New translational metabolomics exploration of Age-related macular degeneration: improving risk assessment and « real-life » patient's follow-up

SCHOUMACHER MATTHIEU

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)
METABOLOMICS GROUP
UNIVERSITY OF LIEGE (ULiège)

SUPERVISORS: DR. PASCAL DE TULLIO AND PROF. BERNARD PIROTTE

April 2023

Acknowledgements — Remerciements

Si ce travail résume mes 7 années de thèse passées au sein du service pharmaceutique, il ne reflète pas l'aventure humaine incroyable que j'ai eu la chance de vivre. A l'ensemble des collègues, qui sont devenus des amis, des proches, avec qui j'ai eu la chance de partager cette expérience de vie, je vous dédie ces quelques lignes pour vous dire merci !

Merci tout d'abord au Professeur Bernard Pirotte pour m'avoir accueilli au sein du service et pour m'avoir donné la chance de poursuivre mes objectifs. Depuis votre cours M2 de chimie pharmaceutique en passant par mon mémoire et tout au long de ma thèse, vous avez joué un rôle constant dans ma formation et vous avez contribué à faire de moi le scientifique consciencieux que je suis. Merci à vous Chef !

Merci au docteur Pascal de Tullio sans qui je n'aurais certainement jamais commencé cette superbe aventure. Chef, vous m'avez guidé tout au long de ce travail et transmis cette passion pour la recherche et tout particulièrement pour la métabolomique. Merci d'avoir cru en moi et de m'avoir fait confiance toutes ces années. Vous m'avez donné l'opportunité de grandir au sein de l'équipe et de me développer personnellement au travers des diverses tâches qui m'étaient assignées. Merci Chef pour tous ces moments inoubliables passés lors de nos congrès, RFMF et autres, ça va me manquer…

Merci également au Docteur Vincent Lambert pour m'avoir conseillé, formé et guidé tout au long de mon doctorat. Coach, le mot est bien choisi, vous avez été un exemple pour moi et c'est grâce à vous que j'ai appris à ne jamais baisser les bras. Vous avez toujours cru en moi et vous m'avez transmis la force nécessaire pour aller au bout des choses tout en gardant le sourire. Merci pour vos enseignements, tant scientifiques que sur la vie en général, grâce à vous j'ai grandis. Merci coach !

Merci au Professeur Michel Frederich pour sa disponibilité et pour ses conseils avisés lors des nombreux comités de thèses passés. Vos avis m'ont permis d'approfondir mes analyses, d'optimiser mon temps et d'améliorer mes compétences. Merci !

Merci ensuite au Professeur Bernadette Govaerts pour l'aide précieuse fournie sur la statistique et l'utilisation des outils liés à la métabolomique. Merci pour tout le temps que vous m'avez accordé, toujours avec le sourire, merci de m'avoir encouragé à me former et à aller plus loin dans la manière d'aborder les études métabolomiques. Vos précieux conseils m'ont définitivement fait changer la manière de concevoir et de conduire les études métabolomiques. Grâce à vous je suis devenu un scientifique plus complet, curieux et ouvert d'esprit, je vous en remercie.

Merci au Professeur Jean-Charles Martin, vous représentez un des aspects de ma thèse que j'ai le plus apprécié et qui me laissera un souvenir indélébile. Le RFMF, un rassemblement de personnes extraordinaires, à la pointe de leurs domaines respectifs, ayant la volonté de transmettre leurs passions et leurs connaissances, doté d'une sacrée descente (bière, vin, chartreuse, … tout ce qui se boit en fait) et un sens de la fête insatiable. Vous faites partie des personnes incroyables que j'ai eu la chance de rencontrer. Je garderais toutes ces rencontres bien en mémoire, les photos bien cachées, et soyez certains que je continuerais à faire vivre l'esprit du RFMF partout où je me trouverais.

Merci à l'ensemble des collègues du LCP, c'est avec nostalgie que je repenserais à vous. L'humour pourri, l'humeur de chien dès potron-minet ou encore l'enthousiasme sportif inconditionnel de certains d'entre vous me laissera un grand vide le lundi matin. Tout comme les discussions interminables partagées autour d'une tasse de café. Une pensée à mon mentor et ami, celui qui m'a donné ma chance tout au début alors que j'étais son étudiant. Si j'ai pu continuer et me lancer dans cette aventure c'est aussi grâce à toi. Pour finir, une grosse pensée pour tous les membres, passés et actuels, de la « metabolomics band ». Vous étiez des collègues et amis formidables, j'ai passé des moments inoubliables avec vous, tant au labo qu'en congrès ou autres sorties. Je suis arrivé chaque jour le sourire aux lèvres au travail et c'est grâce à vous. Vous avez été une source de motivation pour moi, je n'y serais pas arrivé sans vous. Merci !

A tous mes potes, mes frères d'armes, de l'époque du banc, du collège et de l'arrêt de bus, à mes cousins. Il n'y a pas de place ici pour vous décrire l'impact que vous avez eu dans ma vie. A travers notre histoire, j'ai grandi, appris à ne jamais baisser les bras, à toujours rester moi-même et à représenter les miens quoiqu'il en coute. Je suis fier du chemin que nous avons parcouru. J'ai vécu chacune de vos réussites comme une réussite personnelle et aujourd'hui je vous partage la mienne. Merci pour tous ces moments inoubliables tout au long de ces années et pour les nombreux autres à venir. Notre histoire n'est pas finie ! Gloire à nous !

Merci à mes parents et à mon frère, sur qui j'ai toujours pu compter depuis le début. C'est grâce à vous que je suis devenu l'homme que je suis. Vous m'avez transmis votre sympathie, votre empathie, votre curiosité et votre ténacité. C'est grâce à ces qualités que j'ai pu mener ce projet à son terme et que je suis devenu un homme accompli. Je suis fier de notre famille, je vous aime. Merci à vous trois !

A Noesy, ma fidèle loutre, et toutes ses dents. La vie ne serait sans doute pas aussi mordante sans ta présence ! Nos balades sont un bol d'air frais qui ont permis d'aérer mon esprit et de me garder en activité tout au long de l'écriture de ce manuscrit.

A toi,

Cette dernière partie est pour toi, toi qui me soutiens, qui me guide, qui me comprend et me supporte au quotidien. Ces quelques lignes ne suffiront pas pour te dire au combien tu es essentielle à ma vie. Petit love, je te remercie pour la personne que tu es et pour la personne que je deviens à tes côtés. Tu prends soin de moi et de notre petite loutre au quotidien. Je te remercie pour tout ! Je te dédie ces quelques lignes, cette thèse, ma vie… Ti amo tantissimo

« Je ne sais où va mon chemin, mais je marche mieux quand ma main serre la tienne » (A. de Musset)

New translational metabolomics exploration of Age-related macular degeneration: improving risk assessment and « real-life » patient's follow-up

In this work, an NMR-based metabolomics approach was applied to the study of age-related macular degeneration (AMD) in way to explore several challenges linked to patient's management: (1) patient's stratification; (2) evaluation of disease progression and patient's responses; and (3) early diagnostic. To achieve these goals, blood derived samples coming from two different cohorts were analyzed. The first cohort consist of plasma samples collected from AMD patient in active and inactive phase of the disease and from healthy volunteers. NMR-based metabolomics analysis of these samples will provide us information about disease status and patient stratification. For the second cohort, a follow-up of 32 advanced AMD patients under treatment over a period of two years (269 time points) was made. At each time point, clinical data regarding nAMD patients' status was collected and used for biomarker discovery. NMR metabolomics approach was used to correlate changes among the metabolome with AMD patients' evolution. The last cohort is composed of blood derived samples collected from 471 people with no sign of AMD and followed during a period of 7.8 years. During this period, blood samples and clinical data regarding their AMD incidence were recorded. NMR-based metabolomics and lipidomics analyses aimed to highlight spectral signatures of AMD incidence, representing a unique opportunity to improve AMD risk assessment for elderly population. The main objectives of this thesis were to provide new tools able to improve patients' management and identify key metabolites linked to disease occurrence and evolution.

Nouvelle étude métabolomique translationnelle de la dégénérescence maculaire liée à l'âge : Amélioration de la prédiction du risque et suivis du patient.

Ce travail de thèse décrit l'application de l'approche métabolomique par RMN pour étudier la dégénérescence maculaire liée à l'âge (DMLA). Le but de cette étude était d'explorer comment l'approche métabolomique pourrait être utilisée afin de relever différents challenges associés à cette pathologie : (1) diagnostic précoce ; (2) stratification du patient ; et (3) évaluation de la progression de la maladie et de la réponse du patient au traitement. Pour atteindre ces objectifs, des échantillons sanguins provenant de différentes cohortes ont été analysés. La première cohorte de patient comprend des échantillons de plasma prélevé chez des patients atteint de DMLA exsudative en phase active, non active et chez des patients sains. L'analyse de cette cohorte via de nouveau outils et les informations supplémentaires apportées nous permettrons d'identifier les biomarqueurs en lien avec la stratification du patient. Pour la seconde cohorte, un suivi longitudinal de 32 patients en cours de traitement pour leurs DMLA exsudative a été réalisé sur une période de deux ans (269 points de contrôle). A chaque visite, des informations concernant le statu pathologique du patient ainsi que des échantillons sanguins et des mesure biochimiques ont été collectés dans le but de découvrir de nouveaux biomarqueurs permettant d'évaluer la réponse au traitement et la progression de la maladie. La troisième et dernière cohorte se rapporte aux échantillons sanguins collecté auprès de 471 personnes exempte de DMLA et suivie sur une période de 7,8 ans. Durant cet interval de temps, les échantillons sanguins ainsi que les données cliniques se rapportant au développement de leurs DMLA ont été enregistrés. Cette cohorte représente une opportunité unique d'appliquer l'approche métabolomique par RMN afin de découvrir de nouveaux biomarqueurs précoces de la pathologie. L'objectif global de ce travail était donc la mise en place de nouveaux outils afin de mieux diagnostiquer le patient et d'améliorer son suivi au cours de l'évolution de la pathologie. Ceci, dans le but de proposé une thérapeutique alternative et mieux adaptée au besoin du patient.

Abbreviation list

Abbreviation	complete name	Abbreviation	complete name
AI-bucketing	Adaptative Intelligent bucketing	PLS	Partial Least Square
AMD	Age related Macular degeneration	POS	Photoreceptor Outer Segment
AsLS	Asymetric Least Square smooting	PQN	Probabilistic Qotient Normalization
AUC	Aera Under the Curve	PSA	Protstate-Specific Antigen
BER	Balanced Error Rate	PTW	Parametric Time Warping
BM	Bruch's Membrane	ROC	Recieving Operator Characteristic
BMI	Body Mass Index	ROS	Reactive Oxygen Species
CFH	Complement Factor H	RPE	Retinal Pigment Epithelium
CLS	Constrained Line Shape	RR	Relative Risk
CNV	Choroidal NeoVascularization	SNR	Signal to Noise Ratio
COSY	Correlation Spectroscopy	SRF	SubRetinal Fluid
CSN	Constant Sum Normalization	SVM	Support Vector Machine
DA	Discriminant Analysis	TNF	Tumor Necrosis Factor
DCA	Dichloroacetic Acid	VA	Visual Acuity
ELM	External Limiting Membrane	VEGF	Vascular Endothelial Growth Factor
FA	Fluorescein Angiography	VIP	Variable Importance in Projection
FGF	Fibroblast Growth Factor	VLDL	Very Low-Density Lipoprotein
FID	Free Induction Decay		
GA	Geographic Atrophy		
GC	Gas Chromatography		
HDL	High-Density Lipoprotein		
HIFs	Hypoxia Inducible Factors		
HPLC	High Performance Liquid Chromatography		
HSQC	Heteronuclear Single Quantum Correlation		
IDL	Intermeridate Density Lipoprotein		
ILM	Internal Limiting Membrane		
IRC	Intraretinal Cystoid Fluid		
LC	Liquid Chromatography		
LDL	Low-Density Lipoprotein		
MAC	Membrane Attack Complex		
MMPs	Matrix MetalloProteinases		
MS	Mass Spectrometry		
MSI	Metabolomics Standard Initiative		
nAMD	neovascular Age related degeneration		
NIH	National Institutes of Health		
NMR	Nuclear Magnetic Resonance spectroscopy		
NN	Neural Network		
OCT	Optical Coherence Tomography		
OPLS	Orthogonal Partial Least Square		
PCA	Principal Component Analysis		
PCR	Principal Component Regression		
PDGF	Platelet-Derived Growth Factor		
PDH	Pyruvate Dehydrogenase		
PDK	Pyruvate Dehydrogenase Kinase		
PED	Pigment Epithelial Detachment		
PIGF	pPlacenta Growth Factor		

Table of content

Chapter 1:
Introduction

1   Introduction

1.1   Age-related Macular Degeneration (AMD)

1.1.1   Background of AMD

Age-related macular degeneration (AMD) is the leading cause of vision loss among elderly population in the Western countries[1]. The prevalence of AMD increases with aging and, since the growth of our life expectancy, the number of projected cases is likely to rise in a dramatic manner. In 2014, Wong *et al* projected an increment of 47% between 2020 and 2040 reaching 288 millions of cases[2].

Clinically, AMD is classified in three different stages and 90% of the vision loss caused by AMD are due to its last stage, named neovascular Age-related Macular Degeneration (nAMD). Despite recent advance in the management of AMD patient and the development of anti-vascular endothelial growth factor (anti-VEGF) therapy to treat advanced form of AMD, no permanent cure for this disease has been found. Indeed, the complete physiopathology and the age-related change that induce AMD and led to nAMD remains unclear. For these reasons, there is a need to study this complex pathology by using patient driven approaches such as Metabolomics in the context of personalized medicine.

In this chapter, we will first focus on describing the pathology and its evolution through the different stages of the disease. We will put light on the epidemiology and the different risk factors that play key roles in this complex multifactorial pathology and explain the current strategy for prevent AMD. Finally, the current medicinal practice available and the management of patient will be discussed.

1.1.2 Stages of AMD

Age-related macular degeneration is a pathologic condition affecting the macula and is responsible of 8.7% of all blindness worldwide[3]. The macula is the central part of the retina responsible of the central vision. That specific region of the eye has a particular importance due to its high density of cones photoreceptive cells that allow high acuity vision[4]. Macular degeneration leads to a loss of central vision needed for daily activities that requires detailed vision such as driving, reading or face recognition. Peripheral vision may rest intact but the impairments due to AMD conduct to a decrease of independent living since it leads to a loss of ability performing daily and basic task[5]. This makes AMD have a profound impact in quality of life for those who are diagnosed.
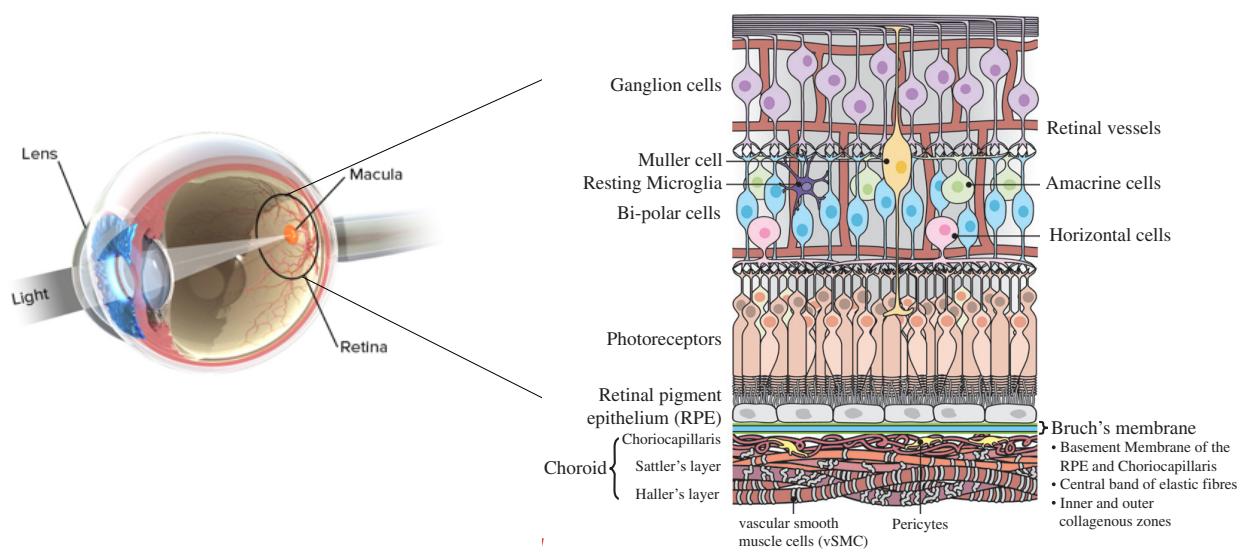
Figure 1. a) Representation of light travel to the eye hitting the central part of the retina called the Macula; b) Representation of the different component in a segment of the retina within Macula, the most affected structure in AMD are: the Retinal Pigment Epithelium (RPE), which is involved in removing waste from photoreceptive cells; the Bruch's membrane (BM) which is the barrier between the retina and the peripheral tissues; the blood vessels within the Choroid that is organized in three layers that vascularize and supply the outer retina with nutrient, and maintain the temperature and the volume of the eye. The peculiarity of the macula compared to the remaining retina is additional layers of ganglion cells and a higher ratio of cone versus rod photoreceptive cells (adapted from[6]).

Clinically, AMD is classified on three distinct stages: (i) the early asymptomatic retinal abnormalities, an intermediate form of the pathology, (ii) the geographic or atrophic form, called also "dry" AMD, and (iii) the neovascular or exudative form (nAMD or "wet" AMD). The last two stages represent the most severe form of the pathology, and although nAMD account for only 10~25% of all AMD patients, 16% of these would become legally blind in 2 years if untreated[7].

In its earliest stage, patients with early AMD exhibits drusen, lipids and proteins deposits, located between the retinal pigment epithelium (RPE) and the Bruch membrane (BM)[8] as it's shown in **Figure 2**. At this stage, the patients are mostly asymptomatic or may notice mild central distortion, difficult dark adaptation or reading difficulties[9]. The prevalence of drusen deposit increases with age; indeed, after the age of 60 approximatively one quarter of the population have drusen deposits while more than the half of adult with AMD are concerned[10]. Drusen are composed of various component including neutral lipids with (un)esterified cholesterol[9], more than 129 different proteins, zinc and iron ions[11]. Despite the fact that their composition remains mostly unchanged, their larger size is used to classify people with early AMD[12].
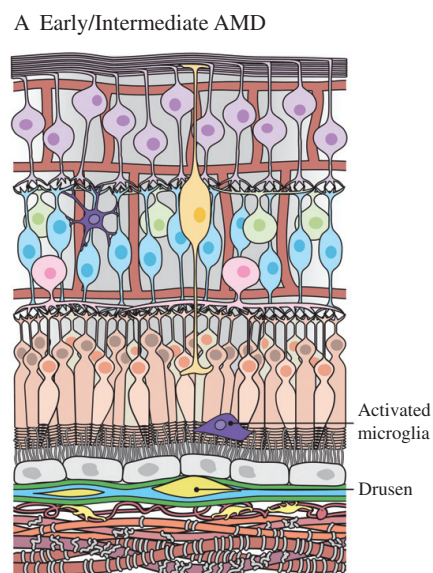
Figure 2. Representation of a retina with early/intermediate AMD. Drusen are accumulation of lipid and protein wastes are the hallmark of early/intermediate AMD. Drusen can disrupt the Bruch's membrane, alter its function, and impede the RPE's capability to transport waste across the Bruch's membrane. Activated microglia present in the outer retina is a sign of local inflammation (adapted from[6]).

The two last stages are associated with severe visual acuity depletion (**Figure 3**), and the two forms responsible for this impairment are the so-called exudative or neovascular AMD (nAMD) and geographic atrophy (GA). Early symptoms of late AMD include distorted vision when driving, watching television or reading, and dark spot in the center of vision that can lead to difficult face recognition abilities[11]. Most of the symptoms might not be perceived until both eyes are impacted by the disease. As late AMD could progress rapidly, especially for the exudative form, early diagnostic and proper patient follow-up are mandatory.

NORMAL   **AMD**
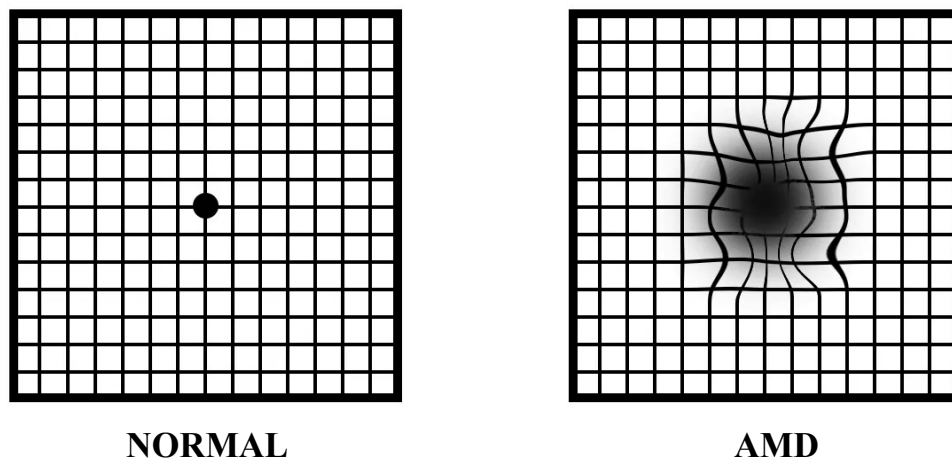
Figure 3. "Amsler chart to test your sight": The Amsler grid is used to test the vision by checking whether the lines look wavy, distorted or whether areas of the visual field are missing. This self-vision test is a common way to detect early changes caused by worsening AMD or other macular disorders. On the right is represented the result of a healthy eye looking at the grip. On the left is shown the result of advanced AMD with all the common symptoms impacting the central vision (image from the "community eye health journal": cehjournal.org).

Geographic Atrophy (GA) accounts for approximatively 35~40% of all last stage AMD cases[13]. This form is characterized by localized deterioration of retinal pigment epithelium (RPE) that led to area in which blood vessels are visible[14]. Some pigmentary alteration may be also present with either hypo or hyperpigmentation surrounding the area of the macular atrophy. A population-based study showed that lesion could be local or multifocal, the latter are more likely to invade the foveal part of the macula located in the center and decrease the visual acuity dramatically[15]. Up to now, no effective therapeutic strategies can prevent or treat patients with GA. Some studies investigated the role of inflammation, mitochondrial dysfunction, oxidative stress, lipid abnormalities and cell death in GA, but the precise mechanism of this pathology remains unclear.

Therefore, the elucidation of the complex mechanistic pathways involved on the development of geographic atrophy from early AMD is mandatory to assess proper prevention and therapeutic strategy. Thereby, in 2019, Handa et al. suggested that the future direction in dry AMD research should integrate omics, pharmacological and clinical data into mathematical model in way to predict disease onset and progression, identify biomarker, establish mechanism, and monitor response to therapy. Developing personalized medicine approach driven by data obtained from large cohort study, should be a cornerstone in the treatment of dry AMD patients.

Neovascular AMD (nAMD) is the last and most severe stage of the disease. Almost 50% of all advanced AMD cases are due to this form, whose rapid evolution and dramatic decrease of visual acuity made of it a significant clinical challenge. nAMD is characterized by abnormal choroidal neovascularization process that led to a disruption of the Bruch membrane, the retinal pigment epithelium (RPE) and the photoreceptors, resulting in focal retinal detachment and vision loss. The retina is vascularized by two independent circulatory systems, the choroid, and the retinal vessels. The inner part receives oxygen and nutrient via the retinal system while the outer part is avascular and is supplied by the choroid.

As depicted in the **Figure 4**, the choroid is composed of three vascular beds: from the bottom to the top, a layer of small arteries and veins that connect the peripheral circulation (Haller's Layer); Sattler's layer that consist of arterioles and venules; and the choriocapillaris, a network of capillaries[6].

Lying on the top of the choroid, the Bruch's membrane (BM) serves as a barrier between RPE and choroid. While BM allows to nutrient pass from choroidal vessels to RPE and photoreceptive cells, RPE allows renewal of photoreceptive cells by removing metabolic waste, absorb the excess of photons and reduce the oxidative damage to the retina[16]. In nAMD, change in the choroidal vasculature disrupt the overlying structure including Bruch's membrane, the RPE and the photoreceptors. This abnormal choroidal neovascularization (CNV) process led to visual impairment that, if not treated, conduct to legal blindness.
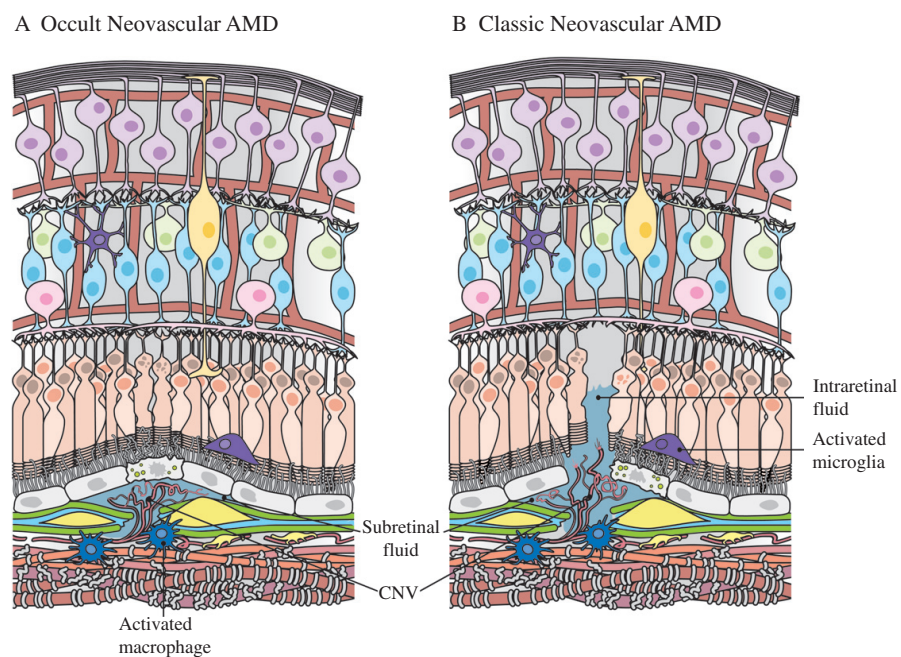
Figure 4. Representation of a retina with a) occult and b) classic neovascular AMD (nAMD). Both share identical pathologic features with the disruption of the Bruch's membrane by growing neovessels and fluids leakage accumulating under the RPE in the subretinal space distorting the local structure and causing stress in the RPE and photoreceptive cells. In the classic expression of nAMD, the neovessels break the RPE layer and invade the photoreceptor layer which result in intraretinal fluid accumulation. RPE and photoreceptor stress and loss are evident and thus this process impact the vision in a dramatic way (adapted from[6]).

1.1.3 Mechanism involved in AMD progression

AMD is a complex and multifactorial disease, it's occurrence and progression through the last stage involves at least four processes named: lipofuscinogenesis, drusogenesis, local inflammation and neovascularization (for the wet form)[17–23].
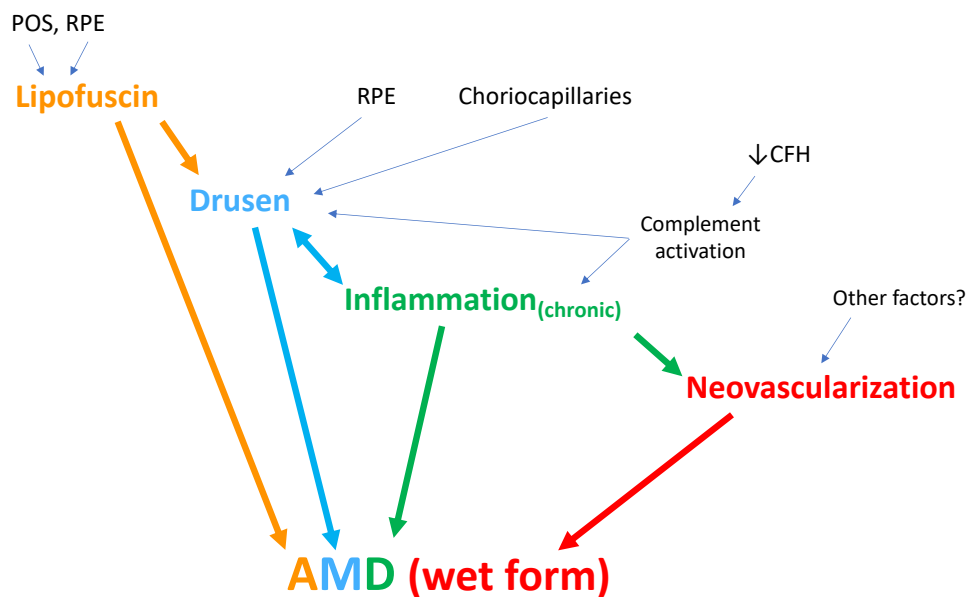
Figure 5. Main processes involved in the occurrence and progression of AMD: lipofuscin formation; drusen formation; local inflammation; and neovascularization. (POS: photoreceptor outer segment; RPE: retinal pigment epithelium; CFH: complement factor H)[24]

RPE have a variety of supportive and metabolic function that are of peculiar importance for retinal photoreceptors such as: maintenance of the blood-retina barrier, participation to the visual cycle (uptake, processing, transport and release of vitamin A derivative), and phagocytic uptake and degradation of excreted photoreceptor outer segments (POS)[25]. Thus, it is not surprising that impairments of RPE cell function is considered as an early and crucial event in the molecular pathway leading to AMD[21,22]. Indeed, age-dependent phagocytic and metabolic insufficiency of postmitotic RPE cells is one the main driving force of its dysfunction. This process leads to a progressive and accumulation of lipofuscin granules, also called "age pigment", mostly composed of lipids and proteins of phagosomal, lysosomal and photoreceptors origins. These components are modified at various levels by oxidative processes resulting of either exposure to light and the presence of high oxygen levels in the eyes[26]; some of them, such as bisretinoid fluorophore, are cytotoxic and potent photo-inducible generator of reactive oxygen species (ROS) that can damage proteins, lipids and therefore impact the RPE function[22].

Drusen are considered as characteristic of AMD. Clinically, they are divided in two types named "hard" and "soft" drusen, depending on their relative shape and size. Herein, the presence of numerous, larger size, hard and soft drusen is considered as a major risk factor for developing advanced AMD forms and losing central vision. Studies demonstrated that photoreceptors overlying drusen undergoes to degenerative changes; the negative impact of drusen on these photoreceptors relies on their physical displacement from the RPE and on their direct influence on the immune system and local inflammation. Indeed, proteomic and immunohistochemical studies performed on drusen of AMD patients indicated an implication of the complement system in the pathophysiology of AMD[17,20,27].

The complement system is part of the innate immune system that acts as a first line of defense fighting against invading pathogens; three different pathways can activate it by leading to the formation of the membrane attack complex (MAC), anaphylatoxin C3a and C5a that further lead to inflammation. In small amount, MAC is sub-lytic (does not lead to lysis) protective during inflammation, in contrast when the number of MAC reaches a given threshold, it becomes lytic and leads to membrane disruption and cell death[28,29]. Prior to MAC formation, the complement system results in the production of powerful pro-inflammatory molecules anaphylatoxin C3a and C5a. These molecules have chemoattractive effect on macrophages that express the respective receptors C3aR and C5aR. Such macrophages has been shown to be implicated in AMD as they exhibits both pro- and anti-angiogenic role in choroidal neovascularization process[27,30–32]. In fine, local inflammation and activation of the complement cascade, with uncontrolled generation of MAC, contribute to drusogenesis, RPE and photoreceptor degeneration, and Bruch's membrane disruption, these events are associated with late stage of AMD.

Choroidal neovascularization (CNV) process is the hallmark of the last stage of the disease called wet/exudative or neovascular AMD (nAMD). CNV process take place within the choroid and can be defined as an abnormal growth of blood vessels. These new neovessels are more permeable and may bleed causing leaking of fluid across BM and the RPE. Indeed, the fluid released form leaky neovessels is accumulating in the retinal subspace, leading to a distortion of the local structure, and causing important stress to the RPE and photoreceptors. Additionally, neovessels can break the BM, the RPE and invade the photoreceptors layer. The resulting intraretinal fluid is accumulating in addition to the subretinal ones and cause RPE and

photoreceptors stress and loss[6]. These events made nAMD patients developing central scotoma and in the worst-case cause legal blindness.
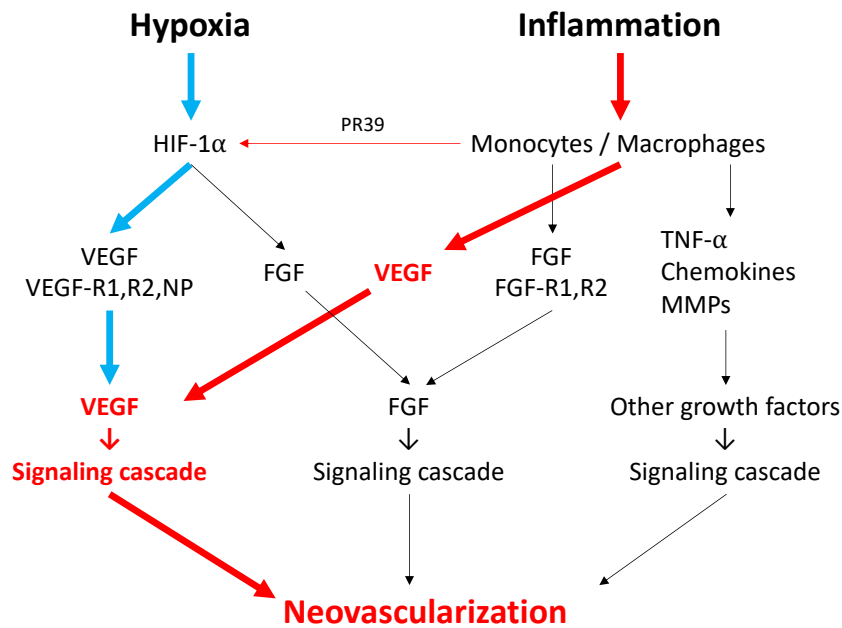
Figure 6. Hypoxia and inflammation processes leading to choroidal neovascularization. By creating a cellular and molecular milieu promoting pro-angiogenic mechanisms, immune reaction and inflammation may play a key role in the development of nAMD. (H-1α: hypoxia-inducible factor-1α; VEGF: vascular endothelial growth factor; VEGF-R1,-R2: receptors of VEGF; NP: neuropilin; FGF: fibroblast growth factor; FGF-R1,-R2: receptors of FGF; TNF-α: tumor necrosis factor; MMPs: matrix metalloproteinases; PIGF: placenta growth factor; PDGF; platelet-derived growth factor: PR39: proline- and arginine-rich peptide)[24]

Two main processes are involved in CNV development and progression (**Figure 6**), angiogenesis and inflammation[33]. Continuous stress or degradation in the RPE and the associated immune response promote the production of pro-angiogenic factors; among these, VEGFA, commonly named VEGF, is found to be a key factor in both animal and human model of nAMD[34–36]. Furthermore, in animal model, VEGF blocking treatment are proved to reduce the laser-induced CNV process in treated mice[37,38]. In physiologic condition, the upregulation of proangiogenic component such as VEGF is involved in the response to tissue hypoxia. This response is mediated by the hypoxia inducible factors (HIFs) transcription regulators and some of these factors has been recently discovered in active CNV patients[39]. In addition to this immune reaction, it's tough that there may be a role for local inflammation as a process that leads to an unbalanced increase in pro-angiogenic activity. In fact, this pathway passes through the recruitment of monocytes, macrophages and neutrophils that together will trigger the release of VEGF and lead to neovascularization.

1.1.4 Risk factors and prevention

The precise mechanism of AMD development going from early stages to pathologic CNV is still poorly understood. However, findings from ongoing studies underlines the fact that the pathogenesis of AMD is resulting from a complex multifactorial interaction between metabolic, functional, genetic and environmental factors[40].

If the impact of age has not to be demonstrated, the genetic predisposition to AMD has been assessed between 46 to 71% in studies conducted on 840 twins[41]. Furthermore, a Genome Wide Association Study published in 2015 reported a study of more than 12 million variants analyzed on 16,144 patients and 17,832 controls. From these, 52 independently associated variant across 34 loci (position on a fixed chromosome) were identified as the first genetic association signal specific to nAMD[42]. These variants were proven to play a role in controlling immune response, inflammatory processes, retina homeostasis, and therefore correlated to dysfunction of these reactions in AMD affected patients[43]. It's evident that risk factors coming from genetics cannot be overcome, nevertheless some others attributed to the lifestyle, the diet and nutrition can be modified in the way to reduce the prevalence of AMD and its development (**Figure 7**).
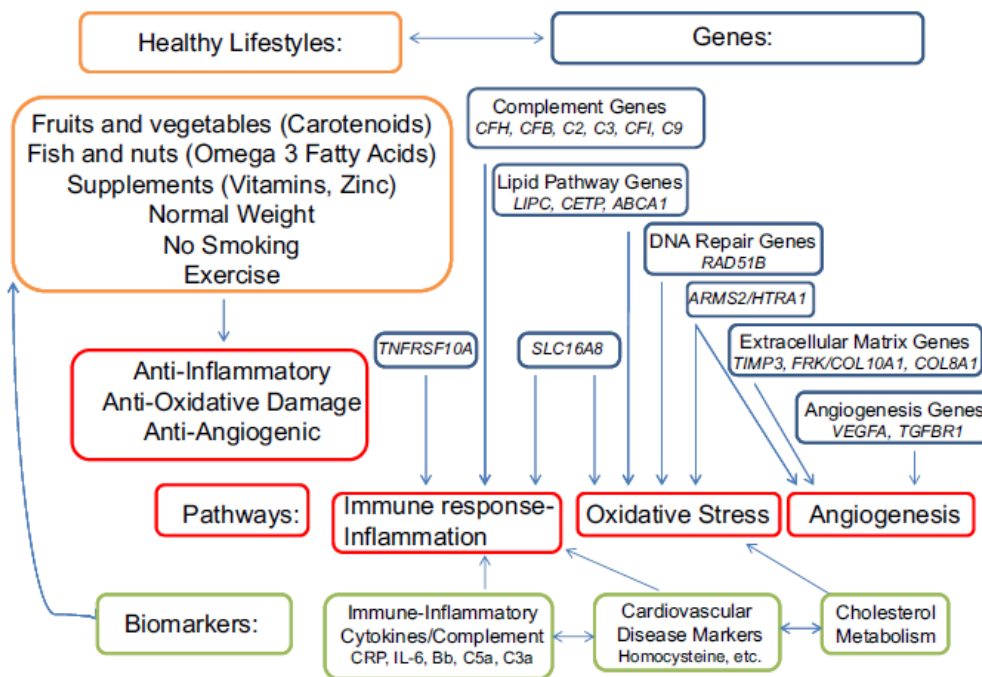

Figure 7. Schematic representation of interconnectivities between environmental and genetic risk factors linked to AMD[44].

Smokers of >40 years old are from two or four time more likely to develop AMD in comparison to nonsmokers of the same age. In fact, after aging, smoking is the strongest risk factor for AMD and progression to its last stage[45]. Other lifestyle risk factors associated with AMD are obesity, physical activity, and diet. Several studies demonstrated that early AMD patients with obesity and/or higher body mass index (BMI) had more chances for progression of AMD to advanced stages compared to patients with lower BMI[46,47]. Similarly, practicing vigorous physical activities three times a week reduced the risk of AMD progression by 25%[47].

In line to these findings, dietary factors including antioxidant and dietary fat intake influence the onset and progression of AMD[44]. On one hand, the intake of carotenoids, such as lutein and zeaxanthin, or other antioxidants from food was investigated in several studies and was associated with a lower risk for AMD[48,49,50]. Since Lutein and zeaxanthin are macular pigments, their relevance for AMD is not surprising because of their physiologic function and location in the retina[51]. Minerals such as zinc or copper may also play a role in antioxidant functions of the retina. Regarding these findings, the impact of high dose supplementation with different vitamins, carotenoids and minerals was evaluated for prevent AMD and vision loss[52]. On the other hand, higher dietary intake of fat was proven to increase risk of AMD. Indeed, there are reported associations between saturated and unsaturated fats, cholesterol, linoleic acid intake and an AMD occurrence[53–55]. However, some studies have pointed out the positive effect of omega-3 fatty acid effect. Omega-3 fatty acid is found in high quantities in fatty fish and nuts and may exert a protective effect on macular degeneration because of its antioxidative, anti-inflammatory and anti-angiogenic properties[56,57]. Omega-3 fatty acids supplementation was studied but shown controversary results[58]. Moreover, a recent study showed that adopting a mediterranean diet, which is characterized by high consumption of plant foods, fish and olive oil as the first source of fat, reduces the risk of incident advanced AMD by 41%[59,60].

