than 50 KTR whose renal function was measured 3 months and 12 months post-transplant. Specifically, by taking advantage of the longitudinal nature of this cohort, urine of KTR patients collected 3 months post-transplantation were analyzed and correlate with the evolution kidney function between 3-and 12-months post-graft. The major strength of this cohort is based on the measurement of GFR, which represent the gold standard in renal function measurement. These data enhanced the phenotyping of this cohort, setting it apart from most studies in the existing literature, where only the estimation of GFR is typically available. Consequently, the addition of mGFR to the clinical data usually available, provided a unique opportunity to establish a direct link between the metabolome of transplanted patients and the precise status of their renal function. Furthermore, the inclusion of GFR measurements for these patients at both time points significantly contributed to the diagnostic assessment of renal function. Specifically, in our study, this aspect not only represented a significant advantage in terms of patient characterization but also played a fundamental role in the development of a predictive model for the decline of renal function.

4.2.2 Methods

Cohort and clinical data

Second morning urine from 56 kidney transplanted patients have been collected at two time points corresponding to 3- and 12-months post-transplantation at Bichat-Claude-Bernard Hospital. At the time of sample collection, mGFR (deindexed=non-indexed by BSA) via urinary clearance of ^{51}CrEDTA was performed in addition to conventional blood and urine assessments (proteins, albumin, Na+, K+, creatinine and Cystatin-C standard measurements). eGFR values (ml/min) were obtained by using MDRD equation (non-indexed by BSA). MDRD « de-indexed » recommended by the KDIGO were computed by multiplying eGFR by each individual's body surface area, using actual body weight, and by dividing this intermediate result by 1.73 m^2.

- MDRD $_{deindexed}$ in mL/min = (eGFR in mL/min/1.73m^2 x BSA) / 1.73m^2

No specific criteria of inclusion/exclusion were applied for patients' recruitment.

Urine sample collection

Urinary samples, corresponding to the second morning urine in fasting conditions was systematically collected and frozen without additives at −80° C.

Patients' stratification

Fifty-six patients transplanted in Paris and whose GFR was measured at 3 and 12 months at Bichat-Claude-Bernard Hospital were included. Samples and clinical measures collected at two consecutive time points enable the stratification of patients based on measured GFR (mGFR) evolution. Particularly, the value considered of mGFR was obtained by measuring the urinary clearance of ^{51}CrEDTA, which is considered a reference method. We used the absolute values of mGFR (non-indexed for body surface area) and calculated the variation of GFR between 3 and 12 months (expressed in %) ***(Figure 51).***

$$\% \text{ renal function} = \frac{\Delta \text{ GFR}}{\text{3month GFR}} * 100$$

Decrease= renal function is between **-7%** and higher negative values
Stable= renal function is between -6.99% and positive values

Figure 51 *Equation used to calculate renal function % at 12 months based on which samples at 3 months (analyzed and used in this study) were stratified.*

Knowing the relative change of mGFR (ΔGFR), we classified the patients as "decrease" or "Stable". "Decrease" were patients whose ΔGFR was declining of more than 7% and other were considered as "stable". The use of 7% as threshold in our stratification method is based on the concept of critical difference obtained at Bichat-Claude-Bernard Hospital were samples and clinical data were collected. Specifically, the critical difference can be defined as the smallest change in results of creatinine measurement which is not due to chance[264].

Sample preparation.

NMR

Aliquots of 500µl of urine samples collected at 3-month were thawed on ice, supplemented with KF (see protocol 3.2.2) and prepared by following an optimized procedure for reduce inter-sample chemical-shift variations [232]. After this procedure urine samples were supplemented with 200µl of deuterated phosphate buffer (DPB, pH 7.4), 100µl of a 5mM solution of maleic acid and 10µl of a 10 mg/ml TMSP D_2O solution for NMR analysis. Urine pH and osmolality were measured before sample preparation and analysis.

LC-MS

An aliquot of 500µl of urine sample collected at 3-month was placed in 10kDa centrifugal filter and centrifuged at 13000g during 30min at 5°C. Once filtered, internal deuterated standards (leucine-5,5,5-d3, L-tryptophan-2,3,3-*d*3, indole-2,4,5,6,7-*d*5-3-acetique acid et 1,14 tetradecanedoic-*d*24 acid) were added to each sample and nitrogen blowdown was done. For each sample pH and osmolality were measured.

Sample measurements.

Metabolomics analyses were performed on urine samples collected at three months in all patients.

NMR

All samples were recorded at 298 K on a Bruker Avance HD spectrometer operating at 700.17 MHz for the proton signal acquisition. The instrument was equipped with a TCI 5-mm cryoprobe with a Z-gradient. Maleic acid was used as the internal standard for quantification and trimethylsilyl-3-propionic acid-*d*4 (TMSP) for the zero for the zero calibration. ^{1}H-NMR spectra were acquired using a 1D NOESY sequence with presaturation. The Noesypresat experiment used a RD-90°-T1-90°-Tm-90°-acquire sequence with a relaxation delay of 4 s, a mixing time (Tm) of 10 ms and a fixed T1 delay of 4 µs. Water suppression pulse was placed

during the relaxation delay (RD). The number of transients is 64 (64K data points). The data were processed with the Bruker Topspin 4.0.8 software with a standard parameter set. Phase and baseline corrections were performed manually over the entire range of the spectra and the δ scale was calibrated to 0 ppm using the internal standard TMSP.

LC-MS

All samples were analyzed on ultrahigh performance liquid chromatography with high-resolution mass spectrometry (UHLPC/MS) by following described method[253] for reversed phase(RP) UHPLC/MS. Tacking advantage of the MS2 capacities of the hybrid quadrupole-orbitrap (Q-Exactive TM) mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) QC samples (i.e. pooled samples) were analyzed, in ESI positive and ESI negative modes, with three cycles of iterative Data Dependent MS^2. The acquisition of the raw data was performed using a full scan mode within the m/z 65-1000 range at a resolving power of 70,000 at m/z 200,000 Da. Compound separation was performed using a Hypersil GOLD-C18 column (1.9 µm, 100 mm x 2.1 mm) from Thermo-Scientific (USA). The column temperature was set at 35°C. The mobile phases were composed of 0.1% of acetic acid in water (solvent A) and in acetonitrile (solvent B). Acetonitrile LC/HRMS grade CHROMASOLV™ LC-MS (Riedel-de Haën), Water LC/HRMS grade CHROMANORM® (VWR Chemicals). The applied gradient (A:B, v/v) was as follows: 95:5 from 0 to 2.4 min, 75:25 at 4.5 min, 25:75 at 11 min, 0:100 from 14 at 16.5 min and 95:5 from 19 to 25 min. The flow rate was set to 0.40 mL/min. The injection volume was 5 µL. All samples were analyzed in one batch without any stopping or recalibration step. The quality control sample (QC) was injected regularly throughout the run after every ten samples approximately. Data acquisition was settled with an automatic gain control of 5.105 and a C-Trap inject time of 20 ms. The acquisition spectrometric parameters were as follows: the spray voltage (+3 kV), the S-Lens RF level (50), the tube lens voltage (+100 V), the capillary temperature (350°C), the heater temperature (300°C), the sheath gas pressure (55 arbitrary units), the auxiliary gas flow rate (10 arbitrary units) and the sweep gas flow rate (0 arbitrary units). Full instrument calibration was performed using a MSCAL5 ProteoMassT LTQ/FT-Hybrid ESI Pos/Neg. In addition, Xcalibur V2.2 (Thermo Scientific®, Bremen, Germany) software was used for the generation of all chromatographic peaks acquired in full scan mode.

Data pre-treatment

NMR

MestReNova (v14.1.1) was used for NMR data pretreatment. Non informative zones were removed from NMR spectra such as water region (4.7 to 5 ppm) and maleic acid (5.6 to 6.2 ppm). Alignment step was done in way to reduce the residual chemical shift effect due to inter-sample pH variations. Spectra were then reduced to integrated regions of equal width (0.02ppm), named "bins", corresponding to the 0.5 to 9.0 ppm region.

LC-MS

Data preprocessing was performed by using workflow4Metabolomics.org (W4M) platform on Galaxy environment. The raw data were at first transformed into a data matrix containing all the peaks present in the samples; centWave algorithm was used for peak detection and a "peak grouping" step was done in way to align the peaks. At this point, undetected ions were integrated with according to m/z and RT through "peak filling" step. Final data matrix table was composed of variables that were repeatable in at least 50% of the samples. Batch correction was done by using Metaboanalyst R package on R environment by using EigenMS as algorithm. At this point, samples were normalized by using dilution factor value calculated through NMR technique and log transformed.

Metabolites identification

NMR

For NMR platform metabolites identification was done using Chenomx profiler 9.0 (Chenomx Inc., Edmonton, AB, Canada), the free web-based tool HMDB (https://hmdb.ca) and tables; 2D NMR (COSY and HSQC) measurements were also done to confirm the identification of metabolites.

LC-MS

For LC-MS platform metabolites identification MS and MS^2 data were used. In MS data isotopologue and adduct were searched by using CAMERA[140] annotation package on W4M. MS^2 data were generated from pool samples with iterative data dependent MS^2 acquisition (iDDA) and processed through msPurity package[254] included in W4M. All features of interest were searched in MS^2 files with 0.0005 filter for m/z and +/-5s for rt. At this point MS2 spectra was compared to external databases for spectra matching(MassBank

https://massbank.eu/MassBank/, HMDB https://hmdb.ca/, GNPS http://gnps.ucsd.edu). Through this process it was possible to reach an annotation level 3 on the Schymanski scale[255].

Data normalization

Data matrices obtained from both NMR and LC-MS platforms underwent normalization based on dilution factor values. Dilution factors for individual samples were determined using the NMR technique. For each ^{1}H-NMR spectrum, the total integral of peaks corresponding to metabolites within the range of 0.5-9.0 ppm was computed. Additionally, the signal of the internal standard (TMSP) at 0 ppm was integrated, and its integral was set to 0. The dilution factor value was computed by calculating the ratio between the integral of all metabolites and the integral of the internal standard (for further details, refer to section 3.4.2).

Multivariate and statistical analysis and data integration

For both NMR and UHPLC-MS, statistical analysis was conduct by using SIMCA-P software (v17.0; Umetrics, Malmö, Sweden), BioStatFlow webtool (biostatflow.org) and GraphPad Prism version 9.4.1 (GraphPad Software, La Jolla, CA,). Samples included in NMR, MS and data integration analysis are resumed in ***Figure 52.***

Principal component analysis (PCA) was used to explore samples without any classification knowledge and spot any separation trend, groups, or outliers. PCA score plot was also used to identify "strong outliers" represented by samples placed outside the 0.95 Hotelling's T2 ellipse; DmodX was used for the detection of samples exceeding the 0.05 cutoff value defined as critical distance of significance. Orthogonal signal correction (OSC) was applied to discriminant model in way to remove the inter-subject variability and to describe maximum separation based on class; its quality was evaluated by the predictability calculated based on the fraction correctly predicted in one-seventh cross-validation (Q^2) by considering model with $Q^2 > 0.5$ as "good" and $Q^2 > 0.9$ as "excellent". Permutation tests were performed to validate models [121]. For both NMR and MS platforms, OSC-PLS models and their loading plot were used to identify relevant metabolites. Variable important projection (VIP) higher than 1 was considered as significant and considered for univariate statistical analysis. Wilcox-Mann–Whitney U test was performed for comparisons between "decrease" and "stable" groups.

The selected features were worn to generate receiver operating characteristic (ROC) curves with the aim to evaluate the performance of biomarker models created through automated feature selection; PLS-DA was used as classification algorithm and univariate T-test as feature ranking method with 2 latent variables. Confusion matrix was used as classifier model for

binary response; this tool allows to visually evaluate the performance of classification through the measurement of accuracy, precision and recall values. The detailed analysis of the metabolic pathways was performed by Metaboanalyst (www.metaboanalyst.ca) using the metabolomic set enrichment analysis (MSEA) tool by using the high-quality SMPDB metabolic pathways as the backend knowledgebase. Over representation analysis (ORA) on list of metabolites was implemented using the hypergeometric test to evaluate whether a particular metabolite set is represented more than expected by chance within the given compound list. One-tailed p values are provided after adjusting for multiple testing.

Data integration of NMR and Orbitrap-MS data was investigated through a high-level approach leading to merge results coming from the different blocks [129,128].

After blocks' integration, chemometric analysis was done by performing OSC-PLS model, multivariate ROC curve and confusion matrix as already done for the single blocks.

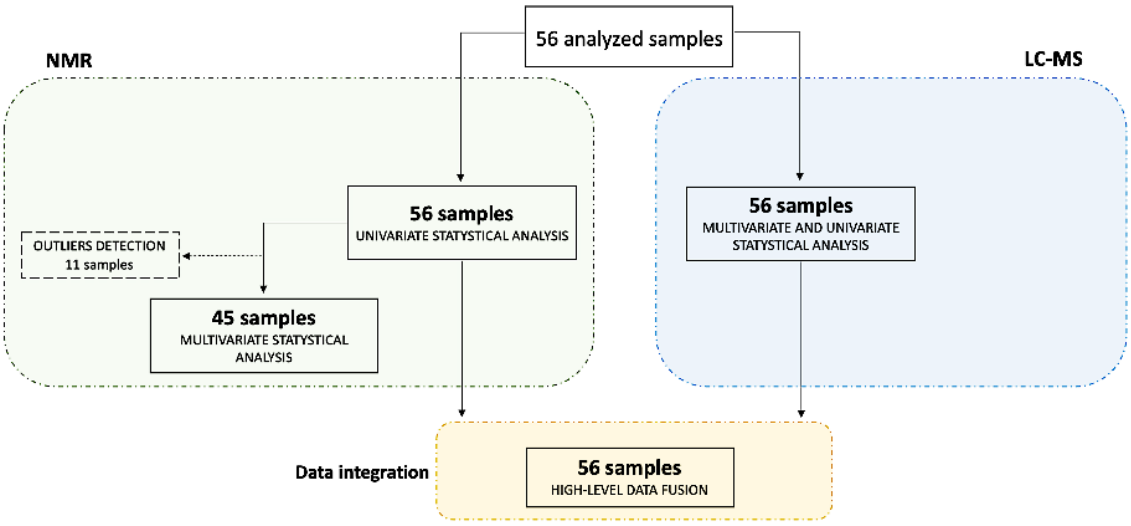


Figure 52 *Scheme of samples involved in NMR, LC-MS and integrated models. For NMR analysis only 45 samples were included in chemometric analysis due to the presence of 11 detected outliers. These 11 samples were excluded from multivariate data analysis due to the presence of exogenous molecules hiding spectral zones of interest. The presence of these signals hampers the building of performant chemometrics model aiming a correct feature selection. However, after feature selection these samples were again included in univariate and integrated predictive models.*

4.2.3 Results

Study population

A preliminary statistical analysis was conducted on the cohort considering the classes "decrease" and "stable" which showed no confounding factors. Patients' characteristics categorized by the type of renal function evolution are presented in ***Table 13.***

	Total (N=56)	Decrease (N=13)	Stable (N=43)	P-value
Age				0.038
Mean (SD)	50.661 (13.576)	43.846 (12.307)	52.721 (13.392)	
Range	19.000 - 77.000	29.000 - 75.000	19.000 - 77.000	
Weight (kg)				0.98
Mean (SD)	74.780 (13.474)	74.862 (13.129)	74.756 (13.729)	
Range	50.000 - 108.000	56.800 - 100.000	50.000 - 108.000	
Height (cm)				0.297
Mean (SD)	170.045 (10.564)	167.346 (10.078)	170.860 (10.686)	
Range	149.000 - 193.000	154.000 - 183.000	149.000 - 193.000	
BSA				0.778
Mean (SD)	1.874 (0.209)	1.859 (0.202)	1.878 (0.213)	
Range	1.540 - 2.310	1.630 - 2.240	1.540 - 2.310	
Ethnicity				0.349
Black	12 (21.4%)	4 (30.8%)	8 (18.6%)	
White	44 (78.6%)	9 (69.2%)	35 (81.4%)	
Sex				0.119
Female	20 (35.7%)	7 (53.8%)	13 (30.2%)	
Male	36 (64.3%)	6 (46.2%)	30 (69.8%)	
CreatPJaffe				0.367
Mean (SD)	138.750 (40.887)	129.692 (36.049)	141.488 (42.247)	
Range	69.000 - 250.000	76.000 - 222.000	69.000 - 250.000	
AlbuU0 (g/L)				0.565
Mean (SD)	85.982 (114.006)	69.846 (95.373)	90.860 (119.651)	
Range	5.000 - 594.000	12.000 - 305.000	5.000 - 594.000	
mGFR 3m (mL/min)				0.329
Mean (SD)	55.491 (15.871)	59.292 (19.851)	54.342 (14.541)	
Range	23.300 - 99.100	33.200 - 99.100	23.300 - 81.000	
mGFR 12m (mL/min)				0.045
Mean (SD)	55.312 (16.237)	47.438 (15.967)	57.693 (15.730)	
Range	24.600 - 89.100	25.600 - 75.100	24.600 - 89.100	

Table 13 (Below) *Descriptive statistics on clinical data of Metarein cohort for the 56 patients included in the analysis; for each clinical mean (standard deviation) and range are reported for the sample collected at 3months visit except for* **$mGFR_{EDTAPdeindex}$ (12m) which was measured as indicated in parenthesis at 12 months.** *(clinical measures included:* ***weigh*** *(in kg);* ***height*** *(in cm) ;****BSA****= body surface area;* ***age*** *(in years);* ***ethnicity****;* ***sex****;* ***CreatPjaffe****= plasmatic creatinine measured though Jaffe assay;* ***Albu U0****= albuminuria measured in second morning urine;* **$mGFR_{EDTAPdeindex}$=** *GFR measured though plasmatic clearance of ^{51}Cr-EDTAand deindexed*

Univariate statistics and chemometric analyses

^{1}H-NMR

In way to detect outliers PCA-X, DmodX and Hoteling T2 models were used on the cohort of fifty-six patients, and eleven samples were excluded from chemometric analysis. The exclusion of these eleven samples was either due to the presence of abnormal signals in the spectral data or because they fell outside the 95% confidence ellipse in the outliers detection models (DmodX and Hotelling T2). A total of forty-five samples were thus included in the multivariate models (decrease=12; stable=33). A discriminant chemometric analysis was performed by generating OSC-PLS model (Q^2=0.775) ***(Figure 53a)***; a permutation test (p-value= 0.0361) was performed by validating the obtained model. A loading plot and a list of VIPs score were generated in way to select relevant features related to the model. Univariate statistical analysis through unpaired Mann-Whitney test was performed on normalized bins by including in the analysis outliers excluded in multivariate data analysis (n=56). A total of 14 metabolites were assessed as significant for separation between decrease and stable groups. However, of these 14 features only 12 were identified and included in the following models (***Table14***).

^{1}H-NMR metabolite	p-value	Variation decrease vs stable
Choline	0.0426	↖
Carnitine	0.0337	↘
Glycine	0.0117	↖
TMAO	0.0393	↘
Dimethylamine	0.0367	↖
N-phenylacetylglycine	0.0332	↘
Serine	0.0001	↖
3-methylhistidine	0.0081	↖
Creatinine	0.0386	↖
2-aminoadipate	0.0285	↖
Mannitol	0.0002	↘
Hippurate	0.0157	↘

Table 14 *List of significant features identified though 1H-NMR platform.*

The selected metabolites were used to perform a multivariate receiver operator characteristics (ROC) curve with increasing number of variables in way to evaluate the performance of our model through automated features selection ***(Figure 53b).*** The most performant model was reached by using 10 metabolites with an AUC value of 0.794 (95% CI 0.65-0.929) and a predictive accuracy of 67.8% . Confusion matrix was used to evaluate the performance of classification model; by using the same set of metabolites employed in multivariate ROC curve,

the confusion matrix ***(Figure 53c***) showed 3 'decrease' patients and 17 'stable' patients being misclassified (accuracy= 0.64; precision= 0.37; recall= 0.77).

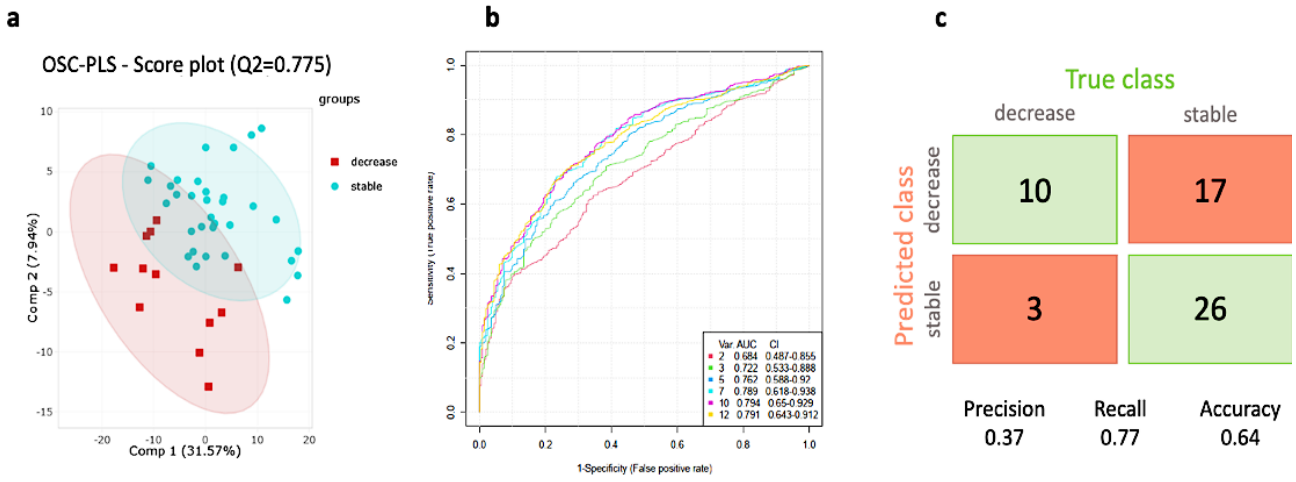


Figure 53 *Chemometric analysis of NMR platform; (a) OSC-PLS score plot of NMR data on decrease and stable groups showed a good performance (Q^2=0.775; p-value=0.03617); (b) multivariate ROC curve based on the 12 features; (c) confusion matrix based on the 12 NMR features with accuracy, precision and recall metrics reported.*

Orbitrap-MS

Positive mode

Fifty-six samples were included in univariate statistics and chemometrics models (n=56; 13 decrease; 43 stable). According to OSC-PLS discriminant analysis, a good discrimination power between decrease and stable groups was highlighted (Q2=0.933), however its overfitting was evident from the results of permutation test (p-value = 0.03073) ***(Figure 54a).*** A VIP score list composed by 753 **(S.I. *TableS2***) features was obtained from the discriminant model by using VIP >1 as threshold. A Wilcoxon-Mann-Whitney test was conduct on VIP by underlining a list of 145 significant features. After annotation only 36 features were associated to known compounds and of these only 17 were not drugs or redundant metabolites (***see S.I. TableS4***). For the following analysis only these 17 selected features (***Table 15***) corresponding to identified endogenous metabolites have been used. A multivariate ROC curve was generated (7 variables; AUC= 0.657; 95% CI 0.351-0.855; predictive accuracy= 55.3%) through automated variable selection ***(Figure 54b)***. In the confusion matrix ***(Figure 54c)***, 17 variables were used, yielding an accuracy of 0.68, with 3 "decrease" and 15 "stable" patients misclassified (precision=0.4; recall=0.77).

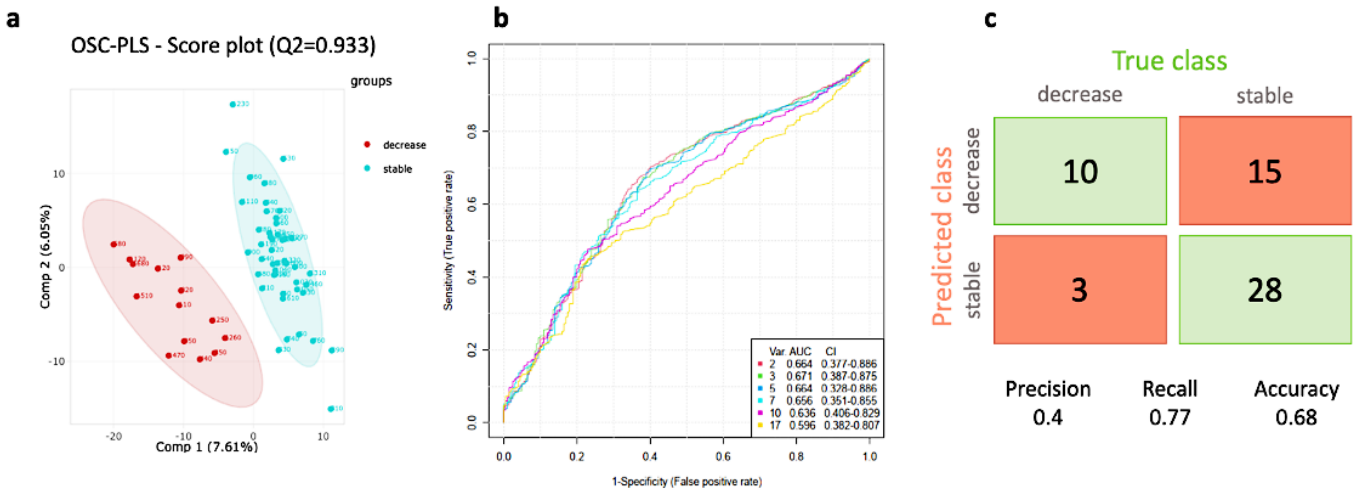


Figure 54 *Multivariate statistical analysis of positive MS platform;* ***(a)*** *OSC-PLS score plot of positive MS data on decrease and stable groups showed a good performance (Q^2=0.933; p-value=0.03073);****(b)*** *multivariate ROC curve based on the 17 known annotated metabolites;* ***(c)*** *confusion matrix based on the positive MS features with accuracy, precision and recall metrics reported.*

Positive mode features	p-value	Variation decrease versus stable
Phenylalanine	0.0027572	↘
Tyrosine (T57)	0.0033591	↘
L-Tryptophan	0.0037193	↘
Kynurenine	0.0072574	↘
Dethiobiotin	0.0075228	↘
Indole	0.01319	↘
Methoxyindoleacetic acid	0.025048	↘
3-ureidopropionate	0.02628	↘
L-3,4-Dihydroxyphenylalanine methyl ester	0.029779	↘
Mannose	0.033364	↘
L-Valine	0.03476	↘
Hippurate	0.035559	↘
5-Hydroxyindole-3-acetic acid	0.037831	↘
2-Hydroxycinnamic acid	0.037939	↘
Aspartame	0.039831	↘
Theophyline	0.040386	↘
Methionine	0.04388	↘

Table 15 *Significant features annotated for LC-MS analysis in positive mode. These 17 metabolites were the only non-redundant and endogenous metabolites identified in databases.*

Negative mode

Fifty-six samples were included in univariate statistics and chemometric analysis (n=56; 13 decrease; 43 stable). A supervised OSC-PLS model ***(Figure 55b)*** was run followed by a permutation test (Q2=0.719; p-value= 0.5381). As done for the positive mode, a list of 537 VIP (VIP > 1) was obtained (**S.I. *TableS3***). Moreover, a univariate t test using Wilcoxon-Mann-Whitney model was performed on peak intensities. A list of 40 features was found to be significant for discrimination between stable and decrease groups but a major part of them were not identified or corresponding to non-endogenous compounds (***see S.I. table S5***) . Only 4 metabolites were unique features corresponding to biological compounds present in human body (***table 16***).

Negative mode features	p-value	Variation decrease versus stable
L-Glutamic acid	0.0077804	↘
Citrate	0.040444	↘
N-acetyl-D-tryptophan	0.0154	↖
Uridine 5'-monophosphate	0.022438	↘

Table 16 *In the table the significant features annotated through LC-MS analysis in negative mode are showed. Starting from the first 40 features detected as significant in Wilcoxon-Mann-Whitney t test only these 4 features found a corresponding non-exogenous metabolite in databases.*

The significant features were used to generate a multivariate ROC curve that show poor model's performance (4 features; AUC= 0.752; 95% CI 0.596-0.91; predictive accuracy= 67.7%) compared to the previous showed results ***(Figure 55b).*** The generated confusion matrix demonstrated, compared to previous models and in line with its ROC curve's result, a lower performance (accuracy= 0.77; precision= 0.5; recall= 0.77) with 3 and 10 patients incorrectly classified ***(Figure 55c).***

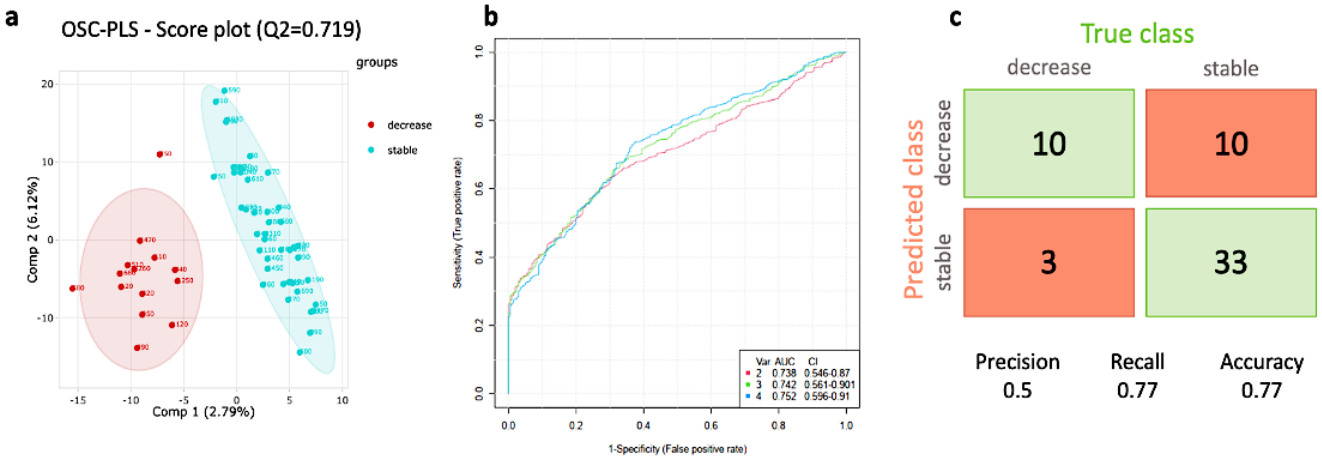


Figure 55 *Chemometric analysis of negative MS platform;* ***(a)*** *OSC-PLS score plot of negative MS data on decrease and stable groups showed a good performance (Q^2=0.719; p-value=0.5381);* ***(b)*** *multivariate ROC curve based on the 4 known annotated metabolites;* ***(c)*** *confusion matrix based on the 4 negative MS features with accuracy, precision and recall metrics reported.*

Data integration

A high-level approach was used for data fusion by allowing the integration of results coming from the different blocks. Samples common to NMR and LC-MS platforms were included in the analysis (n=56; 13 decrease; 43 stable). All the possible combinations of block integration were tested (***see S.I. FiguresS1-S2-S3***); finally, the most performant models were obtained by integration of NMR, Positive and Negative MS modes. A total of 33 variables, coming from the three different blocks (combination of ***Tables 14-15-16***), were integrated and a PCA and OSC-PLS model were performed. The discrimination between the two groups, that was already visible through the PCA score plot (**see S.I. *FigureS4***), was highlight by the high discriminant performance of OSC-PLS model (Q2=0.829; p-value= 0.00214) ***(Figure 56a)***. Like previously done for each separate dataset, a multivariate ROC curve was performed ***(Figure 56b);*** the degree of separability was evaluated (33 variables; AUC=0.845; 95% CI 0.674-0.962; accuracy=71.2%) by showing a more performant power compared to models of single blocks. Only 10 "stable" and 2 "decrease" patients were misclassified in confusion matrix by showing a general increased classification power when coupling NMR and LC-MS selected variables (accuracy= 0.79; precision=0.52; recall=0.85) ***(Figure 56c).***

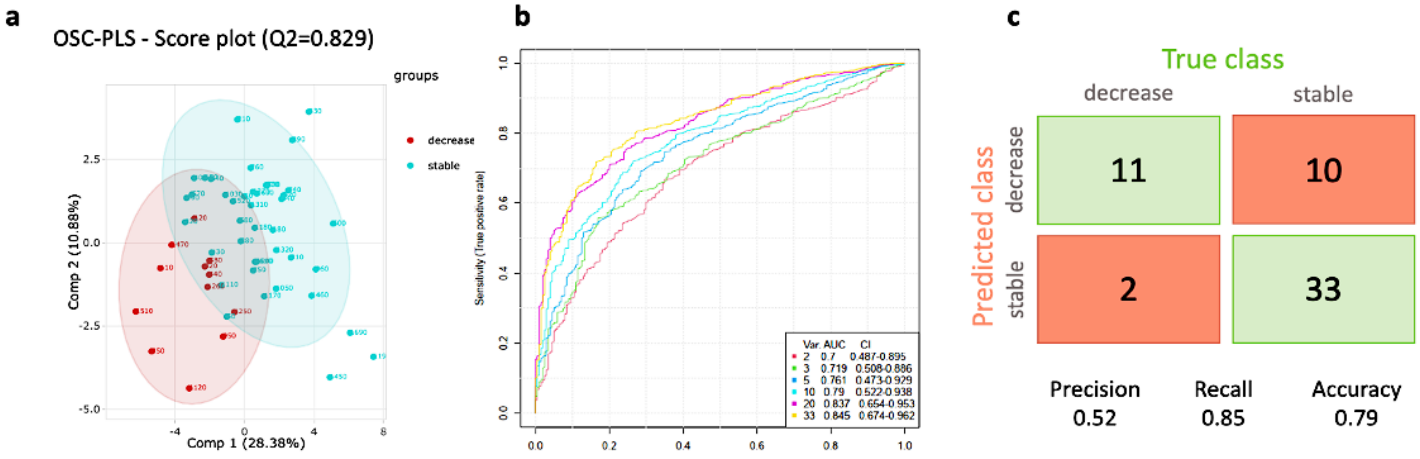


Figure 56 *Chemometric analysis for integrated platform; (a) OSC-PLS score plot for integrated platform on decrease and stable groups showed a good performance (Q^2=0.829; p-value=0.00214); (b)* ***(b)*** *multivariate ROC curve based on the 33 metabolites coming from the three analysis;* ***(c)*** *confusion matrix based on the 33 positive MS features with accuracy, precision and recall metrics reported.*

By comparing the performance of models from individual platforms and their combinations, it becomes evident that the integrated model incorporating NMR, both positive and negative modes (***refer to Table 17***), exhibits superior efficiency ($AUC_{NMR,POS,NEG}$=0.845 *versus* $AUC_{NMR=}$ 0.794). Specifically, when examining the Q2 values of OSC-PLS models, only the model derived from the Positive mode outperformed the one integrating all three platforms. However, despite this, the p-value from the permutation test was lower in the integrated approach, affirming that the observed result is unlikely to have occurred by random chance

alone. Notably, considering the AUC value of the multivariate ROC curve, and the accuracy of the confusion matrix, the integration of NMR and the dual modes LC-MS platform yields the most robust model.

Platform	OSC-PLS (Q^2)	P-value (Permutation test)	ROC curve (AUC)	Confusion matrix (accuracy)
NMR	0.775	0.03617	0.794	0.64
Positive mode (LC-MS)	0.933	0.03073	0.671	0.68
Negative mode (LC-MS)	0.719	0.5381	0.752	0.77
NMR, Positive, Negative	**0.829**	**0.00214**	**0.845**	**0.79**
NMR, Positive	0.768	0.04432	0.814	0.66
NMR,Negative	0.783	0.0196	0.812	0.70
Positive, Negative	0.768	0.00618	0.772	0.62

Table 17 *Comparison of models' performance for integrated and non-integrated models. In the table Q2 value for OSC-PLS models, permutation test p-values a, AUC of multivariate curves, and accuracy of confusion matrix are reported.*

All the 33 annotated metabolites were used to generate a pathway analysis in way to look and the metabolic pathways impacted by our significant metabolites. In particular, the result of this analysis highlights 11 metabolic pathways be involved in renal function decline with a p-value lower than 0.05 threshold (***Table 18).***

Pathways	Total	Hits	p-value
Aminoacyl-tRNA biosynthesis	48	7	2.69E-07
Phenylalanine metabolism	10	3	0.0001256
Phenylalanine, tyrosine and tryptophan biosynthesis	4	2	0.00067098
Tryptophan metabolism	41	4	0.00078284
Glyoxylate and dicarboxylate metabolism	32	3	0.0044684
Histidine metabolism	16	2	0.012418
Pantothenate and CoA biosynthesis	19	2	0.017357
Glutathione metabolism	28	2	0.036209
Alanine, aspartate and glutamate metabolism	28	2	0.036209
Porphyrin and chlorophyll metabolism	30	2	0.041138
Glycine, serine and threonine metabolism	33	2	0.048981

Table 18 *significant metabolic pathways impacted by metabolites coming from the platforms' integration. (**Total**= total number of metabolites composing the pathway; **Hits**= number of metabolites identified in our model)*

4.2.4 Discussion

In the present study, the urine metabolome at 3 months post-KTx was analyzed by using a sample stratification *predictive* of renal function changes at 12 months. Through this novel approach the main aim of this study was to identify a metabolic signature predictive of kidney function decline in early post-KTx. Indeed, after transplantation, the duration of adaptation stage depends on several parameters linked to graft condition, but in clinical practice, this period counts around three months post KTx. In addition to the innovative sample stratification, another primary strength of this cohort reposes on the high quality of clinical data linked to the collected urinary samples. Notably, to our knowledge, this cohort is among the few documented in the literature where the GFR for each patient is not merely estimated but accurately measured. The inclusion of mGFR values for characterizing patients constitutes a significant advantage for a deep and precise understanding of kidney function status. Moreover, the longitudinal nature of the sample cohort has facilitated the establishment of an innovative patient classification.

From a methodological point of view, another focal point of this study was represented by the platform combination and the data integration that allowed an increase of prediction model performance. Particularly NMR quantification character represented an asset for data normalization constituting a crucial step in urine samples analysis; indeed, in this manner, the importance of normalization on creatinine continues to represent an open and fierce debate [240,286,235]. In the context of this study, several preliminary analyses were conducted before to access the validity of normalization on dilution factor value; furthermore, the positive outcomes achieved using the NMR platform convinced us to apply the same normalization procedure to MS data. This approach yielded effective results, thereby validating the robustness of the normalization method. The synergy between these platforms became apparent in our study conclusion, and their application in parallel helps overcome the primary limitation of NMR represented by its lower sensitivity. Hence, through the combination of NMR's quantitative aspect with MS's higher sensitivity, a panel of predictive biomarkers elucidating the deterioration of kidney function was established and this will provide a more encompassing vision of its biochemical processes. From a biological point of view, the lack of numerous common metabolites between the two techniques finds its explication in the list of VIP identified in LC-MS platform; indeed, if we look at these tables the number of common features increase but no significant discrimination were revealed through univariate analysis. These results reflect the major challenge in LC-MS platform represented by variability due to the technique, possible matrix and batches effects.

Concerning the predictive models, by using this new stratification method, it was possible to explore kidney evolution in the months following transplantation and better describe the urinary metabolites that reflects and predict the renal function decline.
The multivariate statistical analysis allowed us to identify through the different platforms a list of putative biomarkers that could be linked to kidney function decline after transplantation. It is of particular interest to note that prior identification of metabolites for each separate platform, predictive models, and specifically confusion matrices had a higher accuracy (***S.I. FigureS5-S6-S7***). The decrease in performances of these models is due to the reduction of features included in the models; indeed, once the identification of metabolites was done, a part of the features corresponding to exogenous compounds, (drugs, pesticide etc.), unknown metabolites and features corresponding to the same biological compound were excluded from the predictive model. At this point, the selected features were used to define predictive and discriminant models for single and integrated datasets. Of relevance is outlined by the model merging the three datasets: NMR, Positive and Negative mode. By evidence, the integration of NMR and MS platforms allows the model to gain in accuracy for groups prediction; indeed, by looking at OSC-PLS model as well as the confusion matrix scores, we can spot the highest discrimination and classification performance of the combined model compared to the single-platform models. Moreover, through the identification of the features on which these models were based it was possible to delineate a panel of putative biomarkers (***Tables 16-17-18***) able to predict the kidney function decline. Even if only Hippurate was find as significant in both platforms, several other common metabolites were in VIP scores (***S.I. TableS1-S2***). Furthermore, of even more interest, was shown how metabolites coming from LC-MS and NMR techniques were sharing the same metabolic pathways allowing to give us complementary information and a global view on metabolome of "decrease" patients. This demonstrates, if proof were needed, the value of combining data from different analytical platform.

This analysis highlights the fact that all the identified features derive from widely various metabolic pathways going from amino acids metabolism (serine, glycine, phenylalanine, valine, hippuric acid etc..) to gut microbiota metabolism (TMAO, choline,) whose importance in CKD frame has been reclaim in the last years [256]. The existing link between decline in kidney function and some metabolites highlighted in this study has already been demonstrated.
For example, TMAO is a well-known metabolite in the context of kidney dysfunction that has been proposed as potential biomarker of CKD in the last decade[257]. Increase in plasmatic

TMAO has been reported as strongly linked to CKD [258,259]and inversely associate with estimated GFR [260,261]. In our study the downregulated TMAO level could be explained by an impaired renal extraction with a consequent depletion of this methylamine in urine and this notion is also supported by some others works[262,263]. Similarly, dimethylamine (DMA)is another amine compound that has found to be downregulated in the present study. This metabolite has already been described in other works as being correlated to medullary damage and acute rejection post transplantation [264,265] .

Among the metabolites whose concentration increased in the "decrease" group were choline, glycine and serine. Changes in choline levels may be directly related to decreased TMAO concentration; indeed, choline is a quaternary amino cationic alcohol that produces through bacterial metabolism TMAO as end-product. The higher level of choline in the present study is consistent with previous study and may signal underlying tubulointerstitial dysfunction [266] or an accumulation due to the increasing formation of TMAO. Another explanation can be found in the increasing levels of glycine which may derive from the choline generated through phosphatidylcholine' metabolism. Indeed, the augmentation of choline with decreasing renal function could contribute, in part, to the increasing circulating concentration and extraction of glycine as also demonstrated in a previous study [267]. Higher urinary serine concentrations were also found in patients whose kidney function declined at 12 months. Serine represents respectively a starting product in glycine biosynthesis and metabolism and its higher level in patients with renal function decrease was already found in a plasma cohort of CKD patients[267].

Concerning amino acids (e.g. phenylalanine, tyrosine, valine and glycine) changes in amino acids levels between patients with different levels of kidney function have already been reported in plasma and urinary samples ; indeed, their concentration were showed to result from inflammatory and acidotic events or changes in their metabolism due to the metabolic disturbances [269,270]. Relative to this, metabolic acidosis is common in transplanted patients due to assumption of calcineurin inhibitor therapy and is responsible of lower citrate excretion in post-graft complications, as showed by Bolen et al.[271]; in line with these results, in our study a downregulation of citrate was demonstrated in patients with kidney function decline post-KTx.

Decreased urinary extraction of carnitine may be directly linked to decreased level of its final product TMAO. Contradictory results are reported in literature, by showing how administration of carnitine in CKD rat model resulted in improvement in renal functions[267] or inversely how

its plasma increased levels are strictly correlated to CKD patients[272,273]. Data from a metabolomics profile of CKD patients included carnitine as one of the urinary metabolites predicting CKD stage [274]. Our results are consistent with this study by demonstrating a downregulated concentration of carnitine in patients with a decreased renal function.

Other well-known metabolites in kidney disease are represented by tryptophan and kynurenine. Tryptophane/kynurenine pathways has gained interest in the last years because of its role in acute injury prediction[275]. In addition to this, several other studies demonstrated decrease in tryptophan related to eGFR in serum of patients with kidney dysfunction but no changes in urine biofluids were reported. Other studies reported upregulation of kynurenine or downregulation of serotonin (other tryptophan derived metabolite) according to eGFR impairment in KTR[276,277]. These observations suggest changes in tryptophan/kynurenine pathways associated with kidney GFR impairment. In our study, we hypothesize that the diminution of tryptophan and kynurenine levels in patients with kidney function decline can be linked to major GFR impairment in those patient compared to "stable" one as also showed by Colas et al.[224] who demonstrated an urinary increase in levels of these metabolites in tolerant KTR compared to non-tolerant ones.

It is known that uremic toxins levels decrease in concordance to CKD severity in urine. In this study 3-methylhistidine and hippuric acid (the last one significant in both platforms) were found to be retained with a consequent downregulation in urine patients with declined renal function. Hippuric acid is a well-known uremic toxin associated to gut microbiome[278]; when its main elimination process, represented by active renal tubular secretion, is dysregulated this metabolite has a lower concentration in urine[279]. Concerning 3-methylhistidine it is another uremic toxin associated with muscle protein breakdown. Several studies have demonstrated the implication of this metabolite in CKD progression and development of CKD-related complication as specifically shown by its plasmatic increase in patients with stage 3–4 CKD[280,281]. These results are confirmed in our study in which the urinary decreased level of 3-methylhistidine could reflect its plasmatic accumulation due to an accelerated loss of muscle mass in patients with renal function decline with a following lower extraction of this uremic toxin. In addition to these, 5-methoxyindoleacetic acid is member of class of 3-indoleacetic acid (tryptophan derived uremic toxin). Studies on this metabolite demonstrate its upregulation in plasma sample of CKD patients according to eGFR decline[282]; since no literature is present concerning their behavior in urine samples, and in particular, since they are supposed to act as

uremic toxin, the downregulation of these molecules in “decrease “ patients in our study is in accordance with reported articles.

Limited literature exists regarding N-phenylacetylglycine, mannitol, and amino adipic acid, particularly in relation to CKD or kidney dysfunction. As other metabolites previously described influenced by gut flora, N-phenylacetylglycine is a product of bacterial metabolism of phenylalanine; however, its role in CKD remains unclear[283]. Notably, this study identifies decreased mannitol levels in the "decrease" group, although the connection to declining kidney function lacks clarification; only one other study in the literature has mentioned this metabolite's decrease during graft recovery[284]. Amino adipic acid, stemming from lysine degradation, has been proposed as an oxidative stress biomarker. Consistent with our research, previous studies have associated it with increased levels in diabetes and renal failure[285].

Considering all the results derived from this study, it appears evident that analyzing the urinary metabolome of KTR during the early post-transplantation period can forecast the evolution of GFR at one year. These discoveries could serve as a novel and valuable resource for clinicians, augmenting patient care during the post-transplantation phase.

Limitations and perspectives

Several limitations in this study should be considered when planning future research. From a clinical point of view, in the stratification process, the “decrease” group is formed based on the GFR’s progression, but there is no available information regarding the presence of eventual comorbidities which may cause this decline in kidney function. Further investigation is warranted to gain a comprehensive understanding of the pathophysiologic events associated with the reduction in GFR within the “decrease” group. Similarly, the significant misclassification rate in the confusion matrix of the stable group may be attributed to this phenomenon. Indeed, the arbitrary classification of patients based on a cutoff value for kidney function decline, without considering the complete clinical picture, may result in incorrect sample stratification. Another aspect of weakness is related to the sample size, which poses a significant constraint. While the number of enrolled patients might be considered as reasonable, the distribution of samples across each category is unbalanced, posing a challenge in identifying reliable biomarkers suitable for clinical application. In addition, the single-site collection and analysis preclude robust result comparison required for routine clinical use; consequently, supplementary studies are imperative to corroborate our preliminary findings.

Even by being aware of limitations, this study represents a valid exploratory analysis in kidney function decline in way to outline biomarkers set able to predict kidney function issues and improve patients' management.

4.2.5 Conclusion

In the context of renal transplantation, the follow-up of transplanted patients is fundamental for enhancing short- and long-term graft outcomes, as well as overall more personalized patients' management. Through the analysis of urine sample collected at 3-month and by using a patient's stratification predictive of kidney function changes at 12-month, we demonstrate in this study the possibility to obtain a panel of predictive biomarkers for decline in graft function with a model that is close to clinics requirements and standard. Specifically, the application of a dual metabolomics approach by using NMR and LC-MS platforms proved valuable in creating a novel workflow for normalizing and analyzing urinary samples. Notably, the combination of NMR and MS permitted to generate a more performant predicting models of mGFR decline and highlights the added value from the combination of these 2 analytical platforms.

Based on these results, a large panel of perspectives can be open. In way to increase the performance of our predictive models, it could be interesting to add a set of pertinent clinical and metabolomics data concerning the donor, the graft and the recipient that may better describe all the aspect of trio donor-graft-recipient. In addition, the enrollment of a multi-center and/or a wider cohort could allow to at first test the validity of our models and following increase its performance. Concerning this point, the recruitment of non-transplant and/or healthy patients could also represent a great challenge and opportunity to evaluate the findings of this work.

While this study is preliminary and several limitations must be considered, these discoveries lay a substantial foundation to enhance patient care following kidney transplantation using metabolomics.

Chapter 5

Metaperfusate project

5 Metaperfusate project

5.1 Background- Metaperfusate cohort

This section of the study concentrated on investigating the biological events influencing the graft both prior to transplantation and during the ischemic period, leveraging a comprehensively described cohort with detailed information on the donor, kidney graft, and recipient.

Thanks to the collaboration with surgeon team of the Department of Abdominal Surgery and Transplantation of CHU of Liege composed by O. Detry, M. Vandermeulen and N. Meurisse, the first 10 milliliters of the cold preservation solution outgoing the kidney right before transplantation was collected and called "perfusate".

A total of 49 perfusate samples originating from the donor's kidney were collected prospectively during renal transplantation (***Figure 57***). These samples were classified according to (1) donor type, distinguishing between donation after brain death (DBD=36) and donation after circulatory death (DCD=13), and (2) the occurrence of delayed graft function (total n=19/49; in DBD n=14/22; in DCD n=5/8). For each sample, some clinical and descriptive data related to the donor, such as age, sex, BMI, cause of death, were gathered. Additionally, some clinical and biological characteristics of the recipients were collected based on graft outcomes, including sex, age, etc.

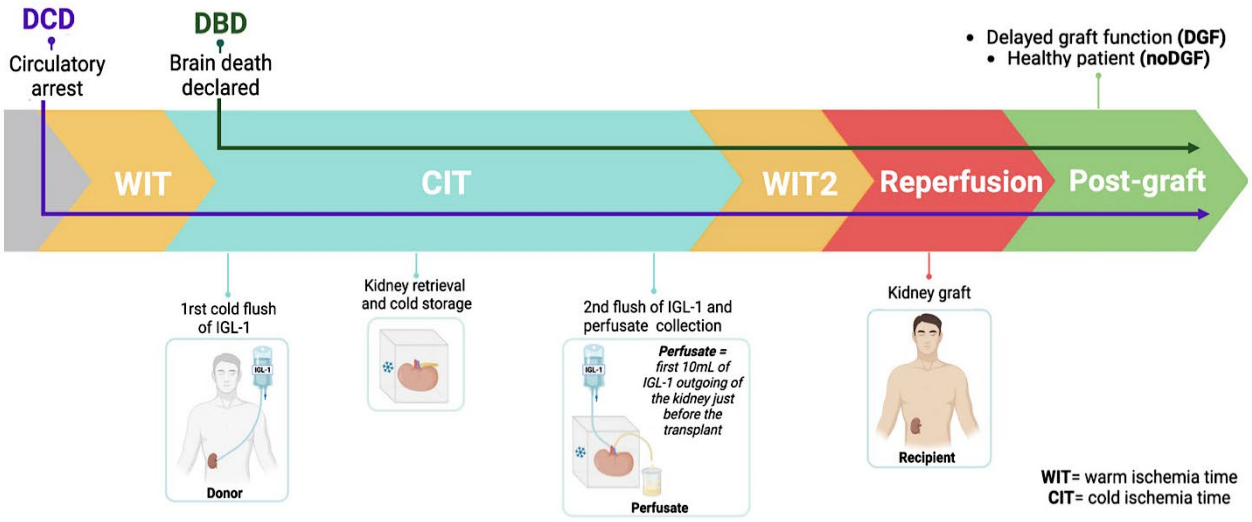


Figure 57 *Timeline of kidney transplantation and perfusate sample collection. (DBD= donation after brain death; DCD= donation after circulatory death; WIT= warm ischemia time; CIT=cold ischemia time; DGF= delayed graft function; noDGF= no delayed graft function)*

In addition to the human cohort, a murine model consisting of 19 perfusates was collected and categorized based on donor type (DBD= 9 /DCD= 10) . The analysis of this experimental model will allow us to simulate the donor's condition in a controlled environment, thereby validating the biochemical pathways linked to donor type observed in the human study. No data about DGF were available in the rat models. For the experimental model (performed by Pinto Coelho T. from Metabolism and Cardiovascular Sciences Group of GIGA) the brain death status was achieved through the increase of intracranial pressure generated by the insertion of a balloon catheter in the extradural space ***(Figure 58)***; for the DCD group, after a previous *tracheotomy and* 6h of mechanical ventilation in rodents, the cardiac death was obtained through an intravenous injection of KCl.

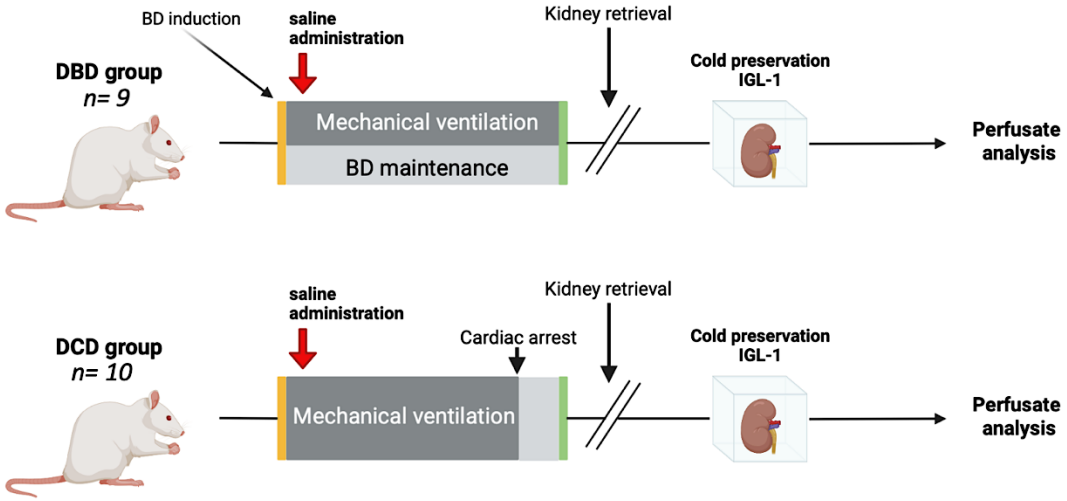


Figure 58 *Experimental model of kidney ischemia performed on rats; brain death or circulatory arrest was mimic and perfusate was collected after kidney retrieval (DBD= donation after brain death; DCD= donation after circulatory death.*

All samples were analyzed though NMR-based metabolomics approach. The multivariate statistical analysis on murine model highlighted a significant difference in perfusate metabolome of grafts derived from DBD versus DCD donors with a higher abundance of 2-hydroxvalerate, valine, isoleucine and alanine in DCD. The same difference was similarly observed in human perfusates, with a higher levels of isoleucine valine and other metabolites. After correcting for the type of donor, the multivariate statistical analysis demonstrated a metabolomics signature typical of human grafts ultimately presenting DGF.

The findings from this study have been published in *Metabolomics* in March 2024, under the title "***Untargeted NMR-based metabolomics analysis of kidney allograft perfusates identifies a signature of delayed graft function.***" An overview of the study, including cohort

composition, experimental design, summarized results, and concise conclusions, is provided in the graphical abstract presented below ***(Figure 59***).

Untargeted NMR-based metabolomics analysis of kidney allograft perfusates identifies a signature of delayed graft function

Study design

Donor type model:

DBD n°= 9

DCD n°=10

DBD n°=36

DCD n°=13

Post-graft status model:

DBD

DGF n°=14

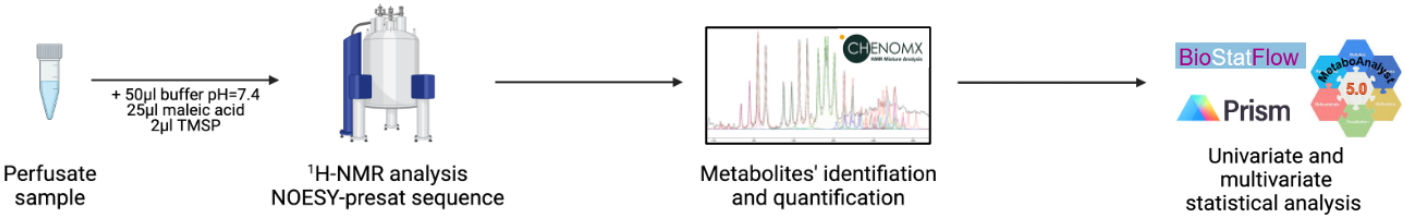
noDGF n°=13

DCD

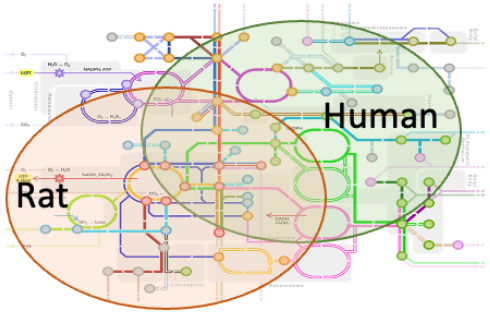
DGF n°= 5

noDGF n°= 8

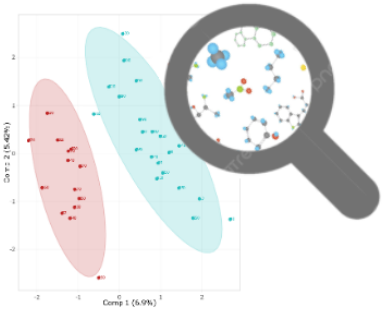
^{1}H-NMR metabolomics workflow



Results



Common metabolic pathways in human and rat study depending on donor type



Panel of biomarkers linked to DGF status as diagnostic tool

Conclusions

The acquired knowledge may help clinicians in the elucidation of the biochemical events occurring between death and transplantation per se to improve the process of organ procurement and reduce the discard rate.

Figure 59 *Graphical abstract for " Untargeted NMR-based metabolomics analysis of kidney allograft perfusates identifies a signature of delayed graft function."*

5.2 Untargeted NMR-based metabolomics analysis of kidney allograft perfusates identifies a signature of delayed graft function.

Cirillo A.[1], Vandermeulen M.[2, 3], Erpicum P. [3,4] , Pinto Coelho T.[3], Meurisse N.[2], Detry O. [2,3], Jouret F.[3,4]*, de Tullio P.[1]*

1 Clinical Metabolomics Group, Center for Interdisciplinary Research on Medicines (CIRM), University of Liege, Liege, Belgium.
2 Department of Abdominal Surgery and Transplantation, CHU de Liege, University of Liege, Liege, Belgium.
3 Groupe Interdisciplinaire de Génoprotéomique Appliquée (GIGA), Metabolism and Cardiovascular Sciences, University of Liege, Liege, Belgium.
4 Division of Nephrology, CHU de Liège, University of Liege, Liege, Belgium.

* Equally supervised the work
Corresponding author e-mail address: Arianna.cirillo@uliege.be (Arianna Cirillo)

Keywords: metabolomics, kidney transplantation, delayed graft function, NMR, perfusate

5.2.1 Abstract

Introduction: Kidney transplantation (KTx) necessarily conveys an ischemia/reperfusion (I/R) process, which impacts on allograft outcomes. Delayed graft function (DGF) is defined as a non-decrease of serum creatinine by at least 10% daily on 3 consecutive days during the first 7 days post-KTx. DGF significantly conditions both short- and long-term graft outcomes. Still, there is a lack of DGF-predictive biomarkers.

Objectives: This study aimed to explore the potential of kidney graft perfusate metabolomics to predict DGF occurrence.

Methods: 49 human perfusates from grafts categorized upon donor type (donation after brain death (DBD)/donation after circulatory death (DCD)) and DGF occurrence and 19 perfusates from a murine model classified upon death type (DBD/DCD) were collected and analyzed by NMR-based metabolomics.

Results: The multivariate analysis of the murine data highlighted significant differences between perfusate metabolomes of DBD *versus* DCD. These differences were similarly observed in the human perfusates. After correcting for the type of donor, multivariate analysis of human data demonstrated a metabolomics signature that could be correlated with DGF occurrence.

Conclusions: The metabolome of kidney grafts is influenced by the donor's type in both human and pre-clinical studies and could be correlated with DGF in the human DBD cohort. Thus, metabolomic analysis of perfusate applied prior to KTx may represent a new predictive tool for clinicians in a more personalized management of DGF. Moreover, our data paves the way to better understand the impact of donor's types on the biochemical events occurring between death and the hypothermic storage.

5.2.2 Introduction

Kidney transplantation (KTx) currently represents the best treatment for patients with end-stage renal disease. Even if 90,000 KTx are performed each year worldwide, kidney transplantation outcome highly depends on graft quality before transplantation and, at present, no reliable tools exist to assess it. In clinical practice, a list of standard classification criteria (SCD) was redacted to guarantee graft quality, including kidney procured in donation after brain death (DBD) conditions [287,288]. The increasing gap between demand and supply for KTx has led to the use of suboptimal organ donors, such as donation after circulatory death (DCD) and extended criteria donors (ECDs) [289,290]. However, kidneys coming from these donors (ECDs excluded) are exposed to a supplementary warm ischemia at procurement leading to higher risk of delayed graft function (DGF) and subsequently poorer graft outcomes [250,291–294]. DGF is a pathological condition, often resulting from ischemic damage and defined as a non-decrease of serum creatinine by at least 10% daily on 3 consecutive days during the first 7 days post-KTx [295]. Aside from early complications following KTx, the occurrence of DGF in transplanted patients translates in lower graft function and worse short- and long-term outcomes [193,296]. This clinical complication currently impacting approximately 25-30% of transplant recipients [194] is associated to several donor-related factors including, among others, donor type (DCD at higher risk compared to DBD) and duration of warm and cold ischemia times (WIT/CIT) [199]. In way to minimize the risks related to ischemic damage, the usage of a cold preservation solution to be flushed in the organ, can minimize these events by decreasing metabolism and slowing the process leading to IRI [297]. In recent studies, the use of IGL-1(Institute Georges Lopez-1, France) cold storage solution has led to the reduction of DGF incidence [186]. However, independently of the types of donors, reliable tools are needed to assess the quality of the graft after procurement and/or to better prevent–DGF in clinical practice [298]. In this context, metabolomics-based approach is particularly suited by placing itself as a solution tool adapted to personalized medicine setting for patient's treatment and follow-up [13,146,150]. This study hypostatized that perfusate solutions represent useful – still poorly explored – biofluids that could inform on what happened to the graft during its cold ischemic period, thereby providing the clinicians with informative data concerning the quality of the graft. Thus, using a Nuclear Magnetic Resonance (NMR)-based untargeted approach, this study aimed to analyze kidney perfusate solutions obtained just before transplantation in search of a metabolomic signature that could predict DGF occurrence in transplant patients. For this purpose, we had access to two types of perfusate samples, one from a human cohort and the other one from an animal model simulating the

donor's type of donation. The parallel use of a human and experimental samples could enable us to explore two main points: (1) the biochemical events occurring between death and hypothermic storage of kidneys related to donor type; (2) the differences in metabolomics content between different donor types.

The final goals of this pilot study were (1) to demonstrate the value of perfusate metabolomic analysis in KTx; (2) to evaluate the impact of donor type on the graft metabolome during the cold ischemia period and (3) to generate a metabolomics model that could be useful as a predictive tool for DGF occurrence.

5.2.3 Materials and Methods

WIT and CIT

In this study, kidney transplantation events were described using the terms: warm ischemia time1 (WIT1), cold ischemia time (CIT), and warm ischemia time2 (WIT2).

WIT1 was measured as the period between the circulatory arrest and the cold perfusion of the kidneys. CIT was defined as the period between the initiation of the cold flush until the removal of the graft from the cold storage to be transplanted in the recipient. WIT2 was defined as the time between removal from cold storage to graft reperfusion in the recipient (suture time).

Surgical models of kidney graft procurements in the rat

This experimental model has been essentially developed to mimic the different types of kidney graft donors. After induction of anesthesia (Isoflurane 2% in O_2 2l/min), a tracheotomy was performed to ensure optimal ventilation using a weight-based autoregulated rodent ventilator (PhysioSuite – KentScientific, Torrington, Connecticut, USA). Arterial and venous catheters (Polyurethane 0.43x0.69m) were placed in the femoral vessels to allow continuous monitoring of intra-arterial pressure (Picco Monitoring Kit and BP-100; CWE inc, Ardmore, Oklahoma, USA) and to have a venous access for intraoperative fluid injection, respectively (Ethic agreement code: 2147).

DBD model

In the DBD group (n = 9), brain death was induced under general anesthesia as previously described by Saat. et al [299]. In summary, after front lateral trepanation, a Fogarty balloon catheter (Edwards Lifesciences, Irvine, California, USA) was introduced and slowly inflated (100uL/min for 4 minutes) into the extradural space. Brain death was confirmed by the absence of corneal and pupillary reflexes, the onset of a hypertensive peak followed by major hypotension and a 60 s apnea test as described [300]. The rats were maintained in a brain-death state for 6 hours. The mean arterial pressure (MAP) was maintained above 60 mm Hg after induction of brain death by intravenous administration of normal saline (1mL/h) and norepinephrine (1mg / mL, Aguettant, Lyon, France) (5-15 μg/h) in case the administration of a 1mL bolus of normal saline would not maintain MAP above 60 mmHg. Then, all animals received a continuous infusion of 1 mL/h of normal saline using an electric syringe pump (Becton Dickinson, Franklin Lakes, New Jersey, USA), previously warmed to body temperature, to compensate for insensible losses. The body temperature was maintained at 38°C

with a rectal probe-controlled temperature pad connected to the ventilator (PhysioSuite – KentScientific, Torrington, Connecticut, USA).

DCD model

In the DCD group (n=10), all animals received 1 mL normal saline/hour, previously warmed to body temperature, to compensate for insensible losses. After 6 hours of mechanical ventilation in a previously tracheostomized rat maintained under general anesthesia, circulatory arrest was induced by an intravenous injection of KCl (150mg/kg) as previously described in rodent DCD models [301]. Circulatory arrest was defined upon cessation of aortic pulsatility and a fall in MAP below 25 mmHg.

Kidney procurement and perfusate collection

After induction of circulatory (DCD group) or brain death (DBD group), a laparotomy was performed to remove the kidneys. Both kidneys were flushed with IGL-1 organ preservation solution through the aorta and then stored immersed in IGL-1 at 4°C for 14 hours, as a model of clinical CIT (left kidney). After 14h of CIT, a second cold IGL-1 flush was made through the renal artery to collect the perfusate, defined as the first 2 mL of effluent evacuated from the renal vein. The perfusate was then stored at -80°C until metabolomic analyses (9 DBD and 10 DCD perfusates).

Clinical kidney transplantation

Informed consent was obtained from all KTx recipients included in the study (Ethics agreement: B707201524484). Clinical data about donor, graft and recipient were collected. Kidney donor risk index (KDRI) and kidney donor profile index (KDPI) were calculated based on donor and transplant factors [176].

Human kidney grafts procured from DBD or DCD donors were cold flushed during the procurement with IGL-1. All kidney grafts were preserved with classical static cold storage at 4°C during variable CIT, according to the clinical situation, and were allocated to a given recipient by Eurotransplant. After preparation of the renal vessels before KTx on the back-table, the kidney grafts were flushed through the renal artery with 1L of IGL1 solution. The first 10 mL of liquid exiting the renal vein during the flush were collected for metabolomics analysis and constituted the "perfusate". A total of 49 perfusate samples were used in this study; 36 from DBD and 13 from DCD donors. Prior to DBD and DCD classification, samples were categorized based on donor criteria classification including SDC (n=40) and ECD (n=9).

Among these 9 ECD donors, 7 were DBD and 2 were DCD. Concerning the post-transplantation status, 7 ECD donors experienced DGF, while 2 were classified as noDGF.

^{1}H-NMR metabolomics

All samples were recorded at 298 K on a Bruker Avance HD spectrometer (Bruker, Billerica, USA) operating at 700.17 MHz for the proton signal acquisition. The instrument was equipped with a TCI 5-mm cryoprobe with a Z-gradient. Maleic acid was used as the internal standard for quantification and trimethylsilyl-3-propionic acid-d4 (TMSP) for the zero for the zero calibration. An aliquot of 200µl of perfusate samples was supplemented with 50 µl of deuterated phosphate buffer (DPB, pH 7.4), 25µl of a 35mM solution of maleic acid and 2 µl of a 10 mg/ml TMSP D2O solution. 1H-NMR spectra were acquired using a 1D NOESY sequence with presaturation. The Noesypresat experiment used a RD-90°-T1-90°-Tm-90°-acquire sequence with a relaxation delay of 4 s, a mixing time (Tm) of 10 ms and a fixed T1 delay of 4 µs. Water suppression pulse was placed during the relaxation delay (RD). The number of transients is 64 (64K data points). The data were processed with the Bruker Topspin 4.0.8 software (Bruker BioSpin, Billerica, USA) with a standard parameter set. Phase and baseline corrections were per-formed manually over the entire range of the spectra and the δ scale was calibrated to 0 ppm using the internal standard TMSP.

Multivariate analysis

Once the spectra obtained, identification and quantification of metabolites were performed through Chenomx profiler 9.0 (Chenomx Inc., Edmonton, AB, Canada). The profiled spectra were used to generate a concentration table that was imported in BioStatFlow webtool (biostatflow.org) for multivariate statistical analysis. Autoscaling normalization was applied to concentration table. Principal component analysis (PCA) was used for looking at outliers and cluster between samples. An orthogonal signal correction-PLS model (OSC-PLS) was performed as discriminant model and its quality was determined by the predictability calculated based on the fraction correctly predicted in one-seventh cross-validation (Q^2). Permutation tests were performed for validate models. Metaboanalyst (www.metaboanalyst.ca) was used for generating multivariate receiver operating characteristic (ROC) curves and confusion matrix to assess the robustness of the models. ROC curves were performed by using PLS-DA models as classification method and univariate AUROC as feature ranking method with 2 latent variables. Confusion matrix was used for calculating model sensitivity and specificity.

Univariate and pathways analysis

Univariate analysis was made using GraphPad Prism version 9.4.1 (GraphPad Software, La Jolla, CA, www.graphpad.com). Mann–Whitney U test was performed for comparisons between two groups. No correction was applied to Mann-Whitney U tests. The detailed analysis of the metabolic pathways was performed by Metaboanalyst (www.metaboanalyst.ca) using the metabolomic set enrichment analysis (MSEA) tool by using the high-quality SMPDB metabolic pathways as the backend knowledgebase.

5.2.4 Results

Study population

A preliminary statistical analysis was conducted on the cohort considering the donor criteria classification including SDC and ECD, which showed no confounding factors. Donor characteristics categorized by the type of death event are presented in **Table 19**.

	DBD (N=36)	DCD (N=13)	Total (N=49)	p-value
Sex				0.233
F	15 (41.7%)	3 (23.1%)	18 (36.7%)	
M	21 (58.3%)	10 (76.9%)	31 (63.3%)	
Age				0.774
Mean (SD)	45.361 (13.495)	46.692 (16.311)	45.714 (14.130)	
Range	19.000 - 68.000	18.000 - 67.000	18.000 - 68.000	
BMI				0.348
N-Miss	0	1	1	
Mean (SD)	25.739 (4.159)	24.502 (3.016)	25.429 (3.912)	
Range	17.700 - 35.900	20.000 - 30.000	17.700 - 35.900	
Serum creat (mg/dl)				0.277
Mean (SD)	0.789 (0.253)	0.703 (0.202)	0.766 (0.242)	
Range	0.310 - 1.400	0.430 - 1.120	0.310 - 1.400	
Weight				0.948
N-Miss	0	1	1	
Mean (SD)	77.111 (13.509)	76.833 (9.437)	77.042 (12.521)	
Range	47.000 - 110.000	65.000 - 94.000	47.000 - 110.000	
Diabetes				0.648
N-Miss	5 (13.9%)	1 (7.7%)	6 (12.2%)	
No	30 (83.3%)	11 (84.6%)	41 (83.7%)	
Yes	1 (2.8%)	1 (7.7%)	2 (4.1%)	
Hypertension				0.831
N-Miss	5 (13.9%)	1 (7.7%)	6 (12.2%)	
No	24 (66.7%)	9 (69.2%)	33 (67.3%)	
Yes	7 (19.4%)	3 (23.1%)	10 (20.4%)	
Cause of death				0.008
anoxia	8 (22.2%)	9 (69.2%)	17 (34.7%)	
cerebral hemor	19 (52.8%)	2 (15.4%)	21 (42.9%)	
trauma	9 (25.0%)	2 (15.4%)	11 (22.4%)	
KDPI %				0.386
N-Miss	1	1	2	
Mean (SD)	39.361 (27.234)	47.308 (30.398)	41.469 (28.007)	
Range	0.000 - 92.000	0.000 - 89.000	0.000 - 92.000	
KDRI				0.228
N-Miss	1	1	2	
Mean (SD)	0.937 (0.277)	1.053 (0.302)	0.967 (0.285)	
Range	0.570 - 1.660	0.640 - 1.540	0.570 - 1.660	

Table 19 : *Clinical characteristics of donors. (BMI=body mass index; Serum creat= serum creatinine mg/dl; KDPI= kidney donor profile index (%); KDRI= kidney donor risk index).*

Statistical analysis on recipient data (DGF= 19; noDGF=30) showed no correlation between DGF status and related clinical data (**Table 20**).

	DGF (N=19)	noDGF (N=30)	Total (N=49)	p-value
Sex				0.326
Female	9 (47.4%)	10 (33.3%)	19 (38.8%)	
Male	10 (52.6%)	20 (66.7%)	30 (61.2%)	
Age				0.606
Mean (SD)	59.000 (12.396)	57.000 (13.562)	57.776 (13.028)	
Range	28.000 - 72.000	22.000 - 73.000	22.000 - 73.000	
CIT				0.377
Mean (SD)	604.158 (281.886)	677.867 (281.889)	649.286 (281.286)	
Range	207.000 - 1069.000	189.000 - 1300.000	189.000 - 1300.000	
WIT1				0.732
Mean (SD)	2.105 (3.710)	2.533 (4.531)	2.367 (4.197)	
Range	0.000 - 11.000	0.000 - 16.000	0.000 - 16.000	
WIT2				0.363
Mean (SD)	37.316 (12.352)	40.000 (8.154)	38.959 (9.956)	
Range	21.000 - 54.000	19.000 - 55.000	19.000 - 55.000	

Table 20: *Clinical characteristics of receivers based on graft outcome (DGF or noDGF). CIT= cold ischemia time; WIT1= warm ischemia time; WIT2= warm ischemia time 2.*

Preliminary analysis on IGL-1 matrix

A preliminary metabolomics analysis allowed to spot a series of peaks coming from IGL-1 matrix represented by lactobionic acid, adenosine, raffinose, and glutathione signals **(S.I. FigureS10-S11-S12-S13)**. These peaks were identified, quantified, and excluded from following analysis on perfusate samples.

Perfusate metabolomes in rat models: differences between DBD and DCD donors

Nineteen perfusate samples were included in the analysis. DBD and DCD kidney donor conditions were mimed in 9 and 10 rats, respectively. After exclusion from statistical analysis of matrix-linked peaks, 44 quantified metabolites that are belonging to the kidney metabolome were used for multivariate statistical models. Supervised statistical investigations using OSC-PLS highlighted a significant discrimination between DBD donor type versus DCD donor type (Q^2 =0.621); the model was validated through a permutation test (p-value=0.0393), confirming the absence of overfitting (**Fig. 60**).

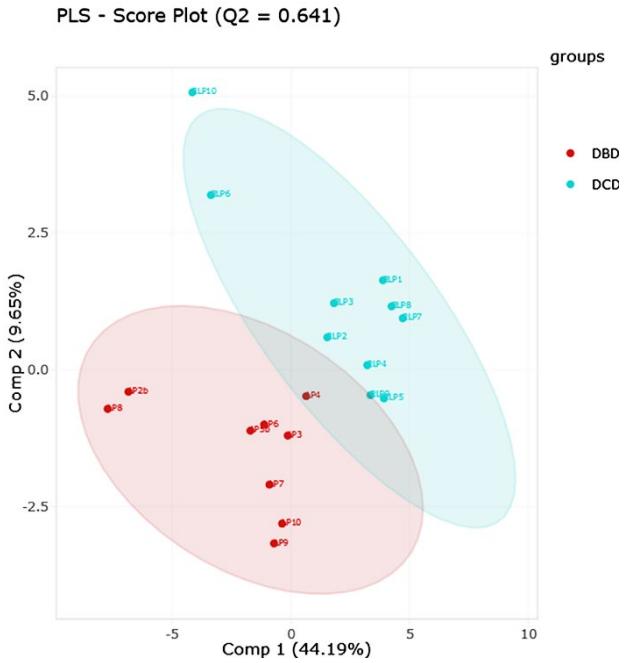


***Figure 60**: OSC-PLS model of rat model showing variation of metabolomics profile according to donor type (Q2= 0.641; cross validation p-value= 0.0393; n total=19).*

A loading-plot was generated in way to underline all the metabolites responsible for the discrimination of groups according to donor type. Based on the results of multivariate models, univariate statistical analysis revealed that the concentrations of 7 metabolites were increased in DCD status compared to DBD status, including 2-hydroxyisovalerate, 2-octenoate, 2-oxocaproate, 3-hydroxyisovalerate, alloisoleucine, creatine phosphate, lysine, o-phosphoethanolamine, taurine and valine (**Figure 61**).