Other risk factors such as higher blood pressure, diabetes, cataract surgeries and light exposure have been also associated with an augmentation of risk to develop AMD. In 2010, a systematic review of the current literature including 18 prospective and cross-sectional studies and 6 case control studies involving 1113,780 persons with 17,236 cases of late AMD estimated the association between advanced AMD and 16 pre-selected risk factors. They've classified the increased age, tobacco uses, previous cataract surgery and family history of AMD as strong risk factors, whereas increased BMI, previous cardiovascular disease, hypertension and higher plasma fibrinogen were found to be moderate risk factors[61].

The prevention strategy for AMD is mainly based on the management of modifiable risk factors[62]. In way to lower risk of AMD (or slow its progression), the National Institutes of Health (NIH), recommends stopping smoking, have a regular physical activity combined with healthy foods diet rich in leafy green vegetables and fish, and take under control blood pressure and cholesterol levels (**figure 8**). In addition to these, the over 60s are recommended to pass annually eye examination in way to allow early diagnosis and better chance to fight the disease.
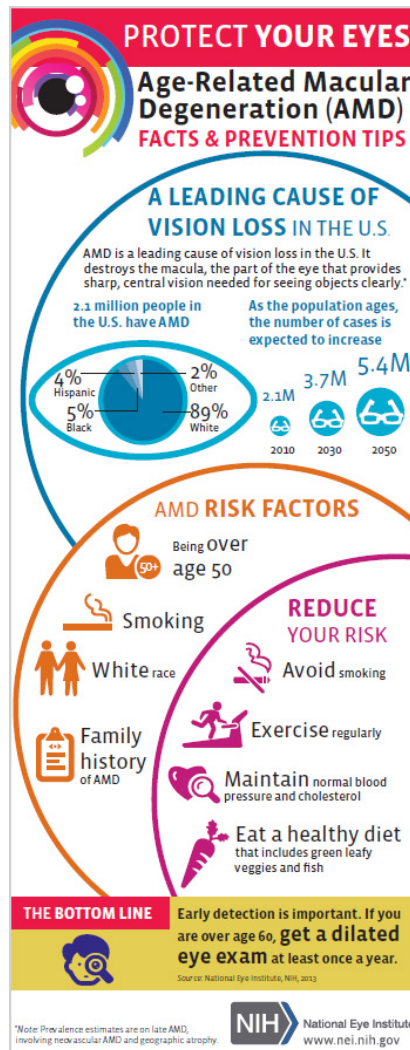
Figure 8. Recommendation for AMD prevention provided by the National Eye Institute (NEI). NIH is an American health institute that aim to prevent, treat, or even reverse vision loss (NEI recommendation updated in June 2021).

1.2   Diagnosis and Current Medicinal Practice for AMD

As mentioned earlier in this work, AMD is a fast-progressing disease, and its onset may be subtle to detect for either patient or physician. The earlier AMD is diagnosed, the later the patient will undergo severe and irreversible visual impairments. Therefore, efforts are put to award individual aged 50 years and older to perform AMD screening regularly. In addition to these general ophthalmologic examination procedures, it is recommended to perform home monitoring routine between visits. Whenever AMD is suspected, advanced diagnostic must be performed in the way to confirm and assess the stage of the disease. This diagnostic procedure includes fluorescein angiography (FA) and optical coherence tomography (OCT), and it is generally assumed that better final outcomes are achieved with better initial visual acuity (VA) and therefore an earlier diagnostic. Unfortunately, nowadays lesions are still usually detected when there are already considerable damages causing severe visual impairments. Herein, when nAMD is diagnosed, the only way for patient to recover VA is intravitreal injection of inhibitor of the angiogenic protein VEGF (**Figure 9**). If the introduction of anti-VEGF therapy was proven to be able to stabilize the vision among nAMD patients, several clinical challenges must be overcome, these will be discussed later in this chapter.
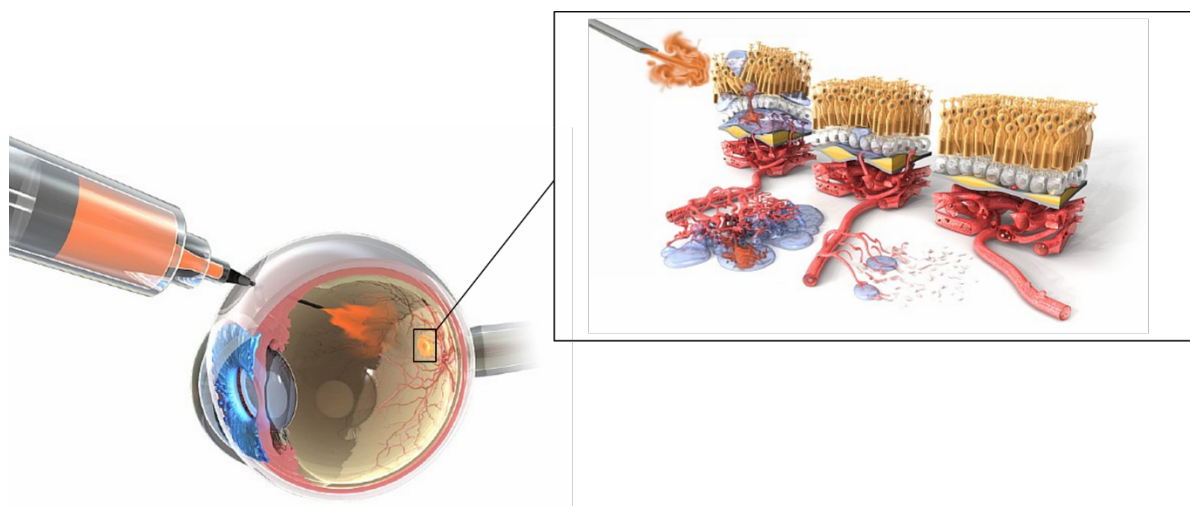
Figure 9. Illustration of intravitreal injection of antiangiogenic drugs and its effect on the neovascularization process within the retinal layer of the macula (illustration reproduced from https://www.scienceofamd.org/fr/treat/).

1.2.1 Anti-VEGF therapy and clinical strategy

Intravitreal anti-VEGF therapy is the only treatment that allows nAMD patient to recover visual acuity. Indeed, the current therapeutic strategies for managing neovascular AMD is a protocol of a fixed monthly dose of antiangiogenic drug for a 96-week follow-up period[63,64]. Despite the risk of overtreatment, and safety risk (infection, retinal detachment, development of GA, systemic side effect), monthly dosing guarantee the maximum efficacy in term of VA benefits[65,66–68]. Some studies investigated the possibility of an individualized dosing regimen based on imaging parameter to overcome these safety problems (system "as needed" named PRN). The result of these less frequent but more adequate regimen were first encouraging but failed in real-case clinical practice (**Figure 10**)[69–72]. This failure could be associated by the fact that OCT-based monitoring is usually applied in a less rigorous manner than in clinical trial. Furthermore, there is a lack of sensitive and robust OCT-based biomarkers that allows a precise individual management of the disease.
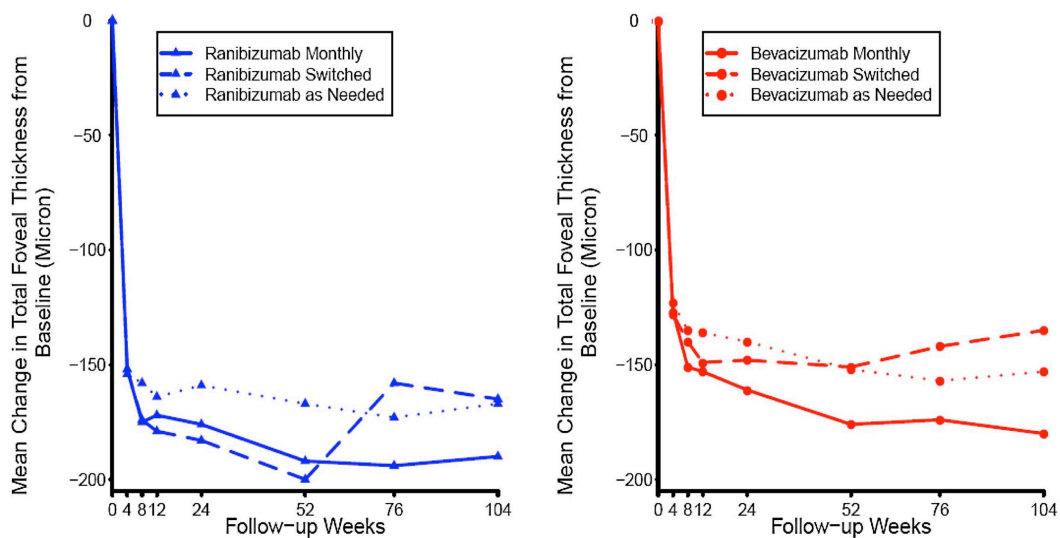
Figure 10. Data from CATT study[65]: Mean change in total foveal (central part of the macula) thickness from the start of the therapy over time by dosing regimen for two antiangiogenic drug groups: a) Ranbizumab and b) Bevacizumab. These data shown greater result for monthly than for PRN (as needed) treatment.

1.2.2 Imaging Technique and patient's follow-up

Since its introduction into Ophthalmology's sciences in 1993[73], Optical Coherence Tomography (OCT) has become the main tool for the diagnostic and the follow-up of AMD patients. This technique allows to identify specific morphological changes that are proven to be relevant for visual function, treatment outcomes and disease management[74]. In this section we will discuss the different imaging biomarkers that are commonly used to assess the status of the patients and that lead to clinical decisions.

Optical coherence tomography is an imaging technique that uses low coherence interferometry to produce cross-sectional images of the retina. Optical scattering from the tissue is captured to analyze spatial details of tissue microstructures. A super-luminescent diode products infrared light that is divided into two parts: the probe beam that is scattered from the targeted tissues, and the reference beam which is reflected from a reference mirror (**Figure 11a**). From the two reflected beams, an interference pattern is measured by a photodetector and then integrated by an interferometer that will analyze the created range echo time delay of given amplitude named A-scan. Several datapoints are integrated to construct a tomogram of the analyzed tissue. For macular scanning, three different types of scan protocol are commonly used in ophthalmology practices: a three-dimensional (3D) scan, radial scan, and raster scan (**Figure 11b-d**).
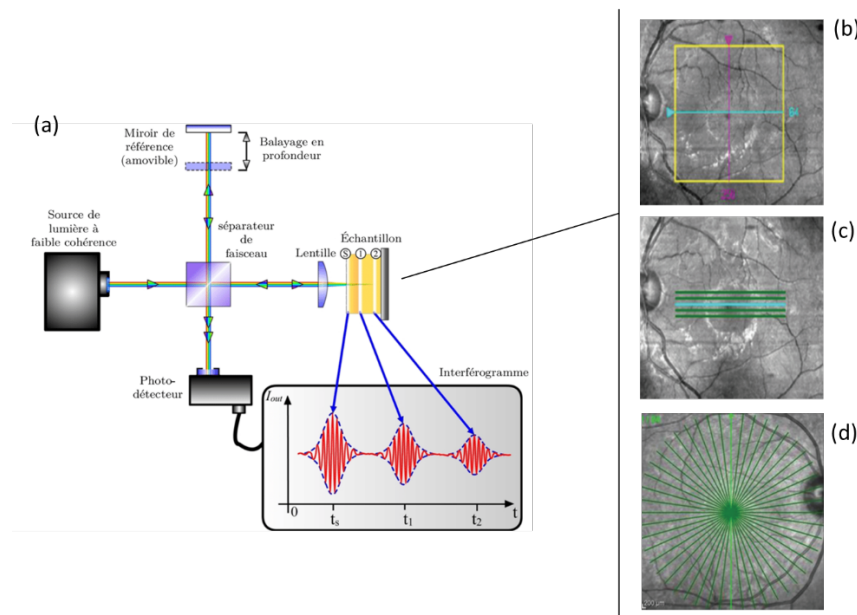
Figure 11. (a) Schematic representation of OCT technology and the three different scan protocols: (b) 3D "macular cube"; (c) radial line scan; and (c) raster scan.

If 3D scans enable the implementation of advanced and complex analysis, in practice radial and raster scan provide detailed view of the macula and give access to information that leads clinical decision. **Figure 12** represent a single radial line scan from an OCT analysis of a healthy macula. On this picture the different retinal layers that compose the macular space are depicted (from the bottom to the top: Choroid, BM, RPE, OPR, IS/OS, ELM, ONL, OPL, INL, IPL, GCL, NFL and ILM). These measurements allow to evaluate the changes within macular subspace and help the clinicians to interpret the damage caused by advanced AMD.
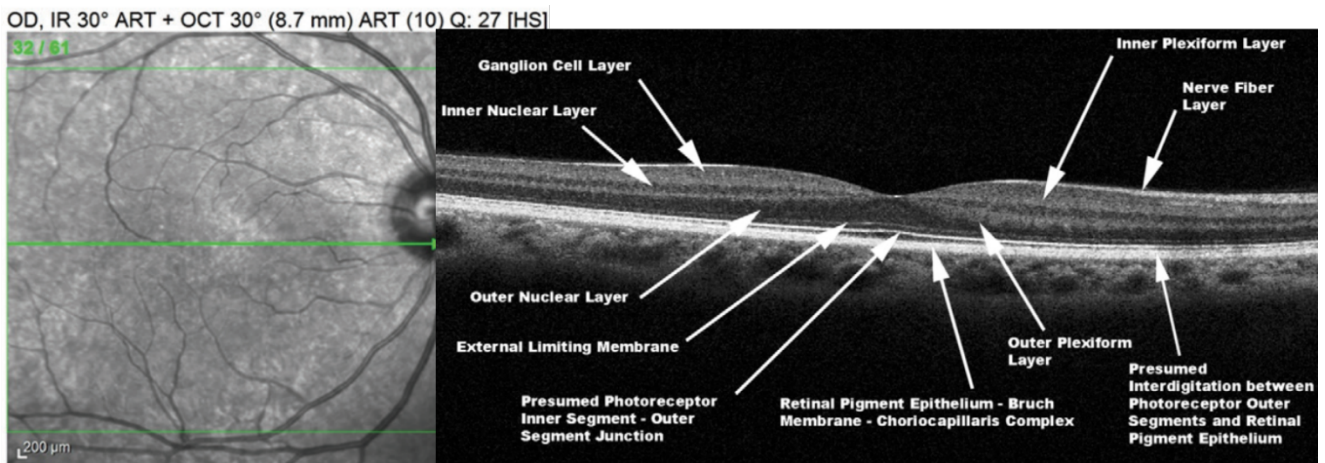
Figure 12. Radial scan of a healthy macula and the different structure detailed on the right (adapted from[75]).

Central retinal thickness is the most obvious and easiest way to quantify retinal change in OCT data measurement. Commercial OCT system can segment the internal limiting membrane (ILM) and the retinal pigment epithelium (RPE) or Bruch's membrane as the inner and outer retinal boundaries (**Figure 13**). Thus can represented thickness map and central retinal thickness value that allows clinical judgment of disease status and treatment opportunities[70,76]. However, due to poor reproducibility, lack of correlation with functional outcomes and low sensitivity to subtle change, the use of retinal thickness measurement for clinical decision is not anymore recommended[77–82].
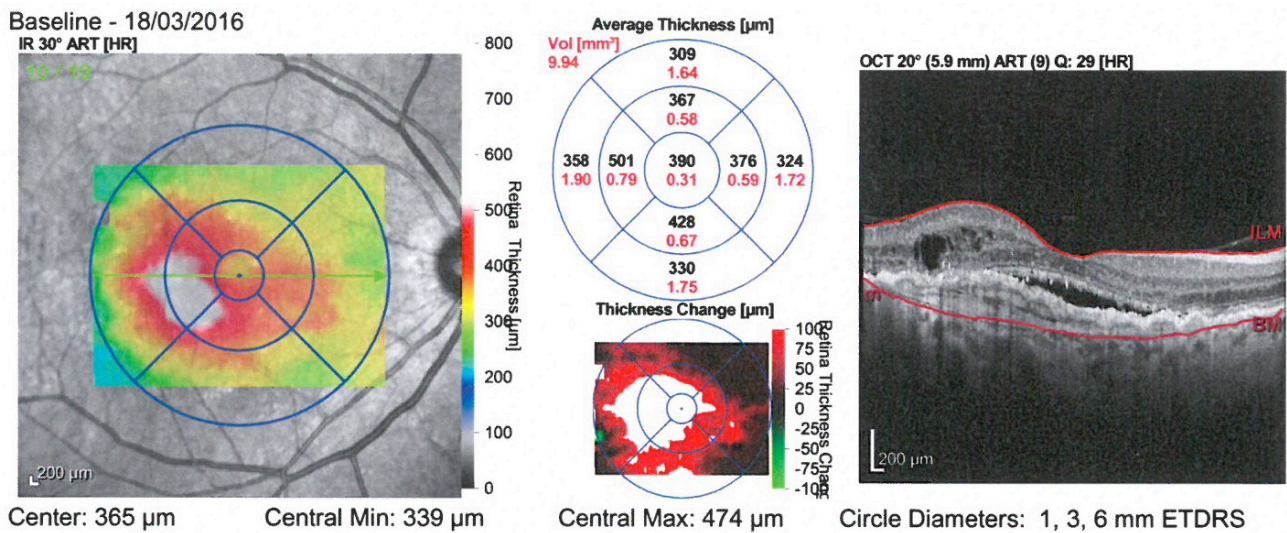
Figure 13. Retinal thickness map: On the left the colored thickness map of the macula showing in red the thickening of the retina in the macular region. The value of the average thickness is showed in 1mm, 3mm and 6mm circles centered in the foveal part of the macula. The left image shows the corresponding spectral domain optical coherence tomography image of the analyzed region.

One of the strongest markers of neovascular AMD events is the disruption of the external limiting membrane (ELM) by invasion and propagation of the CNV lesion. This results in relatively important leakage of fluid into the neurosensory retina. As it is shown in **Figure 14**, intraretinal accumulation of fluid appears as diffuse retinal thickening or as hyporeflective cystoid spaces on OCT images. This intraretinal cystoid fluid (IRC) is the most important factor for visual acuity outcomes and baseline visual acuity in neovascular AMD treatment. Studies showed that approximatively 60 % of the VA in nAMD patients could be explained by examining these changes into the retina. Moreover, IRC appears to be present in all lesion types due to exudative AMD. Usually, studies of IRC in nAMD use a presence/absence approach and very few of them attempt to quantify IRC in OCT data. This time-consuming effort will, in the future, be aided by computer assisted quantification methods. Nevertheless, care should be taken to detect and treat active CNV before the apparition of IRC as they are linked to irreversible damages. Quantification of such morphological changes therefore could be used for individualized prognostic and management.
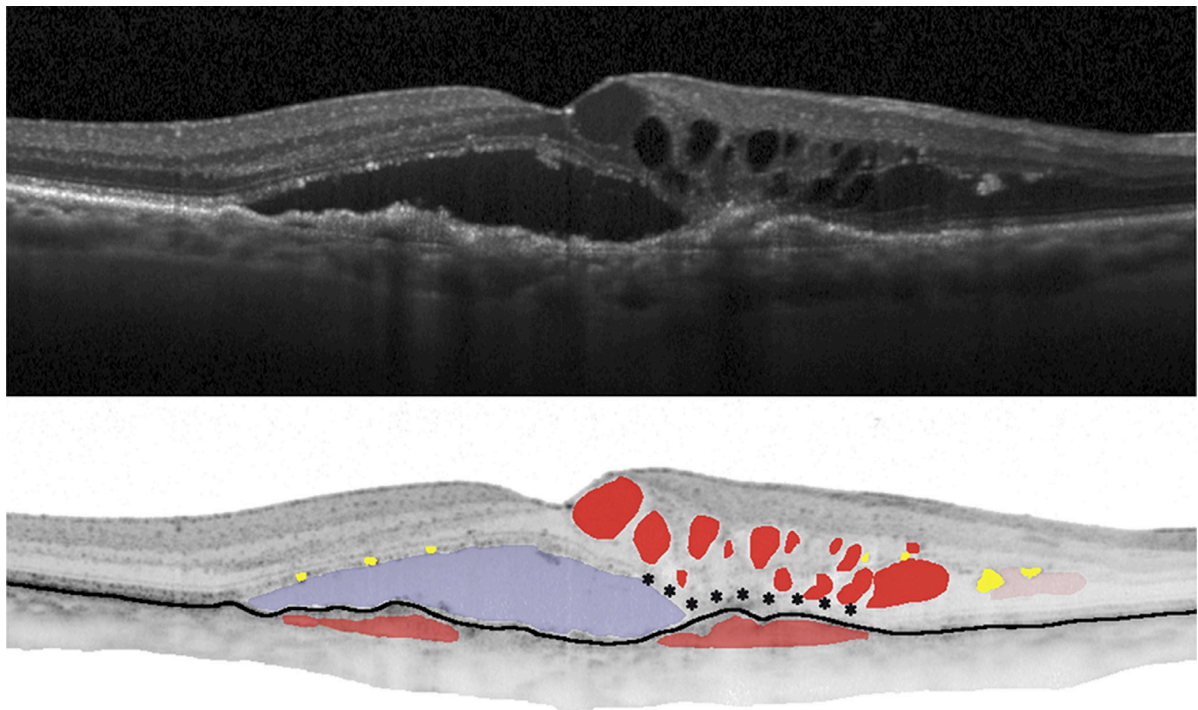
Figure 14. Typical OCT radial scan image from nAMD patient and associated imaging biomarkers. Intraretinal cystoid fluid is highlighted in bright red, subretinal fluid in blue, and pigment-epithelial detachment in dark red. Yellow markings illustrate hyperreflective foci and black asterisks denote alterations in the photoreceptor layers[74].

Another important indicator of neovascular lesion is the presence of pigment epithelial detachment (PED). In different patient's management studies, PEDs was present in almost 80% of patients at the time of enrolement[78,82]. Underlying growth of the PED in treated patient allows the prevention of early, "silent" recurrence of choroidal neovascularization. Therefore, this should indicate retreatment even if the patient is not symptomatic and even the absence of other markers like IRC (**Figure 14**). During follow-up, a precise monitoring of PEDs including quantitative measures should be performed in order to adjust the treatment regimen and prevent long-term vision loss[83].

A third typical sign of nAMD in OCT data is the presence of small hyperreflective dots, named hyperreflective foci, into the neurosensory retina and more specifically located near neovascular lesions[84] (**Figure 14**). In early AMD, presence of hyperreflective foci is considered as a risk factor for progression to advanced form of the disease. Furthermore, poorer VA outcomes in the treatment of nAMD were associated with these particular structures and their resolution was associated with VA gain[85,86]. Therefore, data suggest the use of hyperreflective foci as potential markers for lesion activity and stage.

One particularly interesting imaging biomarker is subretinal fluid (SRF) as studies shown that patients exhibiting SRF derive larger VA gain from anti-VEGF therapy[65,87,88](**Figure 14**). Indeed, it is suggested that presence of SRF is an indicator of perfused neovascular network vessels and/or choriocapillaris layers in the foveal area providing RPE and photoreceptors survival. As SRF in eye of nAMD patients may indicate favorable outcomes in anti-angiogenic therapy, this parameter should be considered in the management of such patients. As a matter of fact, when comparing the outcomes of fixed frequent versus infrequent anti-VEGF injections, patients with SRF had favorable VA gain in both regimen in contrast to the other patients that showed poorer outcomes when treated less frequently than frequently[89]. These together made SRF an interesting marker for the follow-up of nAMD patients.

OCT based biomarker discussed in this chapter are used by clinicians to efficiently personalize prognosis and therapeutic management of nAMD patients. Nevertheless, several efforts must be made in the way to improve their availability and reproducibility in the context of personalized angiogenic therapy. Furthermore, making the link between systemic biomarker and robust imaging biomarker could give access to a more precise representation of the pathological process involved in neovascular AMD. This knowledge could pave the way to new therapeutical strategies and decision-making tools that would help clinicians to better manage nAMD patients.

1.2.3 Resistance to Anti-VEGF therapy.

Although anti-VEGF based therapy represents a cornerstone in neovascular AMD treatment, not all patients respond at the same level. Indeed, some of them have poor or nonresponse to anti-VEGF treatment or exhibit a slow loss of treatment's efficacy after repeated administration over the therapy[90]. Several experts consider the persistence of exudation after a period of 6 month of monthly anti-VEGF injection as an indicator of what is called "Refractory AMD"[91–93]. Refractory AMD is defined as a situation in which the patient doesn't respond to the anti-angiogenic treatment and constitute an important concept for clinical decision making and treatment switching option. In addition, another clinical concept named "Recurrent AMD" is reported and describes a condition in which the patient is suffering from the appearance of new SRF/IRC accumulation after initial resolution of exudates[94,95]. Nevertheless, both conditions result in a diminished therapeutic effect of anti-VEGF agent and/or therapy.

If treatment failure appears early in the therapy, several clinical factors such as genetic predisposition or misdiagnosis may be involved. On the other hand, resistance to anti-VEGF agents and activation of other pathogenic pathways result in a recurrence of exudative lesions after a successful period of treatment. **Figure 15** depicts the different causes of resistance and possible therapeutic approaches that could be applied to achieve good medication[90].

Anti-VEGF therapy is a long and difficult process in which nAMD patient are enrolled with the hope to recover and stabilize their vision. There is a need to investigate changes among patients that lead to anti-VEGF resistance. This would be helpful for clinicians to guide their decision regarding when to swift to other anti-VEGF agent or choose other therapeutic options.
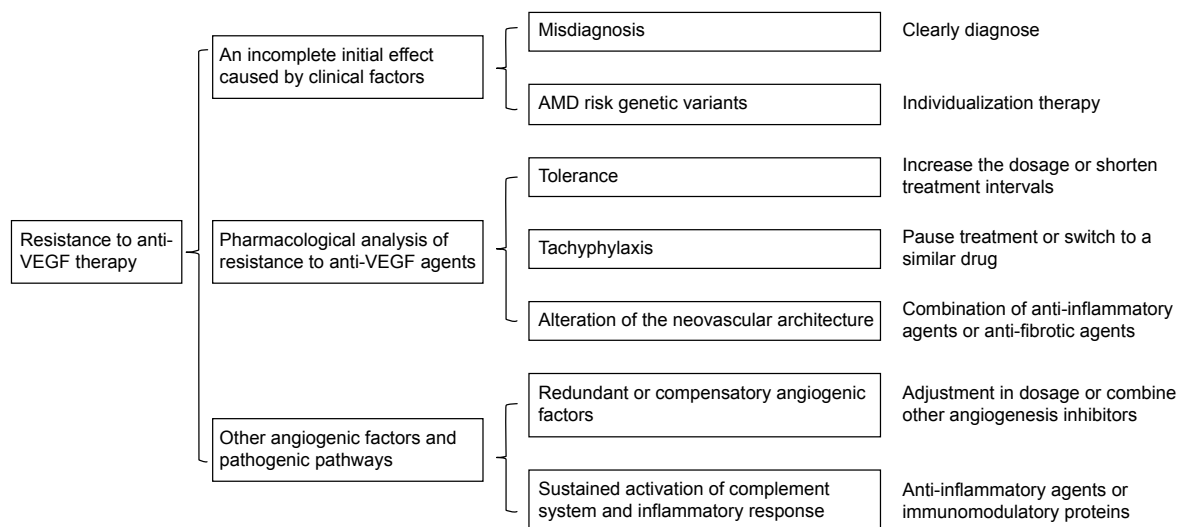
Figure 15. The different causes of anti-VEGF resistance, their causes and possible solution in order to prevent and/or improve patient's healthcare.

1.3 Metabolomics and NMR based Metabolomics

1.3.1 Omics Sciences

Omics sciences refer to specific fields that provide collective information describing a biological system. The aim of Omics is to offer a holistic view of what is made a cell, tissue, or organism and how it behaves in its environment. It studies a specific biological sample in a non-targeted and non-biased manner at a gene, nRNA, proteins and metabolites levels. The peculiar philosophy of these approaches is that a complex system is better understood if considered as a whole and interconnected one. While traditional studies are focused on a chosen set of data that should confirm a reasonable assumption, these hypothesis-generating methodologies acquire and analyze a huge amount of information to define a hypothesis that must be further tested. The exponential growth in technologies and informatics tools enable to generate and integrate large biological data sets pave the way to a paradigm shift in the way to approach scientific questions related to biological individuals. Therefore, these techniques have a broad range of application in numerous fields including toxicology, drug development, food and environmental sciences, personalized medicine, and biomarker discovery.

The main fields of the omics sciences are genomics, transcriptomics, proteomics, and metabolomics and each of them refers to a specific part of the studied system. Genomics is focused on the study of organisms' whole genomes and the development of recent sequencing technologies allows rapid and cost-effective elucidation of an entire genome and the study of all genes simultaneously. This field studies the heredity of an organism and how its characteristics, via DNA, are transmitted from one generation to the next one. Downstream, transcriptomics will study the total RNA content expressed by an organism after the transcription process of its genome. Together these sciences will provide an overview of the genes, expressed or transcribed, of a given organism at a precise time point. The proteome is defined as the snapshot of the whole proteins content of an organism. Thus, proteomics enables the study of the protein inventory of an organism and aimed to be helpful to better understand the metabolism dynamics of an organism. At the end of the omics cascade (**Figure 16**), the global approach that aims to measure all the metabolites in cells, tissues or biological fluids coming from an organism is called Metabolomics. The analysis of the metabolome gives access to unique information about an organism status in a given environment by filling the gap between "what's can happen" and "what's actually happening".
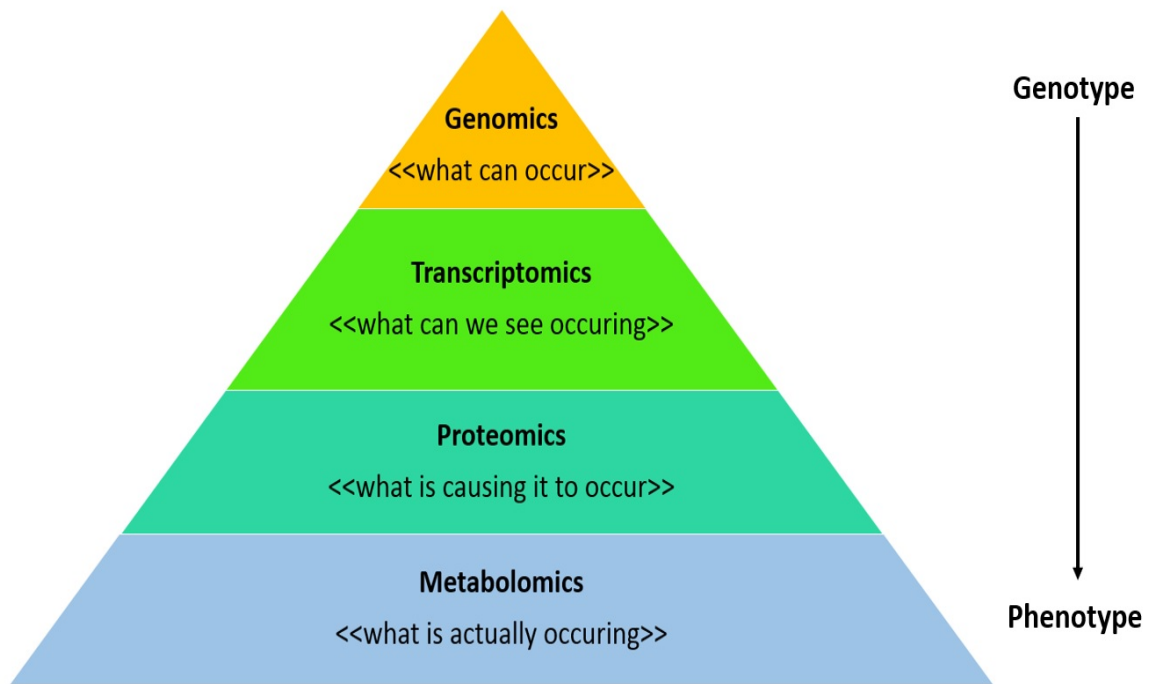
Figure 16. Omics cascade from the study of the genome to the analysis of the metabolome (adapted from[96]).

1.3.2 Metabolomics

1.3.2.1 Generalities

Metabolomics can be defined as the comprehensive study of all metabolites present in a cell, a tissue or a biofluid from an organism at a given time point. Life rises from chromosomal DNA that are transcribed to RNA which are containing the information to translate functional proteins. If studying these events will provide useful information about how differences happen between individuals, it must not be forgotten that what is happening is driven by small molecules metabolites (80-1500 Da). Indeed, these compounds carry out the main work of functioning cells and regulates the activity of the macromolecules in a complex interconnected scheme. This is the interaction of the small molecules with the macromolecules components of the organism that is responsible of how that organism behaves. The so-called metabolome is composed of organic or inorganic molecules, that are involved as reactant, intermediate or product of enzyme mediated biochemical reactions and are present at various levels of concentration. These molecules exhibit various physicochemical properties and biological functions and studying their changes among cell, tissues or biofluid give unique information about an organism status.

The basic idea of studying changes in tissues or biological samples to be able to diagnose disease is not new and goes back at least as far as ancient Greece. Indeed, diagnostic urine charts were used back in the middle age. Such cards were used to make correlation between pathological condition and the color, smells, and taste of urines. Such change in the urine of patients were, of course, caused by change in metabolic events.
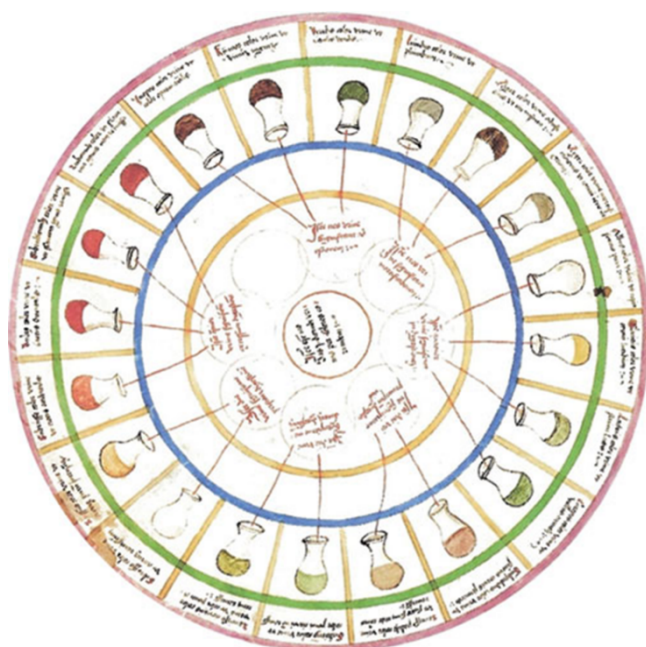
Figure 17. Urine wheel chart, published in 1506 by Ullrich Pinder, that describes the possible color, taste and smell of urine and how to use them to diagnose diseases.

1.3.2.2 Modern Metabolomics

Since its first publication in 1999[97], there is a growing interest for the field of metabolomics among the scientific communities over the past 20 years. Thanks to significant advancements in analytical platforms and bioinformatics tools, the number of publications and applications of metabolomics among a wide range of field increased considerably (**Figure 18**). Nowadays, metabolomics studies are conducted by using advanced analytical tools that provide a broad coverage of the metabolome such as: nuclear magnetic resonance (NMR) spectroscopy, liquid chromatography mass spectrometry (LC-MS) and gas chromatography MS (GC-MS). These techniques, and their associated methodology, provide the analysis of many classes of organic compound including organic acids, amino acids, lipids and others. Each technique has their own advantage and inconvenient and are better to be used to complement each other[98–100]. Nevertheless, metabolomics, conducted by a one or a combination of platforms, is a great tool that enable the study of the dynamics of the metabolome in various biological samples.
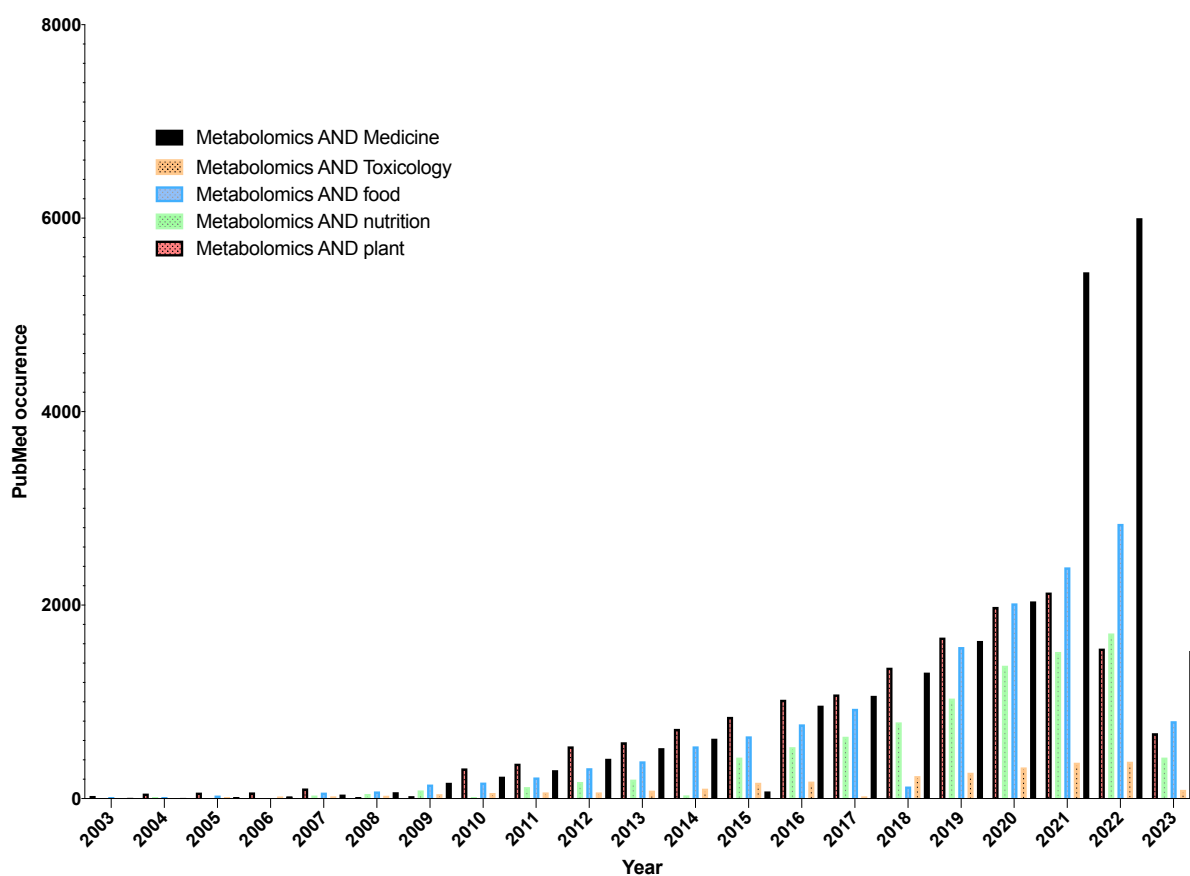
Figure 18. Pubmed occurrence search for "Metabolomics AND Medicine" (dark bars), "Metabolomics AND Toxicology (orange bars), "Metabolomics AND Food" (bluer bars), "Metabolomics AND Nutrition" (green bars) and "Metabolomics AND Plants" (red bars).