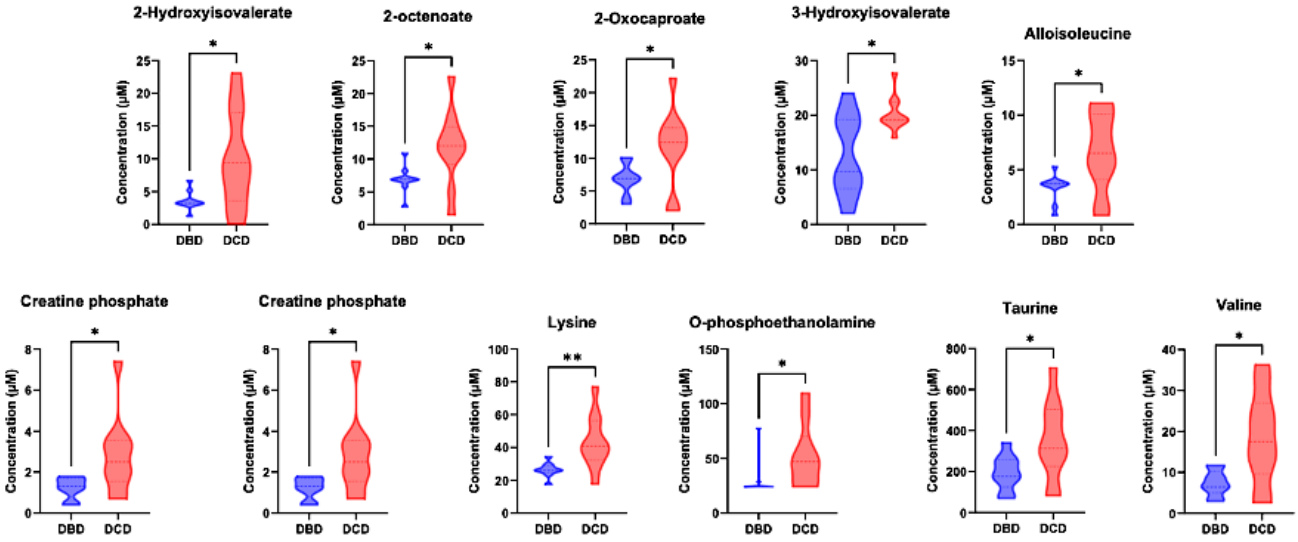


***Figure 61:** Violin-plots of significant metabolites in rat perfusate of DBD (n=9) versus DCD (n=10) . Mann-Whitney U test was used. (*<0.05; **< 0.001).*

Perfusate metabolomes in human cohort: differences between DBD and DCD donors

Forty-nine human perfusate samples were included in the analysis. Multivariate and univariate analyses based on donor criteria (SCD/ECD) indicated no significant statistical impact. Once this confounding factor was excluded, another analysis was conducted on the samples categorized by donor type (DBD= 36; DCD= 13). A total of 54 metabolites were identified independently from IGL-1 matrix and quantified (**S.I. TableS6; FigureS10-S11-S12-S13**) . Multivariate statistical analysis was performed by generating a discriminant OSC-PLS model (Q^2=0.540); a permutation test (p-value= 0.02192) was performed by validating the model obtained (**Figure 62**).

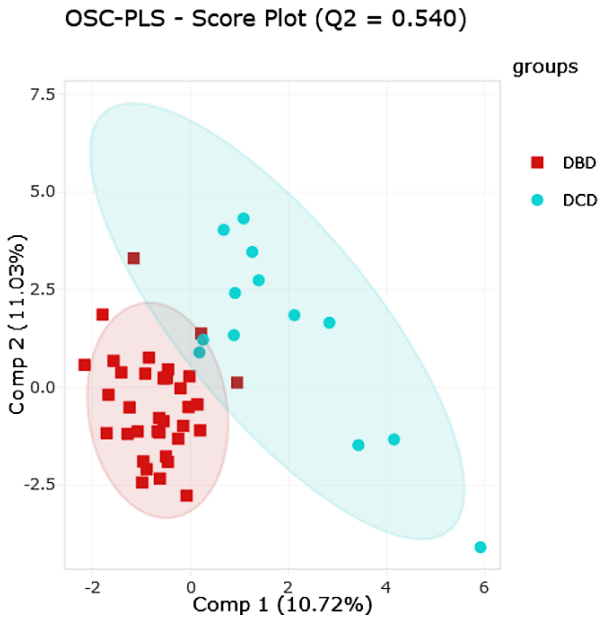


Figure 62: *Chemometric analysis based on donor types in human model (n total=39). OSC-PLS score plot showing a separation between DBD versus DCD donor (Q2=0.540; cross validation p-value= 0.02192).*

The metabolites relevant for discrimination between groups in multivariate models were used to perform univariate statistical analysis. Mann-Whitney U tests highlighted 17 metabolites significantly differing between donor types. Twelve metabolites were increased in DCD donor group (e.g., valine and cystine) while five were decreased compared to DBD group (e.g., lactate and trimethylamine-*N*-oxide) (**Figure 63**).

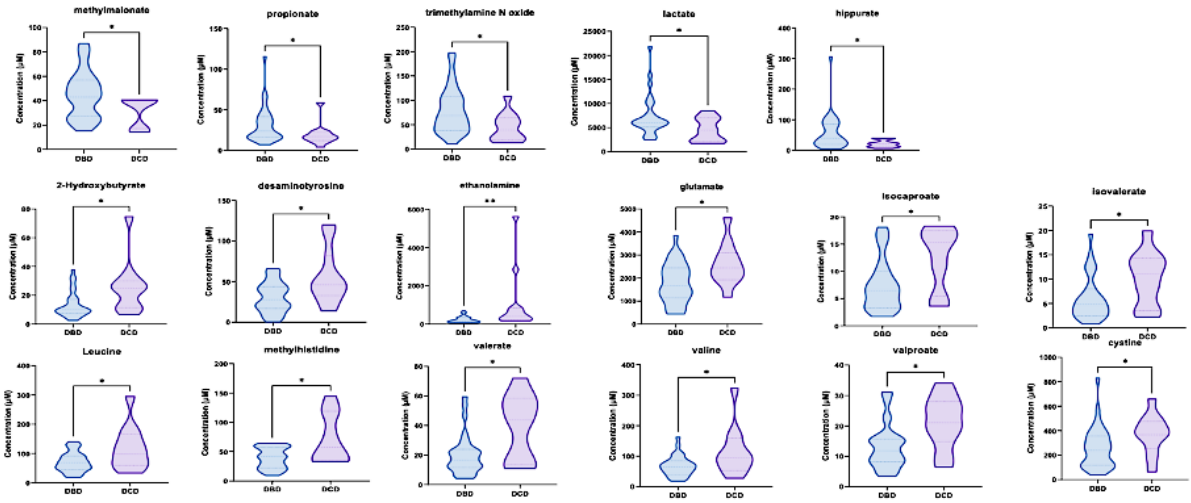


***Figure 63**: Violin-plots for metabolites of human cohort with significant variation according to DBD (n=36) versus DCD (n=13). Comparison between groups was made by using Mann-Whitney U test. (*<0.05; **< 0.001).*

Insights into common metabolic pathways in human and rat perfusates

By comparing the significant metabolites obtained in the two models, it was possible to spot some common features such as valine and isoleucine. In way to better investigate the biochemical events occurring during the hypothermic storage of kidney, a pathway analysis was conducted. MSEA analysis led to the identification of several pathways significantly affected by donor type (**Figure 64a-b**). The most relevant cascades were, in both models, represented by valine/leucine/isoleucine biosynthesis and degradation, aminoacyl-t-RNA biosynthesis, alanine/aspartate/ glutamate metabolism and glycerophospholipid metabolism.

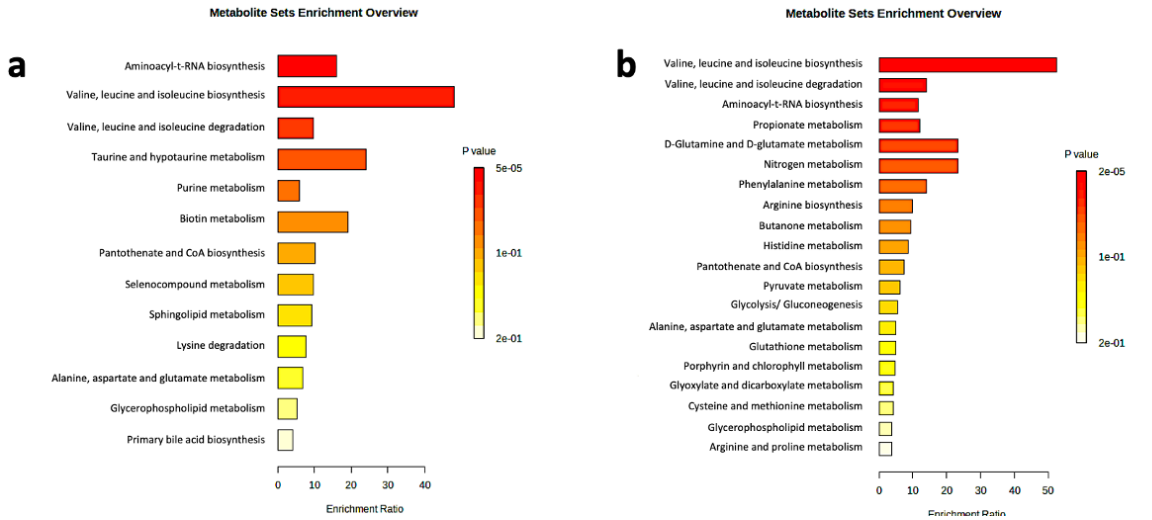


***Figure 64**: Comparison of MSEA analysis in human and rat models showed common biochemical pathways involved in the process of DBD and DCD ; **(a)** MSEA analysis of human cohort based on significant metabolites related to donor type ; **(b)**MSEA analysis of rat model for DBD versus DCD groups.*

Correlation of ^{1}H-NMR-based human perfusate metabolome with DGF incidence post-KTx

By considering DGF outcome as factor (DGF=19 ; noDGF=30), OSC-PLS model was used as a discriminant analysis (Q^2= 0.370; P-value= 0.4387). Moreover, a multivariate receiver operating characteristics (ROC) curve was generated by using the ratios of 54 identified metabolites (73 variables; AUC=0.777 ; predictive accuracy= 71.6%) in way to evaluate the performance of our model through automated feature selection (**Figure 65a-b**). Univariate analysis showed twelve metabolites being significantly increased in noDGF status (such as cystine, leucine, alanine ...), unlike carnosine (**Figure 66**).

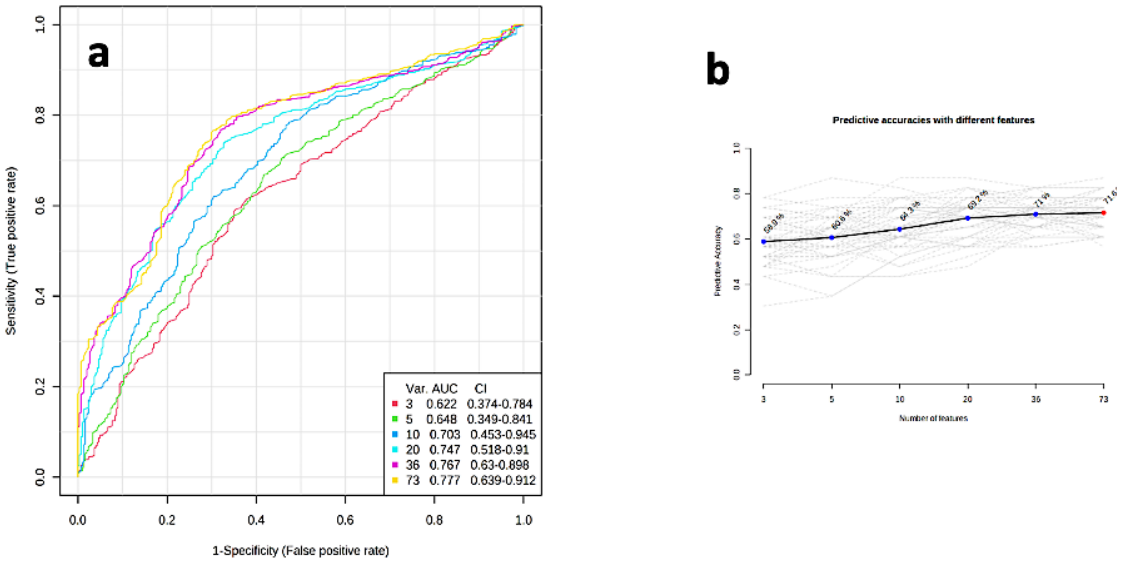


Figure 65 *Predictive models of DGF occurrence in the human cohort: (a) Multivariate ROC curve based on DGF versus noDGF analysis with automated selection of ratio features showed an AUC= 0.777 by considering 73 variables; (b) Plot showing the predictive accuracy with different features by reaching a value 71.6% with 73 variables.*

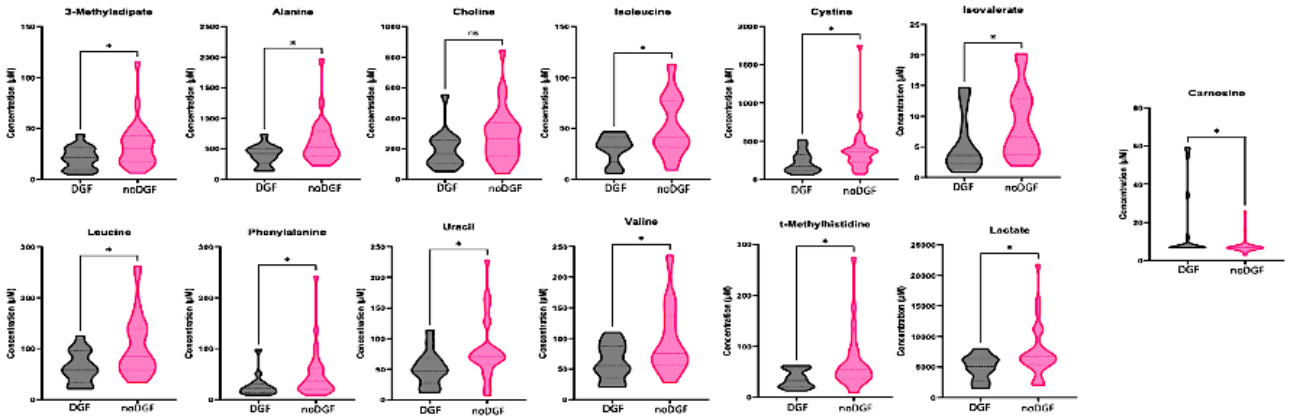


Figure 66 *Violin plots of significant metabolites according to DGF (n=19) versus noDGF (n=30) in human model. Comparison between groups was made by using Mann-Whitney U test. (*<0.05; **< 0.001).*

In the perfusates from DBD kidneys, thirty-six perfusate samples were used for the univariate and chemometric analyses by using DGF status as discriminant factor. A discriminant OSC-PLS model was generated (Q^2=0.747; p value= 0.2642) to look at the separation between the noDGF and the pathological status (**Figure 67a**). Furthermore, the multivariate ROC curve showed a good performance by using the ratios of the 54 identified metabolites generated through Metaboanalyst algorithm (73 variables; AUC= 0.917; predictive accuracy= 83%) (**Figure 67b**). A confusion matrix was additionally performed to cross-validate the model by highlighting all the automated selected ratio features used by the algorithm (**Figure 67c**).

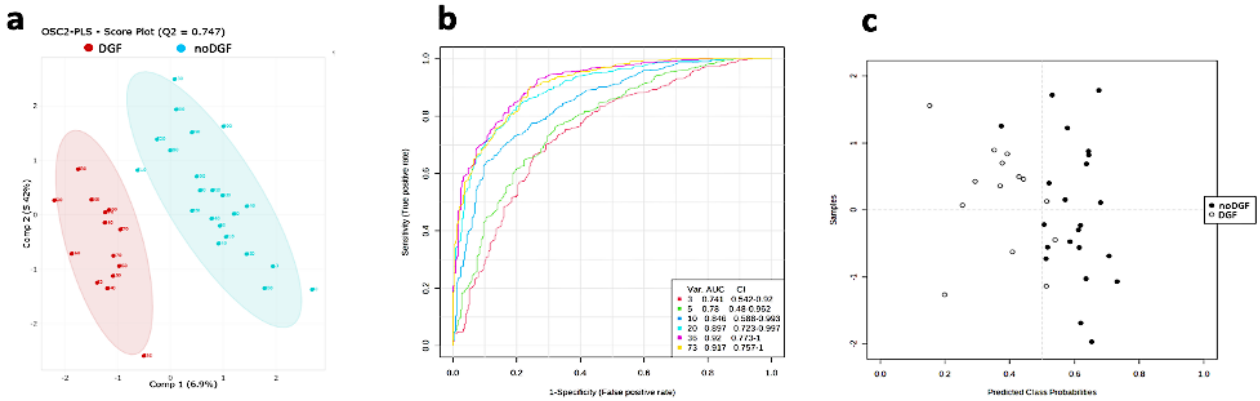


Figure 67 *Prediction of DGF kidney outcome in DBD group (human cohort);* ***(a)*** *OSC-PLS score plot show a separation between groups (Q2=0.747; p-value=0.2642);* ***(b)*** *multivariate ROC curve of DBD group based on DGF versus noDGF analysis with automated selection of ratio features showed an AUC= 0.917 by considering 73 variables;* ***(c)*** *confusion matrix was generated to evaluate model performance.*

Mann-Whitney t-tests highlighted 15 metabolites being significantly increased (e.g., lactate, alloisoleucine, isovalerate...) and ethylmalonate and carnosine being decreased in noDGF condition in comparison to DGF condition (**Figure 68**). To identify biologically meaningful patterns a MSEA analysis was done showing that methylhistine/ beta-alanine metabolism and valine, leucine and isoleucine degradation are the pathways more significantly impacted.

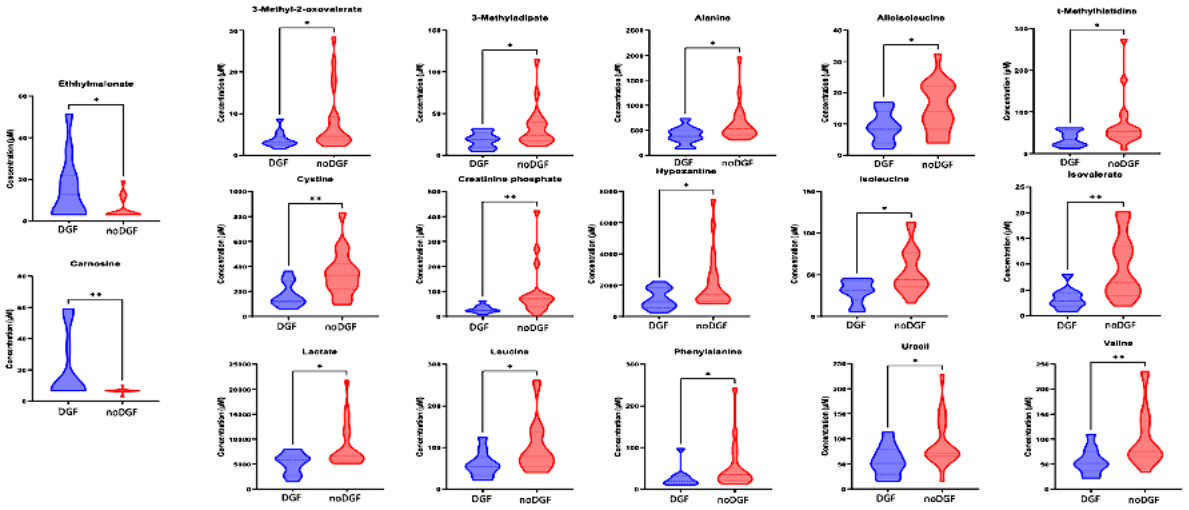


Figure 68 *: Violin-plots of significant metabolites of DBD subgroup according to the analysis DGF (n=14) versus noDGF (n=22) in human model. Mann-Whitney U test was used for comparison.*

A total of 13 perfusates from DCD donors were used for statistical analysis by using graft outcome as discriminant factor (data not shown). The low number of samples for each group (DGF=5; noDGF=8) highlighted the over fitting of the model and a consequent poor performance (Q^2= 0.7333; P-value= 0.4995). Univariate statistical analysis by using Mann-Whitney t-test showed only one metabolite significant for the discrimination of the two groups: creatinine phosphate.

5.2.5 Discussion

DGF represents one major obstacle in allograft survival [302]. Still, no single biomarker or a biochemical signature of graft quality and DGF prediction has been found and used routinely in clinical practice. The metabolomics approach has already been applied in several studies related to renal pathology and dysfunction on urine or plasma samples, leading to the discovery of metabolites linked to pathological status (Humphries et al., 2023; Iwamoto et al., 2022; Wishart, 2006). Its non-invasive nature makes this approach have advantages over biopsy in the context of graft status monitoring. Thus, several metabolomic studies of renal perfusate are described in the literature with particular emphasis on the correlation between perfusion time and metabolic content. [222,223,303]. In addition, other works have investigated perfusates of porcine models through NMR technique and have shown changes in metabolome's profiles of DCD donors between different perfusion strategies [297,304,305]. However, at that time, no relevant metabolomic signature or predictive model was reported about the occurrence of DGF condition.

The first question to ask is whether there is a difference in the graft metabolome depending on the harvesting conditions (DCD or DBD). Using our pre-clinical experimental model, we were able to demonstrate that the metabolomes of the DCD and DBD groups were significantly different. Specifically, perfusate analysis revealed higher levels of 7 metabolites, including lysine, isoleucine, and valine, in DCD samples compared to DBD samples. In the human cohorts, these metabolic trends were confirmed with more than 17 metabolites levels that significantly differs according to the donner type. Some of these metabolites were common in both rat model and human study. More interestingly, several common biochemical pathways such as valine, leucine and isoleucine degradation/ biosynthesis and aminoacyl-tRNA biosynthesis were also identified and could be correlated with some data in the literature. Thus, Hrydziuszko et al. previously observed the implication of these biochemical pathways in the DCD graft in the context of liver transplantation [306]. These similarities, albeit partial, between the metabolomic profiles of the experimental model and human samples demonstrated the value of the animal model for mimicking the graft according to the type of donor. Moreover, in the human study, lactate, trimethylamine-*N*-oxide (TMAO) and methylmalonate were also found to be significantly higher in DBD graft in comparison to DCD graft. Modulations in lactate levels could be explained by the major change in the metabolic oxidative process following brain death by underling a change from aerobic to anaerobic metabolism in the graft [307,308]. Higher levels of TMAO, could be linked to a more important renal medullary damage due to

ischemic events in DBD graft compared to DCD [309,310]. Differences in metabolic profiles related to donor types may also find an explanation in metabolic changes in the kidney due to the I/R process, as already reported in the literature (Jouret et al., 2016), and to the warm ischemia typical of DCD donors.

All these data showed that the biochemical processes that took place within the graft during the period of cold ischemia were different depending on the type of donor and need to be explored more in details. Furthermore, it could be expected that these metabolomic differences could be linked to the occurrence of DGF after transplantation and long-term outcomes in recipients. This shows that, in our view, each type of donor should be studied separately.

This was confirmed when we studied the correlation between the graft metabolome and the occurrence of DGF. Indeed, the results obtained using the data from the whole cohort (DBD and DCD donors together) showed a poor performance in predicting the post-transplantation graft outcome. While focusing our analysis on DBD-derived perfusates improved the quality of the predictive DGF occurrence model. Indeed, the multivariate ROC curve and the matrix of confusion obtained are quite relevant and compatible with a clinical predictive use for DGF occurrence demonstrating the interest of perfusate analysis for the recipients' follow-up.

Moreover, this approach led also to the identification of several biomarkers that differed in concentrations in the perfusates of DGF kidneys compared to noDGF kidneys. Alanine, isoleucine, leucine, valine and other branched chain amino acids were detected to be in significantly lower levels in DGF kidneys. These metabolites were released by the graft which can attest the sustained metabolic activity, ischemic damage, or both. Guy et al. previously observed raised concentrations of amino acids that may be linked to increased cellular breakdown in more ischemically damaged DGF kidneys [226]. However, their results were based on the perfusate's analysis of DCD graft 45 minutes after perfusion with KPS-1 cold storage solution, which hampers direct comparison with our results. Otherwise, Carnosine and ethyl-malonate, were significantly increased in DGF kidneys. The role of carnosine in kidney function has recently been investigated and shown a significantly lower urinary carnosine excretion in kidney transplant recipients compared to healthy group leading to higher risk of graft failure [311]. Potential mechanism of action of this metabolite included lowering of chronic low-grade inflammation, as demonstrated in animal models in which treatment through carnosine was associated with attenuation of I/R-induced renal dysfunction [312,313]. Thus, the higher level of carnosine in the perfusate samples of DGF kidney could be correlated to a possible increase of I/R damage. Levels of ethyl-malonate was increased in the more ischemically damaged DGF kidneys, which would suggest changes in normal function of

tricarboxylic acid (TCA) cycle organic anions (OAs). The biological mechanism linking TCA cycle organic anions with kidney dysfunction is not fully understood, but some studies have reported the relationship of these biomarkers with kidney function [314–316]. A more fundamental study using rat models of kidney transplantation would provide a better understanding of the role of metabolites and associated biochemical pathways in graft function during cold ischemia. This would also allow us to better understand the impact that these metabolic pathways could have on the occurrence of DGF.

Considering all these findings, we are also aware of the limitations of our present pilot study. Our findings need obviously to be validated and confirmed in human by studying a new cohort with an increased number of DCD patients. The development of a rat transplantation models could also be valuable, especially for a more fundamental study of the metabolic pathways implied in DGF events. Another limitation of our study is the lack of complete clinical data and of biofluids from the donors that could allow us to obtain a more complete view on graft and allow to refine and complete the data concerning the graft before transplant.

5.2.6 Conclusion

In KTx, a better understanding of graft quality and a better prediction of short-term renal function in the recipient patient are essential both to improve graft selection and to provide more precise, personalized patient management. By studying the fluid preserved in the graft obtained just before transplantation, we were able to identify, both in an animal model and in a human cohort, some differences in the kidney metabolome depending on the type of donor (DCD or DBD). These differences were directly associated with alterations in the metabolic pathways of the graft, potentially exerting additional influence on kidney function recovery post-KTx. In view of these results, it is obvious that DCD and DBD samples could be analyzed in separate models.

Thus, we have demonstrated that using the perfusate sample in DBD donors could allow to predict the occurrence of DGF status prior to surgery and to identify a panel of metabolites that could be correlated to this status. Our data paves the way to better understand the impact of donor's types on the biochemical events occurring between death and hypothermic storage and to correlate some metabolites of the storage liquid with the occurrence of DGF. This acquired knowledge may help clinicians in the elucidation of the biochemical events occurring between death and transplantation *per se* to improve the process of organ procurement and reduce the discard rate. Obviously, our preliminary results need to be confirmed on a validation cohort that could also include biofluid samples and complete clinical data from both donor and recipients. Moreover, the development of rat models of kidney transplantation from DCD or DBD donors as well as a better exploration of the identified biochemical pathways implied will also allow a better understanding of the ischemic events whose intricate interconnections that play a pivotal role in shaping the graft outcome.

Chapter 6
General conclusions and perspectives

6 General conclusions and perspectives

In the past decades, metabolomics has seen an increase of its success thanks to the variety of applications and its "Swiss knife" aspect. More recently, in the context of personalized medicine developments, the interest of this approach has grown with applications varying between different medicine branches such as oncology, metabolic disorders, cardiovascular disease, or even global routine health monitoring. Analysis of the metabolome allows to get a snapshot of metabolic changes and variations happening when homeostasis is broken, offering insights into how these changes influence the transition from a physiological to a pathological state and/or how it could evolve with diseases. Certainly, it is now widely recognized that pathological conditions result from disturbances in several metabolic pathways rather than being solely attributed to the over or under expression of a single biomarker. Thanks to the acquired knowledge, today, metabolomics serves as starting point for a more in-depth exploration and understanding of biological events that drive disease progression or affect treatment efficacy.

For several years, the Clinical metabolomics group (CliMe), in close collaboration with the department of Nephrology of the CHU of Liège, has been delving into the exploration of ischemia/reperfusion (I/R) events in the kidney using an NMR-based metabolomics approach. The impetus for our interest in this subject originated from the limited understanding of the physiopathological events associated with I/R injury and the substantial repercussions it can have on transplanted kidneys and on kidney function. That is why, in the context of this thesis and following the previous works of the CliMe, our focus was directed towards the kidney transplantation event in its preparatory and subsequent phases. Indeed, while a considerable number of metabolomics publications have explored cohorts with CKD, a noticeable scarcity of studies is reported on transplantation in the existing literature.

For this reason, the major aim of this work was to identify a panel of metabolites related to renal function to build predictive models which may help clinicians in clinical practice using our expertise in renal metabolomics. Indeed, despite significant and ongoing efforts to optimize patient's follow-up, there still exists a degree of imprecision that complicates patient diagnosis; in that light, the development and application of predictive metabolomics models may serve as an innovative diagnostic tool to enhance and personalize patient management and follow-up. Therefore, by using two cohorts in parallel, we decided to investigate three major factors

determining the outcome of kidney transplantation process: (1) donor graft, (2) transplanted recipient and (3) patient follow-up. The combined use of samples collected from the graft after the ischemic period and from graft-recipient may provide complementary information regarding transplantation process enabling the generation of two predictive models for pre- and post-graft events.

Based on these objectives, this work was divided in three fundamental parts whose objectives can be summarized as follow: (1) set-up of an optimized and precise protocol for analysis of urinary samples; (2) identify a metabolic signature predictive of kidney function decline during 1year post-KTx; (3) define a metabolomic signature typical of graft-quality correlating the correct kidney function recovery in early post-KTx.

The first part of the thesis focused on developing a urine preparation protocol to address technical challenges encountered during data preprocessing. Various strategies were implemented to mitigate issues related to chemical shift and data normalization during sample preparation. Both issues were successfully addressed by utilizing KF solution to reduce chemical shift variation and normalization on dilution factor measured though NMR to reduce dilution between samples. As shown through two studies presented in sections 3.3 and 3.4, data pre-processing step plays a crucial role in the context of statistical analysis and biological interpretation. Through this work, we showed the set-up of a robust and adapted sample preparation and pre-processing protocol can impact the performance of statistical models and by consequence results interpretation.

In the second part of this work a cohort of urine samples collected at two time points post-transplantation was analyzed by combining and integrating NMR and MS-based metabolomics approaches. To benefit from the longitudinal character of this cohort an original sample stratification was adopted aiming to forecast kidney function decline at 12-month post-KTx by identification of predictive biomarkers in urine collected at 3 months. Additionally, by taking advantage of the dual-platform metabolomics approach it was possible to highlight the complementarity of NMR and MS techniques allowing an increase in metabolomic coverage and model predictive accuracy for kidney function deterioration . Employing this methodology enabled us to investigate alterations in the metabolomes of patients experiencing a decline in renal function over time and to identify a panel of biomarkers from each utilized platform. This investigation underscores the robust and predictive potential of the metabolomics approach,

which could expand the array of tools available to clinicians in the clinical management of KTR.

As last part of this work, an NMR-based metabolomics approach was applied to a cohort of perfusate samples collected from donor graft just before transplantation. The interest of this study cohort was based on the use of samples collected from *donor before transplantation* coupled with sample stratification of *recipient outcome after transplantation.* Using this approach allowed us to put light the existing relationship between metabolome of the graft and the occurrence of DGF status. Moreover, the use in our cohort of two different types of donors in both animal model and human cohort, enables to explore differences in the kidney metabolome depending on the type of donor potentially exerting additional influence on kidney function recovery post-KTx. Considering these results, we demonstrated the potential of metabolomics approach applied to graft as novel tool to monitor quality graft in pre-transplantation event and how it may serve as diagnostic predictive tools for a more personalized management of DGF in clinics.

By looking back all the obtained results from the two cohorts analyzed in this thesis work, the complementarity of the results coming samples coming from the grafts and the recipients of kidney transplantation is clearly highlighted . However, while the predictive models demonstrate interesting results within our complementary cohorts, caution is advised when extrapolating these findings to a different or broader cohorts. Notably a key limitation of the metabolomics approach applied to both KTx axes in this thesis, is its challenge in generating models that effectively perform across multicenter sampling cohorts, which, in our study, would enhance their robustness and predictive accuracy. In addition to this, the limited sample size,, the lack of supporting literature, make mandatory the analysis of a new prospective cohorts to confirm the founded metabolic profiles describing the donor types, and predicting the delayed graft function status and the decline in kidney functionality in in KTRs.
Despite these limitations, the encouraging results suggest that the utilization of a non-invasive and time-efficient clinical metabolomics approach may holds the potential to advance the development of a novel diagnostic tool for concrete application in clinical practice.

If through the investigation of metabolomes in pre- and post-graft status it was possible to identify changes in metabolomes strictly related to kidney functionality at very early and middle terms after transplantation, several major studies still are needed to obtain an overall vision of the kidney transplantation event. In this sense, in perspective, it could be interesting to study a cohort consisting of matched donor-recipient pairs which may have the interest of combining data from the donor, the graft quality and the recipient before and after the transplantation to establish a predictive model for early post-graft kidney dysfunction or functional decline. Certainly, this study design has the capability to facilitate a more detailed examination of the three crucial elements in the transplantation process: the donor, the recipient, and the ischemic event whose intricate interconnection plays a pivotal role in shaping the graft outcome. Despite the numerous advantages and compelling results that this type of cohort may offer, we are aware of the practical limitations. It is crucial to acknowledge that in clinical practice, donors and recipients are typically not often situated in the same hospital, city, or, in some instances, the same country. In such scenarios, the retrieval of patient data and the collection of samples present challenges that necessitate close and stringent collaboration between multiple medical centers, a collaboration that may not always be easy to set up.

In the context of graft quality control, essential practical considerations need to be addressed: the timing and the setup of the analysis. While perfusate analysis can provide valuable insights to clinicians regarding the graft outcome, the timing presents a critical and limiting factor during surgical interventions. Indeed, during a kidney transplantation, it seems challenging to request the medical team to collect a perfusate sample and wait for the results of its analysis. In many cases, once the organ is available for the kidney recipient, the surgical intervention is promptly carried out without delay. Based on this critical point, if information about graft quality could be obtained soon after the graft intervention, it may be more realistic to utilize this information, for early post-KTx phase. Once predictive results indicating potential kidney dysfunction are available, clinicians could establish preventive measures to improve patient care. Indeed, taking into account the metabolic pathways that could be negatively affected during the ischemic period, another putative approach could be a modification of the composition of the cold preservation solution to avoid these detrimental modulations.

If all these perspectives refer to a prior or early phase of the kidney transplantation process, additional insights can be considered for a mid-post-graft period, such as the 3-month mark after transplantation. To enhance and refine current results regarding GFR evolution, several

options may be explored. Firstly, the utilization of a controlled murine model that mimics the decline of the kidney function in the post-KTx period could provide a more precise panel of biomarkers solely better correlated to GFR reduction over time and that could be related to specific biochemical pathways. Another possibility involves the combination and comparison of biopsies and urine results at 3 months post-KTx to broaden and redefine the panel of predictive biomarkers for kidney function decline. As a final suggestion, considering the advancements in omics sciences, a significant option may be the application of a multi-block approach. This method could not only establish a predictive model based on metabolites but also integrate proteomic or genomic information, thereby enhancing the model's robustness and allowing a comprehensive understanding of kidney function decline throughout the post-KTx period.

By having an overview of all the literature, the commercialized application (such as endometrial carcinoma, cancer screening detection etc..) and the encouraging result of this thesis, we can better realize the impact that metabolomics could have on personalized medicine. Indeed, in a world where our daily gadgets are tailor-made *ad personam* to reflect “who we really are”, and artificial intelligence provides customized solutions to various challenges, the development of personalized and precision healthcare appears evident. Indeed, it seems imperative today to move beyond the "traditional" concept of medicine, where all patients are standardized to a single average subject and cut-off values are defined to categorize them into "control" or "case" boxes. In this scenario, clinical metabolomics stands out as an innovative approach for a new healthcare concept which tailor patient treatment and follow-up to the specific characteristics of each patient. By considering the myriad of factors influencing a patient's health, metabolomics provides access to clinical cases that are no longer defined solely by a single biomarker but by the switch of whole metabolome from at equilibrium to “disease” state.

In summary, all the positive and encouraging results highlighted by this thesis allowed us to demonstrate the possibility to predict GFR evolution and kidney dysfunction in post-KTx phase, by demonstrating that metabolomics may be used in clinics for a more personalized approach on medicine. Finally, as also proved by various clinical applications, we are convinced that the evolution of precision medicine, supported by metabolomics, holds the potential to serve as a crucial diagnostic and prognostic tool, leading in a new era of more personalized and effective healthcare in clinical practice.