1.3.2.3 Metabolomics workflow

The metabolomics pipeline (**Figure 19**) is composed of different stages that reflect the multidisciplinary nature of the field. Indeed, from the setting of the experimental design to the analysis of samples and validation of results, metabolomics requires skills coming from the most advanced methodology in statistic, biology, and analytical chemistry sciences. Final results are driven by the experimental design and each step has to be carefully set in the way to assure reliability of the collected information.
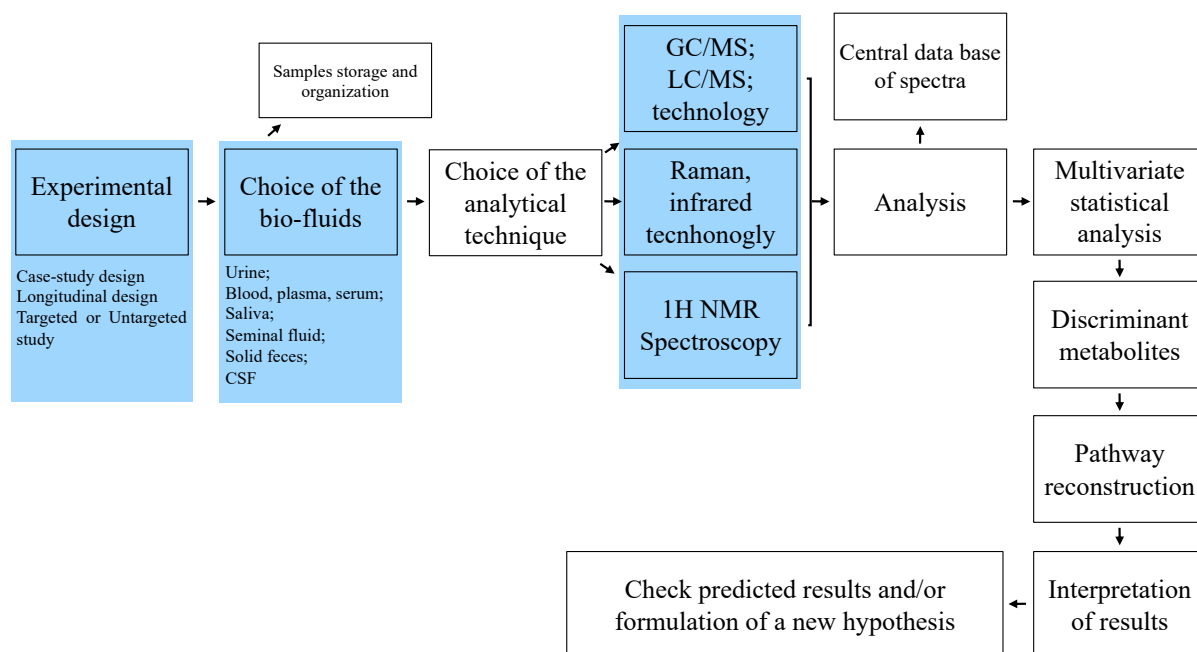
Figure 19. Metabolomics workflow, adapted from [101].

Each study starts with the setting of the experimental design and the definition of which samples will be collected in the way to answer to the biological question. Carefully assess this point will assure the validity and the usefulness of the results. Usually, metabolomics studies are conducted on biofluids such as plasma, serum or urine[102–104]. However, this methodology can be also applied on other fluids like saliva, cerebrospinal fluid and, by using dedicated analysis (mass spectrometry imaging, high resolution magic angle spinning and magnetic resonance imaging), the metabolic content of tissues, biopsies and cells can also be investigated. The use of extraction methods allows the analysis of such sample by using classical tools (ie MS and NMR).

Mass spectrometry and nuclear magnetic resonance spectroscopy are the two main analytical techniques used for metabolomics purpose. Rather than detailing the pro and cons of each analytical platform, that will be the main topic of the following chapter, we will here discuss about the different way to use them, and the specific outcomes provided by these ones. Indeed, if the choice of the samples to study is driven by the biological question, so do the instrumental platform and the way to use it. Herein, studying the metabolome can be achieved in a targeted or an untargeted way. Targeted metabolomics aims to analyze and quantify known metabolites proven to play a role in the biological process investigated. This approach is the one chosen for establishing the baseline metabolite levels of an organism or to define a cut-off value able to make the distinction between healthy vs "perturbed" status. In contrast, untargeted metabolomics is focused on the detection of as many features as possible in a single analysis. Rather than giving access to quantities, this profiling approach aims to discover novel biomarkers or metabolite profile that can be linked to the status of the studied individuals. If the first one has the inconvenient to have a limited coverage of the metabolome and thus enhances the risk of overlooking the metabolic response of interest, the absence of absolute quantification of the later limit the interlaboratory comparison and fail to assess the baseline metabolite response of an "healthy" individuals. The weakness encountered in one metabolomics philosophy is the respective strength of the other, therefore the use of one or another must fit with the goal of the study. In any case, the chosen analytical method has to be carefully optimized in the way to get rid of any unwanted instrumental drift or operator depending on variations.

All the other steps (statistical analysis, pathway reconstruction, validation of biomarker, …) of the workflow depicted in **Figure 19** will depend on the chosen approach. If all the above-mentioned step where carefully conducted, univariate, multivariate analysis will give precious information that will be further validated and integrated to respond to complex biological questions.

1.3.2.4   Analytical Platform for untargeted metabolomics.

The purpose of untargeted metabolomics is to measure and identify as many metabolites as possible in the way to give the most accurate overview of the metabolome at a given time point. From a chemico-physical point of view, the metabolome is a complex mixture of chemical entities exhibiting various properties in term of polarity, charge, pKa, solubility, volatility, reactivity, and stability. Moreover, these compounds are present at various concentrations among the different samples. As an example, the blood's glucose concentration for a healthy adult individual is in the range of mmol/L while for the same individual thyroxin is present in pmol/L of blood. The heterogeneity of chemical entities and the possible 10^6 fold difference in their concentrations make the detection of the whole metabolome a very challenging goal. Therefore, due to the complexity of these samples, no single technique is currently able to provide the coverage of the whole metabolome.

The main analytical tools used in metabolomics for untargeted approach are Mass Spectrometry (MS), usually coupled with a separative technique, Nuclear Magnetic Resonance spectroscopy (NMR), Infrared (IR) and Raman spectroscopy[103,105,106]. Despite this broad change of technique available, LC-MS and NMR are the most used platform for metabolomics purpose. Both techniques have their strength, and both have their weakness (see **table 1**). Over the past decades, advancement in the field of MS-based metabolomics makes this platform the main analytical technique for metabolomics studies and clearly overtakes NMR-based approach (**Figure 20**). However, NMR spectroscopy still offers many advantages over MS and these unique strength makes it part of the future of metabolomics[107–110].
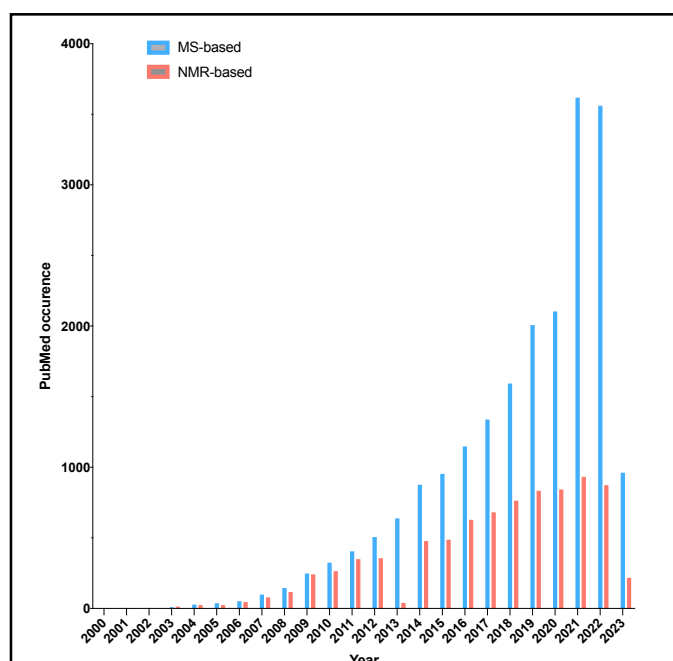
Figure 20. Pubmed occurrence search "Metabolomics AND mass spectrometry (blue bars),
"Metabolomics AND NMR" (red bars).

Mass spectrometry is a highly sensitive method able to detect feature at the nM range. When used with separative chromatographic technique such as liquid chromatography (LC-MS), gas chromatography (GC-MS), supercritical fluid chromatography (SFC-MS) and capillary electrophoresis (CE-MS), MS-based methods offer high resolution data of metabolite mass, fragments and molecular formulae[111]. The most advanced LC-MS platform enable the measurement of tens of thousands of features in one 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)[112]. If all the detected features cannot be directly related to metabolites, advance in metabolites' annotation allows to expand the current knowledge and "illuminate" what is called the "dark metabolome"[113,114]. This have made LC-MS the dominant analytical platform for performing metabolomics.

The major drawback of NMR is the limited sensitivity (sub mM) and even if effort were made to fill the gap with MS, this technique is still a few orders of magnitude less sensitive[115]. Nevertheless, NMR spectroscopy have numbers of characteristic that made it attractive for metabolomics (see **table 1**). First, NMR is highly reproducible, and absolute quantification can be achieved easily using a single internal reference. Secondly, this technique leaves the sample intact and allow multiple analysis to be performed enabling unambiguous identification of unknown metabolites. Furthermore, NMR analysis is independent of the operator or system used and as no separative techniques are used, analytical robustness and therefore inter laboratory reproducibility are enhanced compared to MS.

This thesis is focused on NMR based metabolomics, therefore analytical and technological aspects will be discussed in detail in the next chapter with special attention on the analysis of blood derived biofluids. The choice of NMR as analytical technique for this clinical metabolomics study is driven by its intrinsic quantitative property, its robustness and reproducibility, and its ability to detect a broad range of metabolites of different physico-chemical properties. Indeed, this approach will allow us to analyze an elevated number of samples collected and analyzed over the time and will provide information about lipoprotein macromolecules and small molecule metabolites in the same experiment without any specific sample preparation. Furthermore, if needed, targeted metabolomics through NMR quantification is accessible with little sample preparation. This methodology will allow us to profile the metabolome of patient's as well as study the variations of specific compounds during their follow-up.

NMR	MS
Non-destructive for sample	*Destructive for sample*
Robust intrumentation	*Relatively unstable*
High reproducibility	*Moderate reproducibility*
Small sample preparation	*More complex sample preparation*
No need of chromatographic separation	*Requirement for chromatography*
No need of chemical derivatization	*GC need chemical derivatization*
Predictable spectra	*Spectra difficult to predict*
Allows structure determination	*Allows partial structure determination*
Inherently quantitative	*Not inherently quantitative (need for reference standards)*
Easily automated workflow	*Difficult to automate workflow*
Low cost per sample	*Isotopically labeled reference standards for quantification can be expensive*
Poor to moderate sensitivity (μM)	**Excellent sensitivity (nM)**
Modest metabolite coverage	**Extensive metabolite coverage**
Expensive intrumentation	**Modestly expensive instrumentation**
Few sofware resources	**Many software resources**
Expensive to maintain	**Moderately expensive to maintain**

Table 1. Comparison of analytical pro and cons of NMR and MS-based metabolomics.

1.3.3 NMR-based Metabolomics

As IR or UV spectroscopy, NMR is an absorption spectrometry technique that aims to measure magnetic properties of a given nucleus when placed on an external magnetic field (B). Basically, under appropriate conditions, this nucleus will absorb electromagnetic radiation in the radio frequency (*rf*) domain at a frequency that depends on the intrinsic properties of the studied nucleus. In this chapter we will first describe the basic principle of NMR spectroscopy and details the instrumental platform that allows to record NMR spectra of biological samples. Then we will discuss some particularities of biofluid analysis for metabolomics purpose and focus on NMR data pre-processing. Finally, the part named "data analysis" will explain how key information can be obtained from multi- and uni-variate analysis of NMR pre-processed data. Identification of biomarker and model validation will close this chapter as this step generally end the NMR-based metabolomics workflow.

1.3.4 Principle of NMR spectroscopy

NMR spectroscopy takes advantage of an intrinsic property of nuclei, the so called "spin". All nuclei carry a charge, and for some of them, this charge rotate on the nuclear axis. This circulating charge of the nucleus generate a magnetic dipole along the nuclear axis and the magnitude of this dipole is expressed in term of the nuclear magnetic moment μ. The angular momentum of this moving charge is described in terms of its quantum spin number I that can take values of 0, ½, 1, 3/2, and so on.
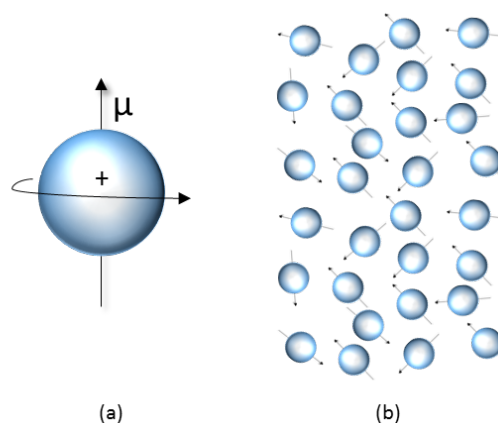
Figure 21. (a) representation of the magnetic dipole generated by the rotating charge of the nucleus and (b) its random orientation in absence of any external magnetic field.

When considering NMR, the spin number is an important feature since this quantum number determine the number of orientation ($2I + 1$) that the nucleus can take when placed in an external

and uniform magnetic field (B_0). For nuclei that have a spin value of 1/2, such as ^1H, when placed in an external magnetic field, there are two energy levels named α and β depending on their orientation regarding B_0. The difference of energy (ΔE) is given by the *equation 1* and according to the Boltzmann distribution, a slight excess of population is observable in the state of lower energy ($N_\alpha > N_\beta$).
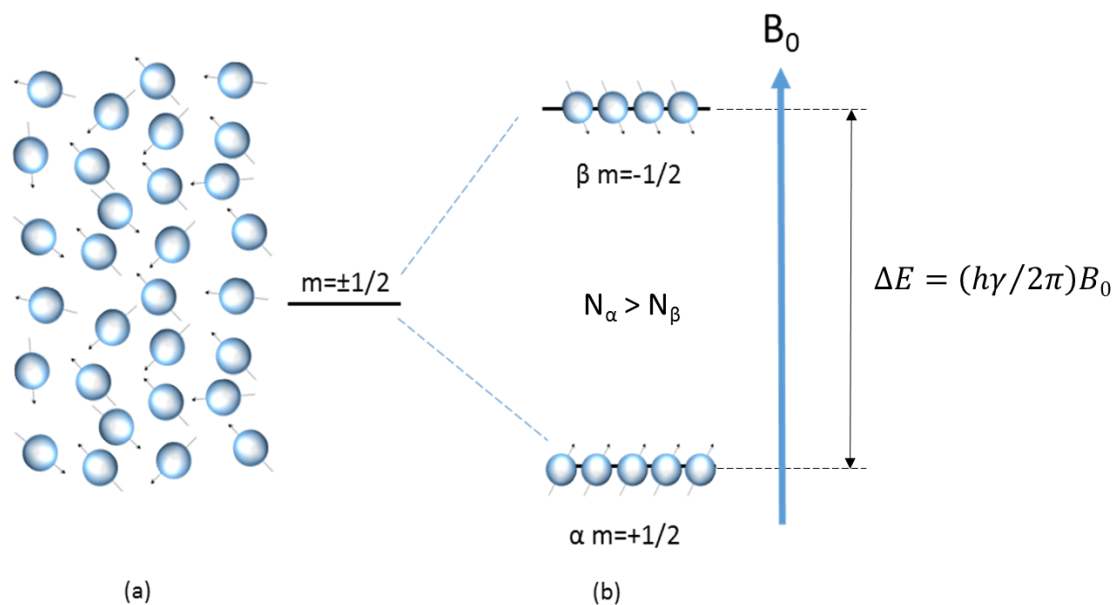
Figure 22. Representation of nuclei in absence (a) and in presence (b) of an external magnetic field B_0. The difference in energy between the two spin states is expressed by the equation on the right (equation 1).

If ΔE is directly proportional to B_0, it's also depending on intrinsic magnetic properties of the studied nucleus as γ represent the magnetogyric ratio. This parameter is proper to each nucleus and is defined as the proportionality constant between the magnetic moment μ and the spin number I (*equation 2*: $\gamma = 2\pi\mu/hI$). Thus, if a well calibrated radiofrequency radiation (ν_I) is used, it is possible to induce transition between the two energy levels in the presence of B_0. This is the basis of NMR, once ν_I is applied, the energy is absorbed by the proton that will raise higher energy level, the system is in resonance and by analyzing the return to the fundamental state, an NMR spectrum can be recorded. The equation 3 is the basic relation that correlate the applied radiofrequency ν_I with the intrinsic magnetic properties of the nucleus, the gyromagnetic ratio γ, and the magnetic field strength B_0.

But how to induce and to record these energetic transitions of protons that are aligned in a magnetic field? By applying a perpendicular radiofrequency (rf) to the static magnetic field and recording the change of magnetization in the same plan. For a better understanding, lets represent the proton in the magnetic field: the spinning proton behaves like a magnet and therefore have a magnetic dipole, in presence of an external static magnetic field (B_0), this magnetic axis will be oriented along B_0 in the axis named z. When considering a bunch of equivalent protons in B_0, their magnetization will be randomly phased along the z axis and the resulting macroscopic magnetization M_0 can be represented as a vector following the z axis. By using a rf oscillator placed in perpendicular to z (in the x plan), the magnetization M_0 can be tilted in the xy plane, and the magnetic component can be recorded by a probe positioned along the y axis. Indeed, when the magnetic field (B_1) generated by the rf oscillator will have the adequate frequency ν_1 that is equal to the processional frequency of the protons (Larmor frequency ν_L), the system will reach the state of nuclear magnetic resonance, the tilt of M_0 is achieved and the resulting magnetization is named M.
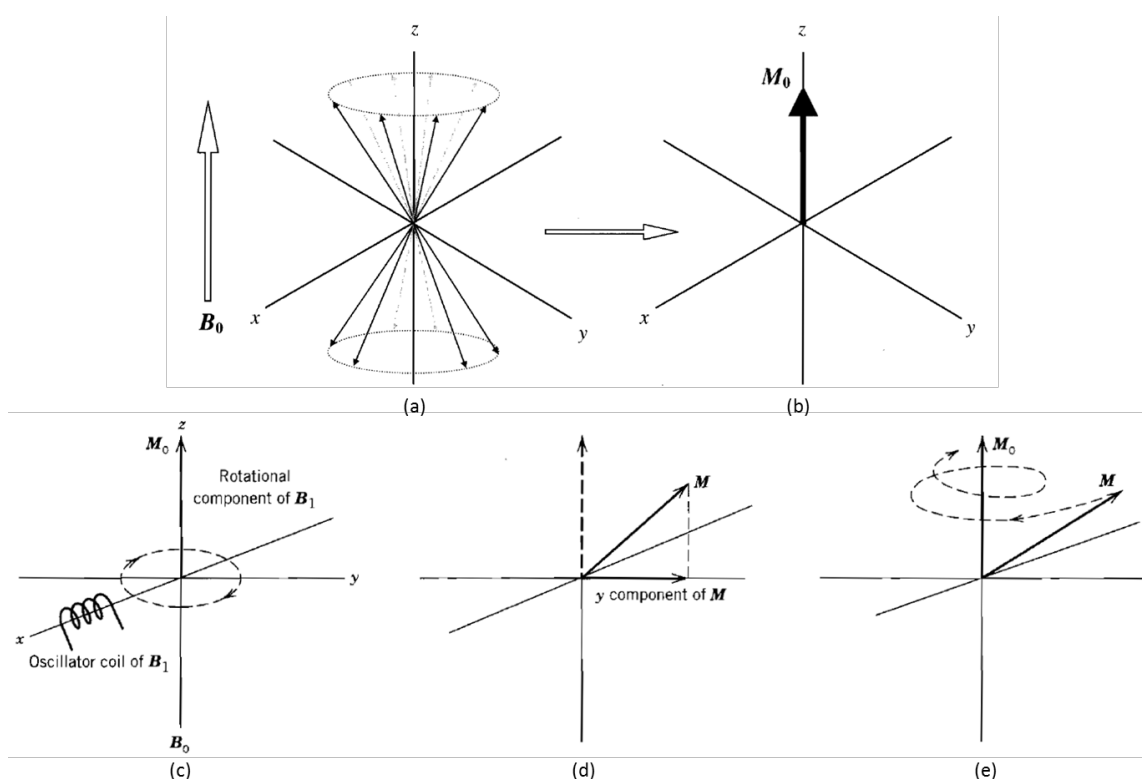
Figure 23. A bunch of equivalent protons in a static magnetic field B_0 (a) and the resulting magnetization M_0 aligned along the z axis (b). The oscillator coil generates the magnetic field B1 on the x plan which results in a tilt of M_0 in this horizontal plane to generate M (c) and (d). The longitudinal relaxation (T_1) is responsible of the relaxation of the M magnetization to M_0 (illustration from Spectrometric Identification of Organic Compounds, 8th Edition).

In practices, a very short and powerful rf pulse (µs) centered on the frequency ν_1 is applied along the x axis and generates the complete frequency range allowing all protons to reach the resonance state simultaneously. The magnetization M_0 is tilted in the xy plane, and a signal can be recorded over a given period (t_2). During this period, the magnetization M will process around the z axis and detected on the y axis. Therefore, right after the pulse, M is on the y axis and the signal is at maximum; it'll evolve around the z axis and reach the x axis where the magnetization along y is null. Then the signal continues its procession movement and goes to the $-y$ axis reaching the minimum and so on until the magnetization relaxes back to the equilibrium. The resulting signal represents a decaying interferogram called FID (Free Induction Decay) that represents the difference between the applied ν_1 and the Larmor frequency of each proton (ν_L). Finally, the time-domain decaying signal is Fourier transformed into a frequency-domain NMR spectrum in which each peak will be present at a given chemical shift (δ) expressed in ppm respectively to the one of the chosen references ($\delta = 0$ ppm).
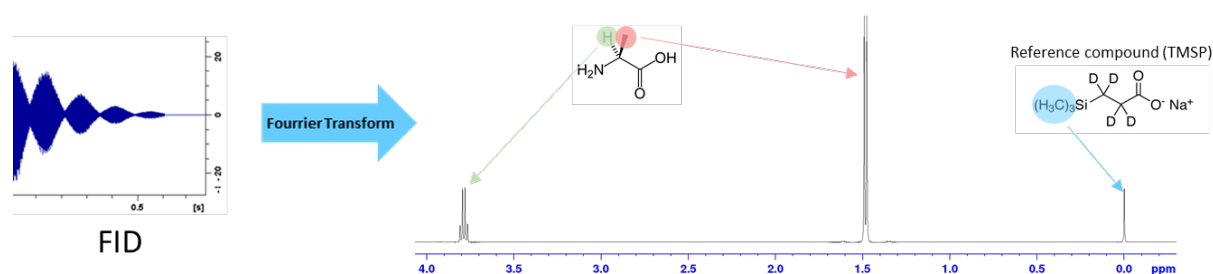
Figure 24. The detected NMR signal on the left and the processed spectra of the studied compound (L-alanine) with the peaks corresponding to the different protons at appearing at specific chemical shifts respecting to the peak of the reference compound (TMSP).

The way the magnetization decreases over the time is governed by two time-components: the longitudinal relaxation (T_1) and the transverse relaxation (T_2). The longitudinal relaxation, also called spin-lattice relaxation, is a monoexponentially process that corresponds to the transitions between quantic states of the spin system allowing the system to return to its thermal equilibrium state. The values of T_1 are characteristic for each proton and are the magnitude of the seconds in liquid. The second phenomenon involves interaction between two neighbors spin systems. The T_2 relaxation, called transversal or spin-spin, is explained by local fluctuation of the magnetic field due to interaction between neighbor spins. Indeed, as interactions between magnetic dipole are generating local magnetic fields (B_L), fluctuations of resonance frequency of individual spins can occur. Herein, after quantic state transition or by movement of spins

systems, these fluctuations of B_L will induce a dephasing and signal loss. Unlike T_1, the time of transversal relaxation is covering a large range of values. For example, for small molecules in liquid states, they can range seconds while for solids, or macromolecules, they can be shorter than a μs. This is due to the fact that in solids, the movement of neighbors' spins are limited and therefore their total component of induced B_L is not null while for small molecules in liquid the magnetization induced by movement is fluctuating fast due to Brownian movement. This resulting a mean effect of B_L that reduces its amplitude and the effect on the relaxation allowing the system to relax slower.

An important concept that allows the differentiation between protons of the sample is the chemical shift and the multiplicity. Till now, we spoke about a single proton without considering the effect of its own electron cloud, its neighbors, or its environment. All these parameters are influencing the way this proton will act in the magnetic field after the rf pulse. Thanks to these, each molecule will give a unique NMR pattern allowing their structural analysis and their discrimination in complex mixture (**Figure 25**).

Atom	δ (ppm)	Integr.	Mult.
4 C			
H	3.79	1	q
6 C			
H3	1.48	3	d
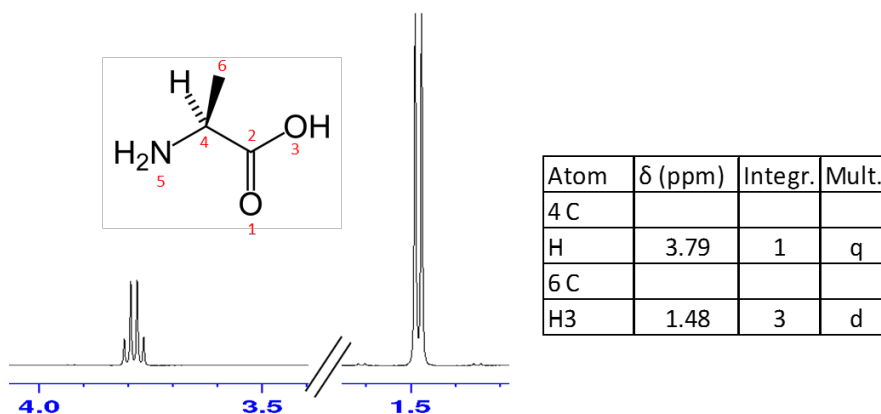
Figure 25. NMR spectra annotation of pure l-alanine compound in D_2O.

Under a magnetic field, circulating electrons generate their own magnetic field opposing the applied magnetic field. Thus, the proton is *shielded* by its electronic cloud which varies with its chemical environment. Therefore, the Larmor frequency can be re-written for all equivalent protons including the constant shield σ that is characteristic for each proton in a given molecule. For organic molecules, the degree of shielding of a proton attached to a carbon atom will depend on the inductive effect of the other groups attached to the carbon atom. Hence, protons that have different chemical environments have different chemical shifts and equivalent protons that have the same environment have the same chemical shift.

This difference of chemical shift (δ in ppm) can be calculated by using an external reference and is expressed as:

$$\nu = \nu 0.(1 - \sigma)$$

Equation 3. Effect of the constant shield on the resonance frequency of a processing proton in a magnetic field.

Therefore, the chemical shift expresses the difference in the absorption position of a given proton in respect to the absorption position of a reference. By convention the peak of the external reference is located on the right of the NMR spectrum axis at 0 ppm. A common reference used in NMR is tetramethylsilane (TMS). This reference is chemically inert and soluble in most organic solvents. The particularity of this compound is that all its protons are equivalent and, due to its silane group, are more shielded than almost all organic protons. Thus, the more the proton is shielded, the effect of the magnetic field on the proton decreases and its signal will be closer to the one of the reference. The more the proton is deshielded, for example when an electronegative group is decreasing the electronic density, the more the proton is exposed and will be affected by the magnetic field; hence their resonance occurs at a higher frequency than the reference and the peak is located on the left of the spectrum.

The multiplicity is describing the phenomenon responsible for the splitting of peaks due to the interaction between spin systems called spin-spin coupling. The spin-spin coupling is a magnetic interaction between two non-equivalent protons that is transmitted through their bonding electrons. Let's consider two vicinal protons (separated by three bonds) in a very different chemical environment, such as HA_2CCX_2H. Each proton is in different chemical environment thus their resonance frequency will differ, and the two-absorption peak will be well separated. If the spin of each proton is slightly influenced by the two different orientations of the spin of the other proton, each will exhibit two absorption peak and the single signal (called a singlet) will be splitted in what is called a doublet. The difference of frequencies within the doublet is independent of the magnetic field but is characteristic of the coupling between the protons and thus is denoted by the coupling constant (J). If the coupling constant gives information about the bonding, the multiplicities give information about the number of neighboring protons on the adjacent carbon. Indeed, as we saw, if the adjacent carbon contains one proton, the signal will be a doublet; a triplet will appear if this carbon contains two protons and a quadruplet if three. The Pascal triangle (**Figure 26**) is representing the *n+1* rule of peak

splitting for *n* equivalent coupling nuclei and their relative intensity. These properties are the basis of NMR identification and structure determination of organic molecules.
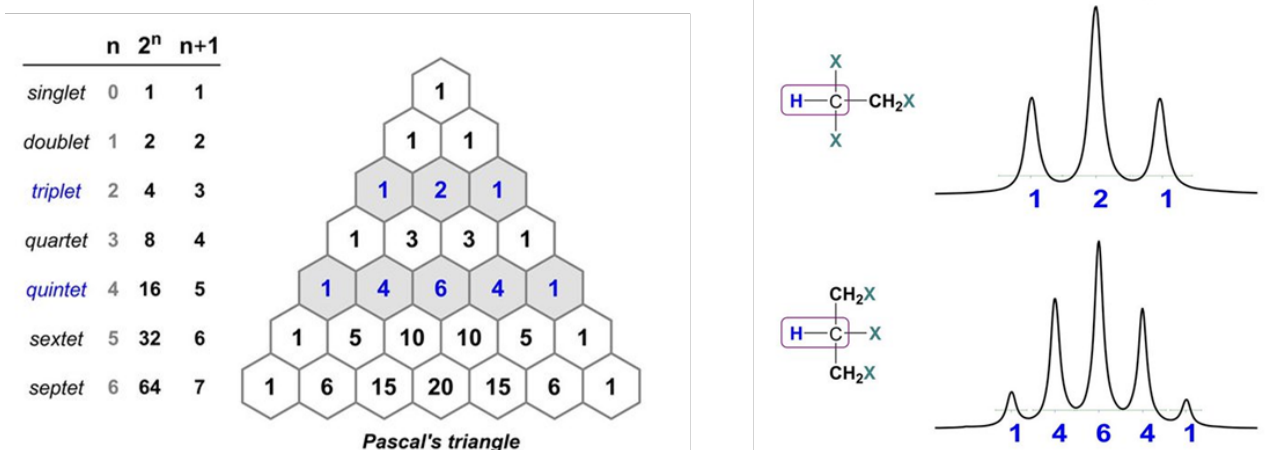
Figure 26. Pascal triangle representing the n+1 rule for neighboring protons and the relative intensities of the resulting peaks.

The last, but not the least, obvious information among the classic NMR spectra is the peak intensity and more precisely its area. Among the NMR spectrum, the peak area for each signal, even if it's a multiplet, will be proportional to the number of protons that are in resonance at a given frequency. Therefore, this information will not only give access to the number of protons present on the studied molecule, but it will also allow the calculation of the concentration on this molecule if an internal standard is used, and the NMR experiment is properly set. Indeed, this property allows the absolute quantification of a molecule among complex mixtures using NMR and represents a major advantage over other technologies. By knowing the relaxation times of both internal reference and analyte, the NMR experiment can be optimized to provide absolute quantification (relaxation delay = $5*T_1$)

1.3.4.1 NMR instrumentation

The instrumental platform required for NMR spectroscopy is composed of three major components: The superconducting magnet, kept under 9.5 K with liquid helium, generating the static magnetic field; a shim coils to assure the homogeneity of B_0; a probe to record the generated NMR signal; and the necessary electronics to transform the measured signal into interpretable NMR spectra.
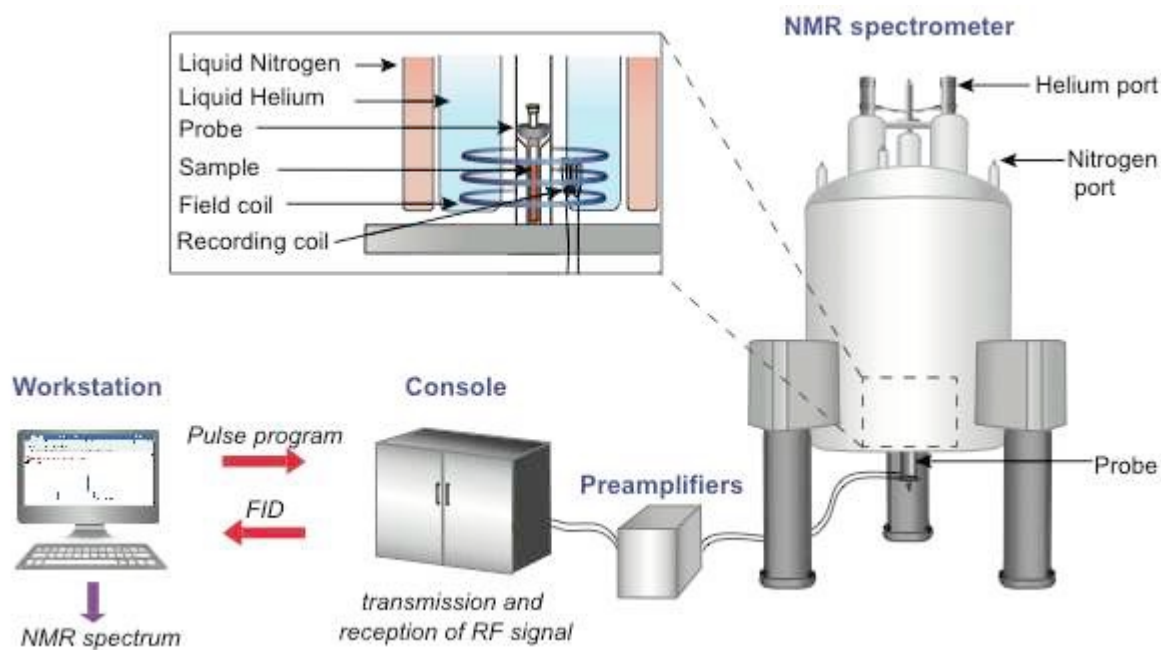
Figure 27. Schematic representation of the instrumental platform for NMR spectroscopy[116].

Magnet

The NMR magnet is the most important part of the NMR spectrometer as the magnitude of the field generated is directly related to the spectral resolution but also to its size and price[115]. Herein, these magnets are usually made of supraconducting wire which are wrapped into a 100 km long wire coil. To ensure the supraconducting state, the coil must be kept under ~10K and thus is placed inside a large Dewar containing liquid He surrounded itself by liquid N_2. Most of metabolomics studies are conducted on magnet operating between 400 and 600 MHz but it's always a compromise between funding and space. Indeed, even if it's clear that the resolution enhance with the strength of the magnetic field[107,117], magnet of 600 MHz are produced at larger scale and thus are cheaper. Therefore, this making them the configuration of choice for routine metabolomics investigations as it's a good compromise between good resolution and cost[117].

The shim coils

The shim coils are a bunch of conducting coil, placed along all three spatial axes, used to adjust the homogeneity of B_0. These are used to correct all local magnetic distortion of the static magnetic field. Indeed, these perturbations can arise as well as from the environment in which the spectrometer is installed or resulting from variation in the sample. Thus, these changes are adjusted by changing the current in one or more shim coils to ensure the stability of the static magnetic field. This process is called "shimming".

The sample probes

The sample probe is the sensor located in the center of the magnet and used both to send rf pulses and to detect the returning signal. Therefore, the choice and the quality of the probe have a direct impact on the spectral resolution of the NMR experiment. As an example, the size of the bore is an important aspect as small volume probes (3 to 1,7mm allowing experiment on the µL scale) give access to better sensitivity but larger volume bore is recommended to analyze higher viscosity samples. Cryogenically cooled probe, commonly named cryoprobe, allows an enhancement of signal to noise ratio by 3-5 time compared to conventional probe. This is therefore an interesting option to gain sensitivity for metabolite profiling and biomarker identification specially to compensate the lack of sensitivity due to the strength of the magnet[118].

Digital filter

Digital filtering applied to NMR data aims to avoid baseline distortions, improves the sensitivity and provides a better dynamic range [119]. Prior sampling, the unfiltered analog data passes through an analog antialiasing filter in order to suppress frequencies higher than the Nyquist frequency. Then the signal is sampled and digitized using an analog to digital converter (ADC) before being filtered and reduced in real time by the digital filter. Finally, the filtered FID is Fourier transformed and provide the spectral data for analysis.

Automation for NMR metabolomics

The metabolomics approach usually involves large cohort studies with several hundreds of samples. The high throughput context of this approach therefore requires simple and robust NMR experiment and a high level of automation. From sample preparation, followed by automatic handling of sample for NMR acquisition and data-processing, all steps of the metabolomic-based NMR experiment have to be carefully set to avoid analytical variation. The analytical platform used for this work is therefore composed of a 500MHz operating NMR equipped with a TCI 5mm liquid nitrogen-cooled cryoprobe and used with a SampleJet sample charger allowing automatization of sample handling. All sample were analyzed using 1D NMR sequences dedicated to the biofluid investigated using IconNMR® tool allowing automatic and reproducible NMR experiments. Recorded data were automatically pre-processed by an automated R package enabling Fourier transform, zero phase order correction, baseline correction and spectral alignment to ensure the consistency of NMR data across the study. All these aspects will be discussed in the following chapters.

1.3.4.2 NMR-based Metabolomics on blood sample

1.3.5 Composition and metabolome coverage

Blood is composed of two different parts, the first one is a cellular component which consists of blood cells (red and white) and platelets; the second one is the plasma, an aqueous solution in which the blood cells are suspended. Plasma, which accounts for 50-55% of blood volume, can be obtained from blood sample, by adding an anti-coagulant to the collection tube and a simple centrifugation step that removes the cellular content. If no anti-coagulant is added to the collection tube, the blood will clot, and the supernatant fluid is called serum. Plasma and serum are very similar and have almost identical endogenous metabolites composition; indeed, the main differences involve compounds linked to the clotting process as fibrinogens and prothrombin, a decreasing viscosity, and a slightly higher overall concentration for serum samples. Both plasma and serum are complex matrixes composed of 95% of water and a great diversity of substances including proteins and peptides (albumins, enzymes, hormones, lipoproteins), metabolites (carbohydrates, amino and organic acids, lipids), electrolytes and other small molecules that are suspended or dissolved in the media. As blood plasma and serum are the primary carrier of small molecules on the body and bathes every organ and tissue, most clinical tests are performed on these samples.

1.3.6 Analysis of intact blood derived samples

When speaking about NMR analysis of intact serum/plasma samples, the most obvious problem is the presence of water in the sample which the resonance will dominate the NMR signal. Therefore, water presaturation sequences must be used to recover the peak of molecule of interest without impacting the intensity or the integrity of the signals that do not resonate close to water frequency. The easiest way to achieve proper water suppression is to use a weak radio-frequency pulse centered on the Larmor frequency of water signal over a period corresponding to the water relaxation time *T1* to selectively saturate its resonance and do not record the signal during the acquisition time (**Figure 28**). This presaturation method is the most convenient choice for metabolomics NMR experiments that involve other complex pulse sequence components. Indeed, this scheme allows good water signal elimination during specific analysis such as *T2* discrimination methods or 2D experiment performed on intact serum/plasma samples. However, for more diluted samples or when a better water suppression is needed to improve baseline or signal/noise ratio, other more advanced pulse sequences are used. We will discuss them later.

The second aspect of intact blood sample to take in consideration is the presence of proteins and lipoproteins. Indeed, without the use of proper methods, the NMR spectra will be dominated by broad resonance peaks coming from these entities and the sharper peaks associated with small molecules metabolites will be poorly resolved. For these reasons spectral filtering techniques such as *T2*-editing cpmg sequence are used to make the small molecules signal more visible by suppressing the signal of high-molecular weight component. The Carr-Purcell-Meiboom-Gill (cpmg) experiment takes advantage of the fact that large molecules such as lipoprotein, protein, and lipids, usually have faster relaxation time (**T2**) than small molecules metabolites. By attenuation of the signal of fast relaxing species, this filtering technique will allow important improvement in the observed metabolite detection and give access to robust spectral data (flat baseline, properly phased spectra, high reproducibility). The pulse sequence consists of a presaturation pulse followed by a 90° excitation pulse and 180° pulse train (**Figure 28**). For each 180°, a spin echo* is generated after a time of the magnitude of 2τ and the amplitude of these echoes will decay exponentially with a time constant that correspond to the transverse relaxation (**T2**) of the spin system. As large molecules exhibit smaller **T2** than small molecules metabolites, their signal will decay, and the acquired signal will be dominated by the slow-relaxing compound. The spectrum resulting from this experience are close to classic 1D spectra since the attenuation of the signal is similar for all the studied metabolites as their NMR signature shows quite narrow linewidth.
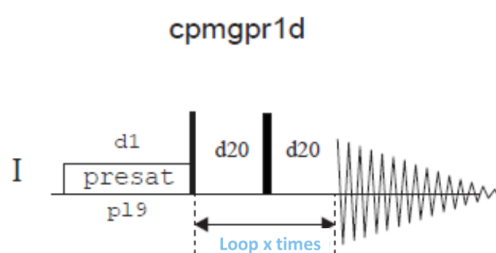
Figure 28. 1D Carr-Purcell-Meiboom-Gill (cpmg) Bruker® pulse sequence with water presaturation. A pulse sequence is a visual representation of the pulses and delays used in each NMR experiment. This sequence starts with a long pulse (white box) during all the relaxation delay d1 to saturate the water signal. After this period, a short 90° pulse (sharp black box) is applied, followed by a d20 delay, a 180° pulse (broader black box) and another d20 delay. This part of the pulse sequence is repeated x times to produce the desired T2-filtering effect.

Spin echo are generated to correct dephasing spins system that have different magnetic properties during the relaxation time. Indeed, by using a 180° pulse after a well-defined time, all the magnetizations will be reversed symmetrically from their original precession following the initial 90° pulse. Hence, the ones that were relaxing faster will be behind the other ones that were relaxing slower and therefore, at the end will have the same phase. This phasing will increase the signal by creating an "echo".

^{1}H-NMR cpmg spectra of intact serum/plasma samples as shown on **Figure 29**, are quite easy to interpret as sharp peaks from metabolites are well resolved and remaining broader peaks of lipids and lipoproteins moieties are simple to recognize. Lipoproteins are amphiphilic macromolecular complexes composed of various lipids and proteins that able lipids to circulate through the human body. Lipids within the lipoprotein structure are responsible of different resonance peaks from glycerol backbone, fatty acyl chains and the polar moiety, choline for example. But if the resonances peaks of these are poorly informative, the methyl and methylene group of various lipid components presents at low frequencies have been found to exhibit some interesting characteristics that will be discussed later in this work.

Aside from these information, sharp peaks coming from different small molecules metabolites are easily identifiable. The most dominant species in the spectra of serum/plasma is glucose, which is present in the blood of healthy individual at a concentration $\approx$ 5mM. This molecule, that is in equilibrium between its two anomeric forms α and β, have intense spectral contribution between 3.2 and 5.25 ppm and can be routinely quantified by ^{1}H-NMR. Next to these peaks, smaller peaks are present and associated to various amino acids that can also be routinely identified and quantified such as glutamine, alanine, tyrosine, or proline. At lower frequencies, other small molecules are present in the spectra of intact serum/plasma. Hence can be found organic acids from TCA cycle such as lactate, formate or citrate, several ketone bodies included 3-hydroxybutyrate, acetone and acetoacetate. Finally, close to broad peaks of lipoproteins, branched chain amino acids are found and can be quantified using proton magnetic resonance spectroscopy.

Figure 29. Cmpg-1D spectra from intact serum/plasma sample and the various molecules identifiable and quantifiable in routine analysis. Overlapped methyl and methylene group of lipoprotein lipids components (a) and (b) signal from various lipids. (adapted from [120]).
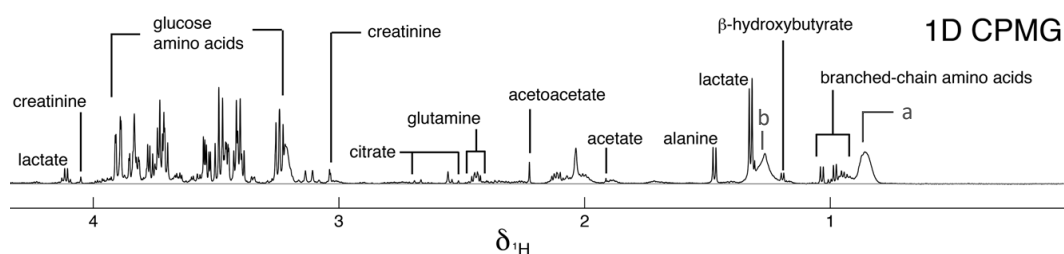
The number of feature and the detection limit is highly dependent of the magnet and the probe used to record spectra, but also of the sample preparation. Indeed, here we discussed the analysis of intact blood samples, but with proper sample preparation, several examples in the literature report the ability of NMR spectroscopy to routinely detect and quantify up to 50-68 metabolites[102,107,121] at a minimal concentration above 2 to 10 µM depending on the multiplicity and the number of protons contributing to the peak resonance[122]. To achieve this, several methods are available to eliminate or separate the protein and lipid content from polar metabolites. From protein precipitation to ultrafiltration and solvent extraction methods, these methodologies use different steps that aim to improve spectral quality, eliminate overlapping broad signals and thus enhance metabolome coverage. For the minimal processing, and so the reproducibility, most of labs still use intact serum/plasma samples, but these strategies are becoming more and more popular. Pro and cons of these methods are summarized on **Table 2** and the choice of the technique is mostly driven by the subject of the study and the quantity of sample available as these are giving access to different information about the sample.

The analysis of intact serum/plasma sample only requires a dilution step with phosphate deuterated buffer, for frequency locking and pH stabilization, and the addition of standard for calibration and quantification (if needed) as TMSP and Maleic acid or formate respectively. If this protocol has the obvious advantage to be fast, simple, and reproducible, it's also the only way to obtain information about the lipoproteins content of the blood sample. However, when the aim is to have the best metabolome coverage or to quantify well identified metabolites, separating polar fraction or getting rid of large molecular weight compound by precipitation or ultrafiltration is the best way to recover greater information. Among the literature, three distinct methodologies exist to prepare serum/plasma samples for optimized ^{1}H-NMR analysis: (i) ultrafiltration using molecular-weight filter in order to retain macromolecules and separate low molecular weight metabolites; (ii) protein precipitation technique by the addition of solvent or acid reagent followed by centrifugation and analysis of the supernatant; and (iii) extraction methods using polar and non-polar solvents in order to recover both fractions and analyze them separately.

Removing high-molecular-weight component from serum/plasma sample by centrifugation of the sample trough commercial 3-10 KDa filters is simple, inexpensive, and well adapted to high throughput metabolomics studies. Indeed, after centrifugation and addition of a solution of reference compound diluted in a deuterated solvent to calibrate the spectra and allow proper lock of the NMR instrumentation, small molecules are well resolved from the rest of the sample

and give access to robust spectral data. This process will provide a better peak resolution and flatter baseline that are essential for quantification through line fitting approach. In these conditions, by using proper NMR analysis, quantification up to 44 identified metabolites coming from various biochemical pathways is achievable[102]. The major drawback of the method is the need to remove glycerol that is present in high quantity in the commercially available filters using extensive washing process (up to 10 washing steps to completely removing the residual signal of glycerol from NMR analysis) and the possible variability in the amount of filtrate generated that could lead to express the quantified feature in relative concentration units. Finally, some metabolites could be retained on the filter as they can bound proteins and therefore not be present in the filtrate or in lower concentration.

Deproteinization of serum/plasma samples is up to date the best methods reported in the literature in order to recover the highest metabolite content from blood samples with high reproducibility. Indeed, even if other several protocols exist[123–125], adding methanol to serum/plasma sample in 2:1 ratio allows proteins to precipitate and lead to the assignation of 67 distinct metabolites in NMR spectra[121]. When comparing the two methods in their paper, Gwonda and Raftery showed that protein precipitation methods exhibited higher concentration in half of the quantified metabolites with a 3 to 4-fold change for some metabolites such as tryptophan, benzoate and 2-oxoisocaproate[126] comparing to the ultracentrifugation method. On the opposite way, the solvent evaporation step during protein precipitation protocol led to a loss or a reduction of some volatile metabolites such as dimethylamine, methanol, acetone, or ethanol compared to ultrafiltrated samples. Furthermore, this additional step decreases the ability of developing high-throughput workflow to analyze big sample cohorts.

	Intact serum/plasma analysis	Solvent extraction/ protein precipitation	Ultracentrifugation
Pros	Quick and easy, minimal processing, widely used, Lipoprotein information volatiles retained	High metabolome coverage, best for aqueous soluble metabolites, Lipids from apolar fraction	No solvant or drying step, volatile retained, good coverage
Cons	Low metabolome coverage, quantifitative accuracy limited	Volatile loss or less reproducible, time consuming	Protein-bound metabolite may be not recovered or less reproducible, cost
Metabolome coverage	Up to 20-30 metabolites, lipoprotein information	Up to 68 metabolites	Up to 67 metabolites

Table 2. Summary of main blood sample preparation protocol for NMR analysis (adapted from [120]).

Until now, if the presented methods have the advantage to be simple and robust or in few steps give access to greater metabolome coverage, none of them can invest changes among the lipids profile of blood samples. Change in lipids and lipids composition are observable in ^{1}H-NMR and even if very few examples are found in the literature, protocols able to separate polar and non-polar fractions are reported[127]. These protocols, based on solvent-solvent extraction involving chloroform and methanol in various ratio, extend the metabolome/lipidome coverage by the analysis of both fractions. Even if the reproducibility of dual-phase extraction remains an obvious challenge, interesting result can be obtained with these methodology as they have a largest metabolome/lipidome coverage than previously described methods. Indeed, spectra coming from polar fractions are comparable to one obtained via ultrafiltration or protein precipitation methods and give access to the quantification of polar metabolites, the NMR analysis of non-polar fractions make possible the identification and relative quantification of major lipid classes. The signal overlap between the different species of lipids is quite important in 1D-NMR spectra and, in order to improve the resolution, different solutions exist from 2D-NMR analysis to the use of deconvolution methods. The later will be discussed in the chapter about the treatment of NMR data. Finally, in a recent publication, Giraudeau *et al* presented a

robust NMR workflow for lipidomic analysis using both 1D- and 2D-NMR spectroscopy and demonstrates that robust analysis can be performed in high-throughput process[128].

Despite all these improvements and the different methods available, investigations are needed to assess if the advantage of alternative sample processing methods prior NMR analysis can make NMR platform more competitive to mass spectrometry. Indeed, improving the metabolome coverage of NMR-based metabolomics cannot be due at the expense of the robustness, the absolute quantitative properties and the non-destructive character of NMR whose are part of the strength of this technique[129].

1.3.7   Analysis of treated blood samples

When dealing with ultrafiltrated, deproteinized or polar fraction of blood samples, the use of NMR pulse sequence able to filtrate the signals is therefore needless. The presence of water in high concentration is the only problem the NMR spectroscopist will face and the best way to overcome this one is the use of 1D-NOESY sequence (also called noesypresat). This method provides NMR spectra with high reproducibility and great ability to reduce water signal without altering the remaining signals. To understand how noesypresat achieve better suppression than classic presaturation methods, we must take a closer look to the sequences. After the equilibrium recovery delay, the sequence begins with a long, lower power saturation period that allows for the selectively of the solvent resonance as its usually done in classic presaturation scheme. While using exclusively this method is not recommended as it led to the elimination of the resonance of the rapidly exchanging protons and can, via spin diffusion or direct NOE effect with water, attenuate the whole spectrum[130], using a limited presaturation period is proven to be useful to achieve optimum solvent suppression. Following this, an inversion of the equilibrium state is achieved by two consecutive 90° pulses followed by a mixing period, a final 90° excitation pulse and the acquisition (**figure 30**).

noesygppr1d
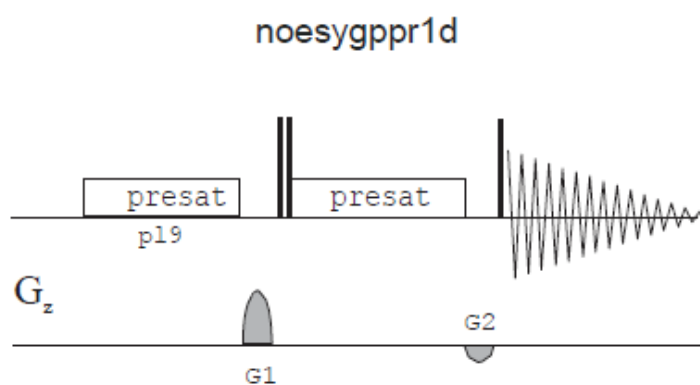
Figure 30. The "1D noesy presat" pulse sequence used in this work to analyze filtrated plasma/serum samples. Compared to the original one, this pulse sequence uses gradient to achieve better water presaturation. This sequence starts with a long pulse (white box) centered on the resonance frequency of water to saturate the water signal. Right before the excitation pulses, the two short 90° pulses (sharp black box), a gradient G1 is applied. Exitation pulses are followed by another water presaturation pulse and gradient G2 is used before a final 90° pulse and signal acquisition period.

The use of this sequence is largely spread among the different research groups as little optimization is needed to achieve high quality solvent suppression. The way the noesy presat suppress the signal of the solvent can be explained by three distinct aspects: the *T1* discrimination; the volume selection; and phase cycling[131]. These mechanisms aim to improve the suppression of the residual water signal that is due to field inhomogeneity in the region at the periphery of the tube and lead spectra with high quality solvent suppression with a good reproducibility. This sequence is one of the most used by metabolomics research groups dealing with urine or filtrated/deproteinized blood samples[102,103].

1.3.8 ¹H-NMR data pre-processing

In metabolomics studies, avoiding analytical and experimental variation that are unrelated to biological variation is of particular importance. Indeed, these variations can be originated from samples collection, storage, or preparation, as well as from the data acquisition and manipulation prior the statistical analysis. When the workflow of data treatment is not optimized, these variations can lead to uninterpretable results or to the discovery of false biomarkers. Therefore, after a well-planned study design and proper preparation and analysis of the samples, a good data pre-processing/pre-treatment is of particular importance as it will determine the interpretability of the acquired data. The objective of pre-processing steps is to minimize instrumental errors, enhance the signal to noise ratio (SNR) and transform the data in interpretable spectra that represents best the samples[132]. For ¹H-NMR data, different solutions exist that aim to give access to high quality spectra and to robust data. Commercial software such as Topspin© and MestReNova© are the most used tools for data pre-processing as they provide simple solution with friendly interfaces. If these tools are designed for the processing and the analysis of raw NMR data, they are rarely designed especially for metabolomics purpose. Thus, they lack advance editing steps and the need of manual adjustment led to operator dependent manipulation and therefore exhibit fewer robustness. Another way to process NMR data is the use of automated open-source software developed for metabolomics studies that provide exhaustive and flexible workflow to deal with raw ¹H-NMR data, starting from the FID to interpretable NMR results.

In this work PepsNMR, a R package developed by M. Martin and B. Govaerts from the UCLouvain[133] in collaboration with our group, was used to pre-process our untargeted ¹H-NMR metabolomics NMR data. Developed in collaboration with our lab, this software provides robustness and flexibility among the complete series of pre-processing functions that can be run independently or automatically. An overview of this workflow is depicted on the **Figure 31** and all the different steps will be briefly discussed in the following section of this chapter.
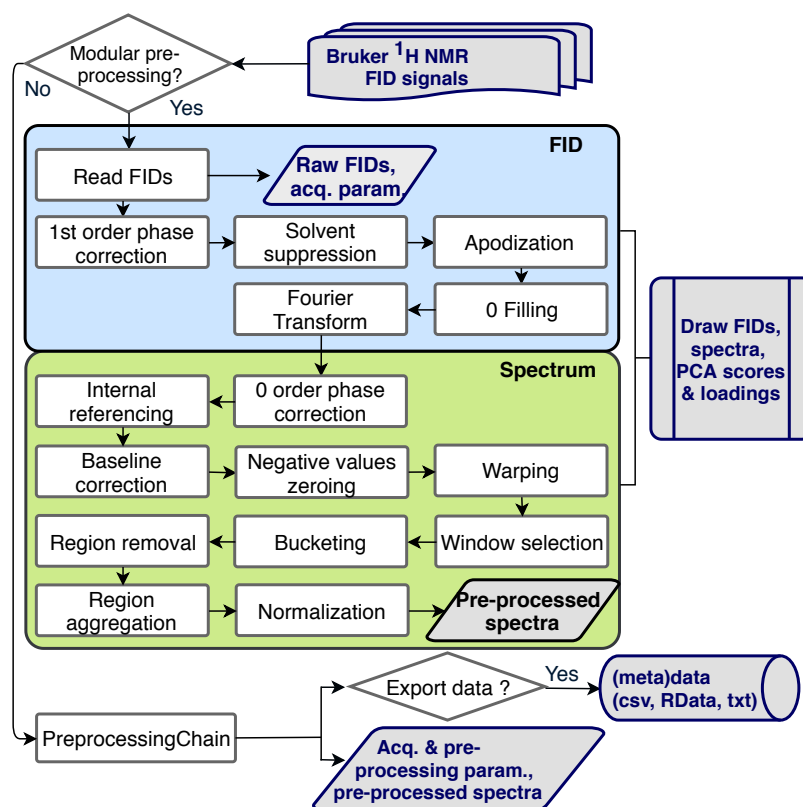
Figure 31. PepsNMR workflow for NMR-based metabolomics[133].

Read FIDs

This function imports raw Bruker files containing two matrices: the complex FID signal matrix and the acquisition parameters needed for the next steps. The FID recorded with the NMR is a complex time-domain decaying signal made of two parts named real (*sx* a cosinus phased signal) and imaginary (*sy*, a sinus phased signal) acquired by the two receptors x and y.

1st order phase correction and Group Delay suppression

The phase of the spectrum is corrected in such a way that the real part of the signal is in absorption mode and the imaginary part is entirely in dispersion mode. The total phase error has a frequency-independent and a frequency-dependent component that are corrected with the zero and first order phase correction. On Bruker spectrometer, a digital filter is used to increase the SNR, avoid baseline distortion, improve dynamic range and reduce signal folding[119]. If this tool has particular importance, it has several drawbacks that complicate the processing of the FID. Indeed, that implies that the first point of the FID is not the first point of the recorder signal as it starts after a time named "group delay". So, the first order phase correction must be applied to the FID after the removal of the signal acquired during this "group delay". With this function, each FID will be circularly shifted after a first Fourier transform, then finally back transformed

with the inverse Fourier transform to recover a processed FID without the signal arising from the group delay.

Solvent suppression

This step aims to suppress the solvent signal based on the assumption that the water, in the case of the analysis of serum/plasma sample, will be the main compound analyzed and therefore dominates the FID signal. The solution proposed by PepsNMR is to subtract the solvent residual resonance signal from the original FID by a smoothing function. The Whittaker smoother function will minimize $V + \lambda R$ where V is the sum of the squared differences between the original and the smoothed signal, R the measure of the roughness of the estimated signal and λ the penalty on roughness used to calculate the smoothed version of the FID[134,135]. PepsNMR allows to tune the λ parameter and different way to adjust it are available in the literature[134,136].

Apodization

To improve the spectral sensitivity and resolution, the FID is multiplied by a an exponentially decaying signal in a step called "apodization". During the acquisition, the NMR signal will decay in intensity over the time whereas the noise will continue to fluctuate randomly with a constant amplitude. Thus, multiplying the FID by a decaying signal will enhance the SNR as this ratio is higher at the start and decline over the time. The classical apodization method used in PepsNMR aims to principally enhance the SNR by using the decaying exponential: *exp(-t(1/T + LB))* where LB is the Line Broadening parameter. This parameter must be carefully assessed as an increase of *LB* will increase the SNR but at the expense of the peak height and thus the spectra resolution. For CPMG experiment the LB values are typically of 0.01 while usual values for the 1D NOESY are 0.3[133].

Zero Filling

The zero-filling function will add points at the end of the FID before Fourier transform to increase the resolution of the spectra. It will improve the data quality by increasing the amount of point in the processed data to better define peaks.

Fourier Transform

The Fourier transform will transform the complex signal into an interpretable spectrum. It will extract each signal and convert it into peaks with specific heights, positions and widths that are depending on the amplitude of the signal, the frequency of the corresponding proton, and its

relaxation time (***T1***) respectively. At the end a matrix is created containing the spectral intensity at specific chemical shift, starting from 0 ppm and increasing toward the left.

Zero order phase correction

As mentioned above, spectrum exhibits zero order phase shift of a certain angle due to the instrumentation. This phase shift is independent to the frequencies and the present function will find the optimal angle that will maximize the positiveness of the spectrum signal.

Internal referencing

This step will calibrate the ppm scale to the resonance frequency of the internal standard, usually Trimethylsilyl-propanoic acid (TMSP) or 4,4-Dimethyl-4-silapentane-1-sulfonic acid (DSS). The chemical shift of such compound is referenced to 0 ppm. The function used in PepsNMR aims to locate the reference compound peak within a range of intensities by selecting the highest intensity on this range. Then the center of this peak is referenced as the 0 ppm value.

Baseline correction

To improve spectral quality and therefore the statistical analysis of the acquired data, baseline correction is an essential step. In PepsNMR, an algorithm is used to estimate the baseline and remove it from the spectra using a nonparametric method known as Asymmetric Least Squares Smoothing (AsLS)[137]. Two main parameters are of particular importance for tuning the baseline correction of the processed spectrum, *p.bc* and *lambda.bc*. The smaller is the first one, the less the estimated baseline will try to follow peaks. The larger is the second one, the smoother the estimated baseline will be thus if *lambda.bc* = 0, the baseline will be equal to the signal and the corrected signal will be 0. By default, these values are set to 0.05 and 1×10^7 respectively.

Negative values zeroing

Since negative signal values have no sense and create problems with statistical analysis, the function Negative value zeroing is used to set the remaining negative intensities to 0 after baseline correction.

Warping

Since samples studied in NMR based metabolomics studies are subjected to several variations (pH modification, temperature variation, salt concentration and modification in concentration of specific ions, etc.) it is therefore not surprising that small peak shifts are observable between identical features from different samples. The solution to obtain more consistent data is to apply

a peak alignment algorithm to increase the similarity between spectra. Several methods are described in the literature and all have their pro and cons[138–140]. The warping/alignment method implemented in PepsNMR is inspired by parametric time warping (PTW) function previously described in the literature[133,141,142]. This method will try to find a warping function (a distortion function in the ppm axis that combine a polynomial term and a penalized B-splines term) between a reference spectrum and a sample. The principle is to find a function that will minimize the distance between the warped spectrum and the reference spectrum. In PepsNMR, the choice of the reference spectrum can be done by different ways: (i) by manually referencing the spectrum; (ii) an automated selection of the spectrum that minimize the sum of squared distances with all other spectra before or after warping (if the choice is "after", the process is slower as it will choose the reference after having been warping all spectra and test each single warped spectrum).

Window selection

This function is used to select informative areas of the spectra and therefore focus the analysis on relevant information.

Bucketing

Depending on the acquisition's parameters, an NMR spectrum is composed of several thousands of points of given intensities at a respective ppm shift. As biofluids only contains more or less than 50 metabolites that could be identified by NMR, it's a nonsense to describe the sample with such number of variables. Therefore, we need to reduce the dimensionality of the data before statistical analysis. The solution often used in untargeted NMR-based metabolomics is a segmentation method, called bucketing, that will divide each spectrum into several regions of the same width called bucket. The total area of each bucket is calculated and at the end of the process each spectrum is represented by a set of 200 to 1000 variables depending on the width of the bucket and the window selection (i.e: 500 for bucket width = 0.02 over 10 ppm window selection). PepsNMR allows not only the selection of the bucket width, but also the choice between two integration options, trapezoidal or rectangular[133].

Region removal

This step is focusing on removing resonance without interest from the spectra such as the water signal (4.5 to 5.1 ppm), the signal of an internal standard (5.8 to 6.2 for maleic acid) or a

region with residual baseline problems. Indeed, these features will add variability in our dataset and thus be problematic during the statistical analysis.

Region aggregation

Especially for sample with high pH variation and that contains ionized such as citrate, lactate, succinate, or taurine that cannot be properly aligned among spectra. This targeted data reduction will aggregate region of the spectra to condense the information in one larger bucket leaving a unique symmetrical triangular peak with null intensities at the border.

Normalization

Normalization is a crucial step as it will make all spectra comparable to each other prior the statistical analysis. It's a row operation in which each bucket from each spectrum is multiplied by a constant term[143]. This will minimize biases introduced by the experimentation during the collection, preparation and analysis of the samples but also will address dilution factor problems that could occur in samples (especially urine samples). PepsNMR provides different normalization methods:

- Mean: refer to constant sum normalization (CSN) in which each spectrum is divided by its mean so that its mean becomes 1. It assumes that the total peak area remains constant across all the samples[143].
- Median: each spectrum is divided by its median, so its median becomes 1.
- First quantile: each spectrum is divided by its first quantile so that its first quantile becomes 1.
- Peak: each spectrum is divided by the value of the peak of a selected compound or reference.
- PQN: This method is specially designed to consider the difference in metabolites concentration that can be quite important in urine samples i.e. In this method a reference spectrum is calculated based on the median spectrum. Then, for each bucket, the quotient of a given spectrum and the reference spectrum is calculated, thus the median of all quotients is estimated. Finally, all buckets are divided by the median quotient[144].

The choice of the normalization methods is not straightforward, and the choice will mostly depend on the type of the sample[145,146]. Applying CSN will be accurate when working on sera/plasma samples coming from human or animal cohort as the homeostasis of the blood media minimize the dilution effect. For urine sample, normalization based on the peak of creatinine can be recommended as creatinine concentration is expected to be correlated to urine

excretion and therefore sample concentration. Finally, for counteracting dilution effect or excretion difference between metabolites, PQN is often recommended[144].

Output data

At the end of the workflow, the R package will return a matrix that contains samples in raw and the buckets corresponding to the NMR intensities in column. This "bucket table" will be used for the subsequent statistical analysis.
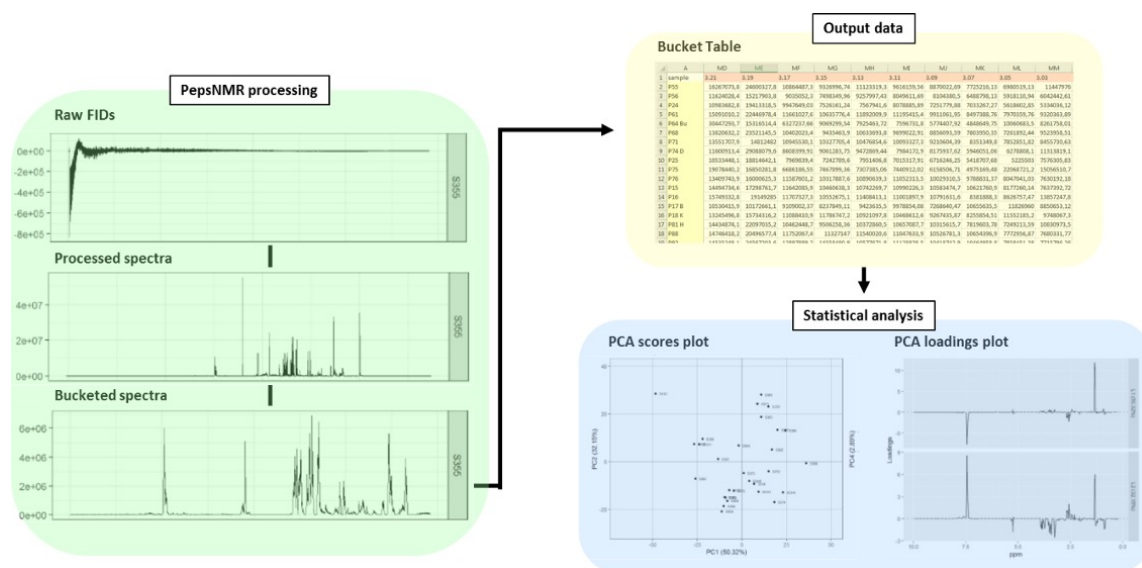
Figure 32. Representation of NMR data processing and analysis workflow with PepsNMR.

Here we presented PepsNMR, a R package that fits perfectly to untargeted NMR-based metabolomics (**Figure 32**). This methodology gives access to a robust pipeline in which manual intervention, and therefore bias, are minimized. Nevertheless, it's not the only way to deal with NMR metabolites profiling data. Indeed, other methods are existing to generate untargeted profiles by using innovative bucketing approaches or alternative strategies.

However, if an equidistant bucketing approach has the advantage to not be sample dependent and highly reproducible, its lack of flexibility can be a problem when samples exhibit high chemical shift variations. Indeed, the information is often divided between several buckets and, if spectra are not well aligned, this approach will lead to inconsistent data analysis. To solve this problem, other methods such as Adaptative Intelligent Binning algorithms (AI-Binning) were developed. This algorithm uses variable bucket in size and determines automatically when to stop before made the next bucket[147]. Hence each single peak will be included on a single bucket and peak splitting will be avoided. Another solution found in the literature to achieve bucketing is the one proposed *Laukens et al* which is based on peak picking algorithm using wavelets to summarize peaks within the spectra[148]. That methodology will return a matrix in which spectra are expressed in term of their peak intensities at given chemical shift rather than their integrals values.

Other alternative is the use of semi-targeted profiling methods that aim to quantify a set of identified metabolites among the sample[99,149]. Indeed, as metabolites resonances in NMR spectra are highly overlapped, bucketing will not give the best representation of metabolites levels and spectral deconvolution approach are the solution of choice. Typically, the set of metabolites will depend either on the biological question than on the tool and the size of the metabolite library available[150]. Indeed, most of the available tools for deconvolution are based on internal databased to work consistently. Hence free-access tool as BATMAN[151], BAYESIL[152], ASICS[153], rDolphin[154] and AQuA[155] provide automated workflow for quantitative profiling of aqueous metabolites previously included in the metabolite library. A solution also exists for lipid profiling as LipSpin was recently published and available via Matlab interface[156]. If no in-house database is available, Chenomx©, a commercially available software, provide access to a database in which each metabolite is modeled using its peak center and J-coupling information[149]. Thus, users are allowed to quantify the identified compound using spectral fitting option and extract metabolites concentrations. Up to 70 compounds where identified and/or quantified in biofluids using this tool[102,103]. However, despite providing robust results for the set of studied compounds, these approaches don't offer the opportunity to identify unknown features. An overview of all these tools is depicted in **Figure 33**.

Data processing/analysis pipeline for NMR metabolomics
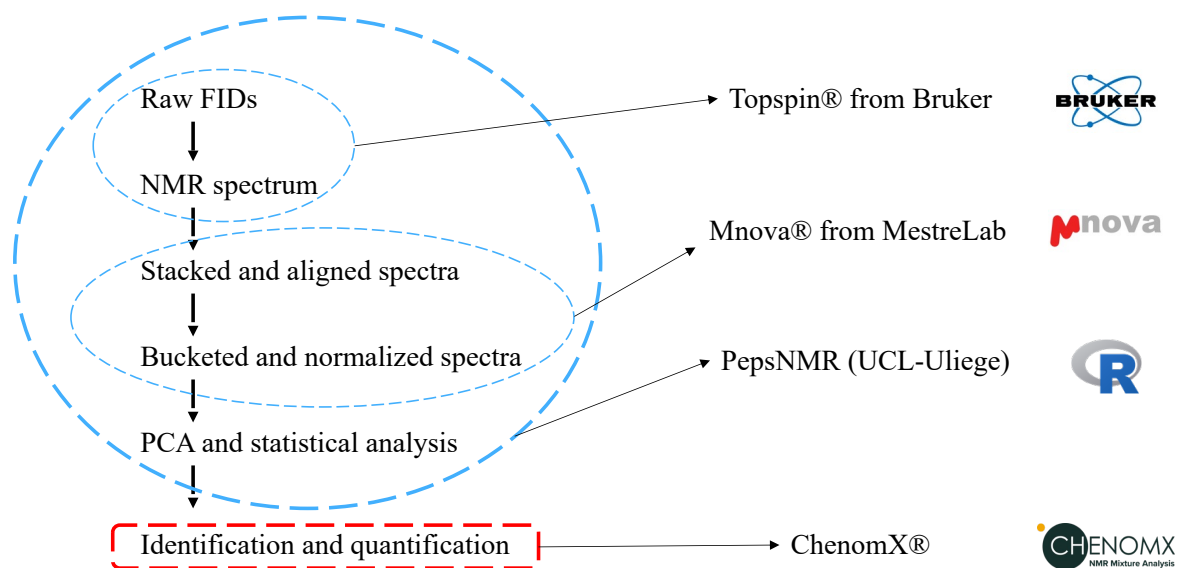
Figure 33. Summary of NMR metabolomics data processing and analysis tools used in this work.

1.3.9 ^{1}H-NMR metabolomics data scaling

In biofluid samples, metabolites are present in various concentration and due to their high concentration, a small number of metabolites can dominate the result of the statistical analysis at the expend of other metabolites. For that reason, scaling metabolites intensities among the dataset is mandatory to avoid bias by rescaling the weight of all variables through a mathematical operation for all samples[145,157]. All the different methods of scaling were reviewed by *Van Den Berg et al.,* they all have their pro and cons and the choice of the method used is driven by the specificity of the dataset. Here some of the most used methods:

- Centering: In this method the mean value is extracted for each variable and the mean of transformed variable is equal to 0. If this method is usually applied in most data prior PCA or PLS regression analysis, it's not sufficient in heteroscedastic data (set of variables exhibiting different variances).

- Autoscaling: In this method each column is centered by subtracting the mean intensity then divided by the standard deviation of each column. As all the variances will be equal, this methodology avoids the domination of variable with high variance in PCA and PLS analysis. However, this strategy will give noisy region as much importance as region of interest. Therefore, this may be a problem for NMR spectra that contain a lot of noisy regions.

- Pareto scaling: This approach will apply mean centering then divide the column by the square root of the standard deviation. This is suitable to reduce the influence of high intensity region while increase the importance of small peak. All the variances will be roughly equal. Particularly relevant for NMR data as it will downgrade the importance of the noise.

- Range scaling: After mean centering each column is then divided by the minimum and maximum range of that variable. This will make all variables roughly equal and is suitable for exploratory analysis. However, when the range between maximum and minimum, this method is quite insensitive and will enhance variable with smaller variability.

1.3.10 ¹H-NMR metabolomics data analysis

As in NMR based metabolomics one individual is described by hundreds of variables and usually such studies potentially include hundreds of individuals, the amount of generated data is considerable. In metabolomics, several chemometrics methods were developed to produce interpretable and robust models that can handle this level of complexity[158]. These methods range from univariate statistical testing to multivariate regression methods as principal component analysis (PCA), partial least square (PLS) or orthogonal projection to latent structure (OPLS), cluster analysis, and finally machine learning or non-linear methods such as support vector machines (SVM) and neural network (NN) respectively. Basically, most of these tools can be classified into three different categories based on their uses for data overview (or unsupervised analysis), classification analysis (or supervised analysis) and identification of potential biomarker.

Unsupervised methods

The goal of unsupervised or data overview methods is to summarize the complex data. These exploratory tools are able to detect data patterns correlated with biological or/and experimental variables[159]. The most common method for metabolomics data overview is principal component analysis (PCA). PCA will summarize the dataset into a set of linearly uncorrelated variables called components[160]. The first component is composed of the variables that represent the most of variance among the dataset and therefore that maximize the dispersion between the individuals. The subsequent components will explain increasingly reduced amounts of variance and the created model will try to minimize the covariance between these components making them as much independent to each other as possible. Hence, after applying PCA analysis, all individuals will be represented in respect to their differences/likeliness to each other in a *score plot* while the weight of each variable within the component are plotted on the *loading pot* (**Figure 34**). Hence the score plot will allow to answer the question "how much are the different samples?" and the loading plot "What is different among the samples?".
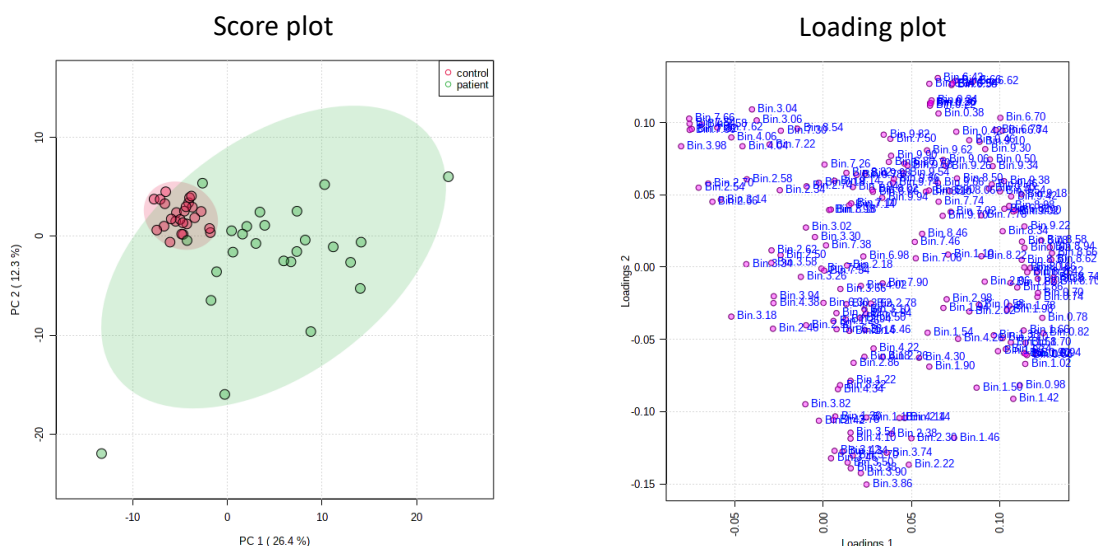
Figure 34. Most common plot of PCA analysis on metabolomics data: (a)a score plot on which each individual is represented on the model; (b) the corresponding loading plot on the right with all the variables used in the model; plot generated using demonstrating dataset of MetaboAnalyst®.

This approach is often used to detect outlier, to highlight analytics/pre-analytics biases or possible grouping of individuals and/or variables without *a priori*[161,162]. For these reasons, unsupervised methods are often used prior to supervised analysis that will focus on differences among class of individuals.

Another statistical method used for unsupervised analysis purpose is Hierarchical Clustering Analysis (HCA). This methodology will define groups of observation that exhibit similarities between them but that are different from the ones of the other groups. The main purpose of such approach is to put in light clusters of observation among the datasets. By using the Euclidian distance as metrics to show similarities between samples, a dendrogram is created. On this dendrogram the two closer observations will form a group, this group will be put in relation with the remaining dataset and will form another group with other close observation and so on (**Figure 35**). If this method does not allow a direct evaluation of the contribution of the variables to class construction, it will provide a good vision of the structure of our dataset and how samples are clustering together.
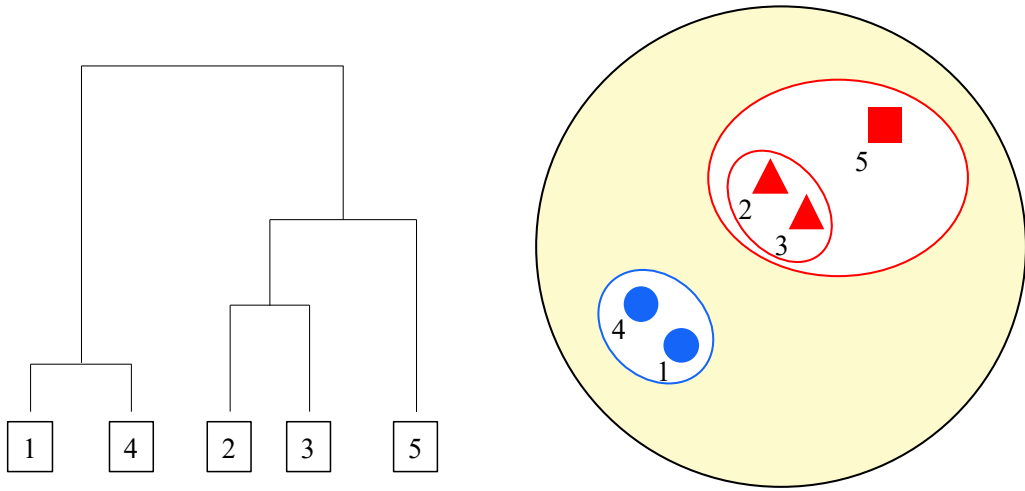
Figure 35. Schematic representation of HCA methodology.