Supplementary information (S.I)

TableS 1 *pH values of urinary samples for the 4 subjects: T0=represents the pH at samples collection time; Phosphate buffer= pH for sample after addition of phosphate buffer; KF= pH of sample aftertreatment with KF protocol; Freeze-drying= urinary pH after freeze-drying process and addition of buffer.*

Subject	T0	Phosphate Buffer	KF	Freeze-dried
1	5.95	6.74	6.95	6.94
2	5.76	6.71	6.80	6.97
3	6.18	7.41	7.47	7.28
4	5.90	6.67	6.72	7.05

TableS 2 *VIP score from LC-MS positive analysis. In the table only the features annotated and for which a match in databases was found are reported. Only 191 of 751 features found a match in compounds databases.*

Positive VIP	Matching compound
M212T67	L-3,4-Dihydroxyphenylalanine methyl ester hydrochloride
M212T49	L-3,4-Dihydroxyphenylalanine methyl ester hydrochloride
M307T38	7'-Hydroxy-[3,4']bichromenyl-2,2'-dione
M197T141	1,9-dimethyluric acid
M197T162_2	1,9-dimethyluric acid
M283T497	2-(5-formyl-3,4-dihydro-2,2-dimethyl-2H-pyran-6-yl)-Benzoic acid
M165T43_1	2-Hydroxycinnamic acid, predominantly trans
M165T38_2	2-Hydroxycinnamic acid, predominantly trans
M165T57	2-Hydroxycinnamic acid, predominantly trans
M227T100	2',3'-Dideoxythymidine
M167T85	3-Methylxanthine
M272T367	N-3-hydroxydecanoyl-L-Homoserine lactone
M154T72_3	3-sulfino-L-alanine
M133T63	3-ureidopropionate
M237T66	3,4-dihydro-2,2-dimethyl-indeno[1,2-b]-pyran-5(2H)-one
M233T65	3',4',5',6'-tetrahydro-6',6'-dimethyl-spiro[isobenzofuran-1(3H),2'-[2H]pyran]-3-one
M233T70	3',4',5',6'-tetrahydro-6',6'-dimethyl-spiro[isobenzofuran-1(3H),2'-[2H]pyran]-3-one
M255T118	3',4',5',6'-tetrahydro-6',6'-dimethyl-spiro[isobenzofuran-1(3H),2'-[2H]pyran]-3-one
M255T147	3',4',5',6'-tetrahydro-6',6'-dimethyl-spiro[isobenzofuran-1(3H),2'-[2H]pyran]-3-one
M246T274	4-Acetamidoantipyrin

M105T38_3	4-aminobutanoate
M224T77	4-Hydroxybenzoylcholine
M235T56	4-methoxy-2-(5-methoxy-3-methylpyrazol-1-yl)-6-methylpyrimidine
M129T48	4-Piperidinecarboxamide
M221T84	5-hydroxy-L-tryptophan
M192T257	5-Hydroxyindole-3-acetic acid
M192T93	5-Hydroxyindole-3-acetic acid
M192T106	5-Hydroxyindole-3-acetic acid
M285T399	5,7-Dimethoxyflavanone
M140T56	6-Hydroxynicotinic acid
M328T297	8-Oxoerythraline_epoxide
M287T277_2	Abacavir sulfate
M266T277_1	Acetochlor-OXA
M330T73	Adenosine 3',5'-cyclicmonophosphate
M348T141	Adenosine 5'-monophosphate
M295T244	Aspartame
M208T152_4	Benzoylcholine
M311T435	Bifonazol
M311T419	Bifonazol
M245T99	Biotin
M266T277_2	Caffeoylcholine
M341T298_1	Cinanserin
M332T435	Ciprofloxacin (Cipro)
M250T221	Conivaptan HCl (Vaprisol)
M114T46_3	Creatinine
M226T38_2	Cyclocytidine
M226T63	Cyclocytidine
M199T94	Cymoxanil
M199T110_2	Cymoxanil
M202T359	DEP_202.1225_16.0
M215T51_1	DETHIOBIOTIN
M215T56	DETHIOBIOTIN
M305T357_1	Diazinon
M223T565_1	Diethyl-phthalate
M203T330	Diphenylsulfoxide
M330T129	Epoxiconazole
M280T340_1	Etamiphylline

M246T48	FEN_246.1101_16.1
M265T277_3	Feruloylputrescine
M196T161	Glucosaminate
M196T178	Glucosaminate
M147T277	Glutamine
M147T72	Glutamine
M173T52_1	Glycyl-L-proline
M303T253	Hesperetin
M303T292	Hesperetin
M180T162_2	Hippurate
M202T252	Hippurate
M118T253	Indole
M193T278	Isoproturon-monodemethyl
M194T278_1	Isoproturon-monodemethyl
M130T312	Isoquinoline
M190T211	Kynurenic acid
M190T195	Kynurenic acid
M209T93_2	Kynurenine
M209T103	Kynurenine
M209T497	Kynurenine
M164T68	L-Ethionine
M120T91	L-Phenylalanine
M205T153	L-Tryptophan
M182T130_3	L-Tyrosine
M182T108	L-TYROSINE
M182T57	L-Tyrosine
M136T142	L-Valine
M338T336	Linezolid
M181T190_1	Mannose
M278T357	Metazachlor
M274T147	Metazachlor-OXA
M150T55_2	Methionine
M206T354	Methoxyindoleacetic acid
M326T149	Midazolam Hydrochloride
M326T141	Midazolam Hydrochloride
M223T124	Mitoxantrone Hydrochloride
M223T118	Mitoxantrone Hydrochloride

M269T50	Diethylstilbestrol
M269T76	Diethylstilbestrol
M269T78	Diethylstilbestrol
M275T177	Forsenazide
M304T159	Fenoterol (hydrobromide)
M160T38	Pargyline hydrochloride
M160T60	Pargyline hydrochloride
M160T79	Pargyline hydrochloride
M160T71	Pargyline hydrochloride
M160T57	Pargyline hydrochloride
M256T197	Phenytolozamine citrate
M357T88	trans- Sulindac
M254T399	(9-oxo-10(9H)-acridinyl)acetic acid
M275T53	Dexchlorpheniramine maleate
M194T99	6-Methyl-2-(phenylethynyl)pyridine
M288T57	Rutaecarpine
M174T49	R-(-)-Desmethyldeprenyl hydrochloride
M174T64	R-(-)-Desmethyldeprenyl hydrochloride
M174T81	R-(-)-Desmethyldeprenyl hydrochloride
M330T244	Cloperastine hydrochloride
M218T54	Securinine
M218T49_1	Securinine
M218T71	Securinine
M190T262	N-Acetyl-DL-glutamic acid
M192T57	N-acetylL-DL-methionine
M256T158	N-alpha-(tert-Butoxycarbonyl)-L-histidine
M256T147_2	N-alpha-(tert-Butoxycarbonyl)-L-histidine
M100T110	N-Methyl-2-pyrrolidone
M100T88	N-Methyl-2-pyrrolidone
M125T38_1	N,N-Dimethylsulfamide
M125T38_2	N,N-Dimethylsulfamide
M404T253	L-Phenylalanine, N-[[(3R)-5-chloro-3,4-dihydro-8-hydroxy-3-methyl-1-oxo-1H-2-benzopyran-7-yl]carbonyl]-
M203T49_2	Pyrrolo[2,1-b]quinazolin-9(1H)-one, 2,3-dihydro-3-hydroxy-
M302T128	2-Phenylethyl beta-D-glucopyranoside
M271T69	Pyran-2-one, 5,6-dihydro-5-hydroxy-4-methoxy-6-(2-phenylethyl)-, (5S,6S)-
M248T82_2	2-(6-hydroxyhexyl)-3-methylidenebutanedioic acid

M293T382_2	3-Benzoxacyclododecin-2,10(1H)-dione, 4,5,6,7,8,9-hexahydro-11,13-dihydroxy-4-methyl-
M289T290	2-[(3S,3aR,5R,7aS)-3a-Acetyl-3-hydroxy-7a-methyloctahydro-1H-inden-5-yl]acrylic acid
M203T359	2-[5-(2-hydroxypropyl)oxolan-2-yl]propanoic acid
M265T381	4-(2,7-dihydroxy-6-methylheptan-2-yl)-3-hydroxybenzoic acid
M329T278_1	Pyrano[3,2-c][2]benzopyran-6(2H)-one, 3,4,4a,10b-tetrahydro-3,4,8,10-tetrahydroxy-2-(hydroxymethyl)-9-methoxy-, (2R,3S,4S,4aR,10bS)-
M567T277	(4S,5Z,6S)-5-[2-[(E)-3-(4-hydroxyphenyl)prop-2-enoyl]oxyethylidene]-4-(2-methoxy-2-oxoethyl)-6-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-4H-pyran-3-carboxylic acid
M365T299	(3S,5Z,8S,9R)-8,9,16-Trihydroxy-14-methoxy-3-methyl-3,4,9,10,11,12-hexahydro-1H-2-benzoxacyclotetradecine-1,7(8H)-dione
M300T440	1-Oxaspiro[2.5]octan-6-ol, 5-methoxy-4-[(2R,3R)-2-methyl-3-(3-methyl-2-buten-1-yl)oxiranyl]-, (3R,4S,5S,6R)-
M249T195	Cyclohexen-1-one, 4,6-dihydroxy-3-[(1E)-3-hydroxy-3-methyl-1-buten-1-yl]-6-methyl-, (4R,6S)-
M283T86	C11H14N4O5
M389T381_1	Cyclopentaneacetic acid, 2-[(2Z)-5-(hexopyranosyloxy)-2-penten-1-yl]-3-oxo-, (1R,2R)-
M180T139	4,6-dihydroxy-3-methyl-2,3-dihydroisoindol-1-one
M291T74	2-(4,8-Dihydroxy-4a,8-dimethyldecahydro-2-naphthalenyl)acrylic acid
M222T287	Pyrrole-3-propanoic acid, 5-acetyl-4-(carboxymethyl)-
M141T183	4-hydroxy-3,6-dimethylpyran-2-one
M142T197	4-hydroxy-3,6-dimethylpyran-2-one
M261T171	8-hydroxy-7-methoxy-3-(2-methylbut-3-en-2-yl)chromen-2-one
M413T253	Pyran-5-carboxylic acid, 3-ethenyl-2-(beta-D-glucopyranosyloxy)-3,4-dihydro-4-(2-hydroxyethyl)-, methyl ester, (2S,3R,4S)-
M289T184	Pyran-4-one, 3-(beta-D-glucopyranosyloxy)-2-methyl-
M257T340	(1R,4Z,9S)-4-(Hydroxymethyl)-11,11-dimethyl-8-methylenebicyclo[7.2.0]undec-4-en-3-one
M257T356	(1R,4Z,9S)-4-(Hydroxymethyl)-11,11-dimethyl-8-methylenebicyclo[7.2.0]undec-4-en-3-one
M243T252	C11H18N2O4
M243T290	C11H18N2O4
M243T273	C11H18N2O5
M247T264	Nefiracetam (Translon)
M247T275	Nefiracetam (Translon)
M298T88	Nelarabine (Arranon)
M267T408_2	Nevirapine (Viramune)
M123T38_3	Nicotinamide

M124T38_1	Nicotinic acid
M124T38_3	Nicotinic acid
M124T55	Nicotinic acid
M209T152_2	Nicotinoylcholine
M209T140	Nicotinoylcholine
M207T253	Norlidocaine
M204T38_3	O-Acetyl-L-carnitine
M205T312	O-Acetyl-L-carnitine
M153T52	Oxypurinol
M153T72_2	Oxypurinol
M152T86_1	Paracetamol
M221T38_1	Thidiazuron
M211T243	Phenazone / antipyrine
M166T68	Phenylalanine
M166T87	Phenylalanine
M166T91_3	Phenylalanine
M239T197	Pirimicarb
M116T54	Proline
M116T48	Proline
M130T410	quinoline
M459T333_1	Raltitrexed
M254T126	Sulfamethoxazole
M268T277_2	Sulfamoxole
M233T440	teasperin
M181T130_4	Theobromine
M181T178_2	Theophyline
M181T42_1	Theophylline
M181T43	Theophylline
M181T184	Theophylline
M181T50	Theophylline
M182T290	Tyrosine
M182T287	Tyrosine

TableS 3 *VIP score from LC-MS negative analysis. In the table only the features annotated and for which a match in databases was found are reported. Only 97 of 538 features found a match in compounds databases.*

Negative VIP	Matching compound
M133T48	(S)-MALATE
M481T429	11-(6-{[2-(3,4-dimethoxyphenyl)ethyl]amino}-4-chloro-1,3,5-triazin-2-yl)-7,11- diazatricyclo[7.3.1.0<2,7>]trideca-2,4-dien-6-one
M213T421_2	2-Deoxyribose-5-phosphate
M159T46_2	2-OXOADIPATE
M211T163	2,4,5-trimethoxybenzoic acid
M211T249	2,4,5-trimethoxybenzoic acid
M168T53_2	3-Methyl-L-histidine
M168T69	3-Methyl-L-histidine
M165T88	3-Methylxanthine
M182T248	4-Pyridoxate
M182T232	4-Pyridoxate
M197T48_2	4,6-dinitro-o-cresol
M463T82	4'-Methoxy-7-O-??-D-glucopyranosyl-8,3'-dihydroxyflavanone
M190T334	5-HYDROXYINDOLEACETATE
M328T85	Adenosine 3',5'-cyclicmonophosphate
M328T70	Adenosine 3',5'-cyclicmonophosphate
M229T94	Arabinose 5-phosphate
M289T310_1	Argininosuccinate
M109T155	Catechol
M191T69	Citrate
M191T86	CITRATE
M191T49_3	Citrate
M191T46_2	Citrate
M191T85	CITRATE
M178T76	Cyclamate
M261T340	D-Sorbitol-6-phosphate
M249T108	Dimethachlor OXA
M250T310	Dimethachlor OXA

M165T330	Ethylparaben
M165T345	Ethylparaben
M331T471_2	Gibberellin A20
M249T58	Heptenophos
M249T135	Heptenophos
M197T387	IQ
M173T44	Isocitrate
M173T58_2	Isocitrate
M141T81	Kojic acid
M159T127	L-alanyl-L-alanine
M146T46	L-Glutamic acid
M146T50	L-Glutamic acid
M145T278	L-GLUTAMINE
M145T49	L-LYSINE
M407T153	L-Tryptophan
M181T75	MANNITOL
M220T445	Metaxalone
M193T396	Monoethyl phthalate
M319T509_5	Mycophenolic acid
M319T424	Mycophenolic acid
M245T370	N-ACETYL-D-TRYPTOPHAN
M245T369	N-ACETYL-D-TRYPTOPHAN
M245T389	N-ACETYL-D-TRYPTOPHAN
M254T158	N-alpha-(tert-Butoxycarbonyl)-L-histidine
M189T77	N-Carbamylglutamate
M503T38	Hypericin
M503T53	Hypericin
M397T425	5-Benzofuranpropanoic acid, 4-(beta-D-glucopyranosyloxy)-6-methoxy-
M369T243	7-hydroxy-6-methoxy-8-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxychromen-2-one
M337T442	(2S,3S,4aS)-2,3,7-Trihydroxy-9-methoxy-4a-methyl-2,3,4,4a-tetrahydro-6H-benzo[c]chromen-6-one
M477T363	(5S)-1,7-bis(3,4-dihydroxyphenyl)-5-[(2S,3R,4S,5R)-3,4,5-trihydroxyoxan-2-yl]oxyheptan-3-one
M371T279	(E)-3-[2-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyphenyl]prop-2-enoic acid

M371T343	(E)-3-[2-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyphenyl]prop-2-enoic acid
M411T103	NCGC00380243-01!
M439T355	2-Hydroxycyclohexyl 6-O-[(2E)-3-(3,4-dihydroxyphenyl)-2-propenoyl]-beta-D-glucopyranoside
M439T408_2	(E)-3-[1-[(E)-but-2-en-2-yl]-9-hydroxy-10-(hydroxymethyl)-3-methoxy-4-methyl-6-oxobenzo[b][1,4]benzodioxepin-7-yl]but-2-enoic acid
M444T497	Pyrazino[2,1-b]quinazoline-3,6(1H,4H)-dione, 1-methyl-4-[[(2S,9S,9aS)-2,3,9,9a-tetrahydro-9-hydroxy-2-methyl-3-oxo-1H-imidazo[1,2-a]indol-9-yl]methyl]-, (1S,4R)-
M359T53	1-[(Z)-but-2-en-2-yl]-8-chloro-3,9-dihydroxy-4,7-dimethylbenzo[b][1,4]benzodioxepin-6-one
M397T88_1	(1R)-2-chloro-1,7-dihydroxy-3,9-dimethoxy-1-methylbenzo[c]chromene-4,6-dione
M465T496	1-[(2E,4E)-6,7-Dihydroxy-2,4-octadienoyl]prolyl-N-methylvalyl-N~2~-methylalaninamide
M465T498	[(2E,4E)-6,7-Dihydroxy-2,4-octadienoyl]prolyl-N-methylvalyl-N~2~-methylalaninamide
M511T347	1-[(2E,4E)-6,7-Dihydroxy-2,4-octadienoyl]prolyl-N-methylvalyl-N~2~-methylalaninamide
M275T96	Spiro[cyclopent-4-ene-1,1'(3'H)-isobenzofuran]-3,3'-dione, 2,4'-dihydroxy-6'-methoxy-5-methyl-
M275T75	Spiro[cyclopent-4-ene-1,1'(3'H)-isobenzofuran]-3,3'-dione, 2,4'-dihydroxy-6'-methoxy-5-methyl-
M441T324	3-Hydroxy-3-(hydroxymethyl)-4-methylpentyl 6-O-[(2S,3R,4R)-3,4-dihydroxy-4-(hydroxymethyl)tetrahydro-2-furanyl]-beta-D-glucopyranoside
M369T525	4-hydroxy-7-methoxy-2,3,8-trimethyl-3-(4-methylpent-3-enyl)-2H-benzo[g][1]benzofuran-6,9-dione
M302T152	Nicotianamine
M150T46	Paracetamol
M150T322	Paracetamol
M223T424	PHIP
M164T57	Phthalamic acid
M164T60	Phthalamic acid
M166T51_2	Quinolinic acid
M166T46	Quinolinic acid
M375T273	Riboflavin
M173T94	Shikimic acid
M341T86	Sucrose

M264T277_1	Thiamine
M338T87	Topiramate
M180T71	Tyrosine
M180T73	Tyrosine
M180T287_2	Tyrosine
M323T60	Uridine 5'-monophosphate
M323T64_1	Uridine 5'-monophosphate
M343T420	Usnic acid
M167T52_1	vanillic acid
M167T67	vanillic acid
M167T330	vanillic acid
M204T153	Xanthurenic acid

TableS 4 *Significant features derived from LC-MS positive analysis. The features here listed correspond to drugs or exogenous compound significant after Wilcox-Mann-Whitney t-test but not used for built prediction models.*

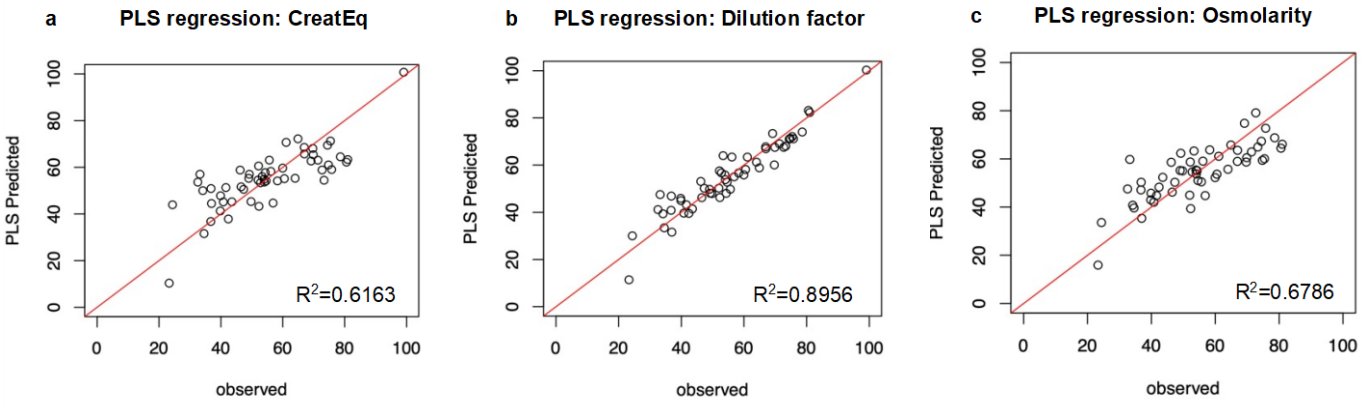
POS Feature	Matching compound	Description
M272T367	N-3-hydroxydecanoyl-L-Homoserine lactone	/
M233T65	3',4',5',6'-tetrahydro-6',6'-dimethyl-spiro[isobenzofuran-1(3H),2'-[2H]pyran]-3-one	/
M255T147	3',4',5',6'-tetrahydro-6',6'-dimethyl-spiro[isobenzofuran-1(3H),2'-[2H]pyran]-3-one	/
M246T274	4-Acetamidoantipyrin	drug metabolite of metamizole (analgesic and antipyretic) part of the human exposome
M226T63	Cyclocytidine	pyrimidine nucleoside exhibiting antiviral and anticancer chemotherapeutic activities.
M330T129	Epoxiconazole	Fungicide active substance used on a significant portion of cultivated surfaces in France
M304T159	Fenoterol Bromide	A synthetic adrenergic beta-2 agonist that is used as a bronchodilator and tocolytic.
M160T71	Pargyline hydrochloride	a monoamine oxidase inhibitor with antihypertensive properties.
M218T49_1	Securinine	plant-derived alkaloid that has previously been used as a therapeutic for primarily neurological related diseases
M256T147_2	N-alpha-(tert-Butoxycarbonyl)-L-histidine	/
M404T253	L-Phenylalanine, N-[[(3R)-5-chloro-3,4-dihydro-8-hydroxy-3-methyl-1-oxo-1H-2-benzopyran-7-yl]carbonyl]-	(OcratoxinA) mycotoxin with carcinogenic, nephrotoxic, teratogenic, immunotoxic and possibly neurotoxic properties.
M248T82_2	2-(6-hydroxyhexyl)-3-methylidenebutanedioic acid	/
M261T171	8-hydroxy-7-methoxy-3-(2-methylbut-3-en-2-yl)chromen-2-one	natural product derived from hydroxycumarin
M413T253	Pyran-5-carboxylic acid, 3-ethenyl-2-(beta-D-glucopyranosyloxy)-3,4-dihydro-4-(2-hydroxyethyl)-, methyl ester, (2S,3R,4S)-	Secologanol is a natural product found in Gentiana verna, Eucnide bartonioides, and Curtia tenuifolia
M247T275	Nefiracetam (Translon)	used in trials studying the treatment of Alzheimer's Disease.
M221T38_1	Thidiazuron	plant growth regulator (pesticide)

TableS 5 *Significant features derived from LC-MS negative analysis. The features here listed correspond to drugs or exogenous compound significant after Wilcox-Mann-Whitney t-test but not used for built prediction models.*

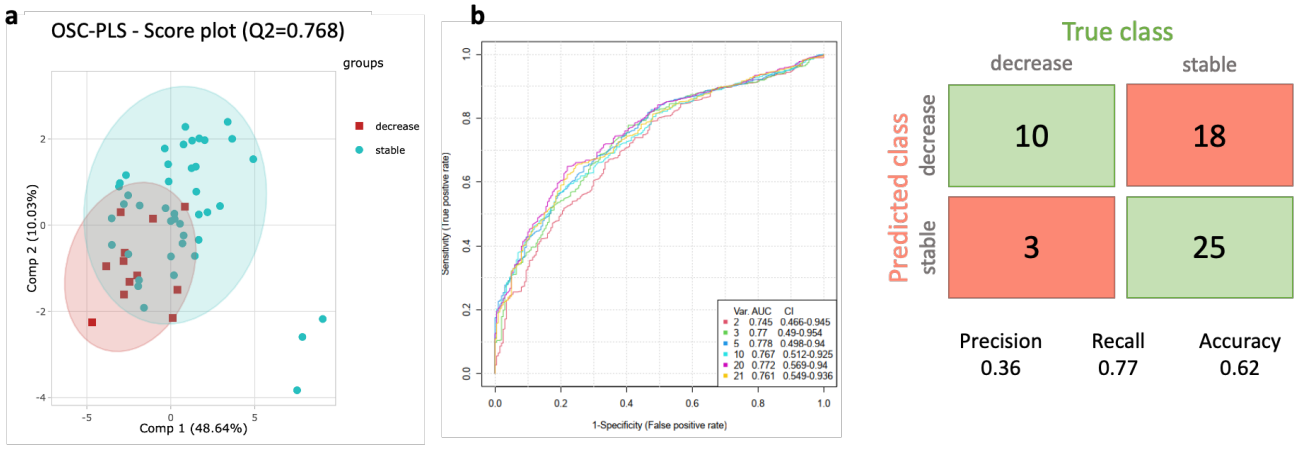
NEG Feature	Matching compound	Description
M211T163	2,4,5-trimethoxybenzoic acid	used to alleviate, treat, and prevent inflammatory diseases
M369T243	7-hydroxy-6-methoxy-8-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxychromen-2-one	Fraxin is a coumarin glycoside that has been found in Fraxinus and has anti-inflammatory activity.
M444T497	Pyrazino[2,1-b]quinazoline-3,6(1H,4H)-dione, 1-methyl-4-[[(2S,9S,9aS)-2,3,9,9a-tetrahydro-9-hydroxy-2-methyl-3-oxo-1H-imidazo[1,2-a]indol-9-yl]methyl]-, (1S,4R)-	Fumiquinazoline A is a natural product found in Aspergillus fumigatus
M359T53	1-[(Z)-but-2-en-2-yl]-8-chloro-3,9-dihydroxy-4,7-dimethylbenzo[b][1,4]benzodioxepin-6-one	/
M465T496	1-[(2E,4E)-6,7-Dihydroxy-2,4-octadienoyl]prolyl-N-methylvalyl-N~2~-methylalaninamide	/
M150T322	Paracetamol	analgesic and antipyretic

TableS 6 *Quantified metabolites through Chenomx Suite software with mean (μM) ± standard deviation values reported for all the 49 perfusate samples of human cohort.*

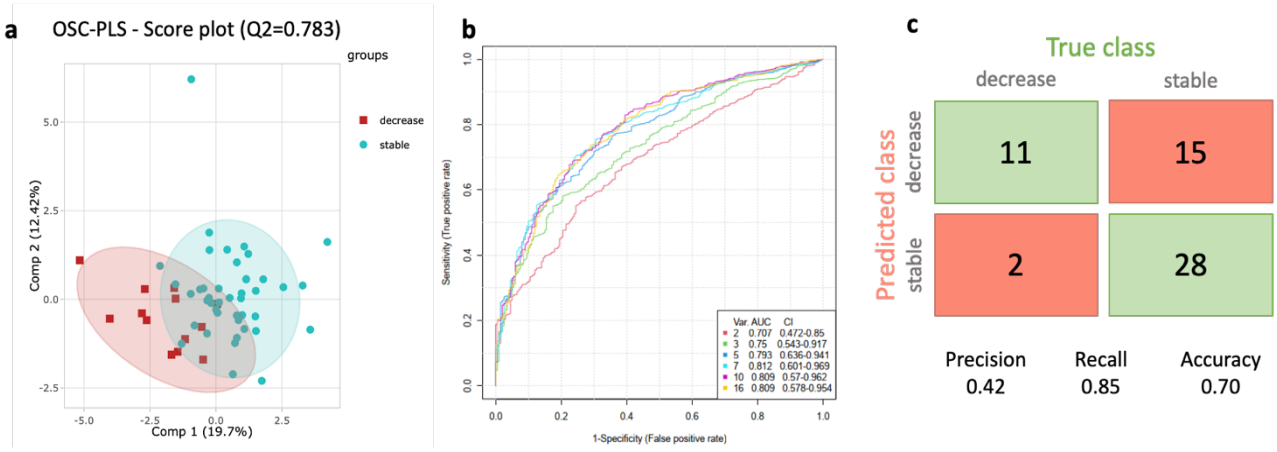
Metabolites	Mean ± SD
2-Hydroxybutyrate	17.910 ± 14.261
2-Oxocaproate	20.984 ± 18.128
2-Phenylpropionate	23.098 ± 17.322
3-Aminoisobutyrate	142.504 ± 64.853
3-Hydroxybutyrate	84.260 ± 53.913
3-Methyl-2-oxovalerate	7.920 ± 9.939
3-Methyladipate	30.118 ± 21.958
Acetoin	3.468 ± 6.117
Acetone	17.128 ± 9.432
Alanine	545.654 ± 356.556
Alloisoleucine	14.332 ± 10.349
Aspartate	112.610 ± 115.076
Butyrate	15.876 ± 16.424
Carnosine	13.142 ± 20.035
Choline	254.612 ± 162.980
Creatine	94.212 ± 98.976
Creatine phosphate	83.238 ± 135.711
Creatinine	271.512 ± 342.862
Cysteine	336.962 ± 251.807
Cystine	329.096 ± 270.555
Desaminotyrosine	41.568 ± 32.839
Ethanolamine	434.718 ± 854.103
Ethylmalonate	10.200 ± 11.516
Glucose	1316.544 ± 1686.962
Glutamate	2036.124 ± 1000.358
Glutamine	30.600 ± 16.788
Glutarate	9.600 ± 9.618
Glycylproline	27.400 ± 33.286
Hippurate	49.528 ± 53.569
Hypoxanthine	1695.448 ± 1382.189
Inosine	1895.966 ± 1130.915
Isocaproate	10.730 ± 9.084
Isoleucine	46.170 ± 32.105
Isopropanol	37.204 ± 87.465
Isovalerate	8.260 ± 6.198
Lactate	6718.058 ± 4061.611
Leucine	94.332 ± 65.011
Lysine	42.712 ± 84.769
Malonate	85.806 ± 116.664
Methylmalonate	44.176 ± 25.664
O-Phosphoethanolamine	252.170 ± 208.821
Pantothenate	15.212 ± 13.094
Phenylalanine	41.596 ± 41.832
Propionate	29.332 ± 22.800
Pyroglutamate	265.690 ± 205.703
Succinate	221.866 ± 217.853
Taurine	853.250 ± 623.938
Trimethylamine N-oxide	86.628 ± 81.698
Uracil	78.232 ± 66.355
Valerate	29.684 ± 28.000
Valine	88.648 ± 62.103
Valproate	17.178 ± 10.723
myo-Inositol	729.604 ± 527.485
τ-Methylhistidine	55.980 ± 46.844



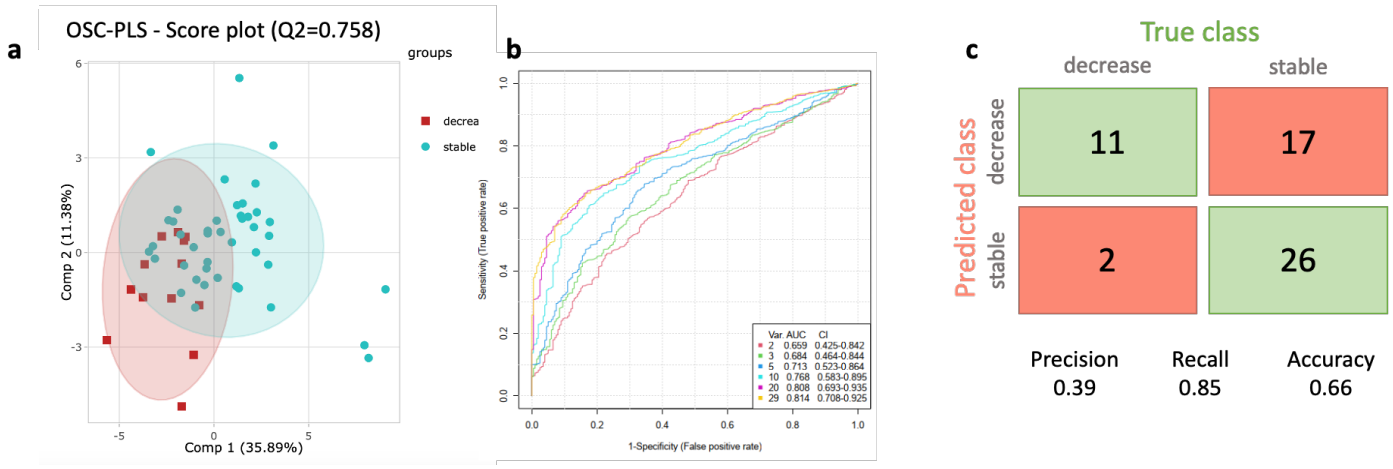
FigureS 2: *PLS regression plots for other used factors;* ***(a)*** *dataset normalized on CreatEq (R2=0.6163);* ***(b)*** *dataset normalized on dilution factor (R2=0.8956);* ***(c)*** *dataset normalized on osmolarity (R2=0.6786)*



FigureS 1: *Chemometrics analysis for integrated* ***Positive and negative mode of MS platform****;* ***(a)*** *OSC-PLS score plot for integrated platform on decrease and stable groups showed a good performance (Q^2=0.768; p-value=0.00618);* ***(b)*** *multivariate ROC curves based on the grooving number of metabolites coming from the POS and NEG analysis analysis;* ***(c)*** *confusion matrix based on the 21 positive and negative MS features with accuracy, precision and recall metrics reported.*



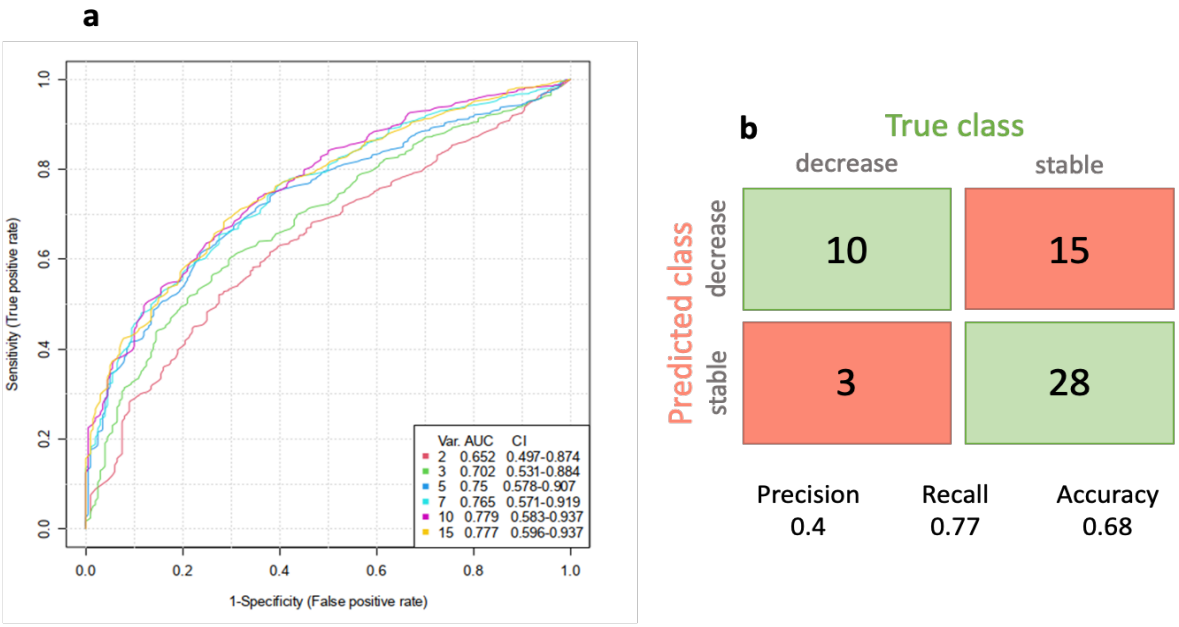
***FigureS 3**: Chemometrics analysis for integrated **NMR and negative MS mode platforms**; **(a)** OSC-PLS score plot for integrated platform on decrease and stable groups showed a good performance (Q^2=0.783; p-value=0.0196); **(b)** multivariate ROC curve based on the 16 metabolites coming from the two analysis; **(c)** confusion matrix based on the 16 NMR and negative MS features with accuracy, precision and recall metrics reported.*



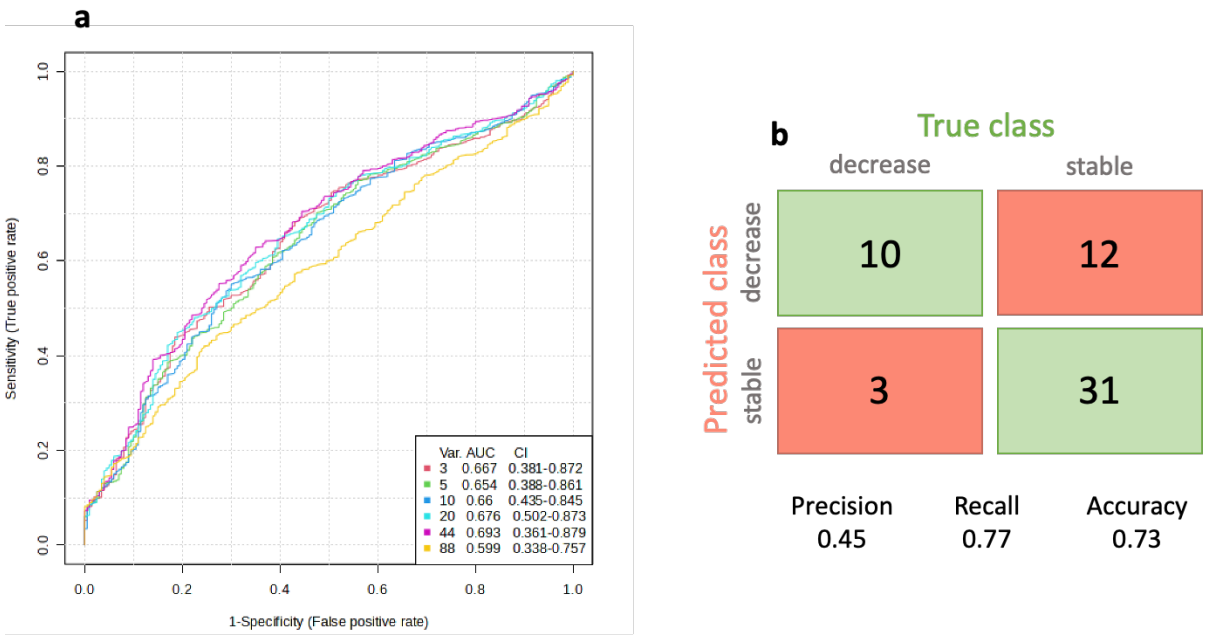
*FigureS 4 : Chemometrics analysis for integrated **positive MS mode and NMR platforms; (a)** OSC-PLS score plot for integrated platform on decrease and stable groups showed a good performance (Q^2=0.758; p-value=0.04432); **(b)** multivariate ROC curve based on the 29 metabolites coming from the two analysis; **(c)** confusion matrix based on the 21 NMR and positive MS features with accuracy, precision and recall metrics reported.*