Supervised analysis

The goal of supervised analysis is to find the variables able to classify samples according to their group, in other word to find a correlation between the X matrix and the Y response matrix. This can be achieved using linear methods such as PLS(DA) and OPLS or non-linear methods like SVM and NN[158] and the Y matrix can be quantitative (age, drug doses, …) or a qualitative binary response (treated/control, hill/healthy, …)

Partial Least Square (PLS) is the most used supervised analysis in metabolomics studies[159]. At the opposite of PCA, PLS will not maximize the variance of the dataset but the covariance between the response of interest and the metabolomic data[163]. Therefore, the loading plot of PLS component will represent which features are contributing to the discrimination of the different sample groups. The major drawback of this modeling method is that features uncorrelated to the variable of interest can influence the model and lead to less interpretable model. It is for solving this problem that orthogonal Partial Least Square (OPLS) was developed[164]. This model will factorize the data variance into a set of components: the first ones are correlated with the factor of interest, and the others contain variables uncorrelated with this factor and are therefore orthogonal. The features related to the separation of the sample groups are then summarized on one unique component while the others are not considered. This makes loading a score plot more easily interpretable as the effect of the studied factor is summarized in one component.

Despite their usefulness, supervised methods must be used carefully as they can lead to overfitted models and thus to uninterpretable results. To avoid overestimated result and to produce robust model, each performance evaluation and the validation of each model coming from supervised analysis is mandatory.

Performance assessment and validation methods

The main objective of metabolomics, and therefore of applied classification methods, is the identification of biomarkers or use of spectral profiling as fingerprint. The use of supervised analysis allows to identify features/profile that are explaining the biology behind the dataset. In order to assure the usefulness of the identified feature, performance assessment and validation of the generated model are a crucial step.

Performance assessment tests will measure if the predicted outcomes are matching to the real outcomes. These tests will measure the percentage of well classified samples (predictive accuracy), the percentage of true positives correctly classified (sensitivity) and the percentage of true negative correctly classified (specificity)[165]. The most common performance assessment methods are the Receiving Operator Characteristic (ROC) curve. The ROC curves show how the specificity and the sensitivity change as the classification boundary is varied among the range of identified biomarkers. The ROC curve is often summarized by the area under the curve (AUC) metric. The AUC indicate the probability that a classifier will rank a randomly chosen positive sample higher than a randomly chosen negative one[165]. Thus, for a perfect classifier, that always classify samples in the good group, the AUC = 1. For a random classifier the value will be 0.5 while a value greater than 0.7 is considered as the minimal performance to reach for a clinically relevant biomarker[165].

The validation step will assess how the model will succeed when applied to a new cohort and is especially important when the model is based on a small number of samples[159]. The main tools used for this purpose are the permutation testing and the cross-validation[166].

The permutation test will assess the probability of observing equal or better performance for model obtained by permuting class labels and/or randomly assigning them to different subjects. If the model is good, none of the randomly generated model will exhibit higher performance than the tested one[166].

The cross-validation method is another common tool in metabolomics. It's based on two parameters to assess the model's performance, R^2 (x and Y) and Q^2Y. The R^2 represent the explained variance of the matrix of the x and y variables, while the Q^2Y is representing the predictive quality of the model[166]. The closer to 1 these values are, more the model is predictive and the result of the separation significant. The cross-validation method will estimate these parameters using an iterative approach[159]. During this test, the total sample is split into two group, a training group and a testing group[166]. The training group is used to produce a predictor model using a specific set of parameters and the performance of this model is evaluated using the remaining group (testing group). This procedure is repeated numerous times so at the end

of the process each sample have been used once in the testing group. The average results will represent an unbiased estimation of the performance of the tested model (**Figure 36**)[159,166].

Figure 36. (a) OPLSDA score plot and (b) the corresponding permutation test; plot generated using demonstration dataset from MetaboAnalyst®.

Important features and metabolites identification

Once the supervised models are validated and the performance of the classifiers assessed, the important variables are represented in the variable importance in projection (VIP) plot (**Figure 37**). In this plot, the importance of each variable that explain the discrimination between groups is estimated. Typically, VIP-values larger than 1 indicates important variables while unimportant variables exhibit values lower than 0.5. As the important variables are corresponding to the buckets that are relied to the different spectral zone, a careful investigation of these zones is needed to identify the important metabolites and provide a correct biological interpretation of the data.
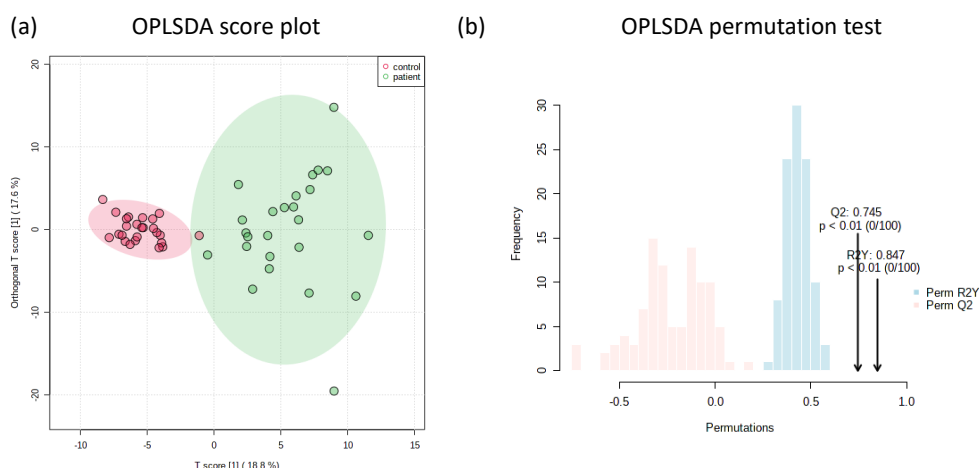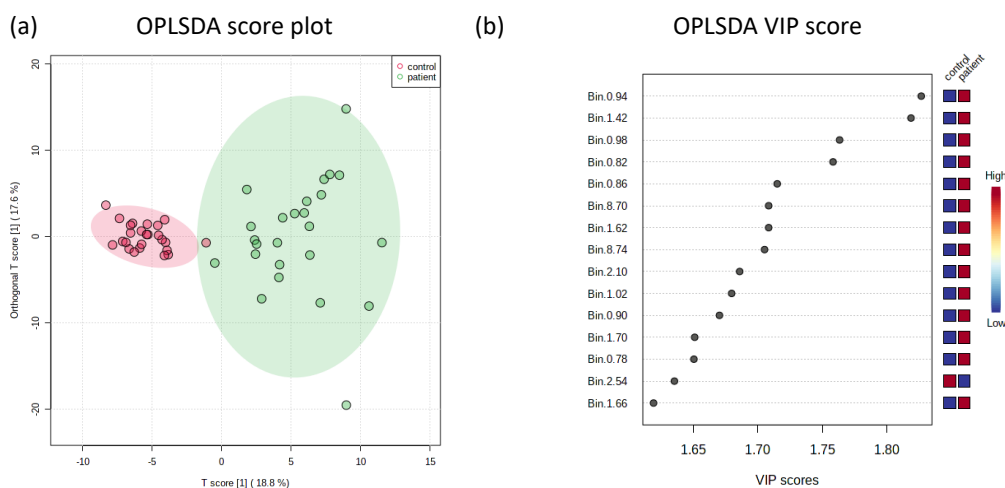

Figure 37. (a) OPLSDA score plot and (b) the corresponding VIP score plot; plot generated using demonstration dataset from MetaboAnalyst®.

In first intention, the use of chemical shift list coming from literature for a given biofluid or tissue[102,103] allow to visualize all the metabolites having spectral component on the zone of interest. Despite this information, identifying the metabolite remains a difficult task due overlapping peaks in most part of the 1D NMR spectra. To solve this problem, we use specific 2D NMR experiments that allows to study connectivity between neighboring spins systems and thus overcome the overlapping problems (**Figure 38**). The most common 2D NMR pulse sequence is the Correlation Spectroscopy (COSY) that studies homonuclear ^{1}H correlations through bond interaction. COSY experiment allows to study the structure of the molecules and therefore reveals specific connection able to individualize some metabolites. Another frequently used experiment is the Heteronuclear Single Quantum Correlation (HSQC) spectroscopy that allows to visualize connectivity between the ^{1}H spins and their directly attached ^{13}C spins. It provides unique signatures that allows identification of metabolites and is particularly useful to annotate singlet resonance signal coming from isolated spins systems. As for 1D NMR, web databased such as HMDB (www.hmdb.ca) give access to 1D and 2D spectra as well as theoretical concentration of metabolites among the different biofluids. An alternative to supplement NMR experiment for identification of unknown metabolites is the use of dedicated software like ChenomX© that allow the assignment and the identification of signals from an 1D NMR spectrum by matching peaks coming from an internal database.
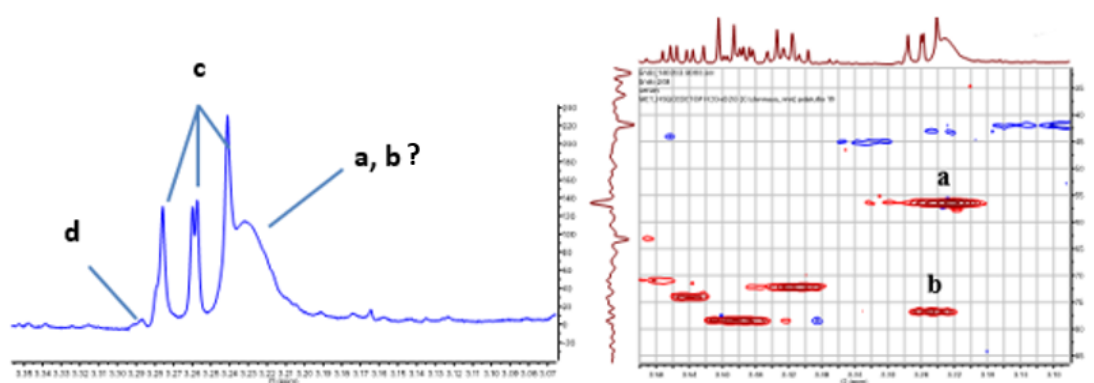
Figure 38. Overlapping peak (a) and (b) on the 1D spectra are resolved in the ^{1}H-^{13}C HSQC experiment and can be identified and/or quantified.

Taken together, all these tools will provide successful identification of important feature spotted during the supervised statistical analysis. This makes possible the identification of biomarkers of interest and a better understanding of the biochemistry behind the design of experiment. Indeed, when metabolites are identified, the reconstruction of metabolic networks

is possible and constitute the final step for biological interpretation of the data. Pathway analyses will provide a visualization of the interactions among genes, proteins or other metabolites that relate to the identified features. This can be made using database such as KEGG (Kyoto Encyclopedia of Genes and Genomes), a web interface, or dedicated software such as Metaboanalyst[167], Cytoscape[168] or Metexplore[169]. It must be noted that several guidelines exist to publish new biomarkers in dedicated journals of metabolomics. Hence, the authors are encouraged to report the level of confidence (table 3) in metabolite identification following the guidelines defined by the Metabolomics Standard Initiative (MSI)[170]. Furthermore, some journals as Metabolites require the submission, in an open-access database, of all the information about the design experimental and the spectral data acquired in order to gain in transparency and reproducibility[171].

Level 1	Identified compounds (see below).
Level 2	Putatively annotated compounds (e.g. without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/com- mercial spectral libraries)
Level 3	Putatively characterized compound classes (e.g. based upon characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class)
Level 4	Unknown compounds—although unidentified or unclassified these metabolites can still be differentiated and quantified based upon spectral data.

Table 3. Level of metabolite identification in published metabolomics literature. For NMR "identified compounds" relate for spectral matching with an authentic spectrum of the reference compound[170].

1.4   Metabolomics and personalized medicine

Personalized medicine is an emerging and promising clinical practice that relies on new technologies to help clinicians regarding prediction, prevention, diagnosis, and treatment of disease. These technologies aim to give access to the individual's unique characteristics at the molecular, physiological, environmental, and behavioral levels that have a significant impact on their disease processes and therefore to the treatment they need for a maximized health care. In this context, metabolomics could provide a precise characterization of metabolic phenotypes and thus can be useful for underlying metabolic changes liked to disease, for the discovery of new therapeutic targets and biomarkers that can be used either for early diagnosis, patients stratification or treatment response evaluation[172,173].

The NIH[174] define a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention." In a recent review, *D.K. Trivedi et al.* summarized several key metabolomics studies that have identified new biomarker candidates for diverse diseases[175] and pointed out that metabolomics should not only be for pathological cure but also for preventive screening of healthy individuals. Indeed, early biomarkers have particular importance to prevent disease and may be useful in directing dietary or lifestyle change prior more radical intervention. For many diseases, screening healthy individuals make sense as any change in biomarker(s) levels is personalized. As an example, Prostate-specific antigen (PSA) is used for early detection of prostate cancer but people exhibiting enlarged prostate will already have higher PSA level[176]. Therefore, early screening and follow-up of individuals could be used in complement to current clinical practice and could dramatically improve patient's health care.

As concluded in a recently published white paper, the use of metabolomics data in clinical practice is a powerful tool for personalized medicine. Indeed, it will provide overview of the changes among the patient's metabolism at baseline, prior to treatment, during the treatment and post treatment and can inform about treatment outcomes and variation in drug responses[177].

1.5 AMD and metabolomics

As previously described, biomarkers identified through metabolomics studies have particular importance for diagnosis, patient's stratification and follow-up, drug responses evaluation. This approach aims to better understand the physiopathology of disease and can lead to new drug target discovery and improved therapeutic strategies. In the context of AMD several studies were conducted in order to investigate metabolome changes among diverse biofluids of AMD patients at different stages of the disease[178]. These approaches identified different compounds belonging to the oxidative stress and energetic pathway, to inflammatory processes or to lipid metabolism[179–181]. In this chapter we will describe some of them and discuss their role an implication into the different stages of AMD as well as their potential clinical application[182].

Oxidative stress is a condition that occurs within cells or tissues when exposed to an excess of molecules containing free radicals such as reactive oxygen species (ROS) and reactive nitrogen species. Even if antioxidative mechanisms exist to maintain the homeostasis, continuous presence of reactive oxygen species (ROS) can lead to oxidative modification of such cells or tissues[183]. ROS led to the degradation of all types of biological molecules such as proteins, DNA or lipids and can cause various pathological condition[184]. Because of its constant exposure to light, its high metabolic activity and the presence of oxidable species (Polyunsaturated fatty acids, PFUA, a content of membrane photoreceptive cells), the macula is highly exposed to oxidative stress. Indeed, photoreceptive shedding, light exposure, and environmental factors such as smoking, or alcohol consumption, are responsible of ROS production and accumulation and are proved to play a role in AMD pathogenesis. Therefore, it's not surprising that potentially valuable biomarkers for AMD incidence or progression are related to this metabolic pathway. The following table (**table 4**) provides a non-exhaustive list of potential biomarkers investigated and the conclusions of the studies:

Compound	Reference	Conclusion	Type of AMD	Matrix
Malondialdehyde (MDA)	185–188	Up	nAMD/Early AMD/Any AMD	plasma/serum
Total Oxidation Status (TOS)	184,186	Up	nAMD	plasma/serum
Oxidized LDL (Ox-LDL)	189–191	Up	nAMD	plasma
Nitric Oxide (NO)	185,192,193	Up/Down	Any AMD/Early AMD	plasma
Homocysteine (Hcy)	194–198	Up/N.D	nAMD/Early AMD/Any AMD	plasma/serum
Total Antioxidant Capacity (TAC)	184,186,188	Down/N.D	nAMD/Early AMD/Any AMD	plasma/serum
Thiol content (tSH)	184,190	Down	nAMD	plasma/serum
Glutathione (GSH)	190,199,200	Down	nAMD	plasma/serum
Carotenoids	182	Down	nAMD	plasma/serum
Lutein	48,201,202	Down/N.D	nAMD/Early AMD/Any AMD	plasma/serum

Table 4. List of potential biomarkers related to the oxidative stress pathway; Up and Down indicate the fold change in concentration for the case group in comparison the the control group; N.D is for no significative difference.

As discussed above, life light exposure participates in the accumulation of oxidative damage within the retina that trigger age-related degenerative disease such as AMD. These damages will occur at different levels and will induce a pro-inflammatory response[203]. If inflammation caused by tissue damage is considered as an essential response of the immune system, chronic inflammation caused by recurrent event is a pathological condition. Indeed, it is well known that chronic inflammation plays a role in many age-related diseases such as cancer, Alzheimer's disease and AMD[204,205]. In both wet and dry AMD, inflammatory immune response was associated with drusen as histopathological studies demonstrated the presence of complement cascade component in their composition[17,206].

Drusen, the major hallmark of AMD, are composed by at least 40% of lipids and many associations were made between AMD and lipid-linked genes such as CETP, ABCA1, APOE and LIPC[42,206–208]. Therefore, it's not surprising that numerous studies have investigated lipid metabolism changes among serum or plasma of AMD patients. These studies included Triglycerides (TG) and cholesterol, phospholipids (PL), several fatty acids (DHA, EPA, ALA, DPA, AA, LA, PA, OA, SA), lipoproteins and apolipoproteins moieties. All these compounds and the conclusions of the linked studies are reported in the following table (**table 5**). Together these studies show that significant progress was made to identify interactions between lipids

and AMD. To fully understand this association, supplemental studies are needed using omics approach such as lipidomics, metabolomics, proteomics, and transcriptomics. These can help to clarify whether the concerned lipids are coming from the diet or produced locally by specific cells or tissues during AMD processes[209].

Compound	Reference	Conclusion	Type of AMD	Matrix
Docosahexaenoic acid (DHA)	210–213	Up/N.D/Down	Any AMD/Dry AMD/nAMD	plasma/serum
Eicosapentaenoic acid (EPA)	212,214	Down	nAMD	RBMC, serum
α-Linolenic acid (ALA)	211,212	N.D/Down	Dry AMD/Any advanced AMD	Serul/plasma
Docosapentaenoic acid (DPA)	211,212	N.D	Any AMD/Dry AMD	plasma/serum
Arachidonic acid (AA)	210,211,213	N.D/Up	Any AMD/nAMD	plasma/serum
Linoleic acid (LA)	210,211	N.D/Down	Any AMD	plasma/serum
Oleic acid (OA)	210,211	Down/N.D	Any AMD/Dry AMD	plasma/serum
Palmitic acid (PA)	210,211	Down/N.D	Any AMD/Dry AMD	plasma/serum
Stearic acid (SA)	210,211	N.D	Any AMD/Dry AMD	plasma/serum

Table 5. List of potential biomarkers liked to lipids pathways ; Up and Down indicate the fold change in concentration for the case group in comparison the the control group; N.D is for no significative difference.

Lipoproteins are of particular interest as they aim to transport lipids trough the organism to achieve their role in energy storage, cell signaling or in the composition of cell membranes. This group of particles express different biochemical and physico-chemical properties and is divided into five classes according to their density[209] **(Figure 39)**.
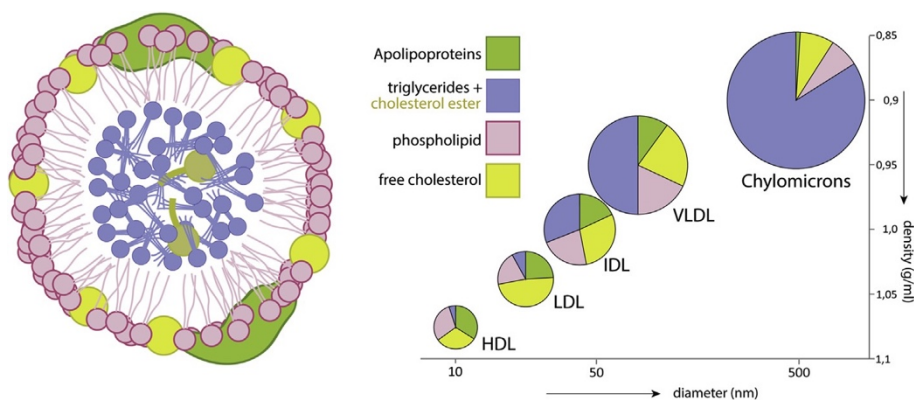
Figure 39. Lipoproteins and their lipids distribution for the 5 main classes based on their density (adapted from[209]).

The pathologic condition in which the level of circulating lipoproteins level is altered is called dyslipidemia and is associated with cardiovascular diseases. Therefore, measurement of circulating lipoproteins levels, as well as circulating cholesterol and triglycerides, is part of routine clinical practice for risk prediction and monitoring treatment efficiency[215]. For AMD, study suggested an increased risk for AMD for elderly patients with higher HDL concentration[216]. Moreover, HDL dysfunction seems to be implicated in AMD pathogenesis as change in their composition could lower their anti-inflammatory properties and lead to LDL oxidation[217,218]. Oxidation of lipoproteins content is thought to increase oxidative stress, inflammation and permeability of RPE cells, this induces dysfunction of outer blood retinal barrier and promotes VEGF production, the hallmark of nAMD[219]. Finally, higher levels of LDL and/or oxidized-LDL were found associated with nAMD in several studies[190,220,221] showing the potential usefulness of such entities as biomarkers linked to the severity of the pathology. It must be noted that if several studies find association between lipoproteins levels and AMD, other large population-based studies did not. Difference in result can be partially explained by the uses of different measurement methods across studies. Indeed, if these can be measured directly by different methods including analytic ultracentrifugation, gradient gel electrophoresis, HPLC or ^{1}H-NMR[222], lipoproteins levels are often estimated using the Friedwald equation[223] and the collected data are therefore less consistent[224].

In conclusion, data available seem to indicate the potential role of lipid metabolism and lipoprotein balance/composition in the occurrence/development of AMD. However, further investigations are needed to identify putative biomarker and highlight key mechanisms of the pathology.

1.6 AMD and NMR-based metabolomics

As discussed in the first chapter, Age-related macular degeneration (AMD) is the leading cause of vision loss among elderly population in developed countries. This degenerative disease is evolving trough different stages and 90% of blindness due to AMD result from the exudative form of this pathology (nAMD). Up to now, diagnosis only relies on ophthalmologic exams and treatments of the most aggressive form are based on anti-angiogenic drug targeting the vascular endothelial growth factors (VEGF). Despite these, major clinical challenges must be overcome in order to improve patient's care with early diagnosis methodology, refined patients' stratification and improving evaluation of treatment responses. For this purpose, we decide to apply an NMR-based metabolomics approach based on both AMD patients and nAMD experimental mice model cohort. This work started few years before the beginning of this thesis and has set the foundation of this work. All the results reported below were published in 2020[225]. All the metabolomics' related analysis reported here were done during this thesis. Either for human and mice cohorts, these analyses were performed on already acquired ^{1}H-NMR data to confirm precedent assumption and refine generated information.

NMR based metabolomics discriminate patients with late AMD[225]

Our first study was based on a cohort of unrelated European Caucasian individuals (>59 y.o) affected with nAMD (*n*= 72) and healthy volunteers without sign of any AMD or known family history (*n*= 50). Plasma and serum samples were collected to perform NMR-based metabolomics analysis, lactate quantification and blood analysis (blood cell counts, leucocyte differential, and CRP measurement). For the NMR analysis, all samples were recorded at 298K on a Bruker Avance spectrometer operating at 500.13 Hz for proton acquisition. Maleic acid was used as internal standard for quantification and trimethylsilyl-3-propionic acid-*d*4 (TMPS) for signal calibration. ^{1}H-NMR spectra were recorded using a 1D-CPMG relaxation editing sequence with water presaturation. Data were processed using Bruker Topspin 3.1 software with standard parameter set to prepare spectra for bucketing steps. Optimized ^{1}H-NMR spectra normalized to total intensities and reduced to integrated regions of equal width (0.04 ppm) over the region comprised between 0.5-10 ppm. The former bucketing steps were conducted using AMIX software, provided by Bruker, and unwanted region were deleted from the processed spectra (water between 4.7-5 ppm and maleic acid 5.6-6.2 ppm) returning a matrix containing samples in row and buckets integration in column. The corresponding data matrix is then imported to SIMCA for statistical analysis. After Pareto scaling of the variables, non-supervised and supervised discriminant analysis were performed (PCA, PLSDA and OPLSDA). While

PCA was used for outlier detection and determination of intrinsic cluster, PLS-DA and OPLS-DA were used for group comparison and biomarker detection.

The prospective cohort of patients was sub-divided into 3 groups: active AMD (patients with exudative late AMD); inactive AMD (patients with stabilized AMD); and healthy volunteers. Supervised analysis of NMR metabolomics data allowed partial discrimination between all AMD patients ($n= 72$) and healthy donors ($n= 50$). Moreover, when focused on the 2 subgroups of AMD patients, discrimination between patients exhibiting active AMD and those in the inactive phase of the disease was noticeable (**Figure 40a-b)**. Corresponding loading plots allowed the identification of the spectral zone responsible and metabolites for these separations (**Figure 40c)**. Among these, lactate levels and lipoprotein profile changes appeared to be linked to the active phase of the pathology. Subsequent ^{1}H-NMR and enzymatic dosage confirmed the increase of lactate in the blood of AMD patients in active phase compared to patients with inactive AMD and healthy donors (**Figure 40d**). This study highlights lactate and lipoprotein profile as marker for exudative AMD and could potentially be used as a marker of the active phase of the pathology. Further investigation has to be made in order to assess the changes among the lipoprotein profile and the possible pathologic role of lactate in the neovascularization processes.
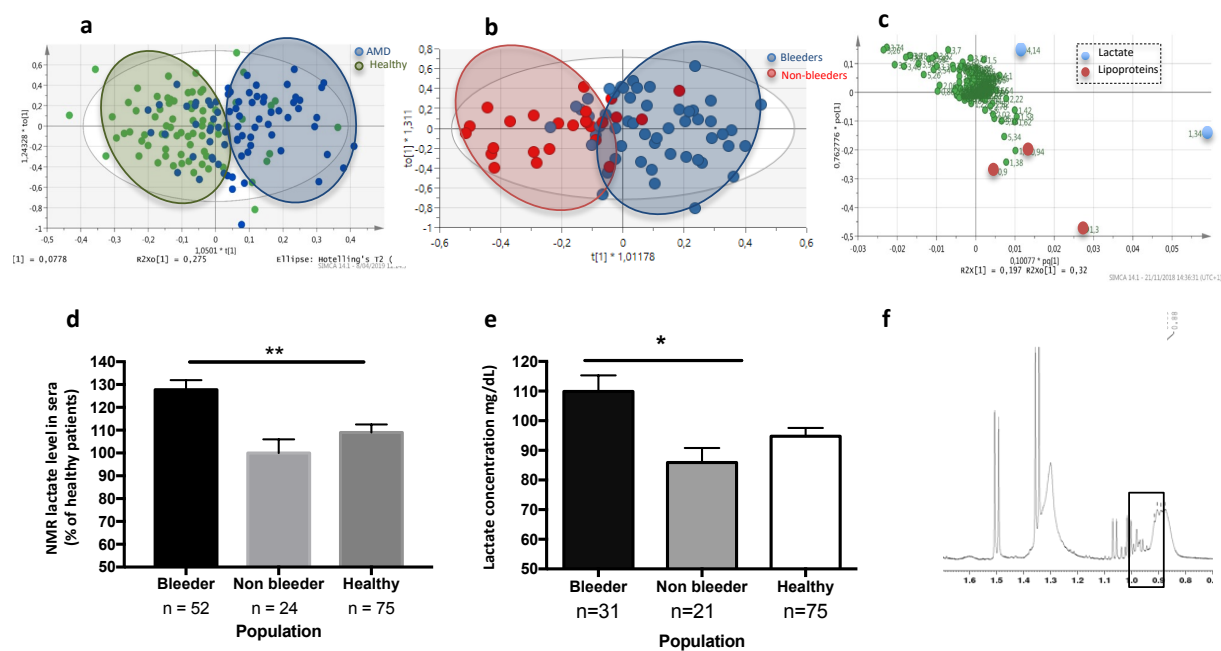
Figure 40. Lactate and lipoproteins are the main increased discriminant metabolites in serum of patients with nAMD. (a) Score plot derived from an OPLS-DA of spectral data (3 components, R2 = 0.527, Q2 = 0.151) collected from patients with nAMD (blue dots, n = 72) and healthy volunteers (green dots, n = 50). Each data point represents an individual patient. (b) OPLS-DA score plot (2 components, R2 = 0.468, Q2 = 0.21) of spectral data collected from patients with active (red dots, n = 49) and inactive (blue dots, n = 23) nAMD. (c) Loading plot of spectral data collected from patients with active and inactive nAMD highlighting lactate and lipoproteins as biomarkers of active status. (d) NMR and (e) biochemical dosages of blood lactate in the serum of healthy volunteers and patients with active and inactive forms of AMD. Data are expressed as the percentage of healthy donors. (f) NMR spectrum of human serum highlighting the lipoprotein profiles (from 0.88 to 0.92 ppm)[225].

CNV is induced in 8-week-old C57BL/6J mice by laser impact as previously described. Neovascular lesions appear at day 5 post laser-burn and are the largest at day 7, collection of blood samples allows kinetics studies of CNV formations[226]. To study the kinetics of CNV formations, blood samples were collected from scarified mice at day 3, 5 and 7 post laser-burn. CNV were quantified through FITC-dextran labelled flat-mounted choroid and metabolic profile measured using ^{1}H-NMR. Longitudinal NMR-based metabolomics analyses spotted a good concordance between changes in the metabolome and CNV progression. Indeed, most significant discrimination between control and induced mice occurred at day 5 and 7 post-laser burn concomitantly with CNV progression quantified on flat-mounted choroids. In accordance with the conclusion obtained in the clinical study, lactate and lipoprotein profiles were the major discriminant features responsible for the separation between groups. An increase of approximatively 15% of blood lactate concentration was measured by NMR in CNV mice at day 5 and the major changes among the lipoprotein profile were noted at day 7. These data suggest that lactate and lipoprotein profile are strong indicator for CNV occurrence and progression (**Figure 41**).
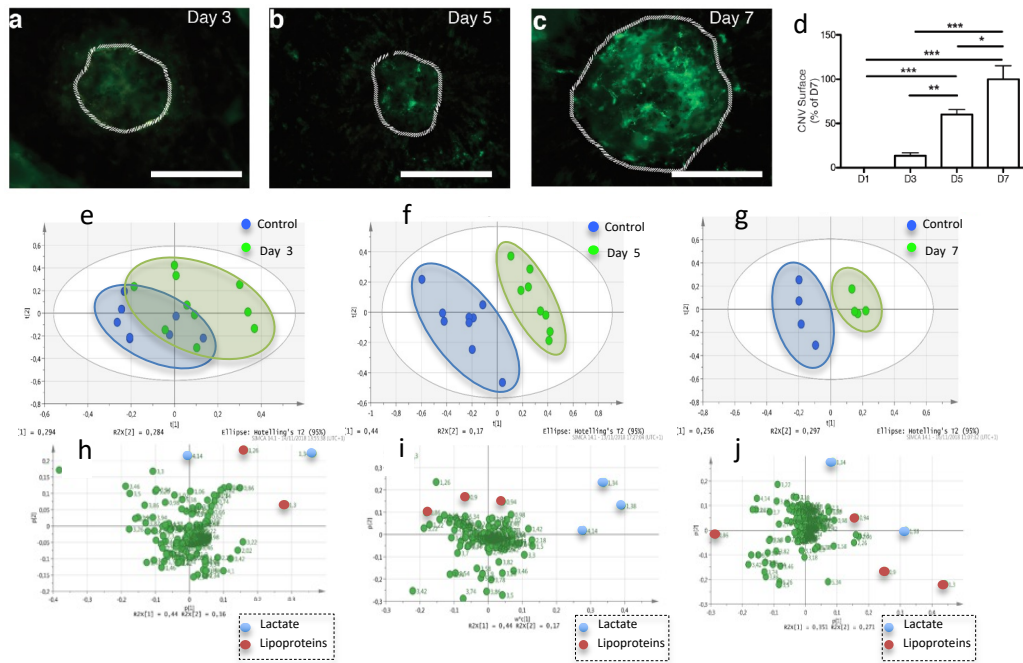
Figure 41. Lactate and lipoproteins are the main increased discriminant metabolites in the serum of mice subjected to CNV. Mice were subjected (CNV) or not (CTL) to a laser burn. FITC-dextran-labeled flat-mounted choroid observed at day 3 (a), 5 (b) or 7 (c) after laser induction. Dashed lines delineate the lesion. Scale bars, 100 μm. Quantification of fluorescent neovessel area with ImageJ software (n ≥4 mice/group, n ≥ 12 laser impacts/group) at days 1, 3, 5, and 7 (d). *P < 0.05; ***P < 0.001. Error bars indicate SEM. Score plot resulting from a PLS-DA analysis of spectral data performed at day 3 (e), day 5 (f), and day 7 (g) after laser induction (control vs J3: 2 components, R2 =0.579 and Q2 = 0.119; control vs J5: 3 components, R2 = 0.669 and Q2 =0.841; control vs J7: 3 components, R2 = 0.793 and Q2 = 0.734). Laser-induced mice (green dots) were distinguishable from non-induced mice (blue dots) at day 5 and day 7. Each data point represents an individual mouse (n ≥ 4 mice/group). Loading plot resulting from a PLS-DA analysis of spectral data performed at day 3 (h), day 5 (i), and day 7 (j) after laser induction. Lactate (blue dots) and lipoproteins (red dots) are the main discriminant metabolites[225].

An interesting point of our study is that the normalization of blood lactate level trough dichloroactetic acid (DCA) treatment led to a reduction in CNV formation in our mice nAMD model. Indeed, DCA is a known inhibitor of pyruvate dehydrogenase kinase (PDK) and will modulate lactate level trough the deactivation of the pyruvate dehydrogenase (PDH) involved in pyruvate conversion into Acetyl-CoA in the mitochondria (**Figure 42**). Data recorded on treated mice demonstrated the functional implication of lactate in CNV progression and paved the way to a new therapeutic axis based on blood lactate level normalization. Interestingly, on our experimental mice model, lactate level normalization led to a reduction of CNV and a normalization of lipoprotein profiles with the same efficiency as anti-VEGF drug did.
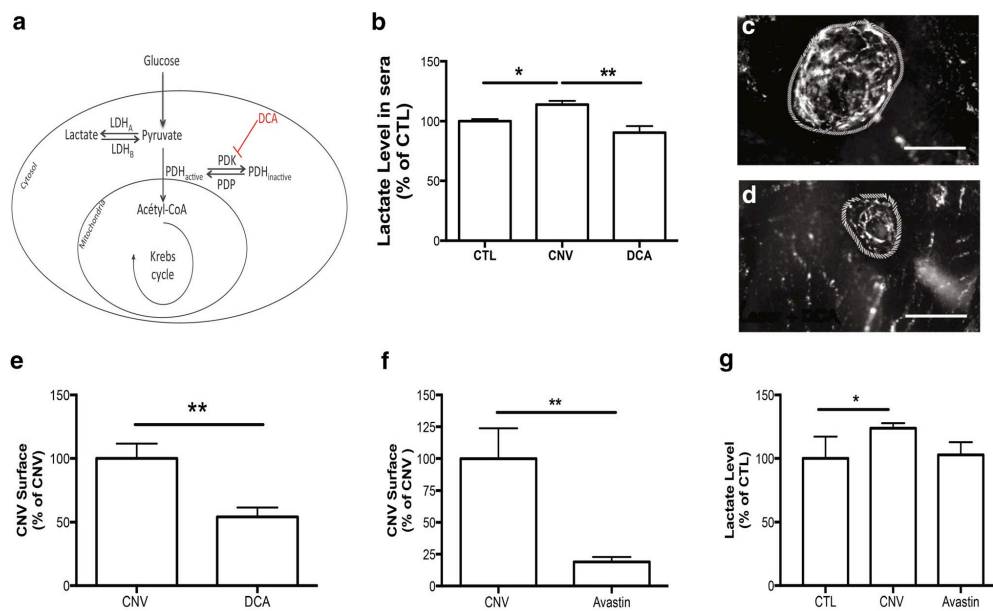
Figure 42. DCA treatment normalizes lactate level and reduces CNV surface. (a) Schematic overview of how DCA can impact lactate levels. DCA inhibits mitochondrial PDK activity, thereby maintaining PDH in its (unphosphorylated) active form and facilitating the decarboxylation of pyruvate to acetyl-CoA. As the flux of pyruvate is accelerated, the equilibrium between lactate and pyruvate is unbalanced towards pyruvate. (b–e) Mice subjected to laser-induced CNV were treated or not with DCA (3 mg DCA/day/mouse) (n $\geq$ 5 mice/group). Untreated mice (laser) were used as controls. bNMR dosage of serum lactate level at day 7 after DCA treatment (n $\geq$ 5 mice/group). The results are expressed as the percentage of laser-induced mice without treatment. *P < 0.05; **P < 0.01. Error bars correspond to SEM. Flat-mounted choroid of (c) an untreated mouse and (d) a DCA-treated mouse at day 7: Dashed lines delineate the lesion. Scale bars, 100 μm. (e) Quantification of CNV after DCA treatment at day 7 (n $\geq$ 6 mice/group, n = 26–27 laser impacts/group). The results are expressed as the percentage of laser-induced mice without treatment. **P < 0.01. Error bars correspond to SEM. (f) Quantification of CNV after Avastin treatment at day 7 (n $\geq$ 6 mice/group, n = 26–27 laser impacts/group). The results are expressed as the percentage of laser-induced mice without treatment. (g) NMR dosage of serum lactate level at day 7 after Avastin treatment; the results are expressed as the percentage of laser-induced mice without treatment (n $\geq$ 6 mice/group). **P < 0.01. Error bars correspond to SEM[225].

To determine the cellular source of lactate present at day 5, we measured by NMR lactate concentrations in different organs (eye, bone marrow, spleen and liver) at day 3 and 5 post-laser burn (**Figure 43**). At day 3, lactate increase is only noticed locally in the injured eye and no significant modification was detected neither in serum, bone marrow, spleen, or liver. Interestingly, increased lactate level was measured in the bone marrow at day 5 indicating a two phases release of this metabolite.

Indeed, these data reveal a two-site production of lactate at two-time point: early and locally in the eye, followed by a systemic release in blood circulation promoted by bone marrow derived inflammatory cells. Lactate is known to be implicated in angiogenesis and inflammation processes[227,228], two underlying causes of AMD. Moreover, in this study we provided evidence that lactate control the recruitment of macrophage, which are recognized as key cellular regulator of CNV formation[229,230]. Indeed, lactate was found to modulate the macrophage recruitment and influence the M1/M2 balance in favor of M2 macrophage by converting M1 macrophage to M2-like type (**Figure 44**). M2 macrophages are considered as pro-angiogenic and their implication in different pathologies including cancer and AMD is well documented[231,232].