R^2=0.72 n component= 5

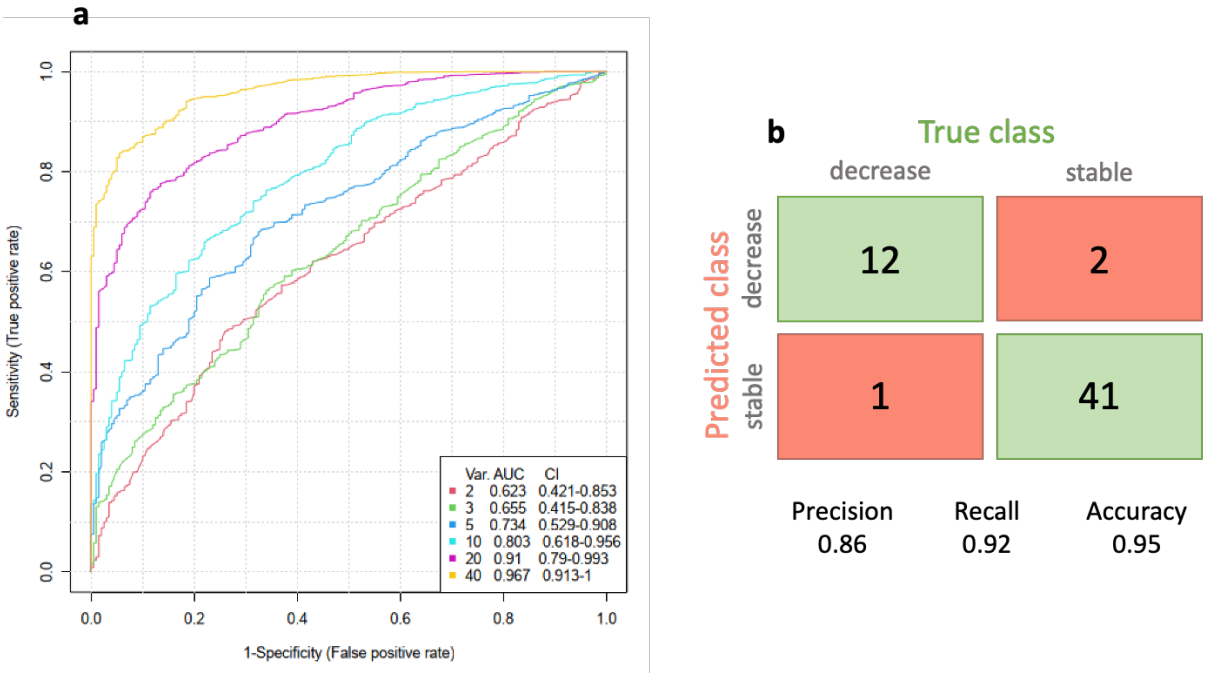
***FigureS 3**:PCA-X for integrated platform model. As we can see, the separation between the two groups can see already spotted through this unsupervised method without a priori concerning samples group membership.*



FigureS 6 *Models generated with significant features known and unknown coming from NMR platform* ***(a)*** *Multivariate ROC curve with increasing number of variables;* ***(b)*** *confusion matrix with reported precision, recall and accuracy.*

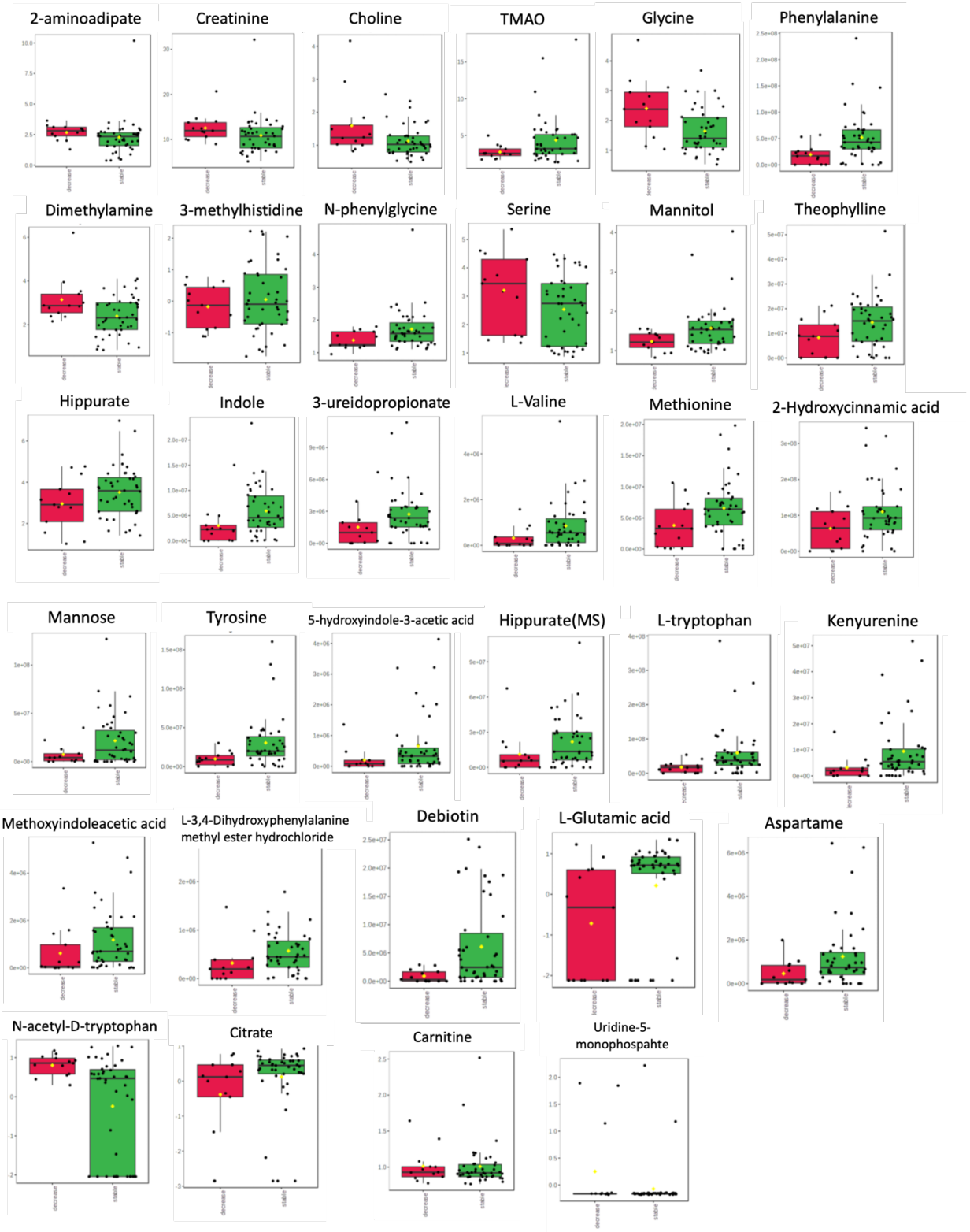


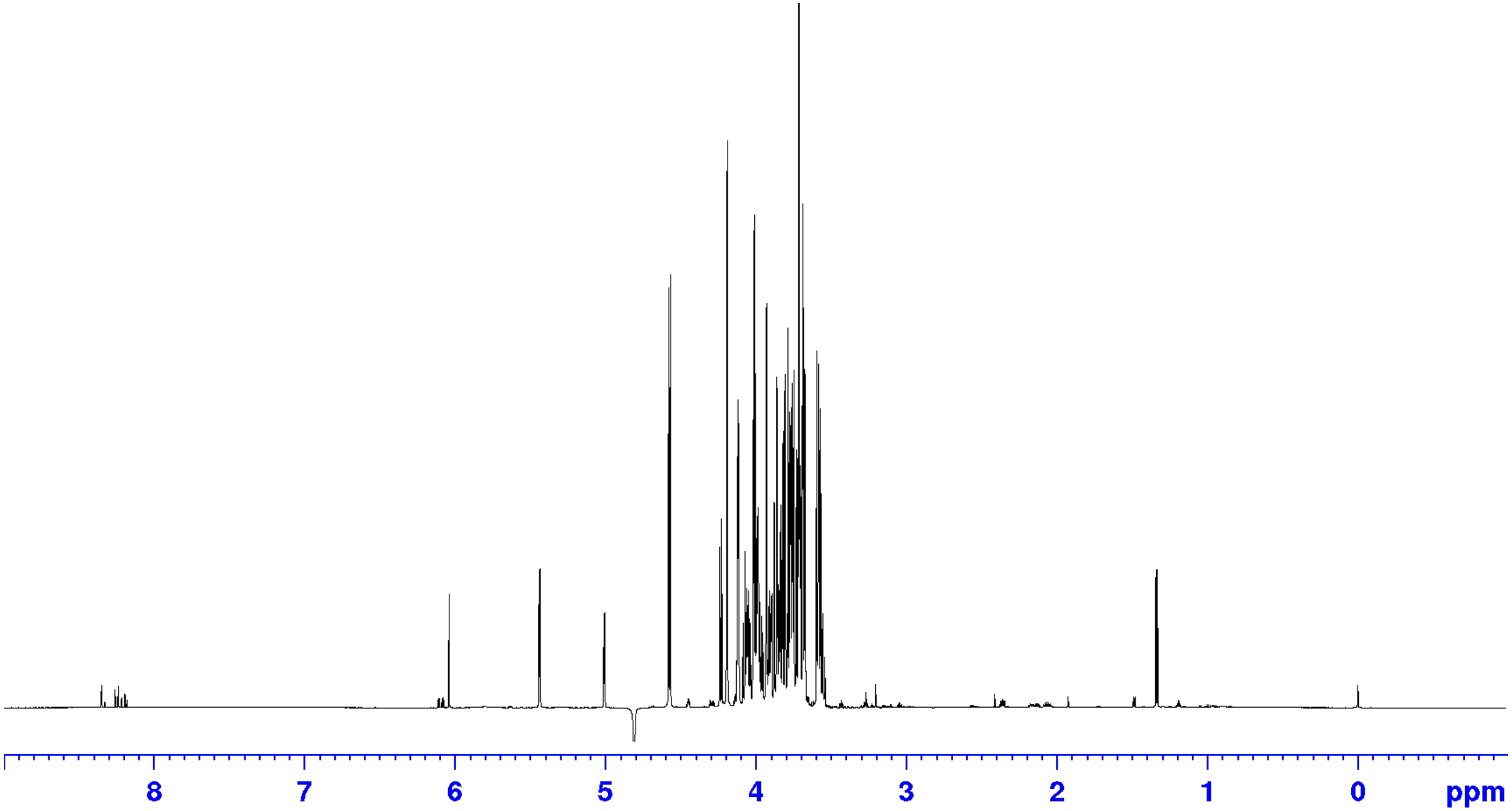
FigureS 7 *Models generated with significant features known and unknown coming from Positive MS mode ;* ***(a)*** *Multivariate ROC curve with increasing number of variables;* ***(b)*** *confusion matrix with reported precision, recall and accuracy.*



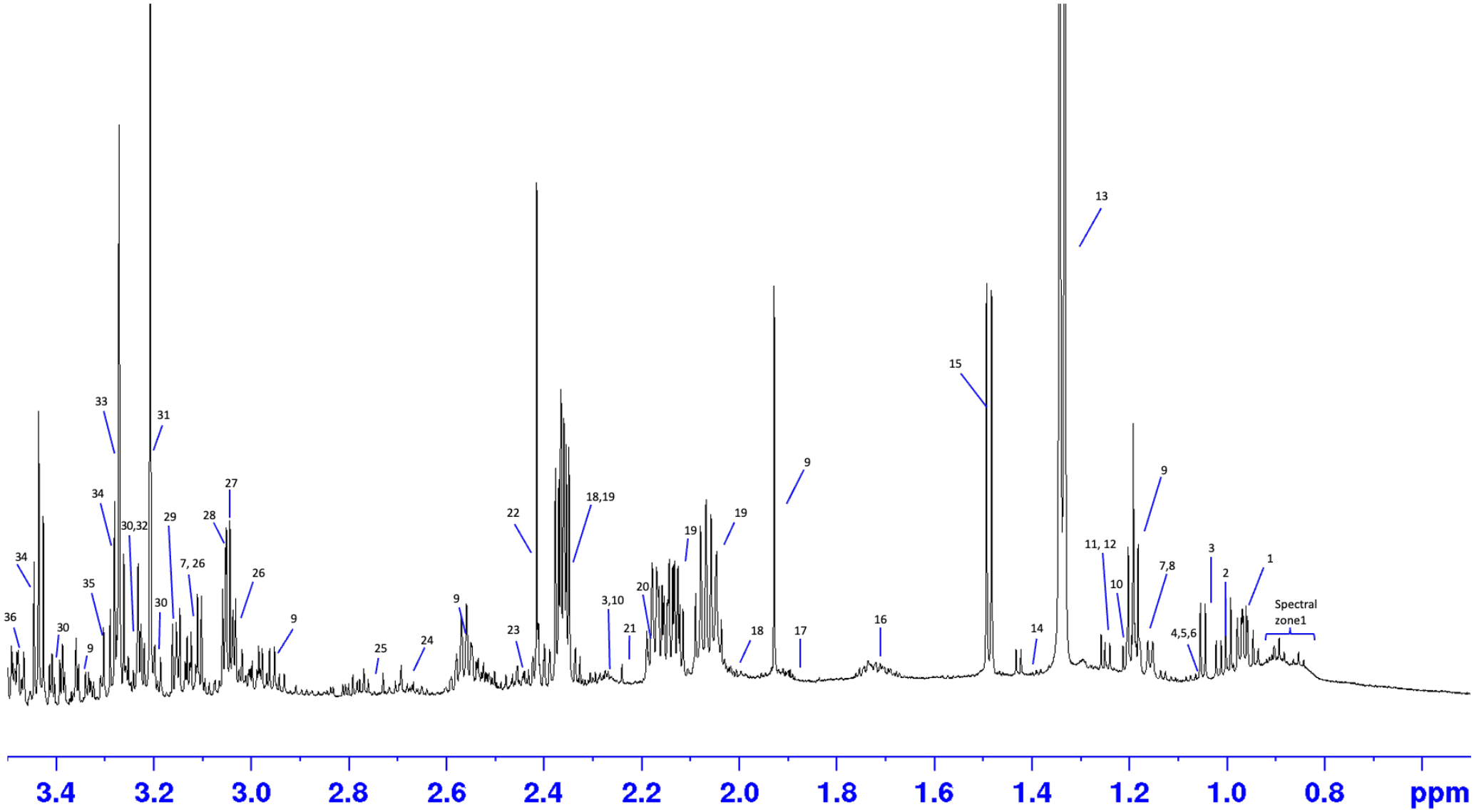
FigureS 8 *Models generated with significant features known and unknown coming from Negative MS mode* ***(a)*** *Multivariate ROC curve with increasing number of variables****; (b)*** *confusion matrix with reported precision, recall and accuracy.*

FigureS 9 *Box plots of the significant identified metabolites coming from the three platforms.*

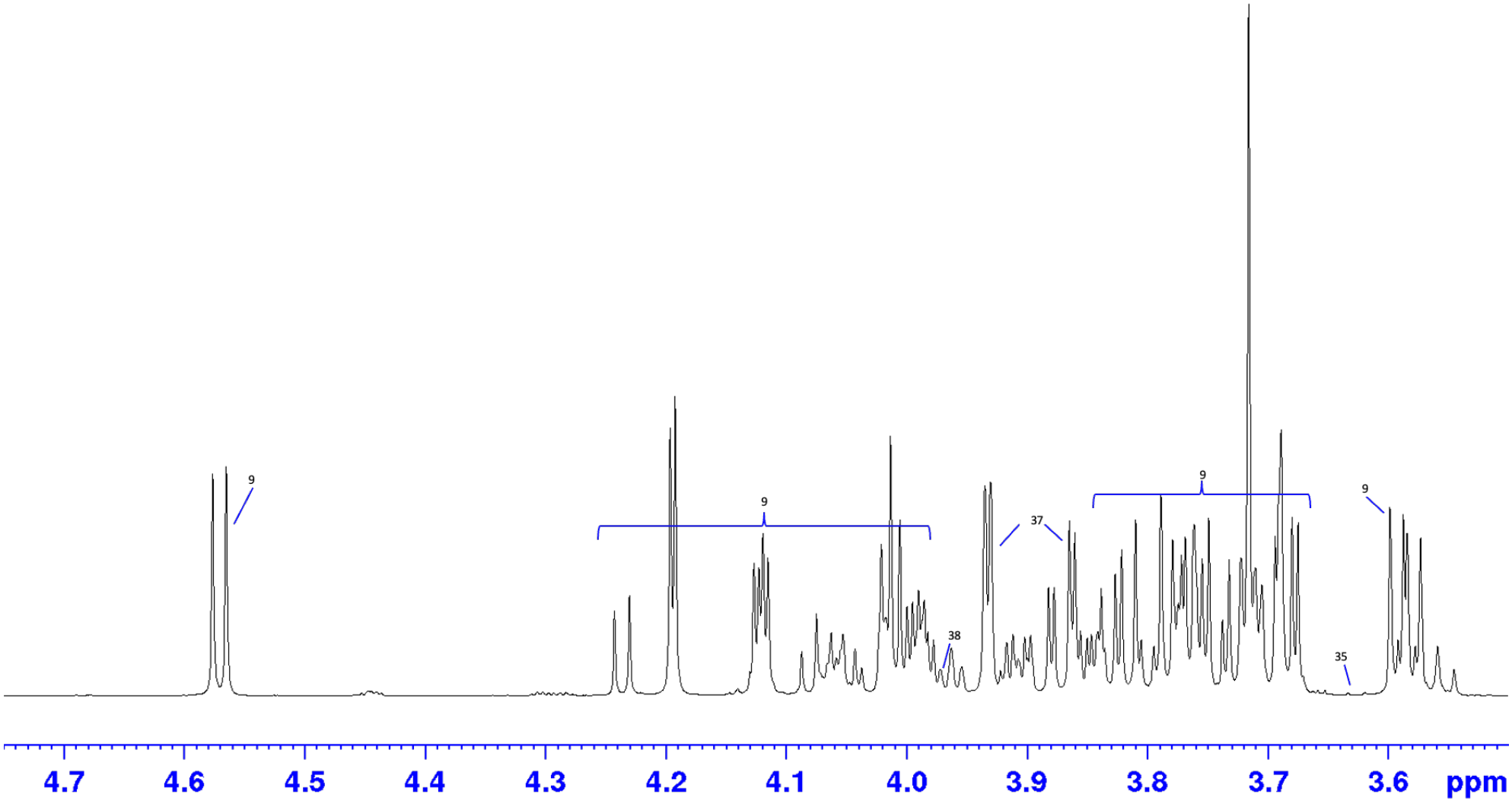




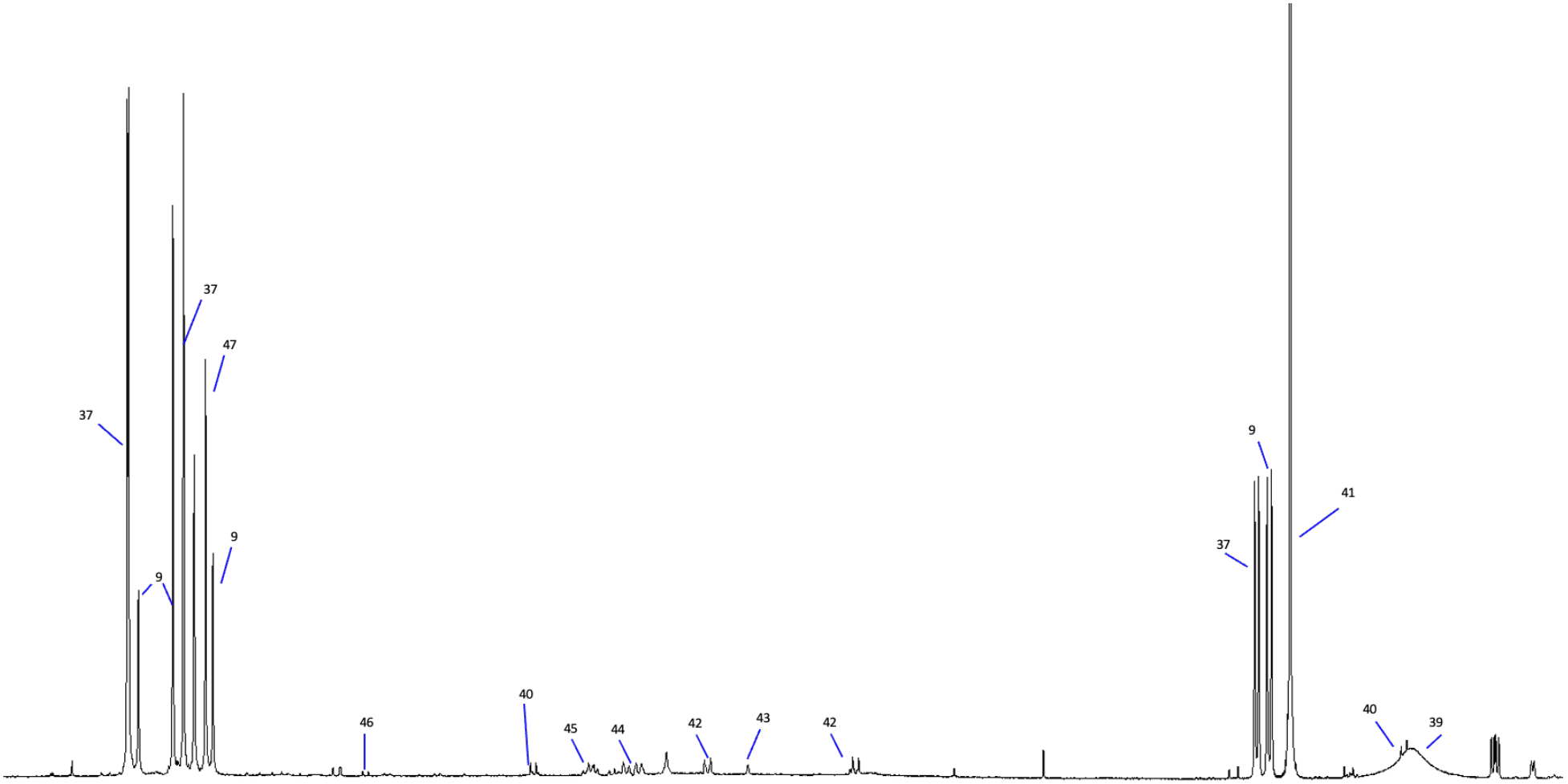
FigureS 10 *Perfusate 1H-NMR spectrum of a random sample from 0 to 9ppm recorded on Bruker Advance HD spectrometer at 700.17 MHz.*



FigureS 11 : *Zoom on spectral zones between 0.5 at 3.5ppm of ^{1}H-NMR spectrum of perfusate sample. The metabolites are numbered accordingly: Spectral zone 1: Isocaproate, Valproate, Butyrate, Panthothenate, isovalerate, 2-Hydroxybutyrate, 2-Methyl-2-oxovalerate, 2-Oxocaproate, 3-Methyladipate, Alloisoleucin; 1: Leucine; 2: Isoleucine; 3: Valine; 4: Propionate; 5: 3-methyl-2-oxovalerate; 6 fumarate; 7: 3-aminoisobutyrate; 8: isopropanol; 9: IGL-1 matrix; 10: 3-hydroxybutyrate; 11: Methylmalonate; 12: valproate; 13: Lactate; 14: 2-Phenylpropionate; 16: Leucine; 17: Lysine; 18: Pyroglutamate; 19: Glutamate; 20: Butyrate; 21: Acetone; 22: succinate; 23: Desaminotyrosine; 24: Aspartate; 25: 2-oxocaproate; 26: Cysteine; 27: Creatinine; 28: Creatine; 29: Ethanolamine; 30: Cystine; 31: Choline; 32: O-Phosphoethanolamine; 33: Trimethylamine-N-oxide; 34: Taurine; 35: Myo-Inositol; 36: Glucose; 37: Inosine; 38: Creatine phosphate; 39: Urea; 40: Uracil; 41: Maleic acid (IS); 42: Desaminotyrosine; 43: Tau-Methylhistidine; 44: 2-phenylpropionate; 45: Phenylalanine; 46: Hippurate; 47: Hypoxantine.*



FigureS 12 *: Zoom on spectral zones between 3.5 at 4.7ppm of ^{1}H-NMR spectrum of perfusate sample. The metabolites are numbered accordingly: Spectral zone 1: Isocaproate, Valproate, Butyrate, Panthothenate, isovalerate, 2-Hydroxybutyrate, 2-Methyl-2-oxovalerate, 2-Oxocaproate, 3-Methyladipate, Alloisoleucin; 1: Leucine; 2: Isoleucine; 3: Valine; 4: Propionate; 5: 3-methyl-2-oxovalerate; 6 fumarate; 7: 3-aminoisobutyrate; 8: isopropanol; 9: IGL-1 matrix; 10: 3-hydroxybutyrate; 11: Methylmalonate; 12: valproate; 13: Lactate; 14: 2-Phenylpropionate; 16: Leucine; 17: Lysine; 18: Pyroglutamate; 19: Glutamate; 20: Butyrate; 21: Acetone; 22: succinate; 23: Desaminotyrosine; 24: Aspartate; 25: 2-oxocaproate; 26: Cysteine; 27: Creatinine; 28: Creatine; 29: Ethanolamine; 30: Cystine; 31: Choline; 32: O-Phosphoethanolamine; 33: Trimethylamine-N-oxide; 34: Taurine; 35: Myo-Inositol; 36: Glucose; 37: Inosine; 38: Creatine phosphate; 39: Urea; 40: Uracil; 41: Maleic acid (IS); 42: Desaminotyrosine; 43: Tau-Methylhistidine; 44: 2-phenylpropionate; 45: Phenylalanine; 46: Hippurate; 47: Hypoxantine.*



FigureS 13: *Zoom on spectral zones between 5.5 at 8.5ppm of ^{1}H-NMR spectrum of perfusate sample. The metabolites are numbered accordingly: Spectral zone 1: Isocaproate, Valproate, Butyrate, Panthothenate, isovalerate, 2-Hydroxybutyrate, 2-Methyl-2-oxovalerate, 2-Oxocaproate, 3-Methyladipate, Alloisoleucin; 1: Leucine; 2: Isoleucine; 3: Valine; 4: Propionate; 5: 3-methyl-2-oxovalerate; 6 fumarate; 7: 3-aminoisobutyrate; 8: isopropanol; 9: IGL-1 matrix; 10: 3-hydroxybutyrate; 11: Methylmalonate; 12: valproate; 13: Lactate; 14: 2-Phenylpropionate; 16: Leucine; 17: Lysine; 18: Pyroglutamate; 19: Glutamate; 20: Butyrate; 21: Acetone; 22: succinate; 23: Desaminotyrosine; 24: Aspartate; 25: 2-oxocaproate; 26: Cysteine; 27: Creatinine; 28: Creatine; 29: Ethanolamine; 30: Cystine; 31: Choline; 32: O-Phosphoethanolamine; 33: Trimethylamine-N-oxide; 34: Taurine; 35: Myo-Inositol; 36: Glucose; 37: Inosine; 38: Creatine phosphate; 39: Urea; 40: Uracil; 41: Maleic acid (IS); 42: Desaminotyrosine; 43: Tau-Methylhistidine; 44: 2-phenylpropionate; 45: Phenylalanine; 46: Hippurate; 47: Hypoxantine.*

Bibliography

Bibliography

1. Giani, A. M., Gallo, G. R., Gianfranceschi, L. & Formenti, G. Long walk to genomics: History and current approaches to genome sequencing and assembly. *Computational and Structural Biotechnology Journal* **18**, 9–19 (2020).
2. Lowe, R., Shirley, N., Bleackley, M., Dolan, S. & Shafee, T. Transcriptomics technologies. *PLoS Comput Biol* **13**, e1005457 (2017).
3. Patterson, S. D. & Aebersold, R. H. Proteomics: the first decade and beyond. *Nat Genet* **33**, 311–323 (2003).
4. Kell, D. B. & Oliver, S. G. The metabolome 18 years on: a concept comes of age. *Metabolomics* **12**, 148 (2016).
5. Jendoubi, T. Approaches to Integrating Metabolomics and Multi-Omics Data: A Primer. *Metabolites* **11**, 184 (2021).
6. Steuer, A. E., Brockbals, L. & Kraemer, T. Metabolomic Strategies in Biomarker Research–New Approach for Indirect Identification of Drug Consumption and Sample Manipulation in Clinical and Forensic Toxicology? *Front. Chem.* **7**, 319 (2019).
7. Kusonmano, K., Vongsangnak, W. & Chumnanpuen, P. Informatics for Metabolomics. in *Translational Biomedical Informatics* (eds. Shen, B., Tang, H. & Jiang, X.) vol. 939 91–115 (Springer Singapore, Singapore, 2016).
8. Trethewey, R. N., Krotzky, A. J. & Willmitzert, L. Metabolic profiling: a rosetta stone for genomics? *Current Opinion in Plant Biology* **2**, 83–85 (1999).
9. Pauling, L., Robinson, A. B., Teranishi, R. & Cary, P. Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography. *Proc. Natl. Acad. Sci. U.S.A.* **68**, 2374–2376 (1971).
10. Nicholson, J. K. & Lindon, J. C. Metabonomics. *Nature* **455**, 1054–1056 (2008).
11. Wishart, D. S. *et al.* HMDB: the Human Metabolome Database. *Nucleic Acids Research* **35**, D521–D526 (2007).
12. Yang, Q. *et al.* Metabolomics biotechnology, applications, and future trends: a systematic review. *RSC Adv.* **9**, 37245–37257 (2019).
13. Letertre, M. P. M., Giraudeau, P. & De Tullio, P. Nuclear Magnetic Resonance Spectroscopy in Clinical Metabolomics and Personalized Medicine: Current Challenges and Perspectives. *Front. Mol. Biosci.* **8**, 698337 (2021).
14. Ibarra-Estrada, E., Soto-Hernández, R. M. & Palma-Tenango, M. Metabolomics as a Tool in Agriculture. in *Metabolomics - Fundamentals and Applications* (ed. Prasain, J. K.) (InTech, 2016). doi:10.5772/66485.
15. Robertson, D. G., Watkins, P. B. & Reily, M. D. Metabolomics in Toxicology: Preclinical and Clinical Applications. *Toxicological Sciences* **120**, S146–S170 (2011).
16. Jones, O. A. H. *et al.* Metabolomics and its use in ecology: Metabolomics in Ecology. *Austral Ecology* **38**, 713–720 (2013).
17. Ramautar, R., Berger, R., Van Der Greef, J. & Hankemeier, T. Human metabolomics: strategies to understand biology. *Current Opinion in Chemical Biology* **17**, 841–846 (2013).
18. Roberts, L. D., Souza, A. L., Gerszten, R. E. & Clish, C. B. Targeted Metabolomics. *Current Protocols in Molecular Biology* **98**, (2012).
19. Gowda, G. A. N. & Djukovic, D. Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. in *Mass Spectrometry in Metabolomics* (ed. Raftery, D.) vol. 1198 3–12 (Springer New York, New York, NY, 2014).
20. Souza, A. L. & Patti, G. J. A Protocol for Untargeted Metabolomic Analysis: From Sample Preparation to Data Processing. in *Mitochondrial Medicine* (eds. Weissig, V. & Edeas, M.) vol. 2276 357–382 (Springer US, New York, NY, 2021).
21. Ivanisevic, J. & Want, E. J. From Samples to Insights into Metabolism: Uncovering

Biologically Relevant Information in LC-HRMS Metabolomics Data. *Metabolites* **9**, 308 (2019).
22. Khamis, M. M., Adamko, D. J. & El-Aneed, A. Mass spectrometric based approaches in urine metabolomics and biomarker discovery: MASS SPECTROMETRY IN URINE METABOLOMICS AND BIOMARKER DISCOVERY. *Mass Spec Rev* **36**, 115–134 (2017).
23. San-Martin, B. S. D. *et al.* Metabolomics as a potential tool for the diagnosis of growth hormone deficiency (GHD): a review. *Archives of Endocrinology and Metabolism* (2020) doi:10.20945/2359-3997000000300.
24. Bayle, M.-L. *et al.* Semi-targeted metabolomic approaches to validate potential markers of health for micronutrients: analytical perspectives. *Metabolomics* **8**, 1114–1129 (2012).
25. Billet, K. *et al.* Semi-Targeted Metabolomics to Validate Biomarkers of Grape Downy Mildew Infection Under Field Conditions. *Plants* **9**, 1008 (2020).
26. Gao, Y. *et al.* Development of simultaneous targeted metabolite quantification and untargeted metabolomics strategy using dual-column liquid chromatography coupled with tandem mass spectrometry. *Analytica Chimica Acta* **1037**, 369–379 (2018).
27. Matthieu Schoumacher. New translational metabolomics exploration of Age-related macular degeneration: improving risk assessment and « real-life » patient's follow-up.
28. Hoffmann, E. de & Stroobant, V. *Mass Spectrometry: Principles and Applications*. (J. Wiley, Chichester, West Sussex, England ; Hoboken, NJ, 2007).
29. Gika, H. G., Theodoridis, G. A., Plumb, R. S. & Wilson, I. D. Current practice of liquid chromatography–mass spectrometry in metabolomics and metabonomics. *Journal of Pharmaceutical and Biomedical Analysis* **87**, 12–25 (2014).
30. Da Silva, R. R., Dorrestein, P. C. & Quinn, R. A. Illuminating the dark matter in metabolomics. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12549–12550 (2015).
31. Zhou, Z. *et al. Metabolite Annotation from Knowns to Unknowns through Knowledge-Guided Multi-Layer Metabolic Network*. http://biorxiv.org/lookup/doi/10.1101/2022.06.02.494523 (2022) doi:10.1101/2022.06.02.494523.
32. Markley, J. L. *et al.* The future of NMR-based metabolomics. *Current Opinion in Biotechnology* **43**, 34–40 (2017).
33. Wishart, D. S. *et al.* NMR and Metabolomics—A Roadmap for the Future. *Metabolites* **12**, 678 (2022).
34. Bouatra, S. *et al.* The Human Urine Metabolome. *PLoS ONE* **8**, e73076 (2013).
35. Walker, L. R. *et al.* Unambiguous metabolite identification in high-throughput metabolomics by hybrid 1D ^{1} H NMR/ESI MS ^{1} approach: Hybrid 1D ^{1} H NMR/ESI MS ^{1} metabolomics method. *Magn. Reson. Chem.* **54**, 998–1003 (2016).
36. Jeppesen, M. J. & Powers, R. Multiplatform untargeted metabolomics. *Magnetic Reson in Chemistry* mrc.5350 (2023) doi:10.1002/mrc.5350.
37. Pochapsky, T. C. & Pochapsky, S. S. Nuclear Magnetic Resonance Spectroscopy. in *Molecular Biophysics for the Life Sciences* (eds. Allewell, N., Narhi, L. O. & Rayment, I.) 113–173 (Springer New York, New York, NY, 2013). doi:10.1007/978-1-4614-8548-3_5.
38. Emsley, J. W. & Feeney, J. Forty years of Progress in Nuclear Magnetic Resonance Spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy* **50**, 179–198 (2007).
39. Spiess, H. W. Nuclear magnetic resonance. in *Structure and Properties of Oriented Polymers* (ed. Ward, I. M.) 234–268 (Springer Netherlands, Dordrecht, 1997). doi:10.1007/978-94-011-5844-2_5.
40. Rabi, I. I., Zacharias, J. R., Millman, S. & Kusch, P. A New Method of Measuring Nuclear Magnetic Moment. *Phys. Rev.* **53**, 318–318 (1938).
41. Filler, A. & Filler, A. The History, Development and Impact of Computed Imaging in Neurological Diagnosis and Neurosurgery: CT, MRI, and DTI. *Nature Precedings* (2009)

doi:10.1038/npre.2009.3267.
42. Antcliffe, D. & Gordon, A. C. Metabonomics and intensive care. *Crit Care* **20**, 68 (2016).
43. Rankin, N. J. *et al.* The emergence of proton nuclear magnetic resonance metabolomics in the cardiovascular arena as viewed from a clinical perspective. *Atherosclerosis* **237**, 287–300 (2014).
44. Zimdahl, A. *Pharmacogenetic Studies of Thiopurine Methyltransferase Genotype-Phenotype Concordance and Effect of Methotrexate on Thiopurine Metabolism*. vol. 1723 (Linköping University Electronic Press, Linköping, 2020).
45. Zheng, X. *et al.* Coupling Front-End Separations, Ion Mobility Spectrometry, and Mass Spectrometry For Enhanced Multidimensional Biological and Environmental Analyses. *Annual Rev. Anal. Chem.* **10**, 71–92 (2017).
46. Orna, M. V. Major Analytical Techniques Based on Color: Volumetric Analysis; Chromatography; Spectroscopy; Color Measurement. in *The Chemical History of Color* 93–110 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013). doi:10.1007/978-3-642-32642-4_6.
47. Injeti, N. Test. *Fut J. Pharm & H. Sci* **3**, 69–79 (2023).
48. Bedair, M. & Sumner, L. W. Current and emerging mass-spectrometry technologies for metabolomics. *TrAC Trends in Analytical Chemistry* **27**, 238–250 (2008).
49. Tang, D.-Q., Zou, L., Yin, X.-X. & Ong, C. N. HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS: HILIC-MS FOR METABOLOMICS. *Mass Spec Rev* **35**, 574–600 (2016).
50. Van Deemter, J. J., Zuiderweg, F. J. & Klinkenberg, A. Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chemical Engineering Science* **50**, 3869–3882 (1995).
51. Awad, H., Khamis, M. M. & El-Aneed, A. Mass Spectrometry, Review of the Basics: Ionization. *Applied Spectroscopy Reviews* **50**, 158–175 (2015).
52. Famiglini, G., Palma, P., Termopoli, V. & Cappiello, A. The history of electron ionization in LC-MS, from the early days to modern technologies: A review. *Analytica Chimica Acta* **1167**, 338350 (2021).
53. Yamashita, M. & Fenn, J. B. Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* **88**, 4451–4459 (1984).
54. Paul Gates. *Esi Source*. https://www.chm.bris.ac.uk/ms/esi-ionisation.xhtml.
55. Banerjee, S. & Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry* **2012**, 1–40 (2012).
56. Tugizimana, F., Steenkamp, P. A., Piater, L. A. & Dubery, I. A. Mass spectrometry in untargeted liquid chromatography/mass spectrometry metabolomics: Electrospray ionisation parameters and global coverage of the metabolome. *Rapid Commun Mass Spectrom* **32**, 121–132 (2018).
57. Kachman, M. *et al.* Deep annotation of untargeted LC-MS metabolomics data with *Binner*. *Bioinformatics* **36**, 1801–1806 (2020).
58. Ntai, I. & Kelleher, N. L. Approaches for Natural Product Detection and Structural Elucidation Using Mass Spectrometry with High Mass Accuracy. in *Methodologies for Metabolomics* (eds. Lutz, N. W., Sweedler, J. V. & Wevers, R. A.) 174–184 (Cambridge University Press, 2013). doi:10.1017/CBO9780511996634.010.
59. Paul, W. Electromagnetic Traps for Charged and Neutral Particles(Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **29**, 739–748 (1990).
60. TRACES. Quadrupoles: How do they work?
61. Kingdon, K. H. A Method for the Neutralization of Electron Space Charge by Positive