Consistently, our finding is in line with a previous study implicating lactate in the M2 macrophage polarization in the context of cancer[233]. Altogether, our data highlight the interest of monitoring lactate level to evaluate disease progression and that targeting lactate metabolism could not only be promising for cancer[234], but also for AMD.
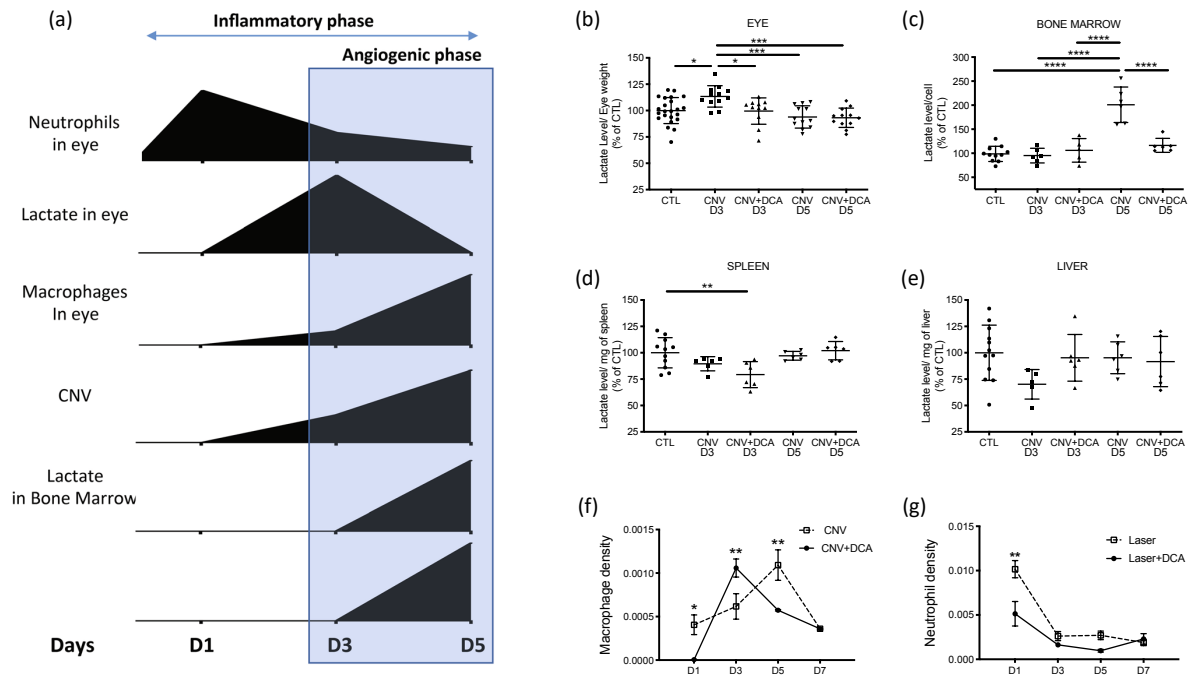
Figure 43. Lactate mediates the recruitment of inflammatory cells during CNV. Mice subjected to laser-induced CNV were treated or not with DCA (3 mg DCA/day/mouse). Summary diagrams of the kinetics formation of inflammation, CNV formation, and lactate levels in eye and sera (a). NMR dosage of the lactate level at day 3 and day 5 in eye (b), bone marrow (c), spleen (d), and liver (e) (n ≥ 5 mice/group). *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001; data are expressed as the mean ± SD. Macrophage/neutrophil density defined as the volume occupied by cells divided by the total laser impact volume with or without treatment with DCA (f, g)[225].
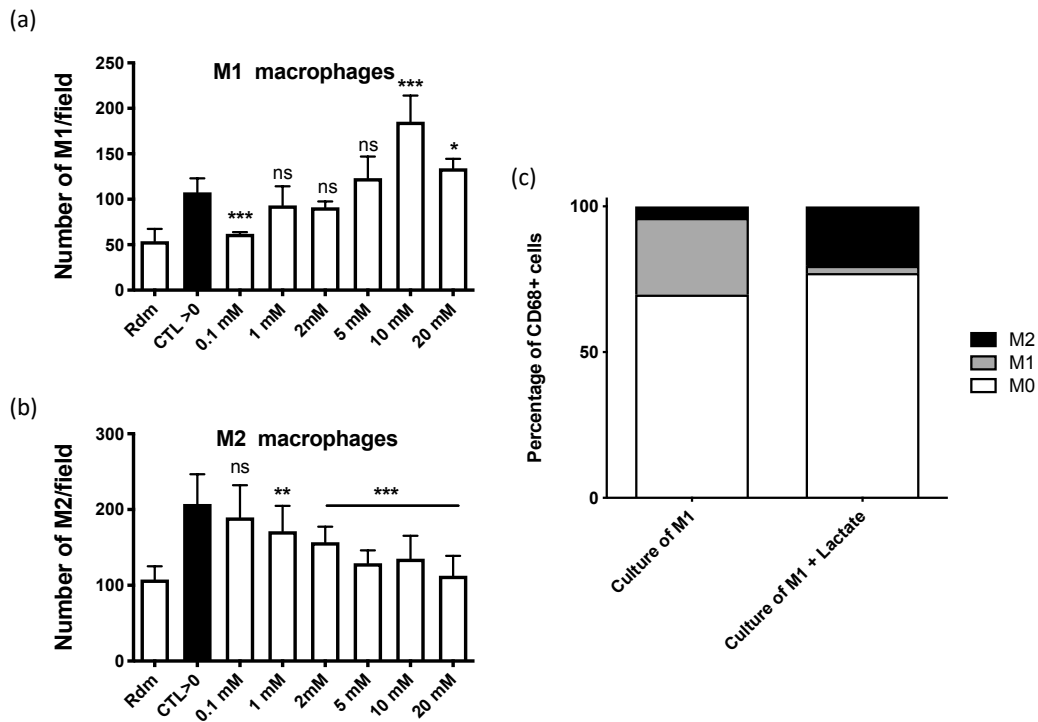
Figure 44. Boyden chamber assay showing the differential migration of human macrophages polarized in vitro into the M1 (a) or M2 (b) subtype (n = 6 wells/condition). M1 cells were attracted by high and M2 by low-lactate concentrations. Rdm corresponds to random cell migration in the absence of chemoattractant and CTL>O to cell migration in response to BSA-containing medium (positive control). Cultures of M1 macrophages were treated or not with 10 mM lactate for 48 h (b). The results obtained after phenotypic analysis showed that lactate induced the conversion of M1 into M2 macrophages (n = 5). *P < 0.05; **P < 0.01; ***P < 0.001; ns, nonsignificant. Data are expressed as the mean ± SEM (i, n, o) or mean ± SD (p, q)[225].

Of great interest is our finding that lactate levels and lipoprotein profiles are normalized and modified upon anti-VEGF and PDK inhibitor treatments. This paves the way for developing a personalized therapeutic approach of patients with late AMD under anti-angiogenic treatment. Indeed, clinicians are lacking tools for rationale decision-making and usually are following general guidelines. These biomarkers could help for patients' stratifications, evaluations of treatment responses and/or to personalize the therapeutic interventions for nAMD by adjusting the frequencies of anti-VEGF injection. In addition, the normalization of lactate level trough oral administration of a PDK modulator is representing a new therapeutic opportunity and/or a complementary treatment to reduce CNV progression.

This work put in light innovative concept of PDK/lactate axis as a functional, traceable, and targetable mediator of CNV and nAMD evolution. Metabolite profiling and lactate/lipoprotein profiles monitoring during anti-angiogenic therapy are of interest for the follow-up of nAMD patients and should be considered in future clinical studies to offer new tools to help clinicians to set up a more personalized approach on their patients. If results about lipoprotein must be refined to assess which lipoprotein classes are involved in CNV processes, a follow-up cohort of nAMD patients is mandatory to evaluate the usefulness of lactate measurement. Indeed, with such cohort we could go deeper in our observation and find out if following lactate levels and lipoprotein profile changes could be helpful for patient's status determination and evaluation of treatment responses. Associating these changes among the metabolome with morphological changes among the retina of nAMD patients under treatment represent a great opportunity to confront clinical NMR metabolomics to real patients' health care.

Chapter 2:
Aims of the thesis

2 Aims of the thesis

Age-related macular degeneration (AMD) is the leading cause of blindness among the elderly population in developed countries. 90% of all vision loss due to AMD result from the exudative form of this pathology, which is characterized by a choroidal neovascularization (CNV). Currently, diagnosis of AMD relies on ophthalmologic exams and treatments of the exudative form are based on the use of anti-angiogenic drug targeting vascular endothelial growth factors. Despite these advances, several clinical challenges must be overcome. Among those, the identification of biomarkers that could allow to refine patient stratification and assess the severity of the pathology, to follow disease progression and evaluate responses to treatment are mandatory. Our previous work has demonstrated the potential of metabolomics to deep insight n-AMD and opened the way for the putative use of this new approach in the clinics and in a personalized approach of this pathology.

Thus, we decide to continue our investigation on n-AMD with the aims using NMR-based metabolomics in order to refine our first results and to explore how these results could be use in a preventive and/or patient's follow-up.
Our works will particularly focus on 3 axes:

i. As previous investigations have indicated that changes among lipoprotein profiles could be correlated to the most advanced stage of the disease (neovascular AMD), we aim to provide more precise information about these changes by a specific and innovative analysis of the lipoproteins NMR data coming from our previous cohort. To achieve this goal, it is necessary to set up an original methodology able to extract lipoproteins' information from old NMR data. Through this approach we aim to provide new insights for AMD patient stratification and evaluation of patients' status.

ii. To evaluate changes among the metabolome linked to the evolution of nAMD over treatments and improve patients' follow-up, we collected plasma samples of patients over a period of 2 years. All patients were diagnosed with pathologic nAMD and undergo anti-VEGF based therapy. During this period clinical data regarding nAMD status of patients were collected as well as several clinical parameters (lactate levels, HDL and LDL-cholesterol levels) and blood samples for NMR metabolomics analysis. These data were used to evaluate the ability of our previously identified

biomarkers for improving patient's healthcare. Furthermore, this work represents an interesting opportunity to explore the applicability of metabolomics in real clinical practices.

iii.   To study metabolic changes that occur during aging and lead to pathologic AMD, plasma samples were collected from 471 peoples with no sign of AMD and followed during a period close to 8 years. During this period, clinical data were recorded regarding their AMD incidence. The incidence of early or advanced AMD was defined by the progression of healthy individuals at the time of enrolment and evolving trough early or advanced AMD during the follow-up. This cohort will help to unravel the metabolic changes that occurs during the development and progression of pathologic AMD. This work will benefit of a close collaboration with INSERM of Bordeaux (Population Health Research Center Unité INSERM U1219) for the statistical analysis and will aim to improve diagnosis and prevention strategies for early AMD.

In this work we will apply proton NMR based metabolomics approach to study the development/progression of Age-related Macular Degeneration and improve patients' monitoring at various levels. From early diagnostic to the evaluation of treatment responses, we aim to provide useful information and demonstrate that metabolomics can play a role in daily clinical practices.

Chapter 3:
Lipoprotein profile for AMD patients' classification

3 Lipoprotein profile for AMD patients' classification

3.1 Introduction

In a previous study, we used proton nuclear magnetic resonance to analyze sera samples collected from a human cohort composed of: (i) bleeder AMD patient that are under treatment, (ii) non-bleeder AMD patient and (iii) Healthy volunteers. In parallel, this methodology was also applied to our experiment laser-induced choroidal neovascularization mice model, a model that mimic the effect of advanced stage of AMD[226]. NMR metabolomics analysis allowed the differentiation between control and AMD patients and between laser-induced and control mice group. Moreover, this study led to the emergence of different putative biomarkers. Among these, lipoprotein signatures are of particular interest since we found an association between lipoprotein levels and active phases of the disease in both human and mice model. In addition, some studies associated high density lipoprotein cholesterol and oxidized low-density lipoprotein with early AMD[216,235] and para-inflammatory process inducing pathologic AMD respectively[236]. Taking together, these information suggest that investigating lipoproteins profile could be a turning point in the comprehension of the pathologic process that occurs during the apparition and/or the development of pathologic CNV process.

Lipoproteins are amphiphilic macromolecular complexes composed of various lipids and proteins that able lipids to circulate through the human body. The core is composed mainly of non-polar lipids (triacylglycerol and cholesterol esters) while on the surface are found polar lipids (free cholesterol and phospholipids) and the protein components named apolipoprotein. This group of proteins is highly diversified and will define the structure, the functionality and how a given lipoprotein class will be metabolized and will interact in the liver and with the different peripherical tissues. They are defined regarding their density and divided into five main classes: chylomicrons; very low density lipoproteins (VLDL); intermediate density lipoproteins (IDL); low density lipoproteins (LDL), and high density lipoproteins (HDL)[237]. For each classis, several subclasses exist but the number is depending on the separation/characterization methods and, since no discrete value corresponding to a given class, lipoprotein density is considered as a continuum and going deeper in the separation between the main classes have poor clinical interest. The core of lipoproteins is composed of hydrophobic lipids such as esterified cholesterol, triglycerides, and some cholesterol, while in surface are founds amphiphilic species as apolipoproteins, phospholipids, and free cholesterol.

Compositions and repartitions of lipids is characteristic of a given class and thus each class exhibit proper physico-chemical properties (**Figure 45**).

Lipids within the lipoprotein structure are responsible of different resonance peaks from glycerol backbone, fatty acyl chain and the polar moiety, choline for example. But if the resonances peaks of these are poorly informative when aim to discriminate the different lipoprotein classes, the methyl and methylene group of various lipid component presents at low frequencies have been found to exhibits some interesting characteristics (**Figure 45**). Indeed, ^{1}H-NMR signals from the methyl (-CH$_3$) and methylene (-CH$_2$) groups in the hydrocarbon chains of lipids moieties shift to lower frequencies for higher density lipoprotein classes (HDL). This effect is due to differences of radii of the particles. Indeed, ratio of the core to surface lipids differs with the particle diameters and as lipids from both shells have different magnetic susceptibilities, a size-related frequency shift is observed[238]. This shift is then varying continuously with the decreasing radii of particles from VLDL to HDL lipoproteins (**Figure 45**).
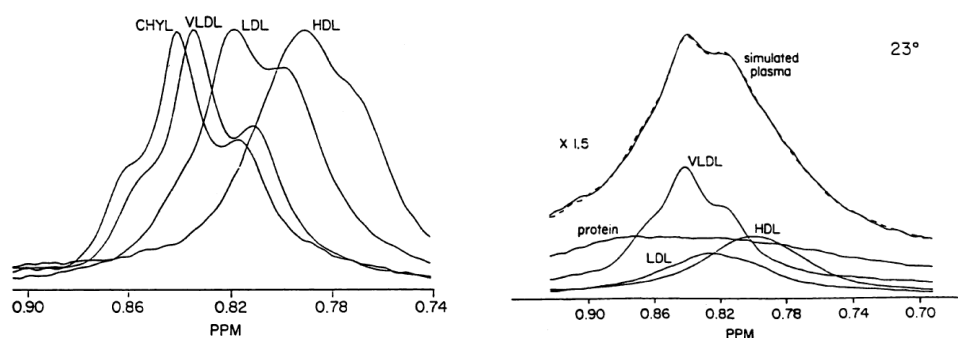
Figure 45. ^{1}H-NMR spectra of pure lipoprotein fractions on the left and on the right how they overlapped to produce the final signal recovered on ^{1}H-NMR analysis of blood samples (adapted from[239]).

To get deeper in the analysis of lipoprotein changes already observed, we need a new tool that could be used on NMR data readily available. Indeed, NMR investigation of lipoprotein content from blood derived samples is well documented in the literature. These methods are mostly based on curves fitting approach that fit the signals coming from isolated lipoprotein moieties analyzed in the same conditions than the samples of interest[240–243]. Other powerful methods using diffusion edited NMR pulses sequence, such as DOSY experiments, aim to resolve peaks of several lipoprotein subclasses based on their diffusion coefficient and allows their quantification[244]. Nevertheless, even though these methods are robust and provide the best representation of lipoprotein content of blood samples, they require to analyze again all samples or use the same NMR instrumentation. This was not possible in our case and therefore we needed a methodology that used already processed NMR data.

In this work we taken advantage of the differences of chemical shift that exist between the different lipoprotein classes. Here we used a peak-picking method that allows the representation of the lipoprotein profile on NMR spectra of sera samples (**Figure 46**). This simple method gives a representation of the shape of the total lipoprotein signals and allows the comparison of profile between individuals. Indeed, the whole lipoprotein signals was fractionated into 4 and 5 signals for human and mice samples respectively. This decomposition is based on the evident maxima of the overlapped signals coming from the various lipoprotein classes. Using this strategy, we were able to differentiate VLDL rich samples and samples exhibiting a higher proportion of HDL moieties. This methodology allowed us to reuse our acquired NMR data providing a valuable information about our human and mice cohorts. All methods and results described in this chapter were published in the Journal of Molecular Medicine[225].
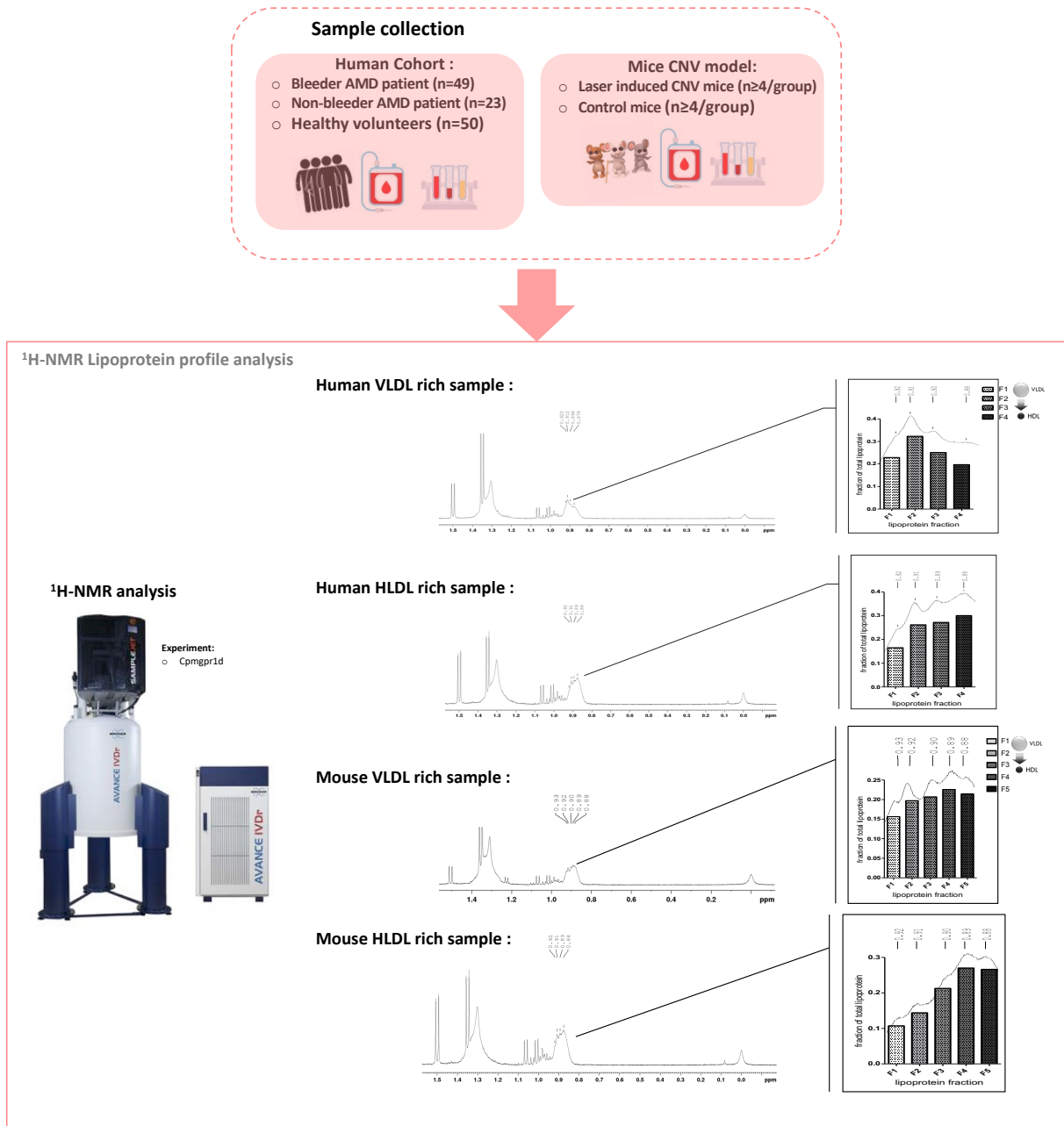
Figure 46. Representation of the cohort used for lipoprotein profile investigation and examples of rich VLDL and HDL samples for human and mice individuals. This representation shows the fragmentation of the overlapped peak of lipoprotein moieties and how we used these data to represent profiles changes among samples.

3.2 Dataset

The dataset used for this section is composed of 122 individuals (49 bleeders, 23 non-bleeders AMD patients and 50 healthy volunteers) for which the NMR lipoprotein profile was evaluated through a peak picking approach (see material and methods). Lipoprotein profile measurements were made on previously acquired NMR data and are representing the 4 main classes of lipoproteins (VLDL, IDL, LDL and HDL). Thus, each human sample are described by 4 values representing each class as the part of the whole lipoprotein contribution to the NMR signal. For mice samples (n≥4/group) coming from our experimental CNV model, the whole profile was subdivided into 5 classes from VLDL to HDL. These data allow us to statically analyze variations among lipoprotein profile of human and mice blood derived samples in different conditions.

3.3 Lipoproteins ^{1}H-NMR profile for nAMD

As described in the previous section, the NMR signal of the main lipoproteins classes was separated into 4 fractions, namely, F1 to F4 representing the shift of the signal from a fraction rich in VLDL (F1), to IDL (F2), LDL (F3) and HLDL (F4). We compared changes among the lipoprotein spectral zones in the 3 patients groups based on their AMD status: healthy, inactive, and active patients. During this analysis we observed a shift toward a VLDL rich profile for patients with active nAMD. Both VLDL and IDL (F1 and F2) proportions were higher, while LDL and HDL (F3 and F4) fractions were reduced in the active group compared to inactive and healthy ones (**Figure 47**).
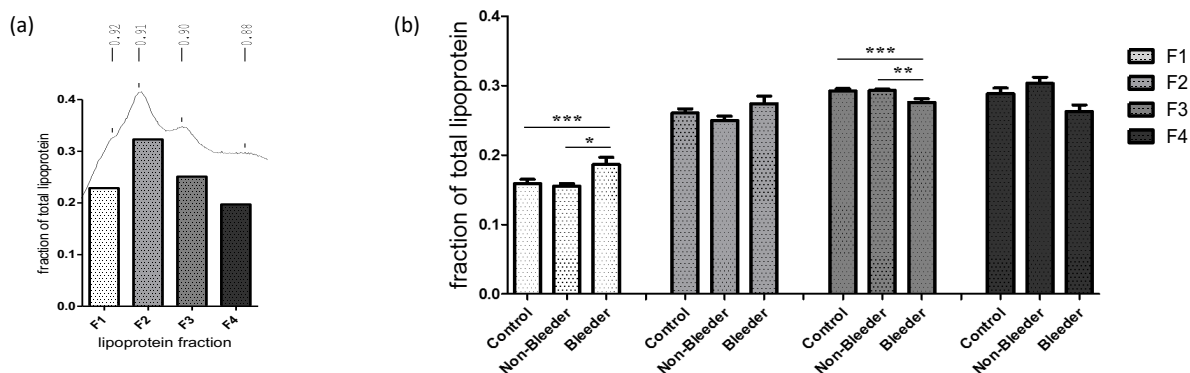
Figure 47. Enlarged view of the lipoproteins NMR spectral zone showing the chemical shift corresponding to the maximum intensity of the signal of the 4 lipoprotein classes (a) and, modification of the lipoprotein profile during CNV development (b). Fraction 1 is mainly composed of VLDL, while fraction 4 is mainly composed of HDL. *P < 0.05; **P < 0.01; ***P < 0.001. Error bars are SEM.

Interestingly, when the same strategies were used on spectra coming from NMR analysis of mice CNV model samples, the same shift of lipoprotein profile was observed (**Figure 48**). Indeed, higher proportion of VLDL rich fraction was noticed in serum sample of mice at day 7 post laser burn compared to control mice in kinetics CNV formation analysis. Of great interest is that these changes in lipoprotein profile are normalized upon anti-VEGF/PDK inhibitors treatment indicating that these modifications are related to the CNV status.
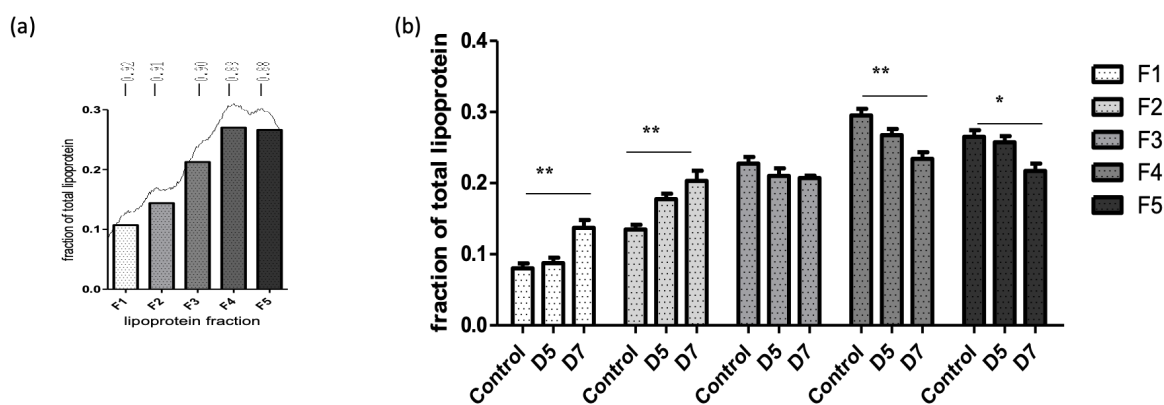
Figure 48. Enlarged view of the lipoprotein NMR spectral zone showing the chemical shift corresponding to the maximum intensity of the signal of the 5 lipoprotein classes (a). Modification of the lipoprotein profile during CNV development (b). Fraction 1 is mainly composed of VLDL while fraction 5 is mainly composed of HDL (n ≥ 5 mice/group). *P < 0.05; **P < 0.01. Error bars are SEM.

These results refine the previous observation in which the spectral zone corresponding to lipoprotein content was identified through discriminant multivariate statistical analysis (see section **1.6**). Shift in the lipoprotein profile toward VLDL rich balance was noted during the active phase of the pathology either in human or mice CNV model cohort. The good correlation between the two metabolomic studies conducted support our concept to use lipoprotein profile changes as indicator of AMD status. Moreover, effect of treatment against pathologic CNV is also reflected by our approach since the switch to a VLDL rich profile is decreased for CNV mice that were treated with anti-VEGF or PDK inhibitors (DCA) compounds. This clearly shows that lipoprotein levels are modified through CNV events and that monitoring changes of this maker during AMD follow-up could help to refine the diagnostic regarding the evolution of the pathology. Together with lactate level, lipoprotein profile could offer a new tool to help clinicians to personalize the therapeutic approach and rationalize patient's nAMD management.

3.4 Conclusion and prospect

In this chapter we refine information previously reported concerning the impact of nAMD on the lipoprotein profile. We demonstrate that active phase of neovascular AMD is associated with a shift in the lipoprotein profile toward rich VLDL balance. Monitoring such modification of the NMR lipoprotein signals could be interesting for assessing patient's CNV status and figure out when treatment must be applied. Moreover, this pattern could be used to predict patients' responses to treatment and set up a personalized therapeutic strategy that fit better to the need of the individuals.

The usefulness of our discovered biomarker must be evaluated in a real case patient's follow-up study. Indeed, switching from case-control study to real life clinical practices represents a turning point that will definitively determine the ability of our methodology to improve nAMD patients' health care. Following lactate levels, monitoring change among lipoprotein profile and NMR metabolome of such patients is mandatory. By this, we aim to evaluate if such information could be correlated with retinal morphologic changes due to disease progression and treatment responses.

3.5 Material and Methods

Patients' selection and samples collection

The study population consisted of unrelated European- Caucasian individuals (>59 years old) affected with nAMD (n = 72) and healthy volunteers without signs of macular disease or a known family history of AMD (n = 50). Patients with AMD and volunteers were not matched for age or sex. Trained ophthalmologists examined all patients with nAMD and divided them into clinically active or inactive subgroups depending on the presence or absence of intraretinal or sub-retinal exudative fluids as assessed by OCT, respectively. A complete medical history of each patient was obtained by using a standardized questionnaire (i.e., lifestyle, pathologies, treatments, BMI, etc.). For all the participants, peripheral blood was collected after ophthalmological exams in: (1) K2 EDTA blood collection tubes for blood cell counts, leucocyte differential, and C-reactive protein (CRP) measurement using CRPLX kit (Cobas®, Roche/Hitachi); (2) serum-separating tubes (Greiner); and (3) sodium fluoride/oxalate tubes (Greiner) for the lactate quantification. Serum samples were routinely taken in the morning with a fasting period of at least 2 h. Samples were treated according to clinical standard processes and stored at − 80 °C after sampling until the metabolomics analysis. Clinical biology analyses

(lactic acid, CRP, red and white blood cells) were also performed on the sera to evaluate patient inflammatory status.

NMR metabolomics analysis

All samples were recorded at 298 K on a Bruker Avance spectrometer operating at 500.13 MHz for the proton signal acquisition. The instrument was equipped with a 5-mm TCI cryoprobe with a Z-gradient. Maleic acid was used as an internal standard for quantification and trimethylsilyl-3-propionic acid-*d4* (TMSP) was used for the zero calibration. Human sera (500 μl) were mixed with D_2O phosphate buffer (100 μl) (0.1 M, pH 7.4), a 35 mM solution of maleic acid (100 μl) (Aldrich, Germany), and TMSP (30 μl) (sodium trimethylsilyl[2,2,3,3-D4] propionate) in D_2O (10 mg/ml). Mouse sera (200 μl) were mixed with D_2O phosphate buffer (400 μl) (0.1 M, pH 7.4), 35 mM solution of maleic acid (100 μl), and TMSP (30 μl) in D2O (10 mg/ml). ^{1}H-NMR spectra were acquired using a 1D-CPMG (Carr- Purcell-Meiboom-Gill) relaxation-editing sequence with presaturation for serum samples. The CPMG experiment used a RD-90-(t-180-t)n-sequence with a relaxation delay (RD) of 2 s, a spin echo delay (t) of 400 ms, and the number of loops (n) equal to 80. The water suppression pulse was placed during the RD. The number of transients was typically 32. The acquisition time was set to 3.982555 s, and a quantity of four dummy scans was chosen. Data were processed with the Bruker Topspin 3.1 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.

Lipoproteins profile ^{1}H-NMR analysis

Estimation of lipoproteins profile modification among spectra collected from patients and mice was performed by using peak picking methods that compare intensities between the different fractions. In human and mouse blood NMR spectra, the global signal of lipoproteins between 0.80 and 0.95 ppm is due to an overlap of several peaks that could be linked to the main classes of lipoproteins: very low-density lipoproteins (VLDL), low-density lipoproteins (LDL), intermediary density lipoproteins (IDL), high-density lipoproteins (HDL), and chylomicron. Then, it could be decomposed into distinguishable signals corresponding to these different classes or to a mixture of 2 classes (4 in humans and 5 in mice).

To evaluate the proportion of each lipoproteins fraction in the samples, a method based on normalized peak intensity calculation was developed. Then, for each class, we determined the chemical shift corresponding to the peak of signal intensity. For human samples, 4 fractions are selected (F1 = 0.92 ppm (mainly VLDL), F2 = 0.91 ppm, F3 = 0.89 ppm, and F4 = 0.88 ppm (mainly HDL). For mouse samples, 5 fractions are selected (F1 = 0.93 ppm (mainly VLDL), F2 = 0.92 ppm, F3 = 0.90 ppm, F4 = 0.89 ppm, and F5 = 0.88 (mainly HDL)). For each sample, the signal intensity at these different chemical shifts was measured and then normalized to the total intensities of all fractions to reduce the impact of the global lipoproteins concentrations that could differ between samples. Therefore, the obtained values represent a fraction of the total signal. This method allows the comparison between lipoprotein profiles issued from the spectra of blood samples collected from patients with AMD, control subjects, and induced/ non-induced mice (**Figure 49**).
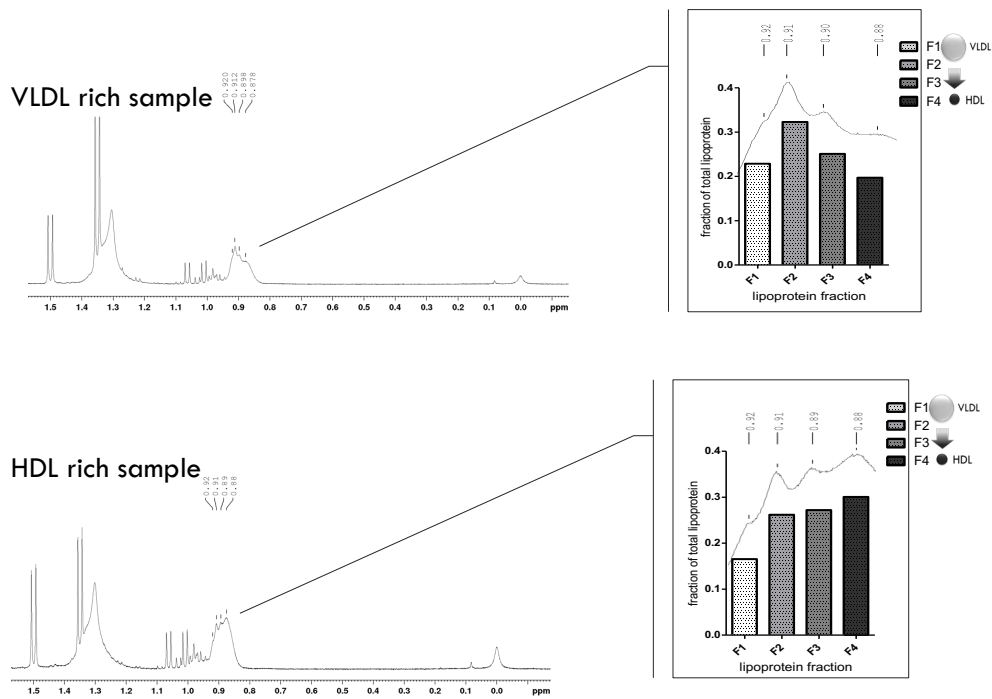
Figure 49. Human lipoproteins profile evaluation trough peak-picking approach of VLDL rich sample (above) and HDL rich sample (bottom). This methodology allows the direct comparison of profile from already processed NMR data.

Statistical analysis

Univariate statistical analyses were performed on each lipoprotein fraction (see above "Lipoproteins profile ^{1}H-NMR analysis") of both mouse and human serum spectra using GraphPad Prism version 7.0. A nonparametric Kruskal-Wallis test with Dunn's multiple comparison was used to compare Fx controls with Fx AMD patients and Fx mice control versus CNV mice (J5 and J7 post-laser) with Fx being the lipoprotein fraction number 1 to 4 for humans and 1 to 5 for mice.

Chapter 4:
NMR metabolomics and nAMD patient's follow-up

4   NMR metabolomics and nAMD patient's follow-up

4.1   Introduction

In the previous study we highlighted a functional role of lactate in AMD and in CNV progression that could be used for patient's stratifications, follow-up, and evaluation of treatment responses. Moreover, changes among the lipoprotein profile toward an increase of VLDL moieties were found to be associated with CNV development events and provide a new tool to monitor the evolution of the pathology. Taken together, these information paved the way to a new patient follow-up and personalized medicine approach for AMD[225]. Indeed, by following lactate concentration's evolution and changes among the lipoprotein balance, we aimed to better characterize the patient regarding its own evolution through the pathology and provide information that leads to a more rationalized treatment.

Indeed, as these tools seems able to assess the CNV status of AMD patient, this could be the key for clinical decision of intra-vitreal injections of anti-VEGF. Adaptative treatment regimen are proved to improve AMD healthcare but failed in practice due to a lack of "handleably" markers of the pathology. Considering this, we are convinced that our biomarkers could fill the gap between scientific evidence and clinical practices[65].

In this chapter, we've followed 32 AMD patients under anti-VEGF treatments over 2 years. By collecting blood samples for NMR analysis and several clinical data and information, we aimed to use our previously discovered biomarkers in a real-life patients' management context. Following the evolution of these marker through the evolution of these patients will assess their usefulness and determine if they can be used for a more rationalized healthcare of AMD.

4.2   Datasets

For this study several datasets on the same cohort were generated in order to answer specific questions. The cohort is composed of 32 patients aged of 62 y/o or older, all diagnosed for their nAMD and followed over 2 years of treatment. Over the 269 time points (an average of 9.6 visits/patient), several clinical parameters from serum analysis, retina measurements and plasma samples for NMR analysis were collected. Hence for each visit the measured parameters included: lactate, glucose, HDL cholesterol and LDL cholesterol levels (calculated), and pH measurements. NMR analyses were performed on plasma samples to generate three distinct way to analyze metabolome related changes: (1) an untargeted metabolic profile of patients composed of hundreds of buckets of the NMR metabolic signature; (2) a targeted metabolic profile that results from the quantification of 60 known metabolites using a dedicated methodology; finally, (3) an NMR lipoproteins profiles (VLDL, IDL, LDL and HDL fractions) obtained through peak picking methods described on the later chapter. The evolution of patients regarding their pathology is represented by clinician's comment about patient's status, evolution of visits schedule, clinician's decision of intra-vitreal injections and responses evaluation, and finally, several measurements of OCT imaging biomarkers of morphological changes among the retina: Intra Retinal cystoid Fluid (IRC), Sub Retinal Fluid (SRF), and Pigment Epithelium Detachment (PED). In particular, IRC is representative of the accumulation fluid leakage from pathologic CNV processes. Therefore, this measure will allows to assess if the studied patient is in active phase of the disease and to monitor its evolution through the treatment[65,74]. PED is also associated with bad prognostic and loss of visual acuity as this morphologic manifestation is a consequence of macular hemorrhagic events. On the other hand, SRF is one of the rare makers of the pathology correlated with positive outcomes. Indeed, these events are considered as a sign of healing retina and are associated with gain of visual acuity[74,81]. Finally, these markers will help to refine patient's status and provide the closest information about the evolution of the pathology.

At the end, we generated 5 datasets, 4 concerning metabolic information/profile of each patient (dataset 2-5) and one dataset focused on their pathology evolution (dataset 1). All these datasets will be used in distinct analysis that aims to put in relation variations of the metabolome with changes among health status of nAMD patients (**Figure 50**).
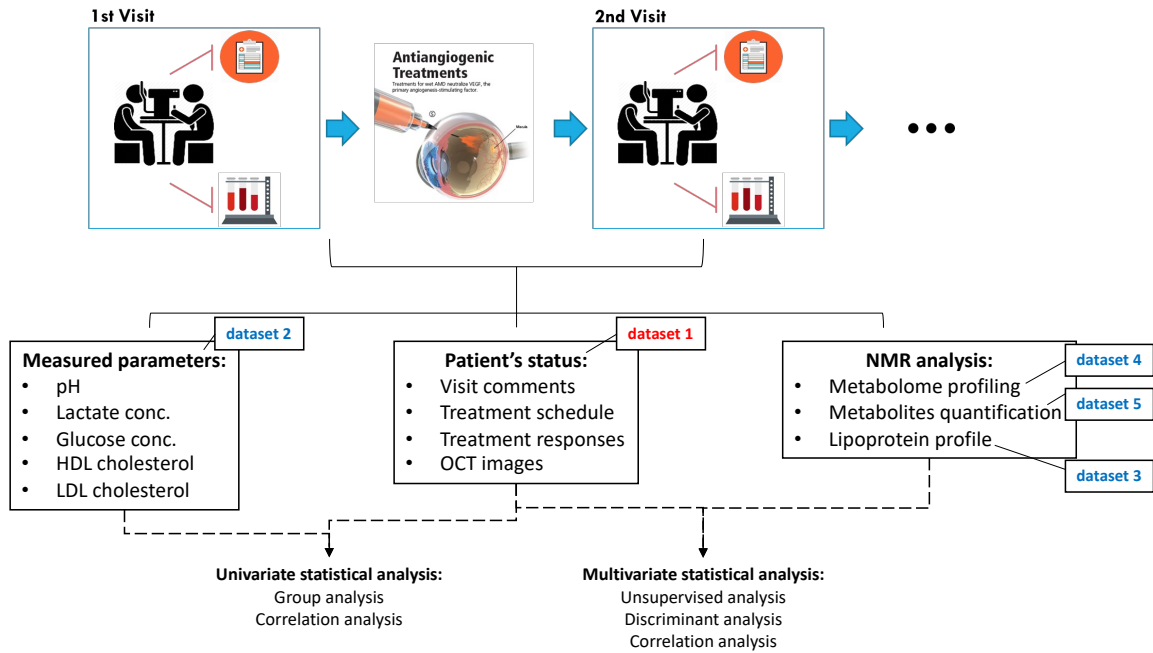
Figure 50. Schematic representation of the generated datasets and their uses. All the data were collected during a two-year follow-up during which each patient had an average of 9.6 visits.

4.3 Lactate and Lipoprotein profile for AMD management

In previous studies, we identified new conceptual insights into the pathogenesis and evolution of advanced AMD. Indeed, our data supported lactate as a new functional, traceable, and targetable mediator of pathogenic CNV process involved in AMD progression. Moreover, metabolites profiling and lactate/lipoprotein level monitoring suggested to be helpful for AMD management as they are linked to CNV events and are normalized upon anti-VEGF treatment. Therefore, our follow-up cohort of AMD patients under treatment represent a unique chance to evaluate the usefulness of our biomarkers (**Figure 51**). For this, we collected blood samples for nAMD patients over a period of two years for both NMR lipoprotein profile analysis (dataset 3) and for the quantification of circulating lactate and glucose levels, pH and HDL/LDL cholesterol measurement (dataset 2). At each visit, parameters regarding patients' status were collected (dataset 1): clinicians' comments; treatment schedule and; OCT biomarker images measurements (IRC: Intra Retinal Cystoid fluid, SRF: Sub Retinal Fluid and, PED: Pigment Epithelium Detachment).
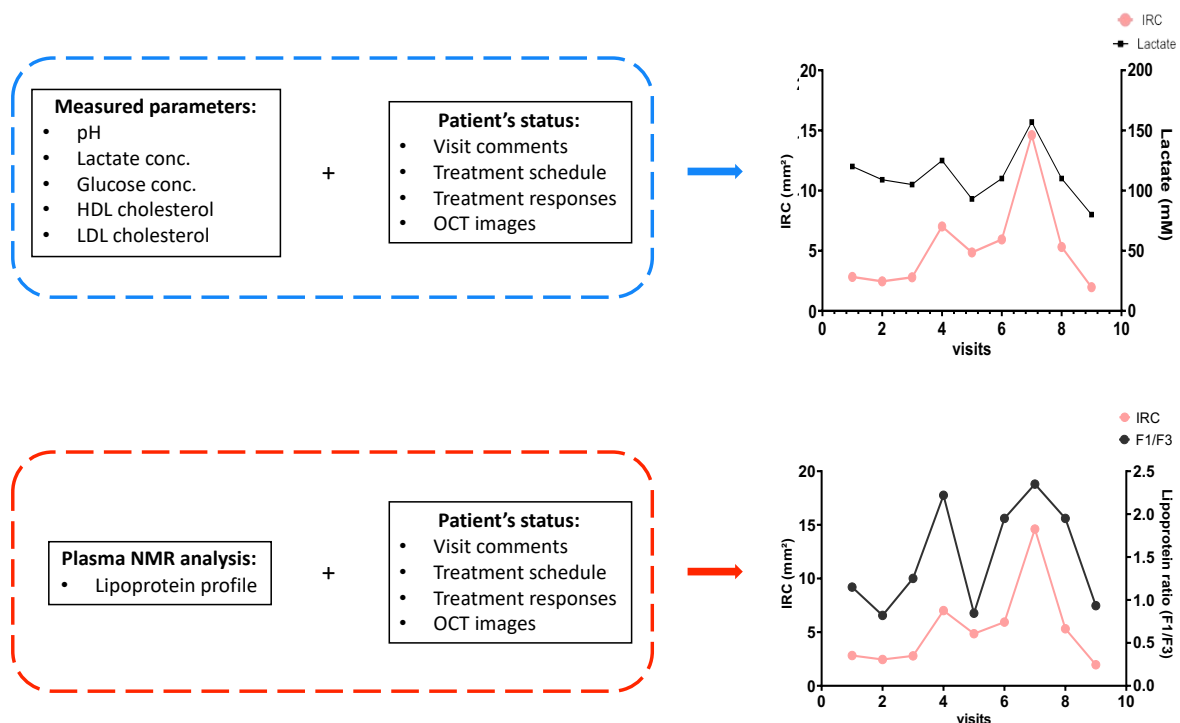
Figure 51. Schematic representation of use of available dataset and expected evolution of markers, these simulated data represent our expectation regarding the evolution of both patients 'status and related biomarkers. In this representation the measured values (measured clinical parameters and lipoprotein profile) are evolving concomitantly with IRC values. As IRC values are the measured parameters that give the best representation of the nAMD status as its directly related to CNV evolution, our previously biomarker should be able to follow its evolution.

In first, we aimed to follow the variations of lactate levels measured from clinical analysis and the lipoprotein profile from the NMR analysis of plasma samples. We will determine if these values can reflect the variation of the ones measured from OCT images (IRC, SRF and PED) and improve disease status characterization. Indeed, by using our tools we will try to determine whether the patient is responding well to the treatment or not. We expected to find a similarity between the evolution of lactate levels and/or the lipoprotein profile and OCT data collected; in particular, in an ideal world, trajectories of lactate values and VLDL rich lipoprotein fraction would followed the IRC surfaces as this marker is the best representation of CNV status.
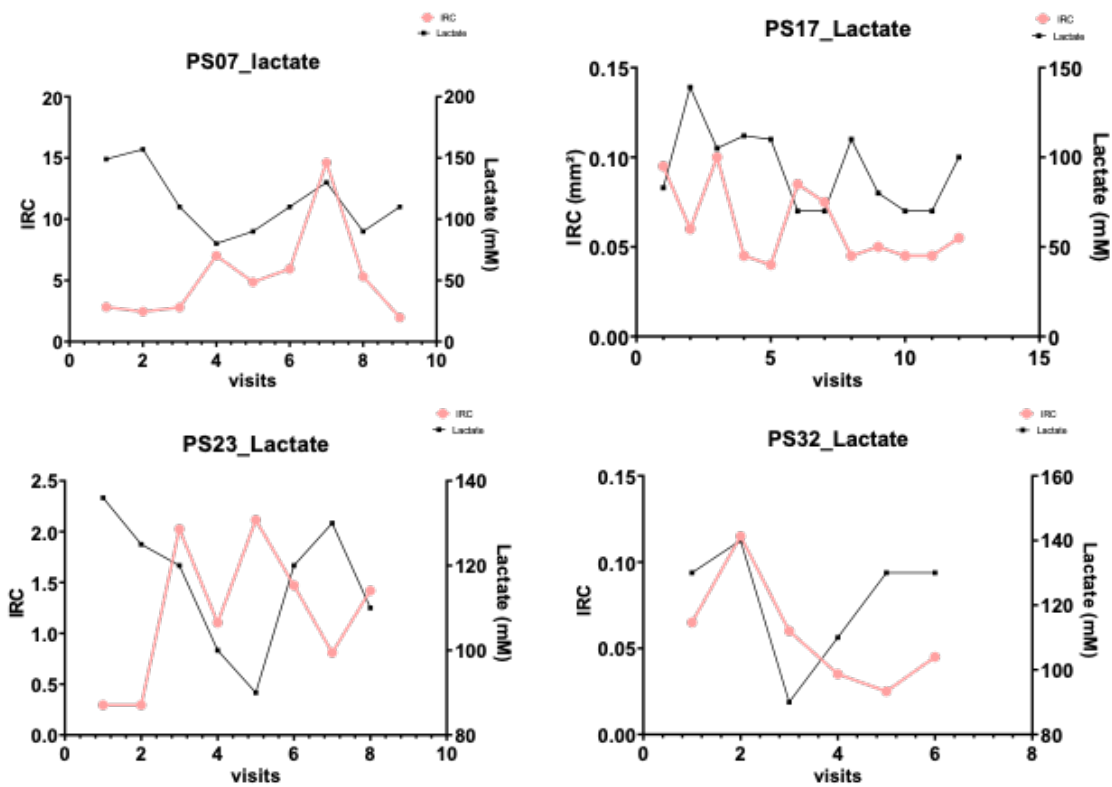
Figure 52. Representation of the evolution of lactate (in black) and IRC values (in red) for a selection of 4 patients.

From the analysis of the data displayed in **Figure 52** it's clear that no relevant information can be extracted. Even if lactate concentration seems to follow at some time points the IRC values for some patients, nothing consistent and systematic can be highlighted for the whole cohort and at all the time point. The same observation can be made from the use of lipoprotein VLDL rich fraction as depicted in **Figure 53** and when other OCT markers are used (data not shown).
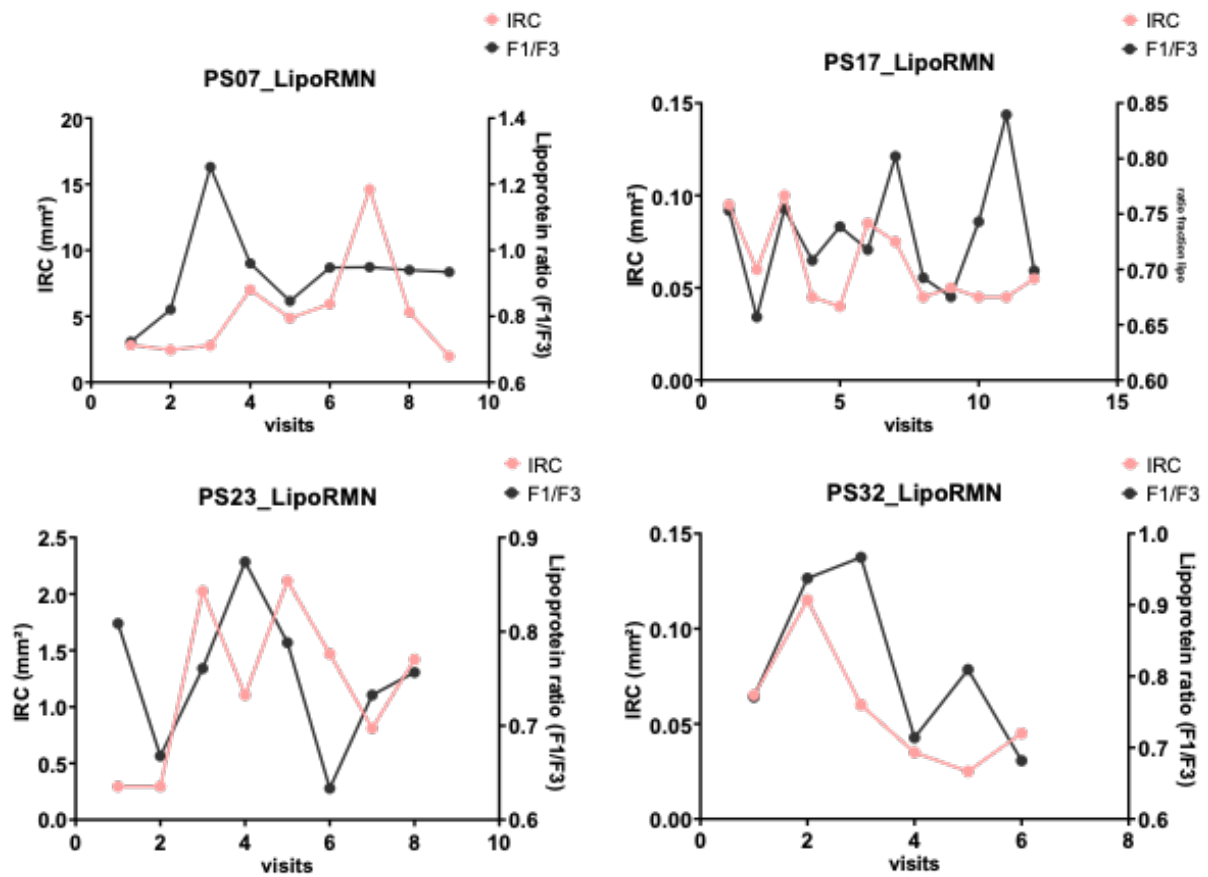
Figure 53. Representation of the evolution of VLDL rich lipoprotein fraction (in black) and IRC surface (in red) for some AMD patient.

From this first observational analysis, it clearly appears that our previously identified biomarkers are not evolving concomitantly with the used markers of the pathology. Therefore, looking at the trajectory of these values will not be helpful for patient monitoring as they aren't providing a good representation of dramatic CNV events that are occurring during patient's management. Maybe our previous discoveries cannot be translated into a less controlled environment than case control studies. Indeed, the time course of patient's follow-up in clinics add complexity to the cohort as all individuals doesn't evolve in the same way regarding their pathology. Moreover, as the disease may evolve differently from one individual to the other, the visit's schedule fixed by the clinician could not be optimized to give the best representation of the pathologic events. We need to change our way to use the data and try to find stronger markers of the pathology.

In first intention, we tried to analyze if any correlations existed between OCT values and the different clinical measurements that were made on the blood samples of patients (Lactate, pH, glucose, HDLc, LDLc, TG-fatty acids and cholesterol). For this we made a Pearson correlogram

in which the correlation between all values across the whole dataset are measured. If any correlation exist, bright red/blue circles for negative and positive correlations respectively appears; thus, strong correlations are easily spotted as big dark red or blue dots. As shown in **Figure 54a**, no correlation can be visualized between clinical parameters and OCT derived values. The same observation can be made when we tried to analyze the correlation between lipoproteins fractions, IRC, SRF and PED values (**Figure 54b**). This analysis only shows correlation among the datasets. For example, strong correlation can be seen between IRC and SRF values that suggests that these pathologic events are linked. This make sense as SRF fluids are sign of retina's responses to the CNV process that led to IRC fluid accumulation. In the same line, correlation between lipoproteins fraction measured by proton NMR spectroscopy also makes sense as negative correlation is observed between VLDL and HDL rich fraction.
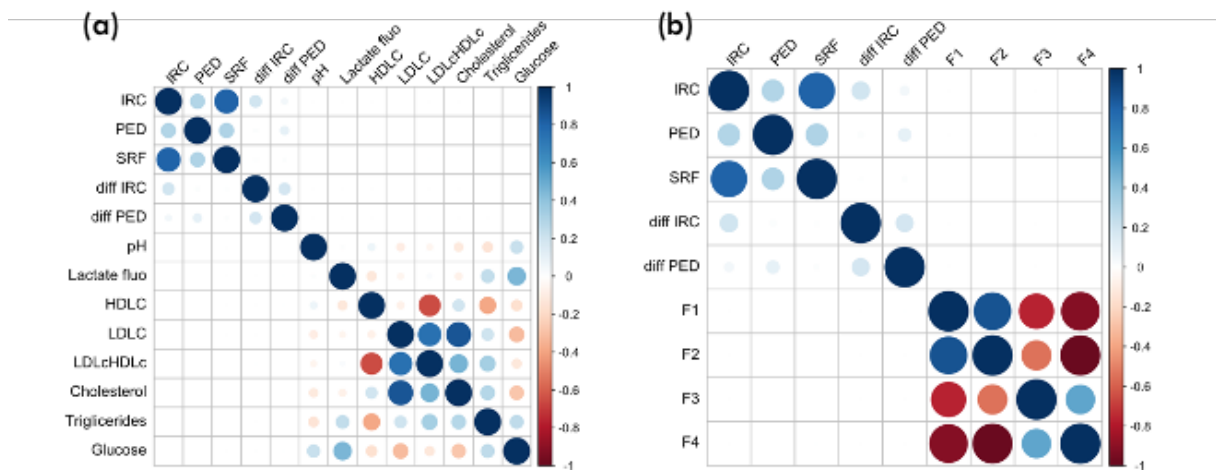
Figure 54. (a) Pearson correlogram of OCT data vs clinical measurement made on blood samples coming from AMD patients, (b) Pearson correlogram of OCT data vs lipoprotein fraction determined from NMR data of plasma sample collected from AMD patients. *Notes: "diff IRC" and "diff PED" are representing the difference between two consecutives visits for IRC and PED values respectively. This was made to represent patient's evolution. "LDLcHDLc" represents the ratio between the two lipoprotein moieties, made to rationalize the values between individuals and focus on lipoprotein balance.*

It appears that none of the measured parameters that we planned to use can explain the pathological OCT markers. These results forced us to reconsider the goal of our research that is no longer the use of previously identified biomarkers but rather to discover new one that can explain/predict the morphological changes among the retina of nAMD patients.

4.4 Biomarker discovery for nAMD management

Since our biomarkers failed in real life practice, this chapter will describe all the analysis that were performed to find new biomarkers that could be used in daily clinical practice. We will first discuss about the NMR analysis and give an overview of the final dataset that will be used. The analytical strategy was set up to answer to several questions than can be asked during a two-year follow-up of AMD patients. By this, we aimed to discover robust biomarkers that gives the best representation of patients regarding their pathology.

4.4.1 The "d20" problem

To capture changes among the metabolome that can be related to the studied pathology, we need the most robust data analysis workflow possible. For this reason, we decide to process our NMR data using PepsNMR software previously described for untargeted NMR metabolomics analysis. After pre-processing, bucketing and normalization of data, the first step is a previsualization of the whole dataset using PCA analysis. This unsupervised analysis will put in light some differences among the dataset and possible "natural" clusters.

Figure 55. PCA analysis of bucketed NMR spectra from AMD patients. In green samples analyzed with the Bruker Avance platform and in blue samples analyzed with the Bruker Neo platform. Strong effect of platform is visible.

Figure 55 shows the score plot of the PCA model generated on our dataset. A clear separation appeared between two groups of samples. Among the two groups are found samples from each individual and no specific distinction regarding their AMD status can explain this separation. It was found that this separation can be explained by changes in NMR sequences parameters. Indeed, during the two years of follow-up, we periodically analyzed our plasma samples in order to recover the information about the lipoprotein profile. The stability of NMR over time normally allows us to compare data acquired at different time points as the result will not depend on experimental condition such as the ones that can affect results in LC-MS analysis. This statement is true if analytical parameters of the NMR sequence remain the same; here our platform has been upgraded at the electronic level during the project.

After the upgrade of the NMR electronics (from Bruker Avance to Neo console), some parameters such as pulses duration, gradients values were re-calibrated and re-adapted to the new instrument. These small changes and adaptations led to differences in the applied sequence. Then, an unexpected scalar coupling problem was noted for strong coupling signals (ie. lactate signals). Scalar coupling issues are known as *J-modulation* and are current in NMR experiment that uses spin-echo. Herein, J-modulation is well documented and several strategies are available in the literature to avoid this effect[245]. This effect mainly arises from improper value of total echo time namely *d20* on Bruker sequences lexicon. Thus, this parameter can be modified to suppress J-modulation and increase spectral quality.

Another parameter that can be responsible of spectral quality is pulse duration. Indeed, well calibrated pulses are mandatory to avoid field inhomogeneity and therefore prevent coherence losses and preserve J-modulations[246]. Thus, to obtain useful spectra, the effect of the spin-echo sequence during cpmg experiment was therefore modified, and the total echo time (d20) has been adapted to the new pulses' durations. These changes had dramatic consequence in our dataset as depicted in **Figure 56**.

Indeed, this parameter has strong effect on our NMR dataset as the efficiency of the T2 filter is directly linked to this increment of time. Thus, signals of lipoproteins and lipids moieties will be upper in samples analyzed with shorter d20. This can be visualized on the spectra colored from VIP predictive score of the OPLS-DA analysis generated on the dataset by setting classes on the used value of d20 (**Figure 56**). This clearly shows which signals are impacted by the changes of this parameter.
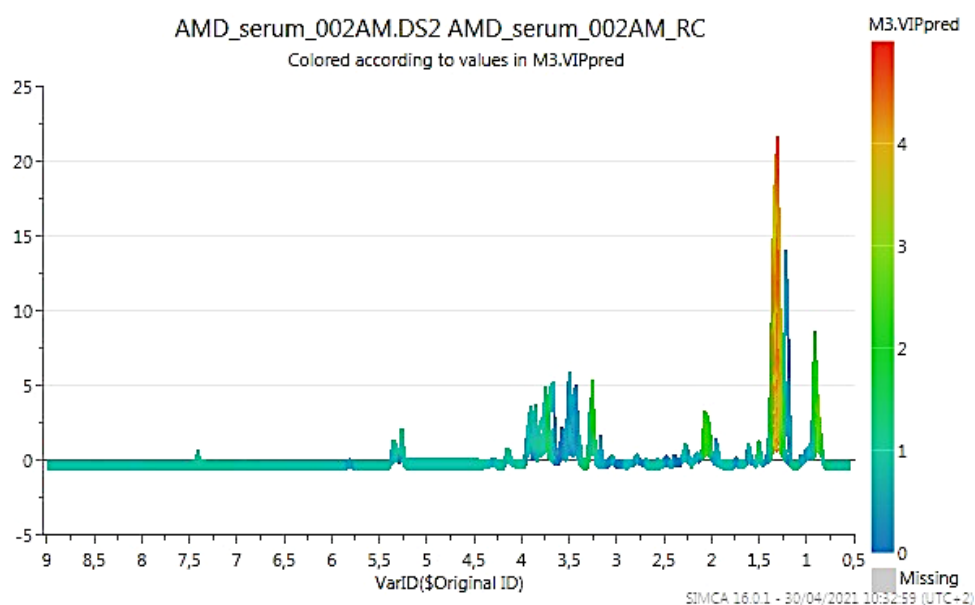
Figure 56. Predictive VIP representation plot from OLPS-DA analysis showing important NMR signals linked to d20 signature.

Fortunately, if this have a strong effect on the intensity of several signals, it has no impact on peak shape and is not depending on the size of molecules. Thus, if the intensity of signals coming from lipoproteins is modified, the effect is almost the same for all fractions. Therefore, as lipoprotein fractions are calculated as a part of whole, no discrimination can be made between the spectra acquired with one or another platform (**Figure 57**). Indeed, non-parametric t-test (generated using GraphPad Prism8®) are unable to discriminate group based on the analytical platform.

Anyway, if NMR lipoprotein profile measures will remain useful, this problem minimizes to zero the chances to perform NMR profiling analysis. For these reasons we decide to take a step back and analyze again our samples. Indeed, avoiding bias induced by instrumentation and batch effect is always challenging in longitudinal studies where samples are collected during a long period and sometimes analyzed at different time point. If different normalization methods allow to correct the batch effect, this relies on QC samples carefully designed regarding the

experimental design. Regarding our experience, we could had avoided the "d20" issue by analyzing all samples at the same time point at the end of the study. To have information about the variability that could come with aging samples during their storage, QC samples based on reference materials should have been included. These QC samples could consist of reconstituted plasma with known metabolites concentrations that could have been stored at different time-points during the study. Following the evolution of known concentration values of metabolites during the storage would give us information about possible bias induced by the stability of plasma samples during their storage.

Figure 57. Mann-Whitney non-parametric t-test generated using GraphPad Prism8® on lipoprotein fraction data obtained 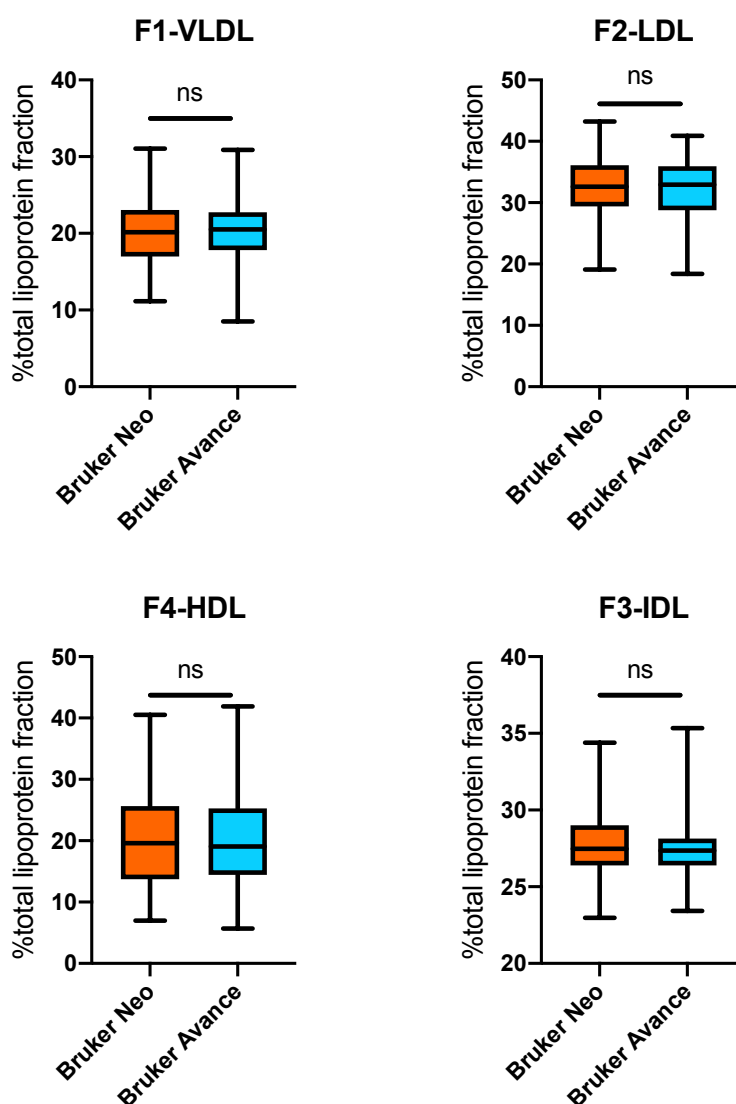through NMR lipoprotein profile evaluation approach. No differences can be made between the two batch of samples acquired on the Bruker® Avance or Neo platform.

4.4.2 Quantification, final dataset, and analytical strategy.

To perform valuable statistical analysis on our longitudinal dataset, we decide to quantify metabolites of our NMR plasma samples. For this, an ultrafiltration step dedicated to protein elimination is mandatory as well as the use of proper NMR pulses sequence (1D-noesypresat) in order to fit with ChenomX© software requirement. This commercially available software allows the identification and the quantification of plasma metabolites through line fitting algorithm allowing spectral deconvolution of resonance peaks. This led to the quantification of 60 metabolites coming from various biochemical pathways (list 1) in all plasma samples coming from nAMD patients involved in the follow-up study.

From the initial 269 visits, 13 have no NMR data because of blood sampling issues and 15 have no OCT data. From the remaining 241 samples analyzed in NMR, 10 were excluded from the analysis because of poor spectral quality. At the end the dataset is containing 231 samples from 29 patients as PS29 and PS19 was removed because having one single time-point, and PS11 was removed because only one value has clinical data from all visits.

From the initial 269 visits, a final number of 231 samples were used for the following analysis (see section Materials and Methods 3.1.2 for more details)

To better represent our cohort and evolution of patients during the study, we made several groups of individuals according to 8 criteria (C1-C8) based either on clinicians' conclusion/decision and OCT data (**Figure 58**). The first two groups are based on the initial injection planning and its evolution, this plan is set by the ophthalmologist regarding the initial status of the patient and its own evolution. Two other groups are based on the clinicians 'conclusions and their evaluation of treatment responses. All other groups are based on the OCT markers and should represent a more rational way to characterize patients. This represents the best way to make clear correlation between the metabolome and the pathologic CNV events occurring during nAMD. All groups and their frequency are represented below:

From Injections plan information

C1 — Analysis based on the initial anti-VEGF injection plan

G1	monthly injected	n=130
G2	injected each 3 month (or more)	n=59
G3	injected each 2 month	n=42

C2 — Analysis based on changes among the injection plan

GA	monthly regiment is kept	n=104
GB	next visit espaced from the initial planning	n=44
GC	espaced regiment is ketp	n=83

From clinicians decisions

C3 — Analysis based on injections's decision following clinician advice

Oui	patient will recieve an injection this day	n=134
Non	patient will not recieve an injection this day	n=88
NA	missing information or patient is never injected	n=9

From informations provided by clinicians

C4 — Analysis based on the general conclusions of the clinician

Pos	visit with positive appreciation	n=90
Neg	visit with negative appreciation	n=77
Stb	visit in which patient is considered as stable	n=53
NA	visit with no conclusion	n=11

C5 — Analysis based on the treatment responses estimated by the clinician

Oui	patient have responded	n=50
Non	patient didn't have responded	n=63
NA	patient was not injected or lack of conclusion	n=118

From OCT images quantification data

C6 — Analysis based on the presence of SRF in OCT images

Oui	sub-retinal fluid are present	n=152
Non	sub-retinal fluid are not present	n=79

C7 — Analysis based on the evolution of IRC surface in OCT images

Up	IRC in augmentation from previous visits	n=127
Down	IRC in diminution from previous visits	n=95
Stabl	IRC stable from previous visits	n=9

C8 — Analysis based on the evolution of PED measures in OCT images

Up	PED in augmentation from previous visits	n=129
Down	PED in diminution from previous visits	n=102
Stabl	PED stable from previous visits	n=0

Figure 58. Representation of all groups formed for statistical analysis and their frequency (n=x). Groups from C1 to C5 were made from information provided by clinicians and groups from C6 to C8 were made on the basis of OCT data and their evolution.

4.4.3 Exploratory analysis of NMR data

In this section we used the dataset in which each sample is described by the concentration of the 60 metabolites measured by NMR (**Figure 59**). We will represent the dataset using unsupervised analysis and color samples according to their identity (PS), their clinical evaluation (comment from the clinicians), their response to treatment or the evolution of the measured OCT markers (IRC, PED and SRF).
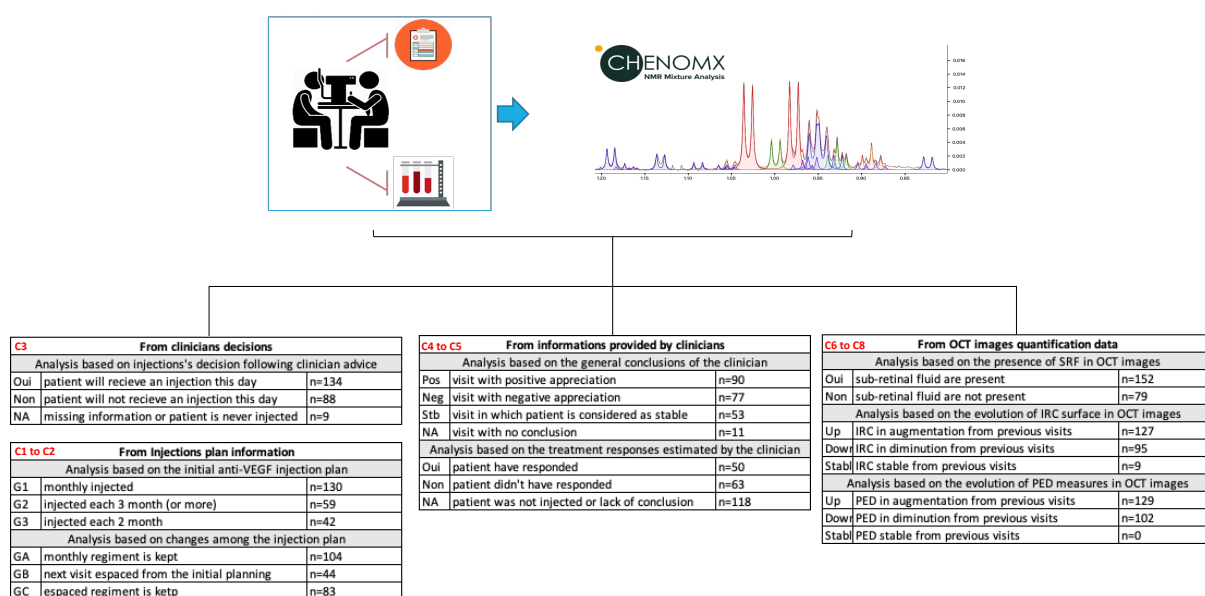

From clinicians decisions

C3	Analysis based on injections's decision following clinician advice	
Oui	patient will recieve an injection this day	n=134
Non	patient will not recieve an injection this day	n=88
NA	missing information or patient is never injected	n=9

From Injections plan information

C1 to C2	Analysis based on the initial anti-VEGF injection plan	
G1	monthly injected	n=130
G2	injected each 3 month (or more)	n=59
G3	injected each 2 month	n=42
	Analysis based on changes among the injection plan	
GA	monthly regiment is kept	n=104
GB	next visit espaced from the initial planning	n=44
GC	espaced regiment is ketp	n=83

From informations provided by clinicians

C4 to C5	Analysis based on the general conclusions of the clinician	
Pos	visit with positive appreciation	n=90
Neg	visit with negative appreciation	n=77
Stb	visit in which patient is considered as stable	n=53
NA	visit with no conclusion	n=11
	Analysis based on the treatment responses estimated by the clinician	
Oui	patient have responded	n=50
Non	patient didn't have responded	n=63
NA	patient was not injected or lack of conclusion	n=118

From OCT images quantification data

C6 to C8	Analysis based on the presence of SRF in OCT images	
Oui	sub-retinal fluid are present	n=152
Non	sub-retinal fluid are not present	n=79
	Analysis based on the evolution of IRC surface in OCT images	
Up	IRC in augmentation from previous visits	n=127
Down	IRC in diminution from previous visits	n=95
Stabl	IRC stable from previous visits	n=9
	Analysis based on the evolution of PED measures in OCT images	
Up	PED in augmentation from previous visits	n=129
Down	PED in diminution from previous visits	n=102
Stabl	PED stable from previous visits	n=0

Figure 59. Exploratory analysis workflow of the dataset generated using ChenomX.

To begin our analysis, we will explore our dataset using principal component analysis. At first this will help us to detect any outliers, to figure out how our data are represented by the model and to determine the major sources of variations. From the analysis of the score plot of the first PCA model (**Figure 60**), we can identify two outliers from the same individual. These samples shouldn't be considered for the rest of the analysis as they present abnormally high concentration values for all metabolites compared to other samples and therefore were excluded from the dataset.
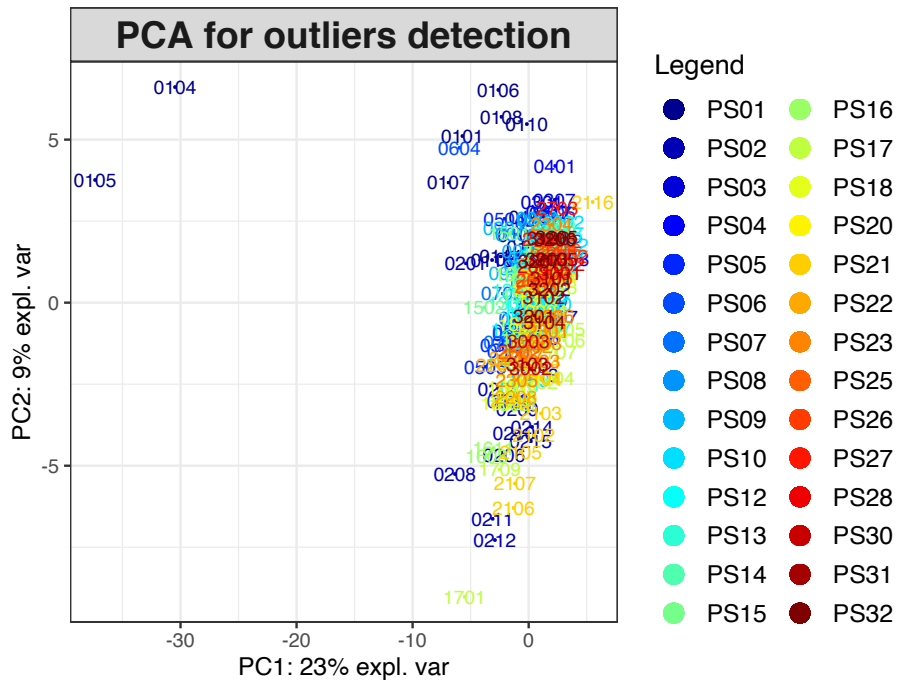
Figure 60. PCA analysis of the studied dataset, data were scaled prior the analysis and each individual are represented by a single color. Samples codes can be read as follow: first two number accounts for the patient ID and the two following number accounts for the visit number (i.e: 0104 code for the fourth visit of the patient PS01)

All the following PCA model were generated using Mixomics R package, an open access R package. The advantage of this package is the integration of the multilevel tool for different multivariate statistical methods. This method was developed by Dr. B. Liquet to take into account the longitudinal structure of complex dataset that involves repeated measurements[247]. In this approach, the "within variation" is split from the "in-between variation" and is called internally to perform the statistical analysis on the desired part of the original dataset (equation 4). This method is based on the "split-up" variation approach developed by Westerhius et.al[248] and enable to extract the stimulation effect from each subject by getting rid of the between-subject effect. Hence, we have a better representation of the changes effect within the subject than by considering all sources of variations. This will avoid the formation of clusters made of samples coming from the same individuals and better represent their evolution during their follow-up.

$$X = X_{...} + X_b + X_w$$

Equation 4. Representation of the split-up approach of the multilevel normalization method. (X= matrix; X...=residuals; X_b=between-subject variation; X_w=within-subject variation).

On **Figure 61** the two PCA score plot, with and without the application of multilevel, are represented and we can clearly see the effect of this operation on the individual PS01. As we see, the samples coming from the different time points of this patient are no longer clustering together. From the explained variance graph of our PCA model (**Figure 61**), we can see that the proportion of variance is mostly captured by PC1 and after PC4 the remaining PCs are explaining less that 5% of the variation among our datasets. Therefore, for the remaining analysis we will focus on the four first components.
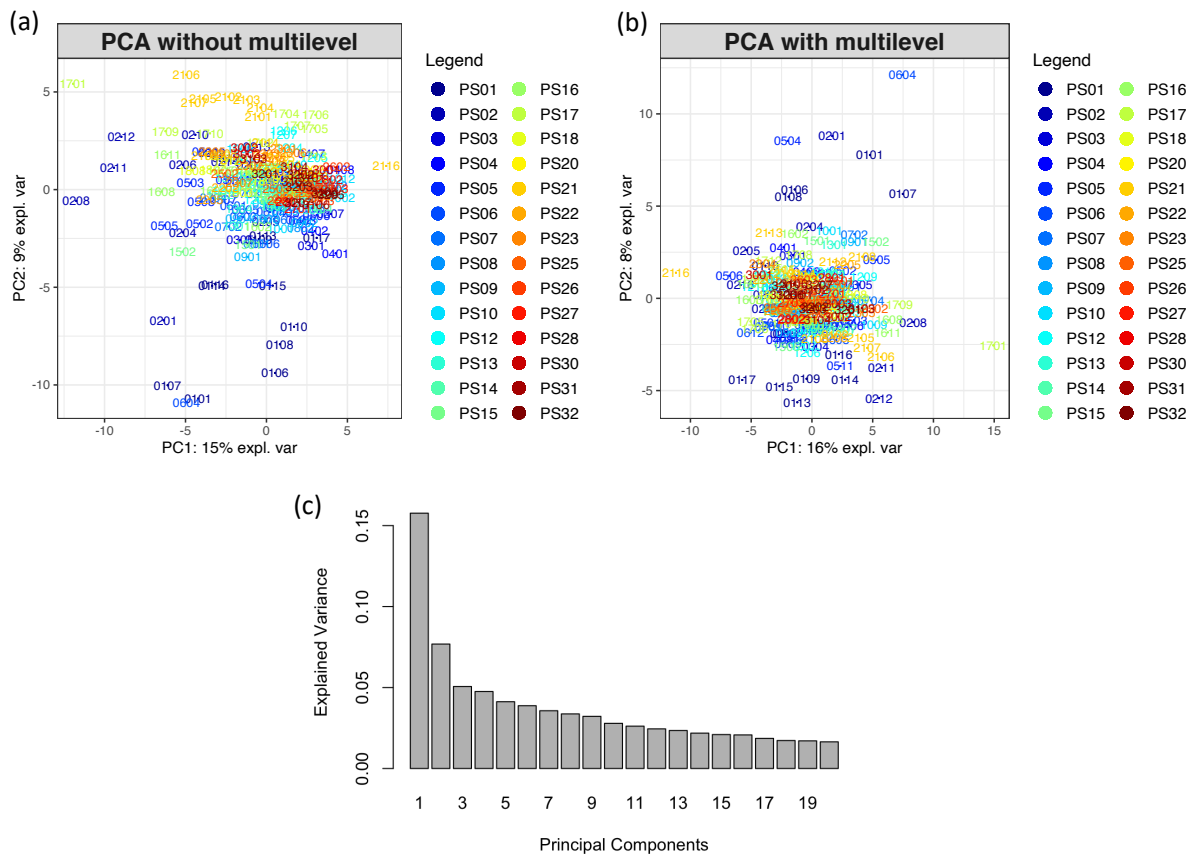

Figure 61. (a) PCA model generated without the use of multilevel; (b) PCA model showing the effect of the multilevel approach on our model; (c) the explained proportion variance of the final model.

To explore our cohort and see if some samples are naturally cluster together and if these clusters are related to some pathological condition, we have generated different score plots that are representing the patients regarding to their AMD status. In the file "Exploratory-Analysis" from *https://github.com/MS28uliege/These_MS_Results* can be found all score and loading plot regarding this analysis, but all conclusions from these remain the same. Here are only shows the results where the samples are represented based on the evolution of OCT markers. The evolution of IRC is calculated using the difference of the measured values between two visits

and samples where then divided in three groups: those for which the IRC values are increasing (IRC up), decreased (IRC down), or remained stable (IRC stb) between visits.