Ionization at Very Low Gas Pressures. *Phys. Rev.* **21**, 408–418 (1923).
62. Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **72**, 1156–1162 (2000).
63. Hardman, M. & Makarov, A. A. Interfacing the Orbitrap Mass Analyzer to an Electrospray Ion Source. *Anal. Chem.* **75**, 1699–1705 (2003).
64. Hu, Q. *et al.* The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* **40**, 430–443 (2005).
65. Scigelova, M. & Makarov, A. Orbitrap Mass Analyzer – Overview and Applications in Proteomics. *Proteomics* **6**, 16–21 (2006).
66. Heiles, S. Advanced tandem mass spectrometry in metabolomics and lipidomics—methods and applications. *Anal Bioanal Chem* **413**, 5927–5948 (2021).
67. Kulyyassov, A., Fresnais, M. & Longuespée, R. Targeted liquid chromatography-tandem mass spectrometry analysis of proteins: Basic principles, applications, and perspectives. *Proteomics* **21**, 2100153 (2021).
68. Lee, D. Y., Bowen, B. P. & Northen, T. R. Mass spectrometry—based metabolomics, analysis of metabolite-protein interactions, and imaging. *BioTechniques* **49**, 557–565 (2010).
69. Sarandi, E., Georgaki, S., Tsoukalas, D. & Tsatsakis, A. M. Metabolomics methodology and workflow: challenges and future prospects. in *Toxicological Risk Assessment and Multi-System Health Impacts from Exposure* 285–293 (Elsevier, 2021). doi:10.1016/B978-0-323-85215-9.00030-1.
70. Contreras, A. V., Cocom-Chan, B., Hernandez-Montes, G., Portillo-Bobadilla, T. & Resendis-Antonio, O. Host-Microbiome Interaction and Cancer: Potential Application in Precision Medicine. *Front. Physiol.* **7**, (2016).
71. Lee, A. Y., Troisi, J. & Symes, S. J. K. Experimental design in metabolomics. in *Metabolomics Perspectives* 27–61 (Elsevier, 2022). doi:10.1016/B978-0-323-85062-9.00002-7.
72. Smith, L. *et al.* Important Considerations for Sample Collection in Metabolomics Studies with a Special Focus on Applications to Liver Functions. *Metabolites* **10**, 104 (2020).
73. Van Der Hooft, J. J. J., De Vos, R. C. H., Ridder, L., Vervoort, J. & Bino, R. J. Structural elucidation of low abundant metabolites in complex sample matrices. *Metabolomics* **9**, 1009–1018 (2013).
74. Ivey, K. L. *et al.* Identifying the metabolomic fingerprint of high and low flavonoid consumers. *J Nutr Sci* **6**, e34 (2017).
75. Dallmann, R., Viola, A. U., Tarokh, L., Cajochen, C. & Brown, S. A. The human circadian metabolome. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 2625–2629 (2012).
76. Slupsky, C. M. *et al.* Investigations of the Effects of Gender, Diurnal Variation, and Age in Human Urinary Metabolomic Profiles. *Anal. Chem.* **79**, 6995–7004 (2007).
77. Thongboonkerd, V. & Saetun, P. Bacterial Overgrowth Affects Urinary Proteome Analysis: Recommendation for Centrifugation, Temperature, Duration, and the Use of Preservatives during Sample Collection. *J. Proteome Res.* **6**, 4173–4181 (2007).
78. Bernini, P. *et al.* Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. *J Biomol NMR* **49**, 231–243 (2011).
79. K Trivedi, D., Jones, H., Shah, A. & K Iles, R. Development of Zwitterionic Hydrophilic Liquid Chromatography (ZIC®HILIC-MS) Metabolomics Method for Shotgun Analysis of Human Urine. *J Chromatogr Sep Tech* **03**, (2012).
80. Saude, E. J. & Sykes, B. D. Urine stability for metabolomic studies: effects of preparation and storage. *Metabolomics* **3**, 19–27 (2007).
81. Emwas, A.-H. *et al.* NMR Spectroscopy for Metabolomics Research. *Metabolites* **9**, 123 (2019).
82. Beckonert, O. *et al.* Metabolic profiling, metabolomic and metabonomic procedures for

NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat Protoc* **2**, 2692–2703 (2007).
83. Féraud, B. *et al.* Two data pre-processing workflows to facilitate the discovery of biomarkers by 2D NMR metabolomics. *Metabolomics* **15**, 63 (2019).
84. Guennec, A. L., Giraudeau, P. & Caldarelli, S. Evaluation of Fast 2D NMR for Metabolomics. *Anal. Chem.* **86**, 5946–5954 (2014).
85. Guo, J., Yu, H., Xing, S. & Huan, T. Addressing big data challenges in mass spectrometry-based metabolomics. *Chem. Commun.* **58**, 9979–9990 (2022).
86. Guo, J. & Huan, T. Evaluation of significant features discovered from different data acquisition modes in mass spectrometry-based untargeted metabolomics. *Analytica Chimica Acta* **1137**, 37–46 (2020).
87. Krasny, L. & Huang, P. H. Data-independent acquisition mass spectrometry (DIA-MS) for proteomic applications in oncology. *Mol. Omics* **17**, 29–42 (2021).
88. Guo, J. & Huan, T. Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography–Mass Spectrometry Based Untargeted Metabolomics. *Anal. Chem.* **92**, 8072–8080 (2020).
89. Giacomoni, F. *et al.* Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. *Bioinformatics* **31**, 1493–1495 (2015).
90. Jacob, D., Deborde, C., Lefebvre, M., Maucourt, M. & Moing, A. NMRProcFlow: a graphical and interactive tool dedicated to 1D spectra processing for NMR-based metabolomics. *Metabolomics* **13**, 36 (2017).
91. Martin, M. *et al.* PepsNMR for 1H NMR metabolomic data pre-processing. *Analytica Chimica Acta* **1019**, 1–13 (2018).
92. Simon, B. & Köstler, H. Improving the sensitivity of FT-NMR spectroscopy by apodization weighted sampling. *J Biomol NMR* **73**, 155–165 (2019).
93. Vu, T. & Laukens, K. Getting Your Peaks in Line: A Review of Alignment Methods for NMR Spectral Data. *Metabolites* **3**, 259–276 (2013).
94. Puchades-Carrasco, L., Palomino-Schätzlein, M., Pérez-Rambla, C. & Pineda-Lucena, A. Bioinformatics tools for the analysis of NMR metabolomics studies focused on the identification of clinically relevant biomarkers. *Brief Bioinform* **17**, 541–552 (2016).
95. Schmid, N. *et al.* Deconvolution of 1D NMR spectra: A deep learning-based approach. *Journal of Magnetic Resonance* **347**, 107357 (2023).
96. Cassiède, M. *et al.* Comparison of computational approaches for identification and quantification of urinary metabolites in ^{1}H NMR spectra. *Anal. Methods* **10**, 2129–2137 (2018).
97. Weljie, A. M., Newton, J., Mercier, P., Carlson, E. & Slupsky, C. M. Targeted Profiling: Quantitative Analysis of ^{1}H NMR Metabolomics Data. *Anal. Chem.* **78**, 4430–4442 (2006).
98. Embade, N. *et al.* NMR-based newborn urine screening for optimized detection of inherited errors of metabolism. *Sci Rep* **9**, 13067 (2019).
99. Spraul, M. *et al.* Mixture analysis by NMR as applied to fruit juice quality control. *Magn Reson Chem* **47 Suppl 1**, S130-137 (2009).
100. Hao, J. *et al.* Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN. *Nat Protoc* **9**, 1416–1427 (2014).
101. Cañueto, D., Gómez, J., Salek, R. M., Correig, X. & Cañellas, N. rDolphin: a GUI R package for proficient automatic profiling of 1D 1H-NMR spectra of study datasets. *Metabolomics* **14**, 24 (2018).
102. Lefort, G. *et al.* Joint Automatic Metabolite Identification and Quantification of a Set of ^{1}H NMR Spectra. *Anal. Chem.* **93**, 2861–2870 (2021).
103. Röhnisch, H. E. *et al.* AQuA: An Automated Quantification Algorithm for High-Throughput NMR-Based Metabolomics and Its Application in Human Plasma. *Anal. Chem.* **90**, 2095–2102 (2018).

104. Castillo, S., Gopalacharyulu, P., Yetukuri, L. & Orešič, M. Algorithms and tools for the preprocessing of LC–MS metabolomics data. *Chemometrics and Intelligent Laboratory Systems* **108**, 23–32 (2011).
105. Schmid, R. *et al.* Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat Biotechnol* **41**, 447–449 (2023).
106. Xia, J., Psychogios, N., Young, N. & Wishart, D. S. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Research* **37**, W652–W660 (2009).
107. Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **78**, 779–787 (2006).
108. Karaman, I. Preprocessing and Pretreatment of Metabolomics Data for Statistical Analysis. in *Metabolomics: From Fundamentals to Clinical Applications* (ed. Sussulini, A.) vol. 965 145–161 (Springer International Publishing, Cham, 2017).
109. Petera, M., Martin, J.-F., Corguillé, G. L., & Workflow4Metabolomics core team,. Mass spectrometry: LC-MS preprocessing with XCMS (Galaxy Training Materials).
110. Wehrens, R. *et al.* Improved batch correction in untargeted MS-based metabolomics. *Metabolomics* **12**, 88 (2016).
111. Reinhold, D., Pielke-Lombardo, H., Jacobson, S., Ghosh, D. & Kechris, K. Pre-analytic Considerations for Mass Spectrometry-Based Untargeted Metabolomics Data. *Methods Mol Biol* **1978**, 323–340 (2019).
112. Van Den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K. & Van Der Werf, M. J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **7**, 142 (2006).
113. Anwardeen, N. R., Diboun, I., Mokrab, Y., Althani, A. A. & Elrayess, M. A. Statistical methods and resources for biomarker discovery using metabolomics. *BMC Bioinformatics* **24**, 250 (2023).
114. MetaboAnalyst. https://www.metaboanalyst.ca/.
115. BioStatFlow 2.9.5 - Statistical Analysis Workflow for 'Omics' Data. http://biostatflow.org/session/ecf8aa3289a1b0c354d6fd3a05126632.
116. Torell, F. Multivariate data analysis of metabolomic multi-tissue samples. (Ume Univeritet).
117. Erdogan Taskesen, Erdogan Taskesen, & Erdogan Taskesen. Outlier Detection Using Principal Component Analysis and Hotelling's T2 and SPE/DmodX Methods.
118. Susanne Wiklund. Multivariate Data Analysis for Omics.
119. Tugizimana, F., Steenkamp, P., Piater, L. & Dubery, I. A Conversation on Data Mining Strategies in LC-MS Untargeted Metabolomics: Pre-Processing and Pre-Treatment Steps. *Metabolites* **6**, 40 (2016).
120. Gromski, P. S. *et al.* A tutorial review: Metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta* **879**, 10–23 (2015).
121. Gavaghan, C. L., Wilson, I. D. & Nicholson, J. K. Physiological variation in metabolic phenotyping and functional genomic studies: use of orthogonal signal correction and PLS-DA. *FEBS Letters* **530**, 191–196 (2002).
122. Alonso, A., Marsal, S. & Julià, A. Analytical methods in untargeted metabolomics: state of the art in 2015. *Front Bioeng Biotechnol* **3**, 23 (2015).
123. Westerhuis, J. A. *et al.* Assessment of PLSDA cross validation. *Metabolomics* **4**, 81–89 (2008).
124. Jared Wilber. *THE PERMUTATION TEST*. https://www.jwilber.me/permutationtest/.
125. Singh, P., Singh, N., Singh, K. K. & Singh, A. Diagnosing of disease using machine learning. in *Machine Learning and the Internet of Medical Things in Healthcare* 89–111

(Elsevier, 2021). doi:10.1016/B978-0-12-821229-5.00003-3.
126. Percival, B., Gibson, M., Leenders, J., Wilson, P. B. & Grootveld, M. Univariate and Multivariate Statistical Approaches to the Analysis and Interpretation of NMR-based Metabolomics Datasets of Increasing Complexity. in *Computational Techniques for Analytical Chemistry and Bioanalysis* (eds. Wilson, P. B. & Grootveld, M.) 1–40 (The Royal Society of Chemistry, 2020). doi:10.1039/9781788015882-00001.
127. Cambiaghi, A., Ferrario, M. & Masseroli, M. Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration. *Brief Bioinform* bbw031 (2016) doi:10.1093/bib/bbw031.
128. Spiteri, M. *et al.* Data fusion between high resolution 1H-NMR and mass spectrometry: a synergetic approach to honey botanical origin characterization. *Anal Bioanal Chem* **408**, 4389–4401 (2016).
129. Boccard, J. & Rudaz, S. Harnessing the complexity of metabolomic data with chemometrics: Metabolomic data analysis with chemometrics. *J. Chemometrics* **28**, 1–9 (2014).
130. Silvestri, M. *et al.* A mid level data fusion strategy for the Varietal Classification of Lambrusco PDO wines. *Chemometrics and Intelligent Laboratory Systems* **137**, 181–189 (2014).
131. Doeswijk, T. G., Smilde, A. K., Hageman, J. A., Westerhuis, J. A. & Van Eeuwijk, F. A. On the increase of predictive performance with high-level data fusion. *Analytica Chimica Acta* **705**, 41–47 (2011).
132. Goodacre, R. *et al.* Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics* **3**, 231–241 (2007).
133. Psychogios, N. *et al.* The Human Serum Metabolome. *PLoS ONE* **6**, e16957 (2011).
134. Wishart, D. S. *et al.* The human cerebrospinal fluid metabolome. *Journal of Chromatography B* **871**, 164–173 (2008).
135. Nagana Gowda, G. A. & Raftery, D. Can NMR solve some significant challenges in metabolomics? *Journal of Magnetic Resonance* **260**, 144–160 (2015).
136. Mahieu, N. G. & Patti, G. J. Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. *Anal Chem* **89**, 10397–10406 (2017).
137. Viant, M. R., Kurland, I. J., Jones, M. R. & Dunn, W. B. How close are we to complete annotation of metabolomes? *Current Opinion in Chemical Biology* **36**, 64–69 (2017).
138. Chaleckis, R., Meister, I., Zhang, P. & Wheelock, C. E. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. *Current Opinion in Biotechnology* **55**, 44–50 (2019).
139. De Jonge, N. F. *et al.* Good practices and recommendations for using and benchmarking computational metabolomics metabolite annotation tools. *Metabolomics* **18**, 103 (2022).
140. Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R. & Neumann, S. CAMERA: An Integrated Strategy for Compound Spectra Extraction and Annotation of Liquid Chromatography/Mass Spectrometry Data Sets. *Anal. Chem.* **84**, 283–289 (2012).
141. Leao, T. F. *et al.* Quick-start infrastructure for untargeted metabolomics analysis in GNPS. *Nat Metab* **3**, 880–882 (2021).
142. Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**, 27–30 (2000).
143. Xia, J. & Wishart, D. S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat Protoc* **6**, 743–760 (2011).
144. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498–2504 (2003).
145. Cottret, L. *et al.* MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Research* **38**, W132–W137 (2010).

146. Jacob, M., Lopata, A. L., Dasouki, M. & Abdel Rahman, A. M. Metabolomics toward personalized medicine. *Mass Spectrometry Reviews* **38**, 221–238 (2019).
147. Wishart, D. S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov* **15**, 473–484 (2016).
148. K. Trivedi, D., A. Hollywood, K. & Goodacre, R. Metabolomics for the masses: The future of metabolomics in a personalized world. *European Journal of Molecular & Clinical Medicine* **3**, 294 (2017).
149. Letertre, M. P. M., Giraudeau, P. & De Tullio, P. Nuclear Magnetic Resonance Spectroscopy in Clinical Metabolomics and Personalized Medicine: Current Challenges and Perspectives. *Front. Mol. Biosci.* **8**, 698337 (2021).
150. Ashrafian, H. *et al.* Metabolomics: The Stethoscope for the Twenty-First Century. *Med Princ Pract* **30**, 301–310 (2021).
151. Turkoglu, O. *et al.* Metabolomics of biomarker discovery in ovarian cancer: a systematic review of the current literature. *Metabolomics* **12**, 60 (2016).
152. Yu, L., Li, K. & Zhang, X. Next-generation metabolomics in lung cancer diagnosis, treatment and precision medicine: mini review. *Oncotarget* **8**, 115774–115786 (2017).
153. Troisi, J. *et al.* Metabolomic Signature of Endometrial Cancer. *J. Proteome Res.* **17**, 804–812 (2018).
154. AminoIndex® | Innovation in Action | Innovation. *Ajinomoto Group Global Website - Eat Well, Live Well.* https://www.ajinomoto.com/innovation/action/aminoindex.
155. Troisi, J. *et al.* A screening test proposal for congenital defects based on maternal serum metabolomics profile. *American Journal of Obstetrics and Gynecology* **228**, 342.e1-342.e12 (2023).
156. AXINON®GFR (NMR). *AXINON®GFR (NMR)* https://www.numares.us/en/products/nephrology/axinonreggfr-nmr.
157. Meeusen, J. W., Stämmler, F., Dasari, S., Schiffer, E. & Lieske, J. C. Serum myo-inositol and valine improve metabolomic-based estimated glomerular filtration rate among kidney transplant recipients. *Front. Med.* **9**, 988989 (2022).
158. Fuhrmann, M. *et al.* Analytical Validation of GFRNMR: A Blood-Based Multiple Biomarker Assay for Accurate Estimation of Glomerular Filtration Rate. *Diagnostics* **12**, 1120 (2022).
159. Soriano, R. M., Penfold, D. & Leslie, S. W. Anatomy, Abdomen and Pelvis: Kidneys. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2023).
160. Ogobuiro, I. & Tuma, F. Physiology, Renal. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2023).
161. Contributor, N. T. Renal system 1: the anatomy and physiology of the kidneys. *Nursing Times* https://www.nursingtimes.net/clinical-archive/renal/renal-system-1-the-anatomy-and-physiology-of-the-kidneys-23-01-2023/ (2023).
162. Matsushita, K. *et al.* Epidemiology and risk of cardiovascular disease in populations with chronic kidney disease. *Nat Rev Nephrol* **18**, 696–707 (2022).
163. Webster, A. C., Nagler, E. V., Morton, R. L. & Masson, P. Chronic Kidney Disease. *The Lancet* **389**, 1238–1252 (2017).
164. Vaidya, S. R. & Aeddula, N. R. Chronic Renal Failure. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2023).
165. Levey, A. S. & Coresh, J. Chronic kidney disease. *The Lancet* **379**, 165–180 (2012).
166. Matovinović, M. S. 1. Pathophysiology and Classification of Kidney Diseases. *EJIFCC* **20**, 2–11 (2009).
167. Tan, S. & Merchant, J. Joseph Murray (1919–2012): First transplant surgeon. *smedj* **60**, 162–163 (2019).
168. Wolfe, R. A. *et al.* Comparison of Mortality in All Patients on Dialysis, Patients on

Dialysis Awaiting Transplantation, and Recipients of a First Cadaveric Transplant. *N Engl J Med* **341**, 1725–1730 (1999).
169. Tonelli, M. *et al.* Systematic Review: Kidney Transplantation Compared With Dialysis in Clinically Relevant Outcomes. *American Journal of Transplantation* **11**, 2093–2109 (2011).
170. Marroquin, C. E. Patient Selection for Kidney Transplant. *Surgical Clinics of North America* **99**, 1–35 (2019).
171. Pham, P.-T., Pham, P.-A., Pham, P.-C., Parikh, S. & Danovitch, G. Evaluation of Adult Kidney Transplant Candidates: KIDNEY TRANSPLANT CANDIDATE EVALUATION. *Seminars in Dialysis* **23**, 595–605 (2010).
172. Abramyan, S. & Hanlon, M. Kidney Transplantation. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2023).
173. Port, F. K. *et al.* Donor characteristics associated with reduced graft survival: an approach to expanding the pool of kidney donors1: *Transplantation* **74**, 1281–1286 (2002).
174. Baskin-Bey, E. S., Kremers, W., Stegall, M. D. & Nyberg, S. L. United Network for Organ Sharing's expanded criteria donors: is stratification useful?*. *Clin Transplant* **19**, 406–412 (2005).
175. Schold, J. D., Kaplan, B., Baliga, R. S. & Meier-Kriesche, H.-U. The Broad Spectrum of Quality in Deceased Donor Kidneys. *American Journal of Transplantation* **5**, 757–765 (2005).
176. Rao, P. S. *et al.* A Comprehensive Risk Quantification Score for Deceased Donor Kidneys: The Kidney Donor Risk Index. *Transplantation* **88**, 231–236 (2009).
177. A Guide to Calculating and Interpreting the Kidney Donor Profle Index (KDPI).
178. Prunster, J. *et al.* Kidney Donor Profile Index and allograft outcomes: interactive effects of estimated post-transplant survival score and ischaemic time. *Clinical Kidney Journal* **16**, 473–483 (2023).
179. Ponticelli, C. E. The impact of cold ischemia time on renal transplant outcome. *Kidney International* **87**, 272–275 (2015).
180. Van Der Vliet, J. A. & Warlé, M. C. The need to reduce cold ischemia time in kidney transplantation. *Current Opinion in Organ Transplantation* **18**, 174–178 (2013).
181. Khan, T. F. T., Ahmad, N., Serageldeen, A. S. & Fourtounas, K. Implantation Warm Ischemia Time in Kidney Transplant Recipients: Defining Its Limits and Impact on Early Graft Function. *Ann Transplant* **24**, 432–438 (2019).
182. Tennankore, K. K., Kim, S. J., Alwayn, I. P. J. & Kiberd, B. A. Prolonged warm ischemia time is associated with graft failure and mortality after kidney transplantation. *Kidney International* **89**, 648–658 (2016).
183. Chen, Y. *et al.* Preservation Solutions for Kidney Transplantation: History, Advances and Mechanisms. *Cell Transplant* **28**, 1472–1489 (2019).
184. Eugene, M. *et al.* Beneficial effects of a low-potassium+ and polyethylene glycol solution on renal function and structure during 48-hour cold storage preservation. *Transplantation Proceedings* **29**, 2360–2362 (1997).
185. Habran, M., De Beule, J. & Jochmans, I. IGL-1 preservation solution in kidney and pancreas transplantation: A systematic review. *PLoS One* **15**, e0231019 (2020).
186. Badet, L. *et al.* Kidney preservation with IGL-1 solution: A preliminary report. *Transplantation Proceedings* **37**, 308–311 (2005).
187. De Beule, J. *et al.* The effect of IGL-1 preservation solution on outcome after kidney transplantation: A retrospective single-center analysis. *American Journal of Transplantation* **21**, 830–837 (2021).
188. McLaren, A. J. & Friend, P. J. Trends in organ preservation. *Transplant Int* **16**, 701–708 (2003).
189. Metcalfe, M. S. *et al.* A CASE-CONTROL COMPARISON OF THE RESULTS OF

RENAL TRANSPLANTATION FROM HEART-BEATING AND NON-HEART-BEATING DONORS: *Transplantation* **71**, 1556–1559 (2001).
190. Matsuoka, L. *et al.* Pulsatile Perfusion Reduces the Incidence of Delayed Graft Function in Expanded Criteria Donor Kidney Transplantation. *American Journal of Transplantation* **6**, 1473–1478 (2006).
191. Naik, R. H. & Shawar, S. H. Renal Transplantation Rejection. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2023).
192. Foroutan, F. *et al.* Risk Factors for 1-Year Graft Loss After Kidney Transplantation: Systematic Review and Meta-Analysis. *CJASN* **14**, 1642–1650 (2019).
193. Oweira, H. *et al.* Risk Factors of Rejection in Renal Transplant Recipients: A Narrative Review. *JCM* **11**, 1392 (2022).
194. Mannon, R. B. Delayed Graft Function: The AKI of Kidney Transplantation. *Nephron* **140**, 94–98 (2018).
195. Ojo, A. O., Wolfe, R. A., Held, P. J., Port, F. K. & Schmouder, R. L. DELAYED GRAFT FUNCTION: RISK FACTORS AND IMPLICATIONS FOR RENAL ALLOGRAFT SURVIVAL1. *Transplantation* **63**, 968 (1997).
196. Tapiawala, S. N. *et al.* Delayed graft function and the risk for death with a functioning graft. *J Am Soc Nephrol* **21**, 153–161 (2010).
197. Basiri, A. *et al.* Living or deceased-donor kidney transplant: the role of psycho-socioeconomic factors and outcomes associated with each type of transplant. *Int J Equity Health* **19**, 79 (2020).
198. Fakhr Yasseri, A. M. *et al.* Living versus deceased kidney transplantation: Comparison of complications. *Urologia* **88**, 185–189 (2021).
199. Ponticelli, C., Reggiani, F. & Moroni, G. Delayed Graft Function in Kidney Transplant: Risk Factors, Consequences and Prevention Strategies. *JPM* **12**, 1557 (2022).
200. Debout, A. *et al.* Each additional hour of cold ischemia time significantly increases the risk of graft failure and mortality following renal transplantation. *Kidney International* **87**, 343–349 (2015).
201. Wight, J. P., Chilcott, J. B., Holmes, M. W. & Brewer, N. Pulsatile machine perfusion vs. cold storage of kidneys for transplantation: a rapid and systematic review: **Review of kidney preservation systems**. *Clinical Transplantation* **17**, 293–307 (2003).
202. Reeves, P. B. & Mc Causland, F. R. Mechanisms, Clinical Implications, and Treatment of Intradialytic Hypotension. *CJASN* **13**, 1297–1303 (2018).
203. Cusumano, A. M., Tzanno-Martins, C. & Rosa-Diez, G. J. The Glomerular Filtration Rate: From the Diagnosis of Kidney Function to a Public Health Tool. *Front. Med.* **8**, 769335 (2021).
204. Levey, A. & Inker, L. Assessment of Glomerular Filtration Rate in Health and Disease: A State of the Art Review: State of the Art Review for Clinical Pharmacology and Therapeutics. *Clin. Pharmacol. Ther.* **102**, 405–419 (2017).
205. Inker, L. A. & Titan, S. Measurement and Estimation of GFR for Use in Clinical Practice: Core Curriculum 2021. *American Journal of Kidney Diseases* **78**, 736–749 (2021).
206. Glassock, R. J., Warnock, D. G. & Delanaye, P. The global burden of chronic kidney disease: estimates, variability and pitfalls. *Nat Rev Nephrol* **13**, 104–114 (2017).
207. Smith, H. W. NOTE ON THE INTERPRETATION OF CLEARANCE METHODS IN THE DISEASED KIDNEY. *J. Clin. Invest.* **20**, 631–635 (1941).
208. Inker, L. A. *et al.* New Creatinine- and Cystatin C–Based Equations to Estimate GFR without Race. *N Engl J Med* **385**, 1737–1749 (2021).
209. Pottel, H. *et al.* Extending the cystatin C based EKFC-equation to children – validation results from Europe. *Pediatr Nephrol* (2023) doi:10.1007/s00467-023-06192-6.
210. Delanaye, P. *et al.* Performance of creatinine-based equations to estimate glomerular

filtration rate in White and Black populations in Europe, Brazil and Africa. *Nephrology Dialysis Transplantation* **38**, 106–118 (2023).
211. Delanaye, P., Cavalier, E., Pottel, H. & Stehlé, T. New and old GFR equations: a European perspective. *Clinical Kidney Journal* **16**, 1375–1383 (2023).
212. Delanaye, P. *et al.* The « race » correction in estimating glomerular filtration rate: an European point of view. *Current Opinion in Nephrology & Hypertension* **30**, 525–530 (2021).
213. Toyohara, T. *et al.* Metabolomic profiling of uremic solutes in CKD patients. *Hypertens Res* **33**, 944–952 (2010).
214. Liu, J.-J. *et al.* Urine Tricarboxylic Acid Cycle Metabolites Predict Progressive Chronic Kidney Disease in Type 2 Diabetes. *The Journal of Clinical Endocrinology & Metabolism* **103**, 4357–4364 (2018).
215. Jia, L., Wang, C., Zhao, S., Lu, X. & Xu, G. Metabolomic identification of potential phospholipid biomarkers for chronic glomerulonephritis by using high performance liquid chromatography–mass spectrometry. *Journal of Chromatography B* **860**, 134–140 (2007).
216. Zhang, J. *et al.* High-Throughput Metabolomics and Diabetic Kidney Disease Progression: Evidence from the Chronic Renal Insufficiency (CRIC) Study. *Am J Nephrol* **53**, 215–225 (2022).
217. Titan, S. M. *et al.* Metabolites related to eGFR: Evaluation of candidate molecules for GFR estimation using untargeted metabolomics. *Clinica Chimica Acta* **489**, 242–248 (2019).
218. Serkova, N., Florian Fuller, T., Klawitter, J., Freise, C. E. & Niemann, C. U. 1H-NMR–based metabolic signatures of mild and severe ischemia/reperfusion injury in rat kidney transplants. *Kidney International* **67**, 1142–1151 (2005).
219. Stenlund, H. *et al.* Monitoring kidney-transplant patients using metabolomics and dynamic modeling. *Chemometrics and Intelligent Laboratory Systems* **98**, 45–50 (2009).
220. Suhre, K. *et al.* Urine Metabolite Profiles Predictive of Human Kidney Allograft Status. *Journal of the American Society of Nephrology* **27**, 626–636 (2016).
221. Blydt-Hansen, T. D. *et al.* Urinary Metabolomics for Noninvasive Detection of Antibody-Mediated Rejection in Children After Kidney Transplantation. *Transplantation* **101**, 2553–2561 (2017).
222. Gagnebin, Y. *et al.* Combining the advantages of multilevel and orthogonal partial least squares data analysis for longitudinal metabolomics: Application to kidney transplantation. *Analytica Chimica Acta* **1099**, 26–38 (2020).
223. Iwamoto, H. *et al.* Metabolomic Profiling of Plasma, Urine, and Saliva of Kidney Transplantation Recipients. *IJMS* **23**, 13938 (2022).
224. Colas, L. *et al.* Urinary metabolomic profiling from spontaneous tolerant kidney transplanted recipients shows enrichment in tryptophan-derived metabolites. *eBioMedicine* **77**, 103844 (2022).
225. Nath, J. *et al.* Metabolomic Perfusate Analysis during Kidney Machine Perfusion: The Pig Provides an Appropriate Model for Human Studies. *PLoS ONE* **9**, e114818 (2014).
226. Guy, A. J. *et al.* Metabolomic Analysis of Perfusate During Hypothermic Machine Perfusion of Human Cadaveric Kidneys. *Transplantation* **99**, 754–759 (2015).
227. Wang, Z. *et al.* Proton Nuclear Magnetic Resonance (^{1}H-NMR)-Based Metabolomic Evaluation of Human Renal Allografts from Donations After Circulatory Death. *Med Sci Monit* **23**, 5472–5479 (2017).
228. Faucher, Q. *et al.* Perfusate Metabolomics Content and Expression of Tubular Transporters During Human Kidney Graft Preservation by Hypothermic Machine Perfusion. *Transplantation* **106**, 1831–1843 (2022).
229. Liu, R. X. *et al.* Untargeted metabolomics of perfusate and their association with hypothermic machine perfusion and allograft failure. *Kidney International* **103**, 762–771 (2023).

230. Kouba, E., Wallen, E. M. & Pruthi, R. S. Uroscopy by Hippocrates and Theophilus: Prognosis Versus Diagnosis. *Journal of Urology* **177**, 50–52 (2007).
231. Emwas, A.-H. *et al.* Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review. *Metabolomics* **11**, 872–894 (2015).
232. Jiang, L., Huang, J., Wang, Y. & Tang, H. Eliminating the dication-induced intersample chemical-shift variations for NMR-based biofluid metabonomic analysis. *Analyst* **137**, 4209 (2012).
233. Gil, R. B., Lehmann, R., Schmitt-Kopplin, P. & Heinzmann, S. S. 1H NMR-based metabolite profiling workflow to reduce inter-sample chemical shift variations in urine samples for improved biomarker discovery. *Anal Bioanal Chem* **408**, 4683–4691 (2016).
234. Féraud, B., Govaerts, B., Verleysen, M. & De Tullio, P. Statistical treatment of 2D NMR COSY spectra in metabolomics: data preparation, clustering-based evaluation of the Metabolomic Informative Content and comparison with 1H-NMR. *Metabolomics* **11**, 1756–1768 (2015).
235. Haslauer, K. E., Hemmler, D., Schmitt-Kopplin, P. & Heinzmann, S. S. Guidelines for the Use of Deuterium Oxide (D_2O) in ^1H NMR Metabolomics. *Anal. Chem.* **91**, 11063–11069 (2019).
236. Comparison of Uncertainties Related to Standardization of Urine Samples with Volume and Creatinine Concentration. *The Annals of Occupational Hygiene* (2004) doi:10.1093/annhyg/meh019.
237. Miller, R. C. *et al.* Comparison of Specific Gravity and Creatinine for Normalizing Urinary Reproductive Hormone Concentrations. *Clinical Chemistry* **50**, 924–932 (2004).
238. Waikar, S. S., Sabbisetti, V. S. & Bonventre, J. V. Normalization of urinary biomarkers to creatinine during changes in glomerular filtration rate. *Kidney International* **78**, 486–494 (2010).
239. Ryan, D., Robards, K., Prenzler, P. D. & Kendall, M. Recent and potential developments in the analysis of urine: A review. *Analytica Chimica Acta* **684**, 17–29 (2011).
240. Warrack, B. M. *et al.* Normalization strategies for metabonomic analysis of urine samples. *Journal of Chromatography B* **877**, 547–552 (2009).
241. 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.* **78**, 4281–4290 (2006).
242. Vasas, M., Tang, F. & Hatzakis, E. Application of NMR and Chemometrics for the Profiling and Classification of Ale and Lager American Craft Beer. *Foods* **10**, 807 (2021).
243. Filzmoser, P. & Walczak, B. What can go wrong at the data normalization step for identification of biomarkers? *Journal of Chromatography A* **1362**, 194–205 (2014).
244. Kohl, S. M. *et al.* State-of-the art data normalization methods improve NMR-based metabolomic analysis. *Metabolomics* **8**, 146–160 (2012).
245. Gagnebin, Y. *et al.* Metabolomic analysis of urine samples by UHPLC-QTOF-MS: Impact of normalization strategies. *Analytica Chimica Acta* **955**, 27–35 (2017).
246. Chen, D.-Q. *et al.* Identification of serum metabolites associating with chronic kidney disease progression and anti-fibrotic effect of 5-methoxytryptophan. *Nat Commun* **10**, 1476 (2019).
247. Port, F. K., Wolfe, R. A., Mauger, E. A., Berling, D. P. & Jiang, K. Comparison of survival probabilities for dialysis patients vs cadaveric renal transplant recipients. *JAMA* **270**, 1339–1343 (1993).
248. Horvat, L. D., Shariff, S. Z., Garg, A. X., & for the Donor Nephrectomy Outcomes Research (DONOR) Network. Global trends in the rates of living kidney donation. *Kidney International* **75**, 1088–1098 (2009).

249. Mirzakhani, M., Shahbazi, M., Oliaei, F. & Mohammadnia-Afrouzi, M. Immunological biomarkers of tolerance in human kidney transplantation: An updated literature review. *Journal Cellular Physiology* **234**, 5762–5774 (2019).
250. Schwartz, G. J. & Furth, S. L. Glomerular filtration rate measurement and estimation in chronic kidney disease. *Pediatr Nephrol* **22**, 1839–1848 (2007).
251. Ebert, N. *et al.* Assessment of kidney function: clinical indications for measured GFR. *Clinical Kidney Journal* **14**, 1861–1870 (2021).
252. Levey, A. S. & Inker, L. A. GFR as the "Gold Standard": Estimated, Measured, and True. *American Journal of Kidney Diseases* **67**, 9–12 (2016).
253. Delanaye, P. & Mariat, C. The applicability of eGFR equations to different populations. *Nat Rev Nephrol* **9**, 513–522 (2013).
254. Lamb, E. J. & Stevens, P. E. Estimating and measuring glomerular filtration rate: methods of measurement and markers for estimation. *Current Opinion in Nephrology and Hypertension* **23**, 258–266 (2014).
255. Pottel, H. *et al.* Cystatin C–Based Equation to Estimate GFR without the Inclusion of Race and Sex. *N Engl J Med* **388**, 333–343 (2023).
256. Fiehn, O. Combining Genomics, Metabolome Analysis, and Biochemical Modelling to Understand Metabolic Networks. *Comparative and Functional Genomics* **2**, 155–168 (2001).
257. Nicholson, J. K., Lindon, J. C. & Holmes, E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **29**, 1181–1189 (1999).
258. Riccio, S. *et al.* New Insights from Metabolomics in Pediatric Renal Diseases. *Children* **9**, 118 (2022).
259. Pereira, P. R. *et al.* Metabolomics as a tool for the early diagnosis and prognosis of diabetic kidney disease. *Medicinal Research Reviews* **42**, 1518–1544 (2022).
260. Dahabiyeh, L. A. *et al.* Metabolomics profiling distinctively identified end-stage renal disease patients from chronic kidney disease patients. *Sci Rep* **13**, 6161 (2023).
261. Kobayashi, T. *et al.* A metabolomics-based approach for predicting stages of chronic kidney disease. *Biochemical and Biophysical Research Communications* **445**, 412–416 (2014).
262. Jouret, F. *et al.* Nuclear Magnetic Resonance Metabolomic Profiling of Mouse Kidney, Urine and Serum Following Renal Ischemia/Reperfusion Injury. *PLoS ONE* **11**, e0163021 (2016).
263. Pan, Z. & Raftery, D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Anal Bioanal Chem* **387**, 525–527 (2007).
264. Delanaye, P., Cavalier, E., Depas, G., Chapelle, J.-P. & Krzesinski, J.-M. New Data on the Intraindividual Variation of Cystatin C. *Nephron Clin Pract* **108**, c246–c248 (2008).
265. Peng, T., Royer, A.-L., Guitton, Y., Le Bizec, B. & Dervilly-Pinel, G. Serum-based metabolomics characterization of pigs treated with ractopamine. *Metabolomics* **13**, 77 (2017).
266. Lawson, T. N. *et al.* msPurity: Automated Evaluation of Precursor Ion Purity for Mass Spectrometry-Based Fragmentation in Metabolomics. *Anal. Chem.* **89**, 2432–2439 (2017).
267. Schymanski, E. L. *et al.* Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **48**, 2097–2098 (2014).
268. Waikar, S. S., Sabbisetti, V. S. & Bonventre, J. V. Normalization of urinary biomarkers to creatinine during changes in glomerular filtration rate. *Kidney International* **78**, 486–494 (2010).
269. Hsu, C.-N. & Tain, Y.-L. Chronic Kidney Disease and Gut Microbiota: What Is Their Connection in Early Life? *IJMS* **23**, 3954 (2022).
270. Zixin, Y. *et al.* TMAO as a potential biomarker and therapeutic target for chronic kidney disease: A review. *Front. Pharmacol.* **13**, 929262 (2022).
271. Bain, M. A., Faull, R., Fornasini, G., Milne, R. W. & Evans, A. M. Accumulation of

trimethylamine and trimethylamine-N-oxide in end-stage renal disease patients undergoing haemodialysis. *Nephrology Dialysis Transplantation* **21**, 1300–1304 (2006).
272. Bell, J. D. *et al.* Nuclear magnetic resonance studies of blood plasma and urine from subjects with chronic renal failure: identification of trimethylamine-N-oxide. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1096**, 101–107 (1991).
273. Stubbs, J. R. *et al.* Serum Trimethylamine-N-Oxide is Elevated in CKD and Correlates with Coronary Atherosclerosis Burden. *Journal of the American Society of Nephrology* **27**, 305–313 (2016).
274. Pelletier, C. C. *et al.* Elevation of Trimethylamine-N-Oxide in Chronic Kidney Disease: Contribution of Decreased Glomerular Filtration Rate. *Toxins* **11**, 635 (2019).
275. Krueger, E. S., Lloyd, T. S. & Tessem, J. S. The Accumulation and Molecular Effects of Trimethylamine N-Oxide on Metabolic Tissues: It's Not All Bad. *Nutrients* **13**, 2873 (2021).
276. Hsu, C.-N. *et al.* Gut Microbiota-Dependent Trimethylamine N-Oxide Pathway Associated with Cardiovascular Risk in Children with Early-Stage Chronic Kidney Disease. *IJMS* **19**, 3699 (2018).
277. Gartland, K. P., Bonner, F. W. & Nicholson, J. K. Investigations into the biochemical effects of region-specific nephrotoxins. *Mol Pharmacol* **35**, 242–250 (1989).
278. Le Moyec, L. *et al.* Proton Nuclear Magnetic Resonance Spectroscopy of Urine and Plasma in Renal Transplantation Follow-Up. *Nephron* **65**, 433–439 (1993).
279. Rhee, E. P. *et al.* A Combined Epidemiologic and Metabolomic Approach Improves CKD Prediction. *Journal of the American Society of Nephrology* **24**, 1330–1338 (2013).
280. Yamaguchi, Y. *et al.* Plasma metabolites associated with chronic kidney disease and renal function in adults from the Baltimore Longitudinal Study of Aging. *Metabolomics* **17**, 9 (2021).
281. Duranton, F. *et al.* Plasma and Urinary Amino Acid Metabolomic Profiling in Patients with Different Levels of Kidney Function. *Clinical Journal of the American Society of Nephrology* **9**, 37–45 (2014).
282. Wannemacher, R., Klainer, A., Dinterman, R. & Beisel, W. The significance and mechanism of an increased serum phenylalanine-tyrosine ratio during infection. *The American Journal of Clinical Nutrition* **29**, 997–1006 (1976).
283. Kopple, J. D. Phenylalanine and Tyrosine Metabolism in Chronic Kidney Failure. *The Journal of Nutrition* **137**, 1586S-1590S (2007).
284. Bolen, E. *et al.* Urine metabolic risk factors and outcomes of patients with kidney transplant nephrolithiasis. *Clinical Kidney Journal* **15**, 500–506 (2022).
285. Ganda, A. *et al.* Plasma metabolite profiles, cellular cholesterol efflux, and non-traditional cardiovascular risk in patients with CKD. *Journal of Molecular and Cellular Cardiology* **112**, 114–122 (2017).
286. Guo, F. *et al.* Renal function is associated with plasma trimethylamine-N-oxide, choline, l-carnitine and betaine: a pilot study. *Int Urol Nephrol* **53**, 539–551 (2021).
287. Luck, M. *et al.* Rule-Mining for the Early Prediction of Chronic Kidney Disease Based on Metabolomics and Multi-Source Data. *PLoS ONE* **11**, e0166905 (2016).
288. Löb, S. & Königsrainer, A. Role of IDO in Organ Transplantation: Promises and Difficulties. *International Reviews of Immunology* **28**, 185–206 (2009).
289. Landsberg, A. *et al.* Non-invasive staging of chronic kidney allograft damage using urine metabolomic profiling. *Pediatric Transplantation* **22**, e13226 (2018).
290. Goek, O.-N. *et al.* Metabolites associate with kidney function decline and incident chronic kidney disease in the general population. *Nephrology Dialysis Transplantation* **28**, 2131–2138 (2013).
291. Mishima, E. *et al.* Evaluation of the impact of gut microbiota on uremic solute accumulation by a CE-TOFMS–based metabolomics approach. *Kidney International* **92**, 634–

645 (2017).
292. Deguchi, T. *et al.* Renal Clearance of Endogenous Hippurate Correlates with Expression Levels of Renal Organic Anion Transporters in Uremic Rats. *J Pharmacol Exp Ther* **314**, 932–938 (2005).
293. Mutsaers, H. A. M. *et al.* Optimized Metabolomic Approach to Identify Uremic Solutes in Plasma of Stage 3–4 Chronic Kidney Disease Patients. *PLoS ONE* **8**, e71199 (2013).
294. Zhang, Z.-H. *et al.* Metabolomic Signatures of Chronic Kidney Disease of Diverse Etiologies in the Rats and Humans. *J. Proteome Res.* **15**, 3802–3812 (2016).
295. Liabeuf, S. *et al.* Difference in Profiles of the Gut-Derived Tryptophan Metabolite Indole Acetic Acid between Transplanted and Non-Transplanted Patients with Chronic Kidney Disease. *IJMS* **21**, 2031 (2020).
296. Wikoff, W. R. *et al.* Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3698–3703 (2009).
297. Stanimirova, I. *et al.* Serum metabolomics approach to monitor the changes in metabolite profiles following renal transplantation. *Sci Rep* **10**, 17223 (2020).
298. Sell, D. R., Strauch, C. M., Shen, W. & Monnier, V. M. 2-Aminoadipic acid is a marker of protein carbonyl oxidation in the aging human skin: effects of diabetes, renal failure and sepsis. *Biochemical Journal* **404**, 269–277 (2007).
299. Metzger, R. A. *et al.* Expanded criteria donors for kidney transplantation. *American Journal of Transplantation* **3**, 114–125 (2003).
300. Wang, Z., Durai, P. & Tiong, H. Y. Expanded criteria donors in deceased donor kidney transplantation - An Asian perspective. *Indian J Urol* **36**, 89–94 (2020).
301. Cooper, J. T. *et al.* Donation After Cardiac Death: The University of Wisconsin Experience with Renal Transplantation. *American Journal of Transplantation* **4**, 1490–1494 (2004).
302. Koffman, G. & Gambaro, G. Renal transplantation from non-heart- beating donors: a review of the European experience. *J Nephrol* **16**, 334–341 (2003).
303. Khbouz, B. *et al.* Kidney-targeted irradiation triggers renal ischemic preconditioning in mice. *American Journal of Physiology-Renal Physiology* **323**, F198–F211 (2022).
304. Ponticelli, C. Ischaemia-reperfusion injury: a major protagonist in kidney transplantation. *Nephrology Dialysis Transplantation* **29**, 1134–1140 (2014).
305. Salvadori, M., Rosso, G. & Bertoni, E. Update on ischemia-reperfusion injury in kidney transplantation: Pathogenesis and treatment. *WJT* **5**, 52 (2015).
306. Wong, G. *et al.* The Impact of Total Ischemic Time, Donor Age and the Pathway of Donor Death on Graft Outcomes After Deceased Donor Kidney Transplantation. *Transplantation* **101**, 1152–1158 (2017).
307. Aydin, Z. *et al.* Randomized Trial of Short-Course High-Dose Erythropoietin in Donation After Cardiac Death Kidney Transplant Recipients. *American Journal of Transplantation* **12**, 1793–1800 (2012).
308. Bahl, D., Haddad, Z., Datoo, A. & Qazi, Y. A. Delayed graft function in kidney transplantation. *Current Opinion in Organ Transplantation* **24**, 82–86 (2019).
309. Peng, P. *et al.* Hypothermic Machine Perfusion Versus Static Cold Storage in Deceased Donor Kidney Transplantation: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Artif Organs* **43**, 478–489 (2019).
310. Reese, P. P. *et al.* Assessment of the Utility of Kidney Histology as a Basis for Discarding Organs in the United States: A Comparison of International Transplant Practices and Outcomes. *JASN* **32**, 397–409 (2021).
311. Saat, T. C. *et al.* Fate and Effect of Intravenously Infused Mesenchymal Stem Cells in a Mouse Model of Hepatic Ischemia Reperfusion Injury and Resection. *Stem Cells International* **2016**, 1–9 (2016).

312. Esmaeilzadeh, M. *et al.* Experimental Rat Model for Brain Death Induction and Kidney Transplantation. *Journal of Investigative Surgery* **33**, 141–146 (2020).
313. Soussi, D. *et al.* Preclinical Modeling of DCD Class III Donation: Paving the Way for the Increased Use of This Challenging Donor Type. *Biomed Res Int* **2019**, 5924101 (2019).
314. Siedlecki, A., Irish, W. & Brennan, D. C. Delayed Graft Function in the Kidney Transplant. *American Journal of Transplantation* **11**, 2279–2296 (2011).
315. Bon, D. *et al.* Analysis of Perfusates During Hypothermic Machine Perfusion by NMR Spectroscopy: A Potential Tool for Predicting Kidney Graft Outcome. *Transplantation* **97**, 810–816 (2014).
316. Darius, T. *et al.* Influence of Different Partial Pressures of Oxygen During Continuous Hypothermic Machine Perfusion in a Pig Kidney Ischemia-reperfusion Autotransplant Model. *Transplantation* **104**, 731–743 (2020).
317. Nath, J. *et al.* Metabolic differences between cold stored and machine perfused porcine kidneys: A 1 H NMR based study. *Cryobiology* **74**, 115–120 (2017).
318. Hrydziuszko, O. *et al.* Mass Spectrometry Based Metabolomics Comparison of Liver Grafts from Donors after Circulatory Death (DCD) and Donors after Brain Death (DBD) Used in Human Orthotopic Liver Transplantation. *PLoS ONE* **11**, e0165884 (2016).
319. Novitzky, T., Cooper, D. K. C., Morrell, D. & Isaacs, S. CHANGE FROM AEROBIC TO ANAEROBIC METABOLISM AFTER BRAIN DEATH, AND REVERSAL FOLLOWING TRIIODOTHYRONINE THERAPY: *Transplantation* **45**, 32–36 (1988).
320. van Der Hoeven, J. A. B. *et al.* Relationship between duration of brain death and hemodynamic (in)stability on progressive dysfunction and increased immunologic activation of donor kidneys. *Kidney International* **64**, 1874–1882 (2003).
321. Hauet, T. *et al.* Noninvasive Monitoring of Citrate, Acetate, Lactate, and Renal Medullary Osmolyte Excretion in Urine as Biomarkers of Exposure to Ischemic Reperfusion Injury. *Cryobiology* **41**, 280–291 (2000).
322. Robert, R. *et al.* A pair analysis of the delayed graft function in kidney recipient: The critical role of the donor. *Journal of Critical Care* **25**, 582–590 (2010).
323. Rodriguez-Niño, A. *et al.* Urinary Carnosinase-1 Excretion is Associated with Urinary Carnosine Depletion and Risk of Graft Failure in Kidney Transplant Recipients: Results of the TransplantLines Cohort Study. *Antioxidants* **10**, 1102 (2021).
324. Fujii, T. *et al.* Preventive effect of l-carnosine on ischemia/reperfusion-induced acute renal failure in rats. *European Journal of Pharmacology* **474**, 261–267 (2003).
325. Kurata, H. *et al.* Renoprotective Effects of *l* -Carnosine on Ischemia/Reperfusion-Induced Renal Injury in Rats. *J Pharmacol Exp Ther* **319**, 640–647 (2006).
326. Geubelle, P. *et al.* Identification and pharmacological characterization of succinate receptor agonists. *Br J Pharmacol* **174**, 796–808 (2017).
327. Gilissen, J., Jouret, F., Pirotte, B. & Hanson, J. Insight into SUCNR1 (GPR91) structure and function. *Pharmacology & Therapeutics* **159**, 56–65 (2016).
328. Lunyera, J. *et al.* Urine tricarboxylic acid cycle signatures of early-stage diabetic kidney disease. *Metabolomics* **18**, 5 (2022).

Appendix

Metabolomics (2024) 20:39
https://doi.org/10.1007/s11306-024-02106-1

ORIGINAL ARTICLE

Untargeted NMR-based metabolomics analysis of kidney allograft perfusates identifies a signature of delayed graft function

A. Cirillo[1] · M. Vandermeulen[2,3] · P. Erpicum[3,4] · T. Pinto Coelho[3] · N. Meurisse[2] · O. Detry[2,3] · F. Jouret[3,4] · P. de Tullio[1]

Received: 27 November 2023 / Accepted: 19 February 2024

Abstract

Introduction Kidney transplantation (KTx) necessarily conveys an ischemia/reperfusion (I/R) process, which impacts on allograft outcomes. Delayed graft function (DGF) is defined as a non-decrease of serum creatinine by at least 10% daily on 3 consecutive days during the first 7 days post-KTx. DGF significantly conditions both short- and long-term graft outcomes. Still there is a lack of DGF predictive biomarkers.

Objectives This study aimed to explore the potential of kidney graft perfusate metabolomics to predict DGF occurrence.

Methods 49 human perfusates from grafts categorized upon donor type [donation after brain death (DBD)/donation after circulatory death (DCD)] and DGF occurrence and 19 perfusates from a murine model classified upon death type (DBD/DCD) were collected and analyzed by NMR-based metabolomics.

Results The multivariate analysis of the murine data highlighted significant differences between perfusate metabolomes of DBD *versus* DCD. These differences were similarly observed in the human perfusates. After correcting for the type of donor, multivariate analysis of human data demonstrated a metabolomics signature that could be correlated with DGF occurrence.

Conclusions The metabolome of kidney grafts is influenced by the donor's type in both human and pre-clinical studies and could be correlated with DGF in the human DBD cohort. Thus, metabolomic analysis of perfusate applied prior to KTx may represent a new predictive tool for clinicians in a more personalized management of DGF. Moreover, our data paves the way to better understand the impact of donor's types on the biochemical events occurring between death and the hypothermic storage.

Keywords Metabolomics · Kidney transplantation · Delayed graft function · NMR · Perfusate

1 Introduction

Kidney transplantation (KTx) currently represents the best treatment for patients with end-stage renal disease. Even if 90,000 KTx are performed each year worldwide, kidney transplantation outcome highly depends on graft quality before transplantation and, at present, no reliable tools exist to assess it. In clinical practice, a list of standard classification criteria (SCD) was redacted to guarantee graft quality, including kidney procured in donation after brain death (DBD) conditions (Metzger et al., 2003; Wang et al., 2020). The increasing gap between demand and supply for KTx has led to the use of suboptimal organ donors, such as donation after circulatory death (DCD) and extended criteria donors (ECDs) (Cooper et al., 2004; Koffman & Gambaro, 2003). However, kidneys coming from these donors (ECDs excluded) are exposed to a supplementary warm ischemia at procurement leading to higher risk of delayed graft function (DGF) and subsequently poorer

F. Jouret and P. de Tullio have equally supervised the work.

A. Cirillo
Arianna.cirillo@uliege.be

[1] Clinical Metabolomics Group, Center for Interdisciplinary Research On Medicines (CIRM), University of Liege, Liege, Belgium

[2] Department of Abdominal Surgery and Transplantation, CHU de Liege, University of Liege, Liege, Belgium

[3] Groupe Interdisciplinaire de Génoprotéomique Appliquée (GIGA), Metabolism and Cardiovascular Sciences, University of Liege, Liege, Belgium

[4] Division of Nephrology, CHU de Liège, University of Liege, Liege, Belgium

Published online: 09 March 2024

graft outcomes (Jouret et al., 2016; Khbouz et al., 2022; Ponticelli, 2014; Salvadori et al., 2015; Wong et al., 2017). DGF is a pathological condition, often resulting from ischemic damage and defined as a non-decrease of serum creatinine by at least 10% daily on 3 consecutive days during the first 7 days post-KTx (Aydin et al., 2012). Aside from early complications following KTx, the occurrence of DGF in transplanted patients translates in lower graft function and worse short- and long-term outcomes (Bahl et al., 2019; Ponticelli et al., 2022). This clinical complication currently impacting approximately 25–30% of transplant recipients (Mannon, 2018) is associated to several donor-related factors including, among others, donor type (DCD at higher risk compared to DBD) and duration of warm and cold ischemia times (WIT/CIT) (Ponticelli et al., 2022). In way to minimize the risks related to ischemic damage, the usage of a cold preservation solution to be flushed in the organ, can minimize these events by decreasing metabolism and slowing the process leading to IRI (Peng et al., 2019). In recent studies, the use of IGL-1 (Institute Georges Lopez-1, France) cold storage solution has led to the reduction of DGF incidence (Badet et al., 2005). However, independently of the types of donors, reliable tools are needed to assess the quality of the graft after procurement and/or to better prevent DGF in clinical practice (Reese et al., 2021). In this context, metabolomics-based approach is particularly suited by placing itself as a solution tool adapted to personalized medicine setting for patient's treatment and follow-up (Ashrafian et al., 2021; Jacob et al., 2019; Letertre et al., 2021). This study hypostatized that perfusate solutions represent useful—still poorly explored—biofluids that could inform on what happened to the graft during its cold ischemic period, thereby providing the clinicians with informative data concerning the quality of the graft. Thus, using a Nuclear Magnetic Resonance (NMR)-based untargeted approach, this study aimed to analyze kidney perfusate solutions obtained just before transplantation in search of a metabolomic signature that could predict DGF occurrence in transplant patients. For this purpose, we had access to two types of perfusate samples, one from a human cohort and the other one from an animal model simulating the donor's type of donation. The parallel use of a human and experimental samples could enable us to explore two main points: (1) the biochemical events occurring between death and hypothermic storage of kidneys related to donor type; (2) the differences in metabolomics content between different donor types.

The final goals of this pilot study were (1) to demonstrate the value of perfusate metabolomic analysis in KTx; (2) to evaluate the impact of donor type on the graft metabolome during the cold ischemia period and (3) to generate a metabolomics model that could be useful as a predictive tool for DGF occurrence.

2 Materials and methods

2.1 WIT and CIT

In this study, kidney transplantation events were described using the terms: warm ischemia time1 (WIT1), cold ischemia time (CIT), and warm ischemia time2 (WIT2).

WIT1 was measured as the period between the circulatory arrest and the cold perfusion of the kidneys. CIT was defined as the period between the initiation of the cold flush until the removal of the graft from the cold storage to be transplanted in the recipient. WIT2 was defined as the time between removal from cold storage to graft reperfusion in the recipient (suture time).

2.2 Surgical models of kidney graft procurements in the rat

This experimental model has been essentially developed to mimic the different types of kidney graft donors. After induction of anesthesia (Isoflurane 2% in O_2 2 l/min), a tracheotomy was performed to ensure optimal ventilation using a weight-based autoregulated rodent ventilator (PhysioSuite—KentScientific, Torrington, Connecticut, USA). Arterial and venous catheters (Polyurethane 0.43 × 0.69 m) were placed in the femoral vessels to allow continuous monitoring of intra-arterial pressure (Picco Monitoring Kit and BP-100; CWE inc, Ardmore, Oklahoma, USA) and to have a venous access for intraoperative fluid injection, respectively (Ethic agreement code: 2147).

2.2.1 DBD model

In the DBD group (n = 9), brain death was induced under general anesthesia as previously described by Saat et al. (2016). In summary, after front lateral trepanation, a Fogarty balloon catheter (Edwards Lifesciences, Irvine, California, USA) was introduced and slowly inflated (100 µL/min for 4 min) into the extradural space. Brain death was confirmed by the absence of corneal and pupillary reflexes, the onset of a hypertensive peak followed by major hypotension and a 60 s apnea test as described (Esmaeilzadeh et al., 2020). The rats were maintained in a brain-death state for 6 h. The mean arterial pressure (MAP) was maintained above 60 mm Hg after induction of brain death by intravenous administration of normal saline (1 mL/h) and norepinephrine (1 mg/mL, Aguettant, Lyon, France) (5–15 µg/h) in case the administration of a 1 mL bolus of normal saline would not maintain MAP above 60 mmHg. Then, all animals received

a continuous infusion of 1 mL/h of normal saline using an electric syringe pump (Becton Dickinson, Franklin Lakes, New Jersey, USA), previously warmed to body temperature, to compensate for insensible losses. The body temperature was maintained at 38 °C with a rectal probe-controlled temperature pad connected to the ventilator (PhysioSuite—KentScientific, Torrington, Connecticut, USA).

2.2.2 DCD model

In the DCD group (n = 10), all animals received 1 mL normal saline/h, previously warmed to body temperature, to compensate for insensible losses. After 6 h of mechanical ventilation in a previously tracheostomized rat maintained under general anesthesia, circulatory arrest was induced by an intravenous injection of KCl (150 mg/kg) as previously described in rodent DCD models (Soussi et al., 2019). Circulatory arrest was defined upon cessation of aortic pulsatility and a fall in MAP below 25 mmHg.

2.2.3 Kidney procurement and perfusate collection

After induction of circulatory (DCD group) or brain death (DBD group), a laparotomy was performed to remove the kidneys. Both kidneys were flushed with IGL-1 organ preservation solution through the aorta and then stored immersed in IGL-1 at 4 °C for 14 h, as a model of clinical CIT (left kidney). After 14 h of CIT, a second cold IGL-1 flush was made through the renal artery to collect the perfusate, defined as the first 2 mL of effluent evacuated from the renal vein. The perfusate was then stored at −80 °C until metabolomic analyses (9 DBD and 10 DCD perfusates).

2.3 Clinical kidney transplantation

Informed consent was obtained from all KTx recipients included in the study (Ethics agreement: B707201524484). Clinical data about donor, graft and recipient were collected. Kidney donor risk index (KDRI) and kidney donor profile index (KDPI) were calculated based on donor and transplant factors (Rao et al., 2009).

Human kidney grafts procured from DBD or DCD donors were cold flushed during the procurement with IGL-1. All kidney grafts were preserved with classical static cold storage at 4 °C during variable CIT, according to the clinical situation, and were allocated to a given recipient by Eurotransplant. After preparation of the renal vessels before KTx on the back-table, the kidney grafts were flushed through the renal artery with 1 L of IGL1 solution. The first 10 mL of liquid exiting the renal vein during the flush were collected for metabolomics analysis and constituted the "perfusate". A total of 49 perfusate samples were used in this study; 36 from DBD and 13 from DCD donors. Prior to DBD and DCD classification, samples were categorized based on donor criteria classification including SDC (n = 40) and ECD (n = 9). Among these 9 ECD donors, 7 were DBD and 2 were DCD. Concerning the post-transplantation status, 7 ECD donors experienced DGF, while 2 were classified as noDGF.

2.4 ^{1}H-NMR metabolomics

All samples were recorded at 298 K on a Bruker Avance HD spectrometer (Bruker, Billerica, USA) operating at 700.17 MHz for the proton signal acquisition. The instrument was equipped with a TCI 5-mm cryoprobe with a Z-gradient. Maleic acid was used as the internal standard for quantification and trimethylsilyl-3-propionic acid-d4 (TMSP) for the zero for the zero calibration. An aliquot of 200 µL of perfusate samples was supplemented with 50 µL of deuterated phosphate buffer (DPB, pH 7.4), 25 µL of a 35 mM solution of maleic acid and 2 µL of a 10 mg/mL TMSP D2O solution. ^{1}H-NMR spectra were acquired using a 1D NOESY sequence with presaturation. The Noesypresat experiment used a RD-90°-T1-90°-Tm-90°-acquire sequence with a relaxation delay of 4 s, a mixing time (Tm) of 10 ms and a fixed T1 delay of 4 µs. Water suppression pulse was placed during the relaxation delay (RD). The number of transients is 64 (64 K data points). The data were processed with the Bruker Topspin 4.0.8 software (Bruker BioSpin, Billerica, USA) with a standard parameter set. Phase and baseline corrections were per-formed manually over the entire range of the spectra and the δ scale was calibrated to 0 ppm using the internal standard TMSP.

2.5 Multivariate analysis

Once the spectra obtained, identification and quantification of metabolites were performed through Chenomx profiler 9.0 (Chenomx Inc., Edmonton, AB, Canada). The profiled spectra were used to generate a concentration table that was imported in BioStatFlow webtool (biostatflow.org) for multivariate statistical analysis. Autoscaling normalization was applied to concentration table. Principal component analysis (PCA) was used for looking at outliers and cluster between samples. An orthogonal signal correction-PLS model (OSC-PLS) was performed as discriminant model and its quality was determined by the predictability calculated based on the fraction correctly predicted in one-seventh cross-validation (Q^2). Permutation tests were performed for validate models. Metaboanalyst (www.metaboanalyst.ca) was used for generating multivariate receiver operating characteristic (ROC) curves and confusion matrix to assess the robustness of the models. ROC curves were performed by using PLS-DA models as classification method and univariate AUROC as feature ranking method with two latent variables.

Confusion matrix was used for calculating model sensitivity and specificity.

2.6 Univariate and pathways analysis

Univariate analysis was made using GraphPad Prism version 9.4.1 (GraphPad Software, La Jolla, CA, www.graphpad.com). Mann–Whitney U test was performed for comparisons between two groups. No correction was applied to Mann–Whitney U tests. The detailed analysis of the metabolic pathways was performed by Metaboanalyst (www.metaboanalyst.ca) using the metabolomic set enrichment analysis (MSEA) tool by using the high-quality SMPDB metabolic pathways as the backend knowledgebase.

3 Results

3.1 Study population

A preliminary statistical analysis was conducted on the cohort considering the donor criteria classification including SDC and ECD, which showed no confounding factors. Donor characteristics categorized by the type of death event are presented in Table 1.

Statistical analysis on recipient data (DGF = 19; noDGF = 30) showed no correlation between DGF status and related clinical data (Table 2).

3.2 Preliminary analysis on IGL-1 matrix

A preliminary metabolomics analysis allowed to spot a series of peaks coming from IGL-1 matrix represented by lactobionic acid, adenosine, raffinose, and glutathione signals (S2, S3, S4 and S5). These peaks were identified, quantified, and excluded from following analysis on perfusate samples.

3.3 Perfusate metabolomes in rat models: differences between DBD and DCD donors

Nineteen perfusate samples were included in the analysis. DBD and DCD kidney donor conditions were mimed in 9 and 10 rats, respectively. After exclusion from statistical analysis of matrix-linked peaks, 44 quantified metabolites that are belonging to the kidney metabolome were used for multivariate statistical models. Supervised statistical investigations using OSC-PLS highlighted a significant discrimination between DBD donor type versus DCD donor type ($Q^2 = 0.621$); the model was validated through a permutation test (p-value = 0.0393), confirming the absence of overfitting (Fig. 1).

A loading-plot was generated in way to underline all the metabolites responsible for the discrimination of groups according to donor type. Based on the results of multivariate models, univariate statistical analysis revealed that the concentrations of 7 metabolites were increased in DCD status compared to DBD status, including 2-hydroxyisovalerate, 2-octenoate, 2-oxocaproate, 3-hydroxyisovalerate, alloisoleucine, creatine phosphate, lysine, o-phosphoethanolamine, taurine and valine (Fig. 2).

3.4 Perfusate metabolomes in human cohort: differences between DBD and DCD donors

Forty-nine human perfusate samples were included in the analysis. Multivariate and univariate analyses based on donor criteria (SCD/ECD) indicated no significant statistical impact. Once this confounding factor was excluded, another analysis was conducted on the samples categorized by donor type (DBD = 36; DCD = 13). A total of 54 metabolites were identified independently from IGL-1 matrix and quantified (S1, S2, S3, S4 and S5). Multivariate statistical analysis was performed by generating a discriminant OSC-PLS model ($Q^2 = 0.540$); a permutation test (p-value = 0.02192) was performed by validating the model obtained (Fig. 3).

The metabolites relevant for discrimination between groups in multivariate models were used to perform univariate statistical analysis. Mann–Whitney U tests highlighted 17 metabolites significantly differing between donor types. Twelve metabolites were increased in DCD donor group (e.g., valine and cystine) while five were decreased compared to DBD group (e.g., lactate and trimethylamine-*N*-oxide) (Fig. 4).

3.5 Insights into common metabolic pathways in human and rat perfusates

By comparing the significant metabolites obtained in the two models, it was possible to spot some common features such as valine and isoleucine. In way to better investigate the biochemical events occurring during the hypothermic storage of kidney, a pathway analysis was conducted. MSEA analysis led to the identification of several pathways significantly affected by donor type (Fig. 5a–b). The most relevant cascades were, in both models, represented by valine/leucine/isoleucine biosynthesis and degradation, aminoacyl-t-RNA biosynthesis, alanine/aspartate/ glutamate metabolism and glycerophospholipid metabolism.

3.6 Correlation of ^{1}H-NMR-based human perfusate metabolome with DGF incidence post-KTx

By considering DGF outcome as factor (DGF = 19; noDGF = 30), OSC-PLS model was used as a discriminant

Table 1 Clinical characteristics of donors

	DBD (N=36)	DCD (N=13)	Total (N=49)	p-value
Sex				0.233
F	15 (41.7%)	3 (23.1%)	18 (36.7%)	
M	21 (58.3%)	10 (76.9%)	31 (63.3%)	
Age				0.774
Mean (SD)	45.361 (13.495)	46.692 (16.311)	45.714 (14.130)	
Range	19.000 - 68.000	18.000 - 67.000	18.000 - 68.000	
BMI				0.348
N-Miss	0	1	1	
Mean (SD)	25.739 (4.159)	24.502 (3.016)	25.429 (3.912)	
Range	17.700 - 35.900	20.000 - 30.000	17.700 - 35.900	
Serum creat (mg/dl)				0.277
Mean (SD)	0.789 (0.253)	0.703 (0.202)	0.766 (0.242)	
Range	0.310 - 1.400	0.430 - 1.120	0.310 - 1.400	
Weight				0.948
N-Miss	0	1	1	
Mean (SD)	77.111 (13.509)	76.833 (9.437)	77.042 (12.521)	
Range	47.000 - 110.000	65.000 - 94.000	47.000 - 110.000	
Diabetes				0.648
N-Miss	5 (13.9%)	1 (7.7%)	6 (12.2%)	
No	30 (83.3%)	11 (84.6%)	41 (83.7%)	
Yes	1 (2.8%)	1 (7.7%)	2 (4.1%)	
Hypertension				0.831
N-Miss	5 (13.9%)	1 (7.7%)	6 (12.2%)	
No	24 (66.7%)	9 (69.2%)	33 (67.3%)	
Yes	7 (19.4%)	3 (23.1%)	10 (20.4%)	
Cause of death				0.008
Anoxia	8 (22.2%)	9 (69.2%)	17 (34.7%)	
Cerebral hemor	19 (52.8%)	2 (15.4%)	21 (42.9%)	
Trauma	9 (25.0%)	2 (15.4%)	11 (22.4%)	
KDPI %				0.386
N-Miss	1	1	2	
Mean (SD)	39.361 (27.234)	47.308 (30.398)	41.469 (28.007)	
Range	0.000 - 92.000	0.000 - 89.000	0.000 - 92.000	
KDRI				0.228
N-Miss	1	1	2	
Mean (SD)	0.937 (0.277)	1.053 (0.302)	0.967 (0.285)	
Range	0.570 - 1.660	0.640 - 1.540	0.570 - 1.660	

BMI body mass index, *Serum creat* serum creatinine mg/dl, *KDPI* kidney donor profile index (%), *KDRI* kidney donor risk index

analysis ($Q^2 = 0.370$; P-value = 0.4387). Moreover, a multivariate receiver operating characteristics (ROC) curve was generated by using the ratios of 54 identified metabolites (73 variables; AUC = 0.777; predictive accuracy = 71.6%) in way to evaluate the performance of our model through automated feature selection (Fig. 6a–b). Univariate analysis showed 12 metabolites being significantly increased in noDGF status (such as cystine, leucine, alanine …), unlike carnosine (Fig. 7).

In the perfusates from DBD kidneys, 36 perfusate samples were used for the univariate and multivariate statistical analyses by using DGF status as discriminant factor. A discriminant OSC-PLS model was generated ($Q^2 = 0.747$; p value = 0.2642) to look at the separation between the noDGF and the pathological status (Fig. 8a). Furthermore, the multivariate ROC curve showed a good performance by using the ratios of the 54 identified metabolites generated through Metaboanalyst algorithm (73 variables; AUC = 0.917;

Table 2 Clinical characteristics of receivers based on graft outcome DGF or noDGF

	DGF (N=19)	noDGF (N=30)	Total (N=49)	p-value
Sex				0.326
Female	9 (47.4%)	10 (33.3%)	19 (38.8%)	
Male	10 (52.6%)	20 (66.7%)	30 (61.2%)	
Age				0.606
Mean (SD)	59.000 (12.396)	57.000 (13.562)	57.776 (13.028)	
Range	28.000 - 72.000	22.000 - 73.000	22.000 - 73.000	
CIT				0.377
Mean (SD)	604.158 (281.886)	677.867 (281.889)	649.286 (281.286)	
Range	207.000 - 1069.000	189.000 - 1300.000	189.000 - 1300.000	
WIT1				0.732
Mean (SD)	2.105 (3.710)	2.533 (4.531)	2.367 (4.197)	
Range	0.000 - 11.000	0.000 - 16.000	0.000 - 16.000	
WIT2				0.363
Mean (SD)	37.316 (12.352)	40.000 (8.154)	38.959 (9.956)	
Range	21.000 - 54.000	19.000 - 55.000	19.000 - 55.000	

CIT cold ischemia time, *WIT1* warm ischemia time, *WIT2* warm ischemia time 2

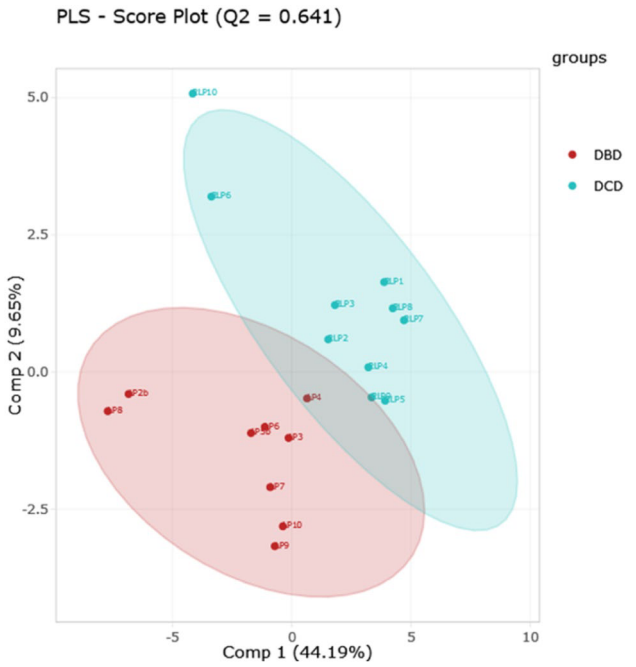


Fig. 1 OSC-PLS model of rat model showing variation of metabolomics profile according to donor type (Q2 = 0.641; cross validation p-value = 0.0393; n total = 19)

predictive accuracy = 83%) (Fig. 8b). A confusion matrix was additionally performed to cross-validate the model by highlighting all the automated selected ratio features used by the algorithm (Fig. 8c).

Mann–Whitney t-tests highlighted 15 metabolites being significantly increased (e.g., lactate, alloisoleucine, isovalerate…) and ethylmalonate and carnosine being decreased in noDGF condition in comparison to DGF condition (Fig. 9). To identify biologically meaningful patterns a MSEA analysis was done showing that methylhistine/beta-alanine metabolism and valine, leucine and isoleucine degradation are the pathways more significantly impacted.

A total of 13 perfusates from DCD donors were used for statistical analysis by using graft outcome as discriminant factor (data not shown). The low number of samples for each group (DGF = 5; noDGF = 8) highlighted the over fitting of the model and a consequent poor performance ($Q^2 = 0.7333$; P-value = 0.4995). Univariate statistical analysis by using Mann–Whitney t-test showed only one metabolite significant for the discrimination of the two groups: creatinine phosphate.

4 Discussion

DGF represents one major obstacle in allograft survival (Siedlecki et al., 2011). Still, no single biomarker or a biochemical signature of graft quality and DGF prediction has been found and used routinely in clinical practice. The metabolomics approach has already been applied in several studies related to renal pathology and dysfunction on urine or plasma samples, leading to the discovery of metabolites linked to pathological status (Humphries et al., 2023; Iwamoto et al., 2022; Wishart, 2006). Its non-invasive nature makes this approach have advantages over biopsy in the context of graft status monitoring. Thus, several metabolomic studies of renal perfusate are described in the literature with particular emphasis on the correlation between perfusion time and metabolic content. (Bon et al., 2014; Faucher et al., 2022; Guy et al., 2015). In addition, other works have investigated perfusates of porcine models through NMR

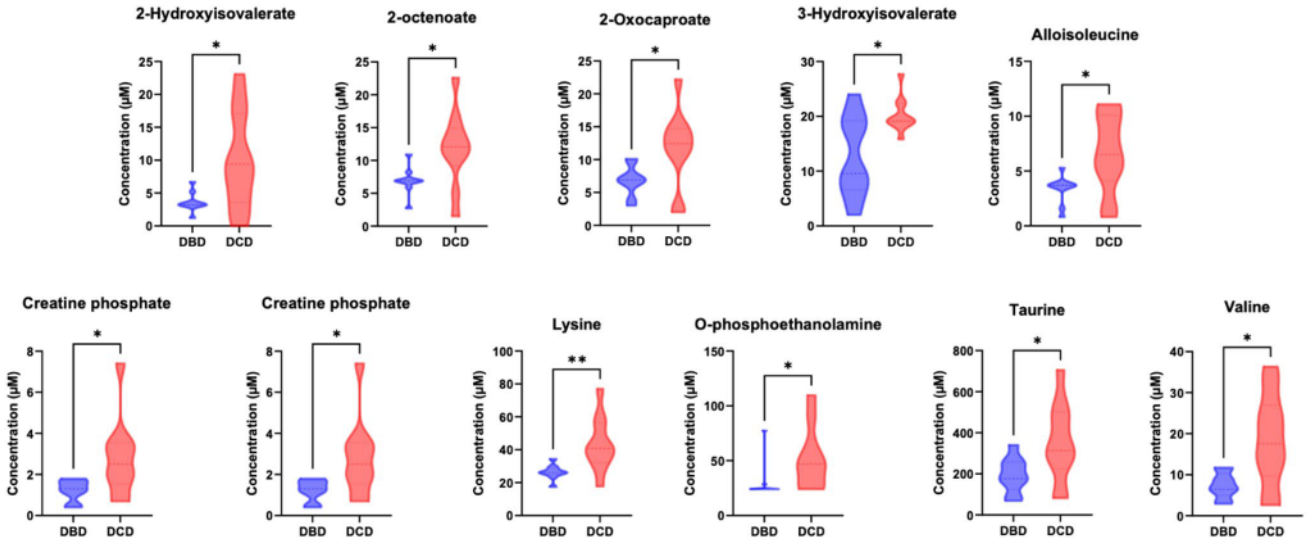


Fig. 2 Violin-plots of significant metabolites in rat perfusate of DBD (n = 9) versus DCD (n = 10). Mann–Whitney U test was used. (* < 0.05; ** < 0.001)

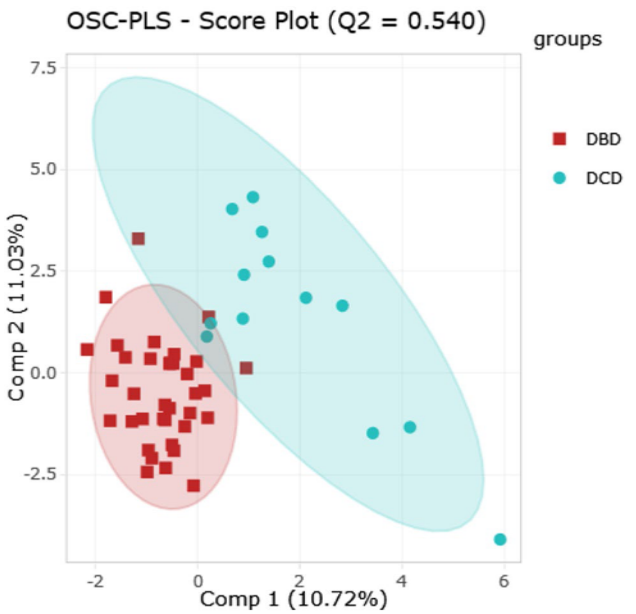


Fig. 3 Multivariate analysis based on donor types in human model (n total = 39). OSC-PLS score plot showing a separation between DBD versus DCD donor (Q2 = 0.540; cross validation p-value = 0.02192)

technique and have shown changes in metabolome's profiles of DCD donors between different perfusion strategies (Darius et al., 2020; Nath et al., 2017; Peng et al., 2019). However, at that time, no relevant metabolomic signature or predictive model was reported about the occurrence of DGF condition.

The first question to ask is whether there is a difference in the graft metabolome depending on the harvesting conditions (DCD or DBD). Using our pre-clinical experimental model, we were able to demonstrate that the metabolomes of the DCD and DBD groups were significantly different. Specifically, perfusate analysis revealed higher levels of seven metabolites, including lysine, isoleucine, and valine, in DCD samples compared to DBD samples. In the human cohorts, these metabolic trends were confirmed with more than 17 metabolites levels that significantly differs according to the donner type. Some of these metabolites were common in both rat model and human study. More interestingly, several common biochemical pathways such as valine, leucine and isoleucine degradation/biosynthesis and aminoacyl-tRNA biosynthesis were also identified and could be correlated with some data in the literature. Thus, Hrydziuszko et al. previously observed the implication of these biochemical pathways in the DCD graft in the context of liver transplantation (Hrydziuszko et al., 2016). These similarities, albeit partial, between the metabolomic profiles of the experimental model and human samples demonstrated the value of the animal model for mimicking the graft according to the type of donor. Moreover, in the human study, lactate, trimethylamine-*N*-oxide (TMAO) and methylmalonate were also found to be significantly higher in DBD graft in comparison to DCD graft. Modulations in lactate levels could be explained by the major change in the metabolic oxidative process following brain death by underling a change from aerobic to anaerobic metabolism in the graft (Novitzky et al., 1988; van Der Hoeven et al., 2003). Higher levels of TMAO, could be linked to a more important renal medullary damage due to ischemic events in DBD graft compared to DCD (Hauet et al., 2000; Robert et al., 2010). Differences in metabolic profiles related to donor types may also find an

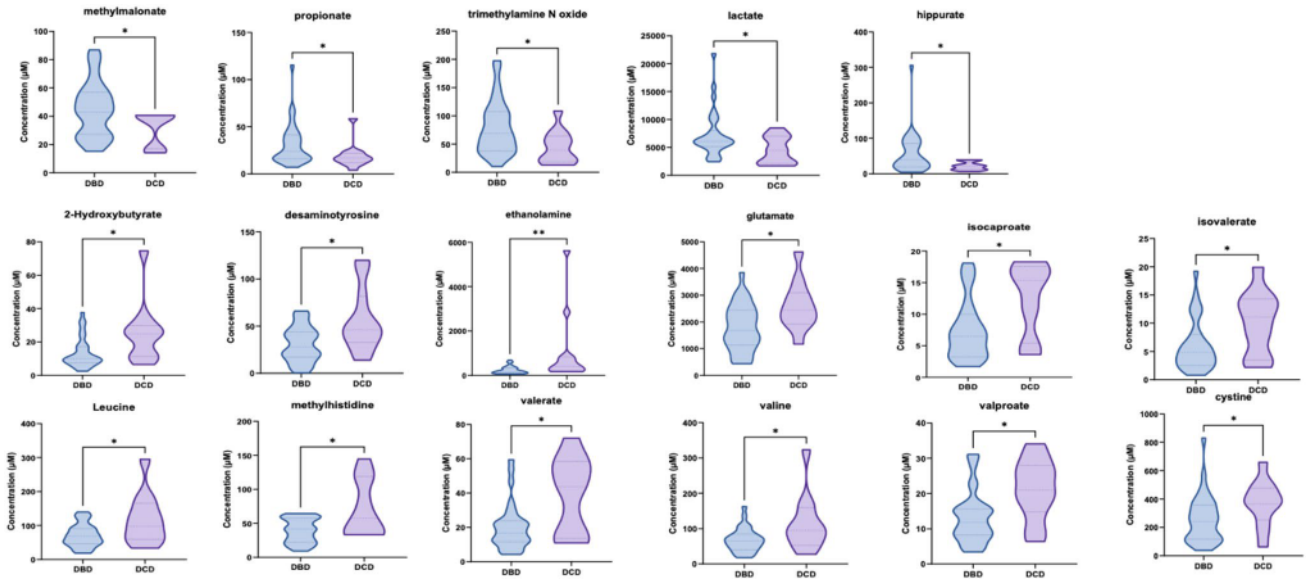


Fig. 4 Violin-plots for metabolites of human cohort with significant variation according to DBD (n = 36) versus DCD (n = 13). Comparison between groups was made by using Mann–Whitney U test. (* < 0.05; ** < 0.001)

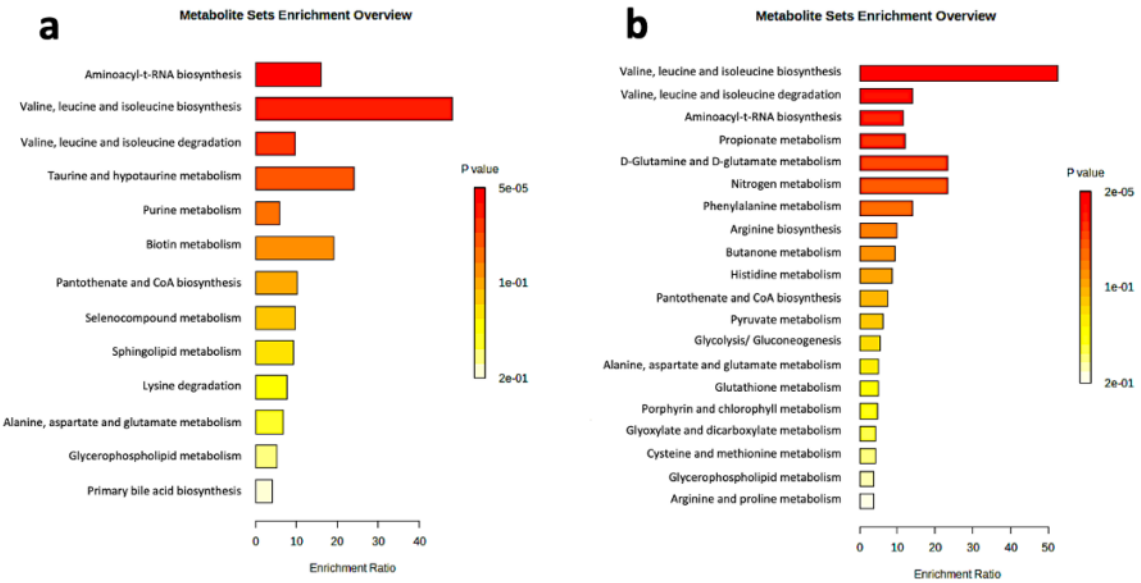


Fig. 5 Comparison of MSEA analysis in human and rat models showed common biochemical pathways involved in the process of DBD and DCD; **a** MSEA analysis of human cohort based on significant metabolites related to donor type; **b** MSEA analysis of rat model for DBD versus DCD groups

explanation in metabolic changes in the kidney due to the I/R process, as already reported in the literature (Jouret et al., 2016), and to the warm ischemia typical of DCD donors.

All these data showed that the biochemical processes that took place within the graft during the period of cold ischemia were different depending on the type of donor and need to be explored more in details. Furthermore, it could be expected that these metabolomic differences could be linked to the occurrence of DGF after transplantation and long-term outcomes in recipients. This shows that, in our view, each type of donor should be studied separately.

This was confirmed when we studied the correlation between the graft metabolome and the occurrence of DGF. Indeed, the results obtained using the data from the whole cohort (DBD and DCD donors together) showed a poor performance in predicting the post-transplantation graft

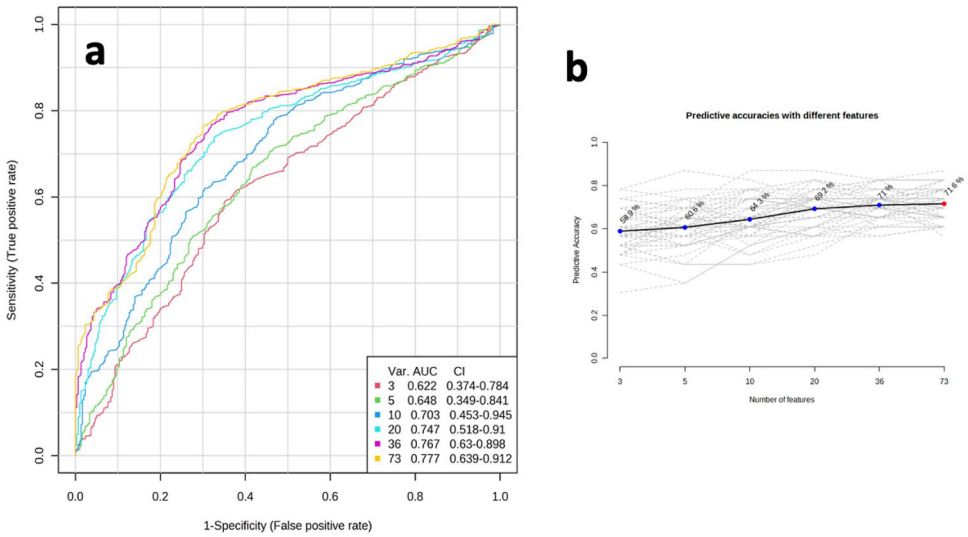


Fig. 6 Predictive models of DGF occurrence in the human cohort: **a** multivariate ROC curve based on DGF versus noDGF analysis with automated selection of ratio features showed an AUC = 0.777 by considering 73 variables; **b** plot showing the predictive accuracy with different features by reaching a value 71.6% with 73 variables

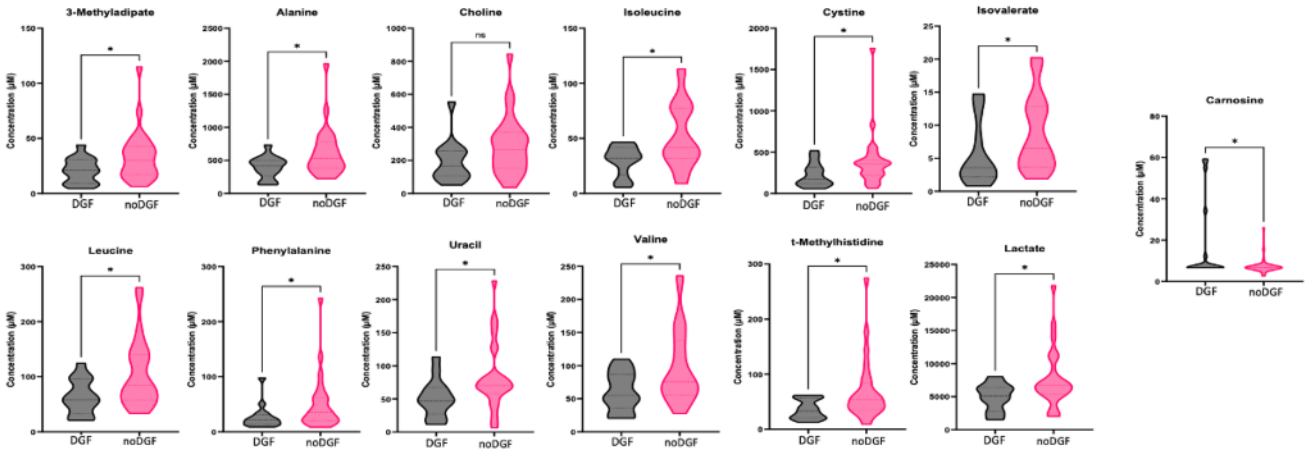


Fig. 7 Violin plots of significant metabolites according to DGF (n = 19) versus noDGF (n = 30) in human model. Comparison between groups was made by using Mann–Whitney U test. (* < 0.05; ** < 0.001)

outcome. While focusing our analysis on DBD-derived perfusates improved the quality of the predictive DGF occurrence model. Indeed, the multivariate ROC curve and the matrix of confusion obtained are quite relevant and compatible with a clinical predictive use for DGF occurrence demonstrating the interest of perfusate analysis for the recipients' follow-up.

Moreover, this approach led also to the identification of several biomarkers that differed in concentrations in the perfusates of DGF kidneys compared to noDGF kidneys. Alanine, isoleucine, leucine, valine and other branched chain amino acids were detected to be in significantly lower levels in DGF kidneys. These metabolites were released by the graft which can attest the sustained metabolic activity, ischemic damage, or both. Guy et al. previously observed raised concentrations of amino acids that may be linked to increased cellular breakdown in more ischemically damaged DGF kidneys (Guy et al., 2015). However, their results were based on the perfusate's analysis of DCD graft 45 min after perfusion with KPS-1 cold storage solution, which hampers direct comparison with our results. Otherwise, Carnosine and ethyl-malonate, were significantly increased in DGF kidneys. The role of carnosine in kidney function has recently been investigated and shown a significantly lower urinary carnosine excretion in kidney transplant recipients compared to healthy group leading to higher risk of graft failure (Rodriguez-Niño et al., 2021). Potential mechanism of action of this metabolite included lowering of chronic low-grade inflammation, as demonstrated in animal models in which treatment through carnosine was associated with attenuation of I/R-induced renal dysfunction (Fujii et al., 2003; Kurata et al., 2006). Thus, the higher level of carnosine in the perfusate samples of DGF kidney could be correlated to a possible increase of I/R damage. Levels of ethyl-malonate was

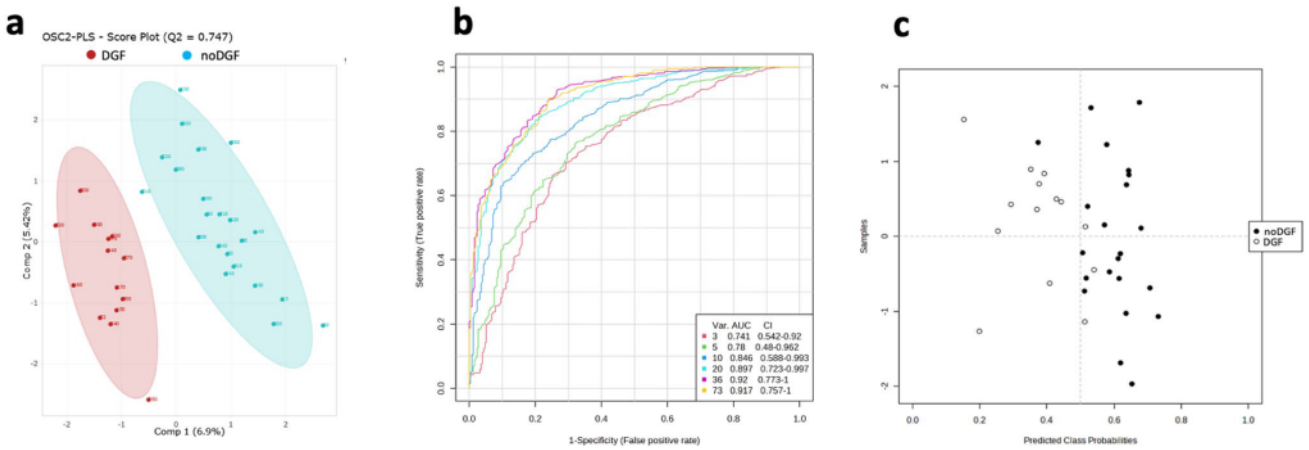


Fig. 8 Prediction of DGF kidney outcome in DBD group (human cohort); **a** OSC-PLS score plot show a separation between groups (Q2 = 0.747; p-value = 0.2642); **b** multivariate ROC curve of DBD group based on DGF versus noDGF analysis with automated selection of ratio features showed an AUC = 0.917 by considering 73 variables; **c** confusion matrix was generated to evaluate model performance

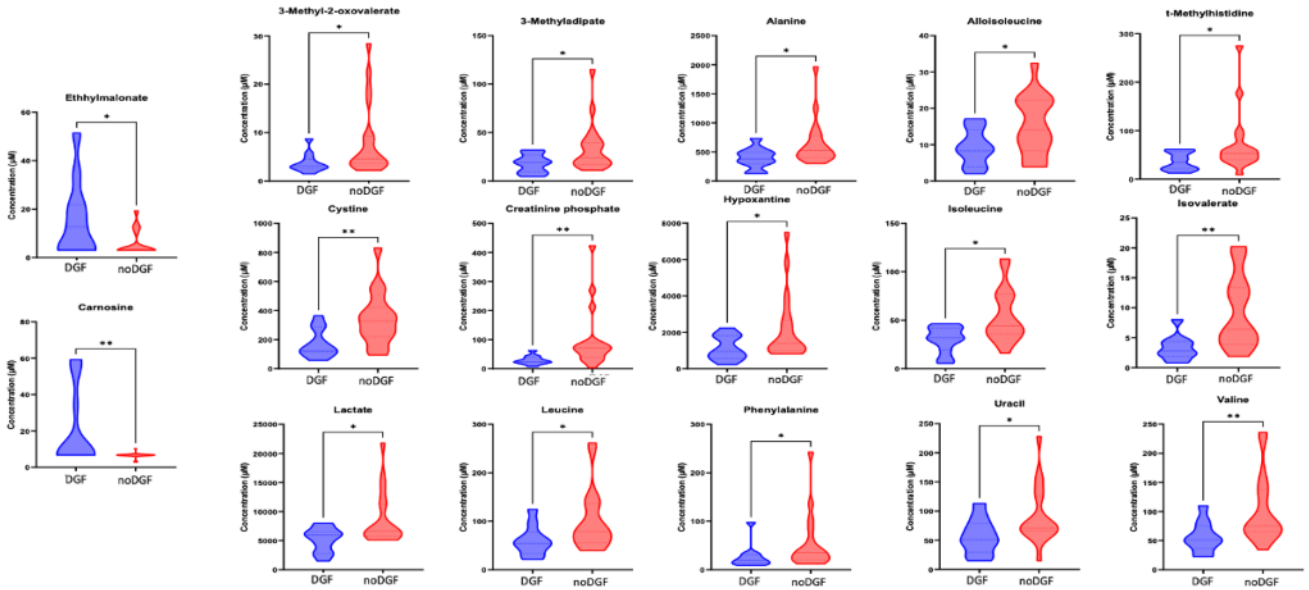


Fig. 9 Violin-plots of significant metabolites of DBD subgroup according to the analysis DGF (n = 14) versus noDGF (n = 22) in human model. Mann–Whitney U test was used for comparison

increased in the more ischemically damaged DGF kidneys, which would suggest changes in normal function of tricarboxylic acid (TCA) cycle organic anions (OAs). The biological mechanism linking TCA cycle organic anions with kidney dysfunction is not fully understood, but some studies have reported the relationship of these biomarkers with kidney function (Geubelle et al., 2017; Gilissen et al., 2016; Lunyera et al., 2022). A more fundamental study using rat models of kidney transplantation would provide a better understanding of the role of metabolites and associated biochemical pathways in graft function during cold ischemia. This would also allow us to better understand the impact that these metabolic pathways could have on the occurrence of DGF.

Considering all these findings, we are also aware of the limitations of our present pilot study. Our findings need obviously to be validated and confirmed in human by studying a new cohort with an increased number of DCD patients. The development of a rat transplantation models could also be valuable, especially for a more fundamental study of the metabolic pathways implied in DGF events. Another limitation of our study is the lack of complete clinical data and of biofluids from the donors that could allow us to obtain a

more complete view on graft and allow to refine and complete the data concerning the graft before transplant.

5 Conclusion

In KTx, a better understanding of graft quality and a better prediction of short-term renal function in the recipient patient are essential both to improve graft selection and to provide more precise, personalized patient management. By studying the fluid preserved in the graft obtained just before transplantation, we were able to identify, both in an animal model and in a human cohort, some differences in the kidney metabolome depending on the type of donor (DCD or DBD). These differences were directly associated with alterations in the metabolic pathways of the graft, potentially exerting additional influence on kidney function recovery post-KTx. In view of these results, it is obvious that DCD and DBD samples could be analyzed in separate models. Thus, we have demonstrated that using the perfusate sample in DBD donors could allow to predict the occurrence of DGF status prior to surgery and to identify a panel of metabolites that could be correlated to this status.

Our data paves the way to better understand the impact of donor's types on the biochemical events occurring between death and hypothermic storage and to correlate some metabolites of the storage liquid with the occurrence of DGF. This acquired knowledge may help clinicians in the elucidation of the biochemical events occurring between death and transplantation per se to improve the process of organ procurement and reduce the discard rate.

Obviously, our preliminary results need to be confirmed on a validation cohort that could also include biofluid samples and complete clinical data from both donor and recipients. Moreover, the development of rat models of kidney transplantation from DCD or DBD donors as well as a better exploration of the identified biochemical pathways implied will also allow a better understanding of the ischemic events whose intricate interconnections that play a pivotal role in shaping the graft outcome.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11306-024-02106-1.

Author contributions VM, OD, NM, TPC: identified the patients and collected the samples. AC: samples and performed NMR analysis and multivariate/univariate statistics. PT and FJ: supervised the study. AC: wrote the first draft. PT, FJ, VM, TPC and PE: critically reviewed and revised the draft. All authors read and verified the manuscript.

Funding AC, PE and PT are respectively a FRIA grantee, a Fellow and a research director of the Fonds de la Recherche Scientifique-FNRS. FJ received support from the University of Liège (Fonds Spéciaux à la Recherche, Fonds Léon Fredericq) and the FNRS (Research Credits 2016 and 2019). TPC received support from the University of Liege (Fonds Léon Fredericq) and from Sociéte Francophone de Nephrologie, Dialyse et Transplantation (SFNDT). AC received support from the University of Liege (Fonds Léon Fredericq).

Data availability NMR data for study in RAW format is freely available in the MetaboLights repository with study identifier MTBLS9015 (www.ebi.ac.uk/metabolights/).

Declarations

Competing interests The authors declare no competing interests.

Ethical approval For the human cohort, ethical approval for the use of perfusate samples and associated metadata in this study was obtained from CHU of Liege (Ethic agreement code: B707201524484). For animal model, all applicable international, national, and/or institutional guidelines for the care and use of animals were followed (Ethic agreement code: 2147).

Research involving human and animal rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References:

Ashrafian, H., Sounderajah, V., Glen, R., Ebbels, T., Blaise, B. J., Kalra, D., Kultima, K., Spjuth, O., Tenori, L., Salek, R. M., Kale, N., Haug, K., Schober, D., Rocca-Serra, P., O'Donovan, C., Steinbeck, C., Cano, I., De Atauri, P., & Cascante, M. (2021). Metabolomics: The stethoscope for the twenty-first century. *Medical Principles and Practice, 30*(4), 301–310. https://doi.org/10.1159/000513545

Aydin, Z., Mallat, M. J. K., Schaapherder, A. F. M., Van Zonneveld, A. J., Van Kooten, C., Rabelink, T. J., & De Fijter, J. W. (2012). Randomized trial of short-course high-dose erythropoietin in donation after cardiac death kidney transplant recipients. *American Journal of Transplantation, 12*(7), 1793–1800. https://doi.org/10.1111/j.1600-6143.2012.04019.x

Badet, L., Petruzzo, P., Lefrançois, N., McGregor, B., Espa, M., Berthillot, C., Danjou, F., Contu, P., Aissa, A. H., Virieux, S. R., Colpart, J. J., & Martin, X. (2005). Kidney preservation with IGL-1 solution: A preliminary report. *Transplantation Proceedings, 37*(1), 308–311. https://doi.org/10.1016/j.transproceed.2004.12.045

Bahl, D., Haddad, Z., Datoo, A., & Qazi, Y. A. (2019). Delayed graft function in kidney transplantation. *Current Opinion in Organ Transplantation, 24*(1), 82–86. https://doi.org/10.1097/MOT.0000000000000604

Bon, D., Claire, B., Thuillier, R., Hebrard, W., Boildieu, N., Celhay, O., Irani, J., Seguin, F., & Hauet, T. (2014). Analysis of perfusates during hypothermic machine perfusion by NMR spectroscopy: A potential tool for predicting kidney graft outcome. *Transplantation, 97*(8), 810–816. https://doi.org/10.1097/TP.0000000000000046

Cooper, J. T., Chin, L. T., Krieger, N. R., Fernandez, L. A., Foley, D. P., Becker, Y. T., Odorico, J. S., Knechtle, S. J., Kalayoglu, M., Sollinger, H. W., & D'Alessandro, A. M. (2004). Donation after cardiac death: The University of Wisconsin experience with renal transplantation. *American Journal of Transplantation, 4*(9), 1490–1494. https://doi.org/10.1111/j.1600-6143.2004.00531.x

Darius, T., Vergauwen, M., Smith, T. B., Patel, K., Craps, J., Joris, V., Aydin, S., Ury, B., Buemi, A., De Meyer, M., Nath, J., Ludwig,

C., Dessy, C., Many, M.-C., Gianello, P., & Mourad, M. (2020). Influence of different partial pressures of oxygen during continuous hypothermic machine perfusion in a pig kidney ischemia-reperfusion autotransplant model. *Transplantation, 104*(4), 731–743. https://doi.org/10.1097/TP.0000000000003051

Esmaeilzadeh, M., Sadeghi, M., Heissler, H. E., Galmbacher, R., Majlesara, A., Al–Afif, S., & Mehrabi, A. (2020). Experimental rat model for brain death induction and kidney transplantation. *Journal of Investigative Surgery, 33*(2), 141–146. https://doi.org/10.1080/08941939.2018.1480677

Faucher, Q., Alarcan, H., Sauvage, F.-L., Forestier, L., Miquelestorena-Standley, E., Nadal-Desbarats, L., Arnion, H., Venhard, J.-C., Brichart, N., Bruyère, F., Marquet, P., & Barin-Le Guellec, C. (2022). Perfusate metabolomics content and expression of tubular transporters during human kidney graft preservation by hypothermic machine perfusion. *Transplantation, 106*(9), 1831–1843. https://doi.org/10.1097/TP.0000000000004129

Fujii, T., Takaoka, M., Muraoka, T., Kurata, H., Tsuruoka, N., Ono, H., Kiso, Y., Tanaka, T., & Matsumura, Y. (2003). Preventive effect of l-carnosine on ischemia/reperfusion-induced acute renal failure in rats. *European Journal of Pharmacology, 474*(2–3), 261–267. https://doi.org/10.1016/S0014-2999(03)02079-X

Geubelle, P., Gilissen, J., Dilly, S., Poma, L., Dupuis, N., Laschet, C., Abboud, D., Inoue, A., Jouret, F., Pirotte, B., & Hanson, J. (2017). Identification and pharmacological characterization of succinate receptor agonists. *British Journal of Pharmacology, 174*(9), 796–808. https://doi.org/10.1111/bph.13738

Gilissen, J., Jouret, F., Pirotte, B., & Hanson, J. (2016). Insight into SUCNR1 (GPR91) structure and function. *Pharmacology & Therapeutics, 159*, 56–65. https://doi.org/10.1016/j.pharmthera.2016.01.008

Guy, A. J., Nath, J., Cobbold, M., Ludwig, C., Tennant, D. A., Inston, N. G., & Ready, A. R. (2015). Metabolomic analysis of perfusate during hypothermic machine perfusion of human cadaveric kidneys. *Transplantation, 99*(4), 754–759. https://doi.org/10.1097/TP.0000000000000398

Hauet, T., Baumert, H., Gibelin, H., Hameury, F., Goujon, J. M., Carretier, M., & Eugene, M. (2000). Noninvasive monitoring of citrate, acetate, lactate, and renal medullary osmolyte excretion in urine as biomarkers of exposure to ischemic reperfusion injury. *Cryobiology, 41*(4), 280–291. https://doi.org/10.1006/cryo.2000.2291

Hrydziuszko, O., Perera, M. T. P. R., Laing, R., Kirwan, J., Silva, M. A., Richards, D. A., Murphy, N., Mirza, D. F., & Viant, M. R. (2016). Mass Spectrometry based metabolomics comparison of liver grafts from donors after circulatory death (DCD) and donors after brain death (DBD) used in human orthotopic liver transplantation. *PLoS ONE, 11*(11), e0165884. https://doi.org/10.1371/journal.pone.0165884

Humphries, T., Vesey, D., Galloway, G., Gobe, G., Francis, R. (2023). Identifying disease progression in chronic kidney disease using proton magnetic resonance spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy, 134-135*, 52–64 https://doi.org/10.1016/j.pnmrs.2023.04.001

Iwamoto, H., Okihara, M., Akashi, I., Kihara, Y., Konno, O., Kawachi, S., Sunamura, M., Sugimoto, M. (2022). Metabolomic profiling of plasma, urine, and saliva of kidney transplantation recipients. *International journal of molecular sciences, 23*(22), 13938 https://doi.org/10.3390/ijms232213938

Jacob, M., Lopata, A. L., Dasouki, M., & Abdel Rahman, A. M. (2019). Metabolomics toward personalized medicine. *Mass Spectrometry Reviews, 38*(3), 221–238. https://doi.org/10.1002/mas.21548

Jouret, F., Leenders, J., Poma, L., Defraigne, J.-O., Krzesinski, J.-M., & de Tullio, P. (2016). Nuclear magnetic resonance metabolomic profiling of mouse kidney, urine and serum following renal ischemia/reperfusion injury. *PLoS ONE, 11*(9), e0163021. https://doi.org/10.1371/journal.pone.0163021

Khbouz, B., Lallemand, F., Cirillo, A., Rowart, P., Legouis, D., Sounni, N. E., Noël, A., De Tullio, P., De Seigneux, S., & Jouret, F. (2022). Kidney-targeted irradiation triggers renal ischemic preconditioning in mice. *American Journal of Physiology-Renal Physiology, 323*(2), F198–F211. https://doi.org/10.1152/ajprenal.00005.2022

Koffman, G., & Gambaro, G. (2003). Renal transplantation from non-heart- beating donors: A review of the European experience. *Journal of Nephrology, 16*(3), 334–341.

Kurata, H., Fujii, T., Tsutsui, H., Katayama, T., Ohkita, M., Takaoka, M., Tsuruoka, N., Kiso, Y., Ohno, Y., Fujisawa, Y., Shokoji, T., Nishiyama, A., Abe, Y., & Matsumura, Y. (2006). Renoprotective effects of *l* -carnosine on ischemia/reperfusion-induced renal injury in rats. *Journal of Pharmacology and Experimental Therapeutics, 319*(2), 640–647. https://doi.org/10.1124/jpet.106.110122

Letertre, M. P. M., Giraudeau, P., & De Tullio, P. (2021). Nuclear magnetic resonance spectroscopy in clinical metabolomics and personalized medicine: Current challenges and perspectives. *Frontiers in Molecular Biosciences, 8*, 698337. https://doi.org/10.3389/fmolb.2021.698337

Lunyera, J., Diamantidis, C. J., Bosworth, H. B., Patel, U. D., Bain, J., Muehlbauer, M. J., Ilkayeva, O., Nguyen, M., Sharma, B., Ma, J. Z., Shah, S. H., & Scialla, J. J. (2022). Urine tricarboxylic acid cycle signatures of early-stage diabetic kidney disease. *Metabolomics, 18*(1), 5. https://doi.org/10.1007/s11306-021-01858-4

Mannon, R. B. (2018). Delayed graft function: The AKI of kidney transplantation. *Nephron, 140*(2), 94–98. https://doi.org/10.1159/000491558

Metzger, R. A., Delmonico, F. L., Feng, S., Port, F. K., Wynn, J. J., & Merion, R. M. (2003). Expanded criteria donors for kidney transplantation. *American Journal of Transplantation, 3*, 114–125. https://doi.org/10.1034/j.1600-6143.3.s4.11.x

Nath, J., Smith, T. B., Patel, K., Ebbs, S. R., Hollis, A., Tennant, D. A., Ludwig, C., & Ready, A. R. (2017). Metabolic differences between cold stored and machine perfused porcine kidneys: A 1H NMR based study. *Cryobiology, 74*, 115–120. https://doi.org/10.1016/j.cryobiol.2016.11.006

Novitzky, T., Cooper, D. K. C., Morrell, D., & Isaacs, S. (1988). Change from aerobic to anaerobic metabolism after brain death and reversal following triiodothyronine therapy. *Transplantation, 45*(1), 32–36. https://doi.org/10.1097/00007890-198801000-00008

Peng, P., Ding, Z., He, Y., Zhang, J., Wang, X., & Yang, Z. (2019). Hypothermic machine perfusion versus static cold storage in deceased donor kidney transplantation: A systematic review and meta-analysis of randomized controlled trials. *Artificial Organs, 43*(5), 478–489. https://doi.org/10.1111/aor.13364

Ponticelli, C. (2014). Ischaemia-reperfusion injury: A major protagonist in kidney transplantation. *Nephrology Dialysis Transplantation, 29*(6), 1134–1140. https://doi.org/10.1093/ndt/gft488

Ponticelli, C., Reggiani, F., & Moroni, G. (2022). Delayed graft function in kidney transplant: Risk factors, consequences and prevention strategies. *Journal of Personalized Medicine, 12*(10), 1557. https://doi.org/10.3390/jpm12101557

Rao, P. S., Schaubel, D. E., Guidinger, M. K., Andreoni, K. A., Wolfe, R. A., Merion, R. M., Port, F. K., & Sung, R. S. (2009). A comprehensive risk quantification score for deceased donor kidneys: The kidney donor risk index. *Transplantation, 88*(2), 231–236. https://doi.org/10.1097/TP.0b013e3181ac620b

Reese, P. P., Aubert, O., Naesens, M., Huang, E., Potluri, V., Kuypers, D., Bouquegneau, A., Divard, G., Raynaud, M., Bouatou, Y., Vo, A., Glotz, D., Legendre, C., Lefaucheur, C., Jordan, S., Empana, J.-P., Jouven, X., & Loupy, A. (2021). Assessment of the utility of kidney histology as a basis for discarding organs in the United States: A comparison of international transplant practices and outcomes. *Journal of the American Society of Nephrology, 32*(2), 397–409. https://doi.org/10.1681/ASN.2020040464

Robert, R., Guilhot, J., Pinsard, M., Longeard, P.-L., Jacob, J.-P., Gissot, V., Hauet, T., & Seguin, F. (2010). A pair analysis of the delayed graft function in kidney recipient: The critical role of the donor. *Journal of Critical Care, 25*(4), 582–590. https://doi.org/10.1016/j.jcrc.2010.02.011

Rodriguez-Niño, A., Pastene, D. O., Post, A., Said, M. Y., Gomes-Neto, A. W., Kieneker, L. M., Heiner-Fokkema, M. R., Esatbeyoglu, T., Rimbach, G., Schnuelle, P., Yard, B. A., & Bakker, S. J. L. (2021). Urinary carnosinase-1 excretion is associated with urinary carnosine depletion and risk of graft failure in kidney transplant recipients: Results of the transplantlines cohort study. *Antioxidants, 10*(7), 1102. https://doi.org/10.3390/antiox10071102

Saat, T. C., Van Den Engel, S., Bijman-Lachger, W., Korevaar, S. S., Hoogduijn, M. J., IJzermans, J. N. M., & De Bruin, R. W. F. (2016). Fate and effect of intravenously infused mesenchymal stem cells in a mouse model of hepatic ischemia reperfusion injury and resection. *Stem Cells International, 2016*, 1–9. https://doi.org/10.1155/2016/5761487

Salvadori, M., Rosso, G., & Bertoni, E. (2015). Update on ischemia-reperfusion injury in kidney transplantation: Pathogenesis and treatment. *World Journal of Transplantation, 5*(2), 52. https://doi.org/10.5500/wjt.v5.i2.52

Siedlecki, A., Irish, W., & Brennan, D. C. (2011). Delayed graft function in the kidney transplant. *American Journal of Transplantation, 11*(11), 2279–2296. https://doi.org/10.1111/j.1600-6143.2011.03754.x

Soussi, D., Rod, X., Thuillier, R., Leblanc, S., Goujon, J.-M., Barrou, B., Hauet, T., & Kerforne, T. (2019). Preclinical modeling of DCD class III donation: Paving the way for the increased use of this challenging donor type. *BioMed Research International, 2019*, 5924101. https://doi.org/10.1155/2019/5924101

van Der Hoeven, J. A. B., Molema, G., Ter Horst, G. J., Freund, R. L., Wiersema, J., van Schilfgaarde, R., Leuvenink, H. G. D., & Ploeg, R. J. (2003). Relationship between duration of brain death and hemodynamic (in)stability on progressive dysfunction and increased immunologic activation of donor kidneys. *Kidney International, 64*(5), 1874–1882. https://doi.org/10.1046/j.1523-1755.2003.00272.x

Wang, Z., Durai, P., & Tiong, H. Y. (2020). Expanded criteria donors in deceased donor kidney transplantation—An Asian perspective. *Indian Journal of Urology: IJU: Journal of the Urological Society of India, 36*(2), 89–94. https://doi.org/10.4103/iju.IJU_269_19

Wishart, D. S. (2006). Metabolomics in monitoring kidney transplants. *Current Opinion in Nephrology and Hypertension, 15*(6) 637–642. https://doi.org/10.1097/01.mnh.0000247499.64291.52

Wong, G., Teixeira-Pinto, A., Chapman, J. R., Craig, J. C., Pleass, H., McDonald, S., & Lim, W. H. (2017). The impact of total ischemic time, donor age and the pathway of donor death on graft outcomes after deceased donor kidney transplantation. *Transplantation, 101*(6), 1152–1158. https://doi.org/10.1097/TP.0000000000001351