Based on the analysis of the score plot, no clustering exists between visits that had the same clinical outcomes even when analyzing smaller components. Moreover, the loading plot of the generated demonstrates the poor variabilities of our dataset. Indeed, the contribution of the variables to the different components is small indicating that values are close for all samples. This doesn't allow us to analyze our cohort in an unsupervised way as we planned. The exploration of our dataset using PCA models will not be able to capture the variability that could exist during AMD patient's follow-up (**Figure 62**).
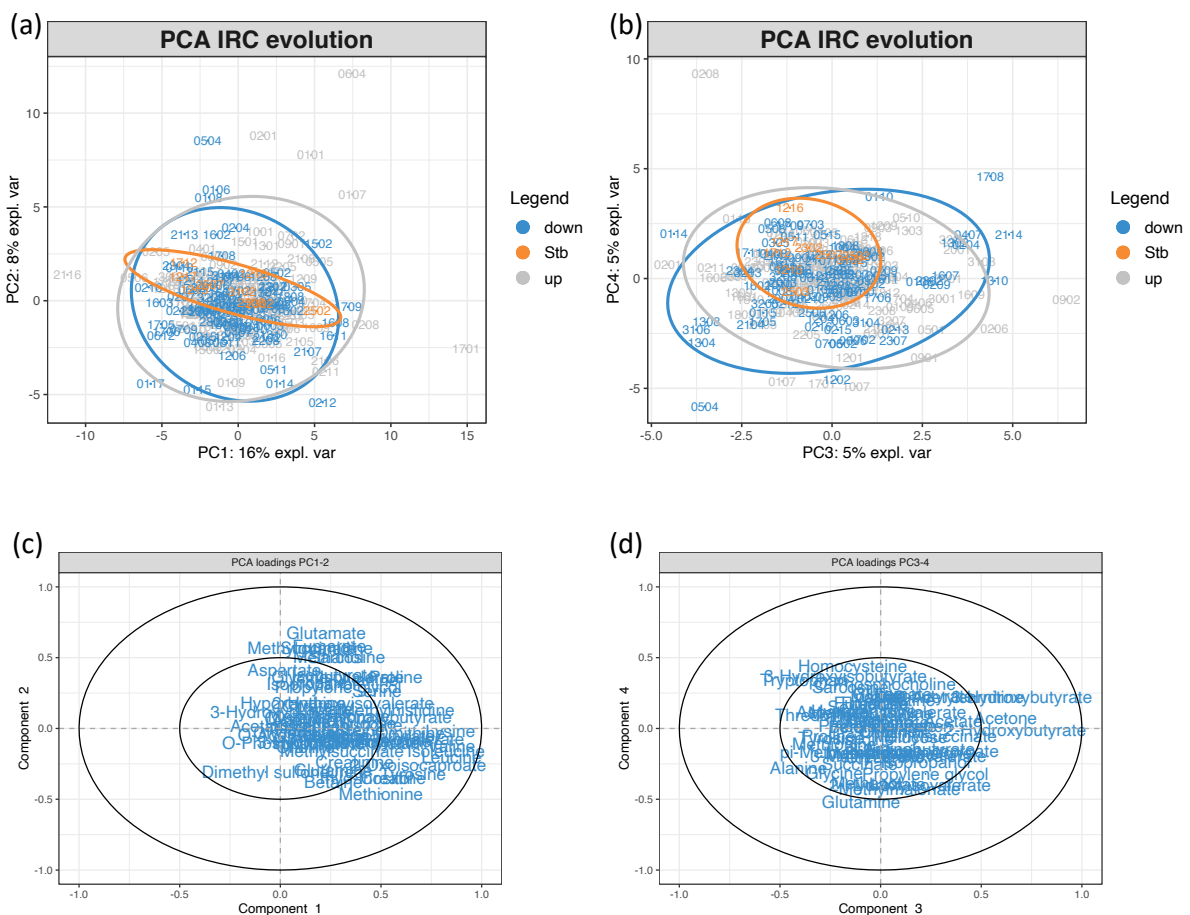

Figure 62. PCA score plot representing the individuals regarding PC1 and PC2 (a); and PC3 over PC4 (b); the corresponding loading plots (c) and (d) (Total variance explained for PC1 and PC2 = 24%; Total variance explained for PC3 and PC4= 10%).

4.4.4 Discriminant analysis

On this part we aim to perform different discriminant analysis to find features able to make distinctions between groups of samples. For all analysis performed, the multilevel correction was applied to consider the repeated measurement of each subject. Groups were formed to answer specific question about AMD evolution and treatment's responses as depicted in **Figure 59**. As no separation between groups were noticed in unsupervised analysis, our primary concern was to not produce over-fitted supervised statistical analysis that would produce non-consistent results. Overfitting can appear when the model is too complex and start to explain the dataset using noise or irrelevant information. This produces a model that seems to describe well the original dataset but will fail on a new one. If a generated model fails to be generalized to other new data, it cannot be used to perform the classification or prediction task that it was created for.

To reduce the risk of over-fitting in our analysis and to optimize the chance to select important variables linked to group separation, we will use MixOmics package. This R package allow the tuning of sparse PLS-DA (sPLSDA) models by returning the optimal number of components and how many features to retain for each component. These parameters are kept for the generation of the final model and the performance will be than evaluated using multivariate ROC curves and permutations test. Sparse PLS-DA performs variable selection and classification in a one-step procedure improving the interpretability of the results[249].

To determine the ideal number of component and variables used to construct them, the Balanced Error Rate (BER) is used as the performance measure (**Figure 63**). For this, a five-fold cross-validation is repeated ten times to calculate the classification error rate that is then averaged to consider the weight of each class. The BER value is from 0 to 1 and lower value indicates accurate model. This process is performed iteratively in the way that each component is tuned one at a time to allow the determination of the ideal number of variables. From the plot below, key parameters (ideal number of components= ncomp; optimal number of variables= keepX) are extracted to build our final model.
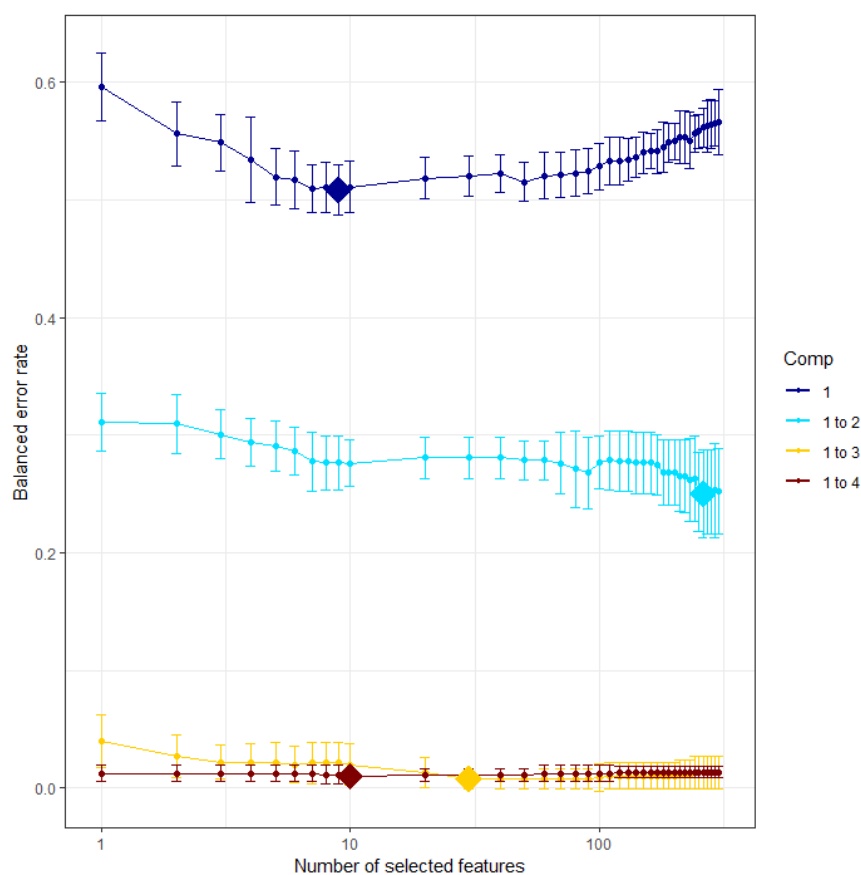
Figure 63. Example of balanced error rate (BER) plot from a simulated dataset: BER is calculated for each component regarding the number of variables taken in consideration. Diamonds indicate the optimal number of variables to select per component. As low BER values indicate accurate model, the lowest diamond will return the ideal number of component (ncomp) and the number of variables to consider (keepX).

The first discriminant analysis on groups was made on the clinicians' conclusions. This was performed using the dataset in which each patient/visit is described by the 60 metabolites obtained from ChenomX. For each visit clinicians left comments about AMD status of patients that can be classified as "positive", "negative" or "stable" regarding their evolution. After tuning, the number of components was fixed at 2 with 5 to 7 variables to consider for the component 2 and 1 respectively. If the loading plot does not show any separation between groups (**Figure 64b**), the ROC curves generated was more optimistic as the AUC calculated was above 0.7 for almost all the three classes (**Figure 64c**). Nevertheless, permutation test showed us the poor predictability of our model as important number of misclassifications is observed among all classes (**Figure 64d**).
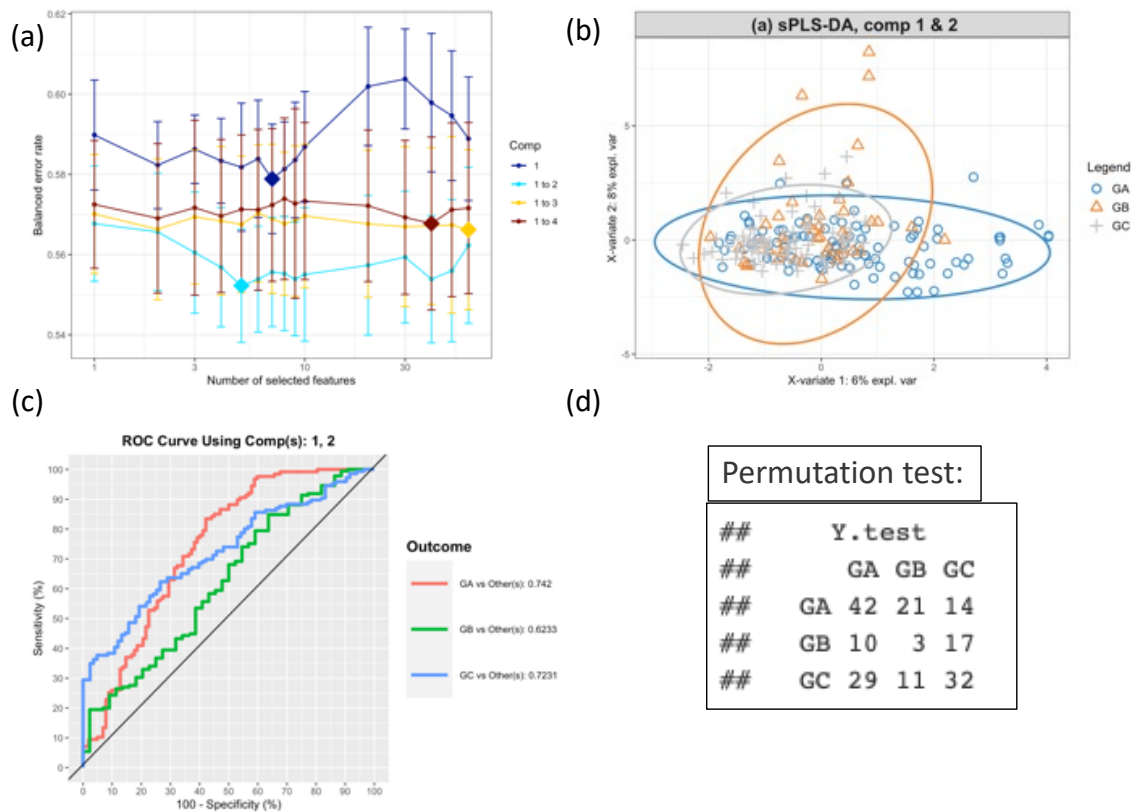
Figure 64. Plot resulting from the tuning of sPLSDA model (a) and the loading plot of the final model (b). ROC curves based on the selected components and their ability to classify the samples regarding their groups (GA for people with positive evolution, GB for negative and GC for patients that are stable) (c) and the permutation test showing the predictability of our model (d).

Results of all discriminant analysis performed are summarized on **Table 6** and conclusions remained the same. None of the model were able to make good classification of patients in any kind of groups based on the different clinical information. The tuning step for some analysis was already a good indicator of how poorly informative the models are. Indeed, for most of them the ideal number of components was 1 and most of the time, the keep X parameter was equal to the number of variables.

We need to change our way to analyze this dataset as this approach is clearly not adapted to our case study. Indeed, as classification of patient is the main problem in AMD management, and since all patients evolves from different starting point regarding their OCT data, it's not surprising that our class analysis was not satisfying. Therefore, changing our way to explore our dataset is essential to be able to refine the way to classify and follow AMD patients.

Groups	ideal ncomp	AUC Roc	Perm.
Initial inj. Plan	4	0,4	Fail !
Adaptation plan inj.	2	0,7	Fail !
Conclusion	4	0,5	Fail !
Injection	1	0,6	Fail !
Response	1	0,5	Fail !
SRF?	1	0,5	Fail !
IRC up/down	1	0,6	Fail !
PED up/down	1	0,7	Fail !

Table 6. Summary of discriminant analysis performed on groups made on clinical information and OCT data. For analysis for which the ncomp= 1, two components were kept to run the entire analysis. For permutation test, considered as *fail* ones that were exhibited an important number of misclassified samples.

Notes: All results from this analysis can be viewed in the file "Discriminant-Analysis" from GitHub repository: https://github.com/MS28uliege/These_MS_Results

4.4.5 Correlation and regression analysis

Since all exploratory and discriminant analyses based on or defined groups failed, we decided to change our approach. In this section, we will focus on the relation that can exist between information about the metabolome and the clinical measures based on OCT images of AMD patients' retina. The measured parameters, namely IRC, SRF and PED, is expected to give us the best representation of the pathology status and of its evolution. Finding biomarkers related to their evolution could therefore be helpful for patient monitoring, classification, and evaluation of treatment responses.

For this purpose, we plan to analyze our dataset using Pearson Correlation test, Partial Least Square regression 2 (PLS2) and Principal Component Regression (PCR) analysis. The objective of such approach is to identify variables coming from our metabolomics dataset (X) that can explain other continuous variables such as OCT measurements (Y). The dataset using here didn't undergoes any kind of normalization to consider the patient's effect (multilevel i.e.). Indeed, these analyses are performed rather to find association between continuous variable than try to discriminate group of samples.

Pearson Correlation test are done to find linear associations or dependency between two quantitative variables. Through this basic approach, we aimed to find metabolites that are linked to some OCT values such as IRC, PED or SRF. On the correlogram represented in the **Figure 65**, positive or negative correlations are represented by bright red or blue circles respectively. Correlation values rages from -1 to 1 and therefore 0 values, that are not represented by colored

circles, indicate independent variables. By analyzing the correlogram generated on our two datasets, none of the variables coming from the metabolome can be linked to the ones coming from OCT images. This indicates that only weak linear associations exist between the two datasets.
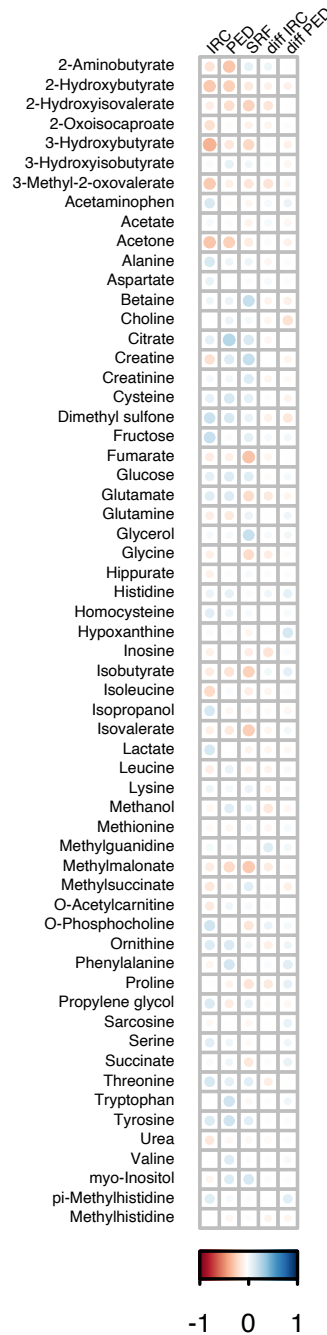

Figure 65. Correlogram of Pearson's correlations tests performed between variables coming from the metabolomics NMR analysis and OCT measurement of AMD patient's retina. Vertical and Horizontal boxes represent area in which interesting correlations could be seen.

If correlation analysis is made to highlight the link between continuous variables, it is not made to handle the complexity of multivariate data. Indeed, all variables coming from the metabolomic analysis are not independent and therefore more robust statistics are needed to better use our dataset. Therefore, we decide to use Partial Least Square regression 2 (PLS2) and Principal Component Regression (PCR) analysis.

PLS2 regression is a multivariate methodology that is used to integrate two datasets measured on the same individuals by extracting correlated information. This model allows the analysis of correlated variables and is suited to handle noisy, collinear, and missing variables. If PCA maximize the variance of components from the original dataset, PLS will maximize the covariance of components. Note that covariance is a measure of the strength of the relationship between two variables and its value has no upper or lower bound. From the PLS model, the score plot (**Figure 66a**) is representing all samples regarding the measured NMR metabolites concentrations (X) and measured parameters in OCT images (Y).

The corresponding correlation circle plots (**Figure 66b**) show the variables correlated with the different components. Since both datasets are represented using the same set of axes, correlation between variables from the two datasets can be highlighted. Some variables seem weakly correlated with IRC and SRF values and could be worth considering (3-methyl-2-oxovalerate, methylmalonate, acetone, 2-hydroxybutyrate, fumarate, isovalerate, 3-hydroxybutyrate).
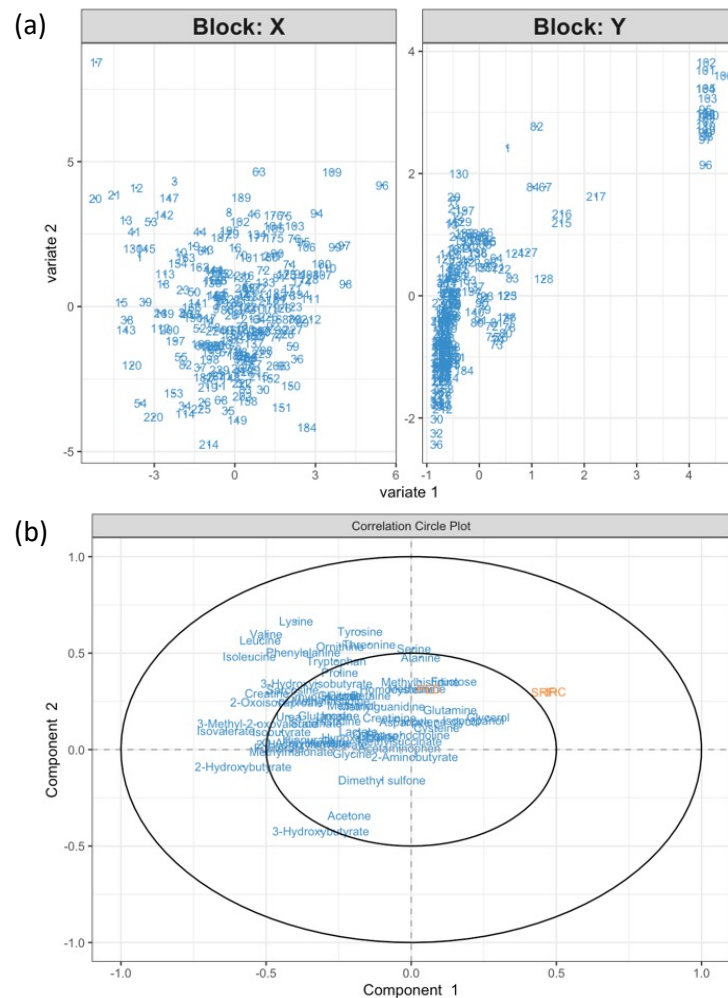
Figure 66. (a) score plot for the PLS analysis of data coming from the measured metabolites (X) and OCT data (Y); (b) Correlation circle plot showing the correlations between variables and components.

From this information a relevance network plot can be generated in which the correlations are represented. In this plot we can identify some features that are correlated to IRC and SRF values (**Figure 67**). It's important to note that for generating the network, the tuning threshold for the relevant associations network had to be lowered. Indeed, no association stronger than 0.31 was found.

Obviously, association below 0.5 are not worth of interest but here are shown to evaluate tendency that could be further validated by other analysis. Indeed, knowing the weakness of our approach (highly unbalanced design, weak variations between visits/individual, missing value), all features that could be identified by different statistical analysis could be interesting.
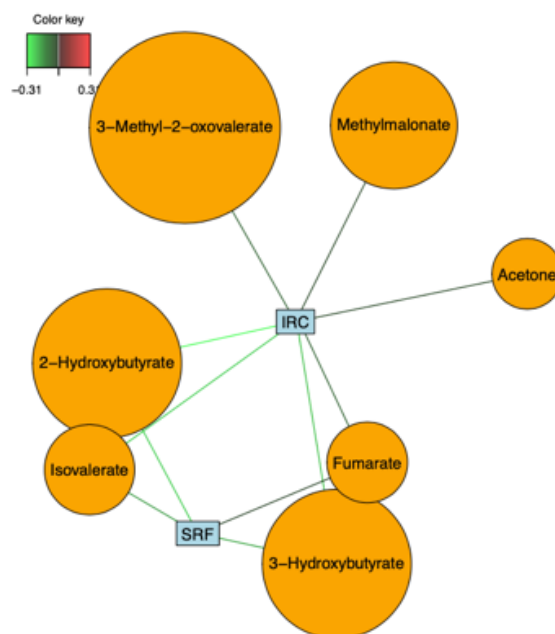
Figure 67. Relevance network representing the interesting correlations between metabolites and OCT measurements. The color of the lines indicate the strongness of the correlation. *Notes: bubbles' dimensions are function of the names variables and have no statistical significance.*

Principal Component Regression (PCR) is a regression technique that uses the principal components coming from classic PCA analysis as the predictor variables. Instead of using all variables for regression, this model only utilizes a subset of the principal components that represent most of the variability. This will reduce overfitting and eliminate multicollinearity leading to better performance than using a standard linear regression model on all original features. This works in three steps, first a PCA is generated on the data, the PCs that explain most of the variance are selected through cross-validation process and a linear regression is fitted on the selected PCs. P-values are returned for each selected component and those <0.05 are worth considering. Loading plot of interesting components are then generated to extract information about the metabolome.

All results from PCR analysis can be inspected on the **Figure 68**. From this analysis the principal component PC3 was identified as interesting to explain IRC and SRF values (p-values < 0.01). The interesting point of this analysis is that some features already identified during the PLS analysis can be retrieved in the PCR approach. Hence, fumarate, isovalerate, 3-hydroxybutyrate and 2-hydroxybutyrate were found to have an important weight on PC3.

From this analysis, some interesting features can be highlighted as metabolites correlated with important OCT markers for AMD evolution. Nevertheless, results must be refined as most of variables selected by the PLS and PCR model are not the same. Moreover, for variable selection with PLS, the cutoff selected for variable selection was lowered. This shows that result could not be that significant as it seems. This could be explained by the fact that the variations between all samples is week and maybe effect of AMD are diluted. Indeed, variations among patients regarding their pathology are small as the variations of their metabolome. Therefore, analyzing all individuals together may not be the best approach. Applying the analytical strategies that was used here could be more helpful by considering one patient at a time. By individualizing the statistical workflow, we would have greater chance to capture the variability that can be correlated with AMD evolution. Then, if these changes are consistent from one patient to another, these changes could be used to build up a new healthcare strategy based on metabolites measurement.
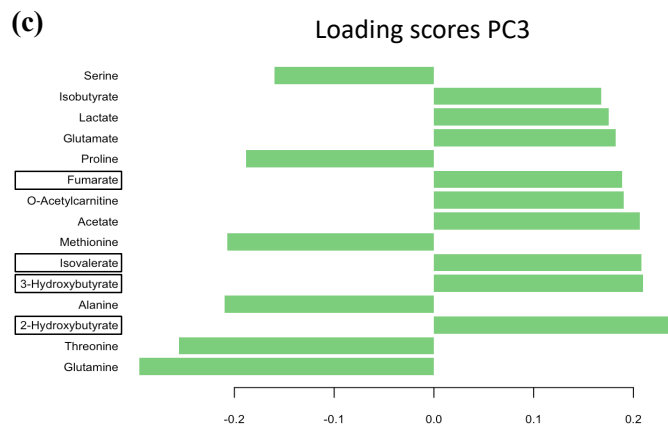

Figure 68. Results from PCR analysis performed to find components able to explain the variation of (a) IRC and (b) SRF values. (c) the loading scores associated with the component PC3.

4.4.6 Individualized approach for biomarker discovery

To better capture the metabolic variability of patients that occurs during their treatments and put it in relation with their pathologic conditions, we aim to analyze patients separately. Indeed, when generating the heatmap of OCT values for all patients, it appears that some people have very higher values than other meaning that they're evolving from different starting points. Hence, improvement or degradation of their health status may not occur at the same level for all patients. Thus, minor but important changes in the measured values between visits of a single patients could be hide by higher and stable values of patients that are constantly in bad conditions (**Figure 69**). Therefore, it seems obvious that the best approach is to deal with a single patient at once and examine if results are consistent from one patient to another.
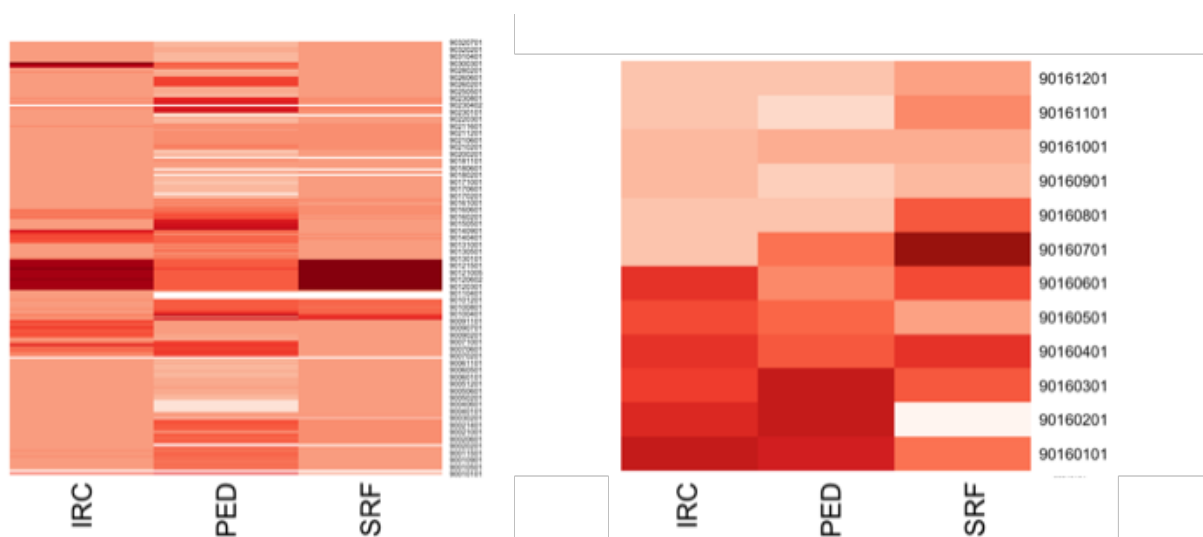
Figure 69. On the left the heatmap of OCT data based on all patients and on the right the heatmap of OCT data from patient PS16. For each visit (abscissa) are represented the measured OCT values (ordinate). On the right panel, we can better observe the positive evolution of the pathologies. This positive evolution is occurring on a small scale and is not reflected when considering all patients on the left panel.

In this section we will use different datasets: the one related to the clinical parameters measured from serum samples collected (**dataset 2**); the dataset which gives information about the lipoprotein profile obtained from NMR analysis of plasma samples (**dataset 3**); and the quantified metabolites coming from the quantification analysis (ChenomX) of NMR data (**dataset 5**); finally, the measured OCT markers will be used to characterize patient's status (**dataset 1**). The statistical workflow (**Figure 70**) applied to each patient's group of samples is as follow: 1) an heatmap of OCT values is generated to analyze the way the patient evolved and possible interesting point in the treatment; 2) a correlation analysis between OCT and clinical data; 3) a correlation analysis between OCT and the quantified metabolites data; 4) PCA on

NMR data on which the evolution of IRC values is represented; and finally, 5) PLS/PCR regression analysis of NMR and OCT data. All these tests are performed for each of the 32 patients involved in the study. In this section, only few results will be discussed as they reflect the results obtained for all the analysis.
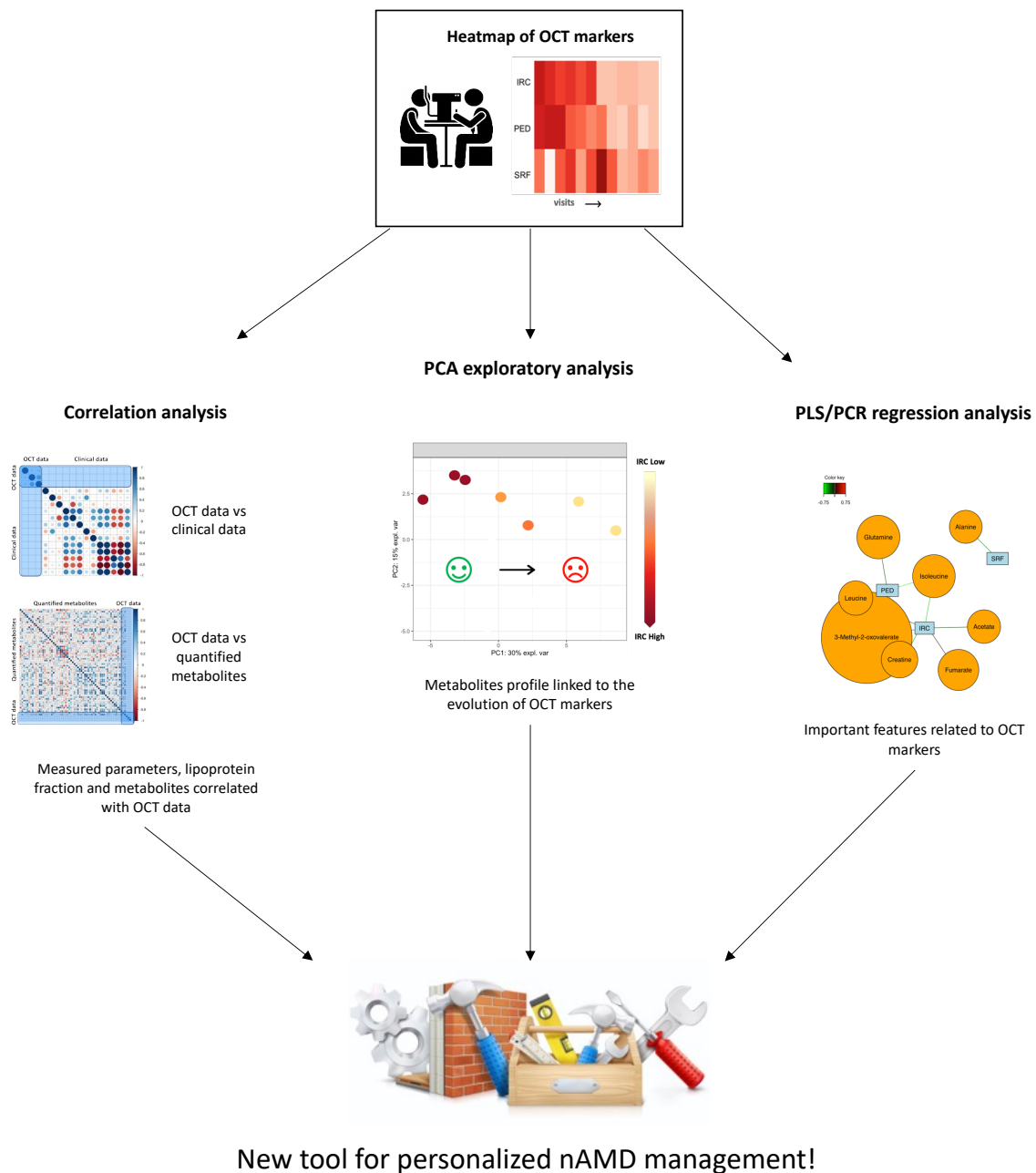
Figure 70. Statistical workflow for important feature identification in the context of nAMD management: Correlation analysis will be made between OCT measured parameters and the clinical parameters measured on serum samples collected at each visit (lactate level, pH, glucose level, HDL cholesterol and LDL cholesterol). The same correlation analysis will be performed between OCT data and the NMR lipoprotein profile obtained from the plasma analysis of patient's samples. The values of the quantified metabolites through ChenomX NMR analysis will be used for exploratory PCA analysis that will represent the variations of the whole metabolome through the follow-up. The same values will be engaged in PLS/PCR regression analysis to find correlation between the whole metabolome and the values of OCT markers collected at each time points.

Here are shown the results obtain on samples coming from the follow-up of patient PS16. The interesting point with this patient is that is AMD status evolves positively during the period of treatment. Therefore, it might be possible to find relationships between the metabolomic profiles and some metabolites and the positive evolution of IRC values or other OCT measures.

PCA analysis performed on the data coming from NMR quantification of metabolites is interesting as it seems representing well the evolution of the patient. Indeed, on the score plot, we can see a distinction between visits in which IRC values are higher along PC1 (**Figure 71**). The more samples are exhibiting a positive score along PC1 (located on the right of the score plot, **Figure 71a**), the better their AMD status seems evolved as the IRC values associated with these samples are lower. Thus, samples on the right of the PCA score plot seem have a positive evolution compared to the samples on the left. The interesting point is that the concentration of some metabolites such as fumarate seems to be increased in visits with higher IRC values, this metabolite was already highlighted in previous analysis concerning all samples. Importantly, none of the previously identified biomarker (lactate level and lipoprotein profile) can explain these changes.
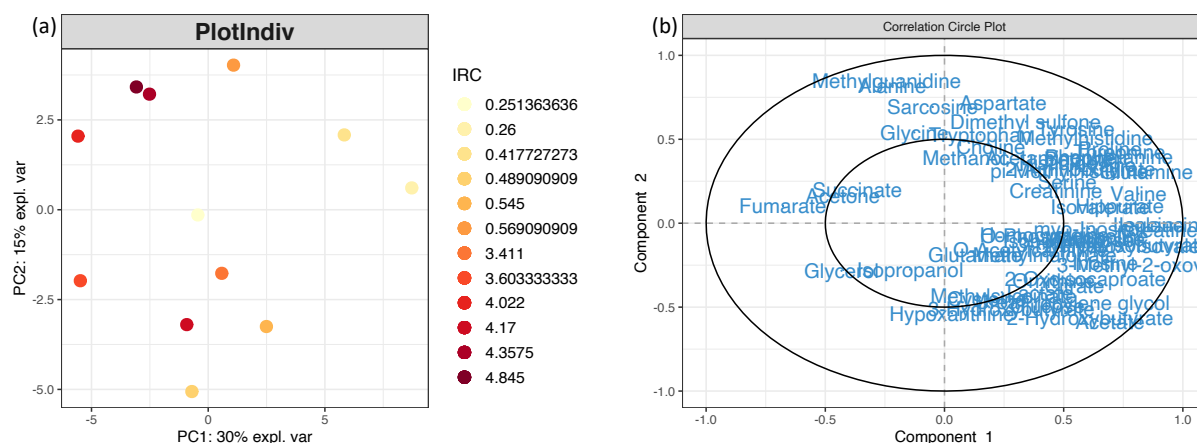
Figure 71. (a) PCA score plot representing PS16 samples colored on the basis of the IRC values and (b) the corresponding loading plot.

Correlation analysis based either on clinical or NMR data (**Figure 72**) seems also provide useful information. These analyses allow us to highlight strong correlation between some variables and markers of the disease (IRC, PED and SRF). Furthermore, it can confirm observations obtained using the PCA score and loading plots since metabolites as fumarate are

also strongly correlated with IRC values. It will be interesting to see if we can observe the same tendency from one patient to the other using the same approach.
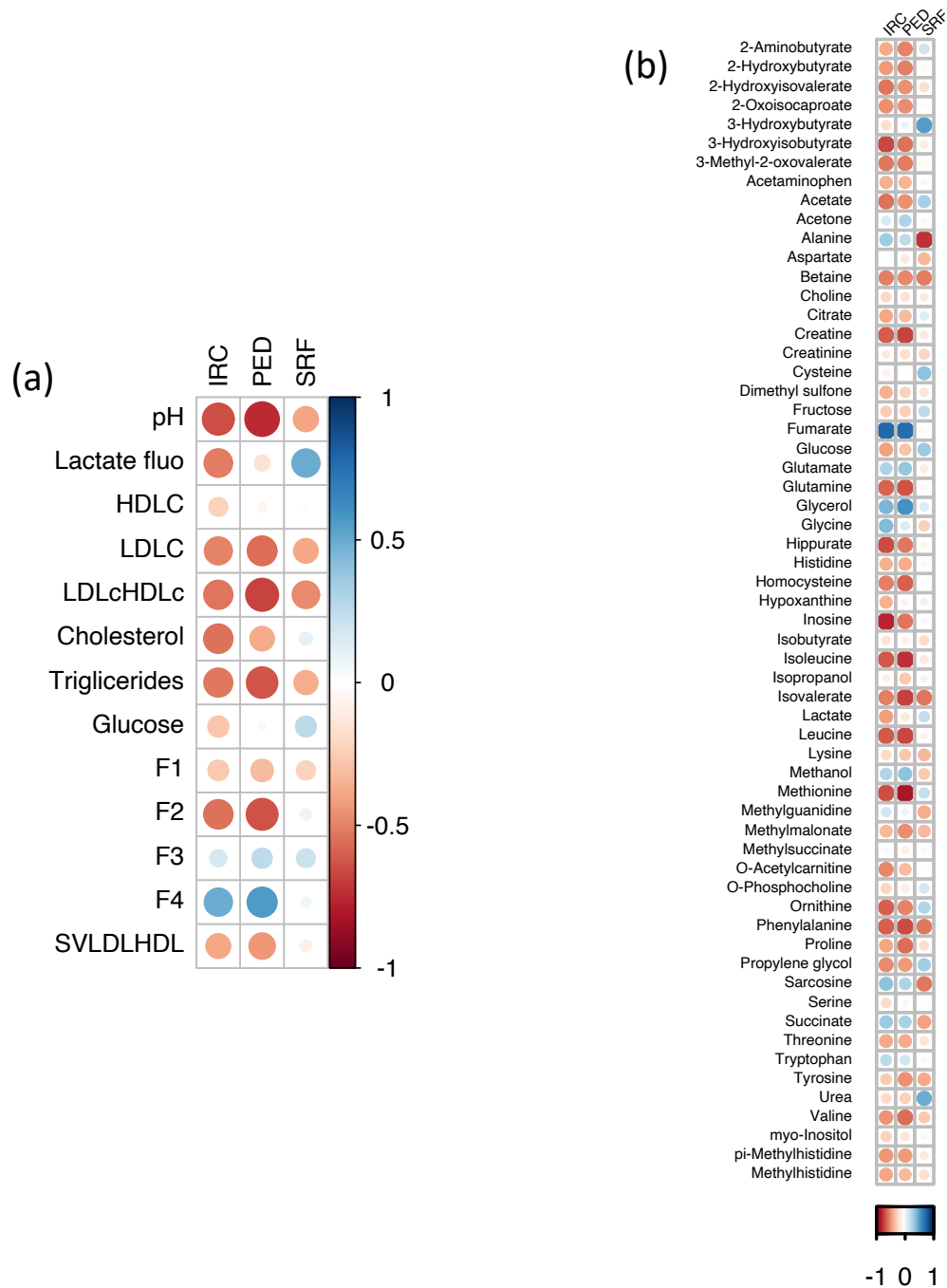
Figure 72. Correlogram of Pearson's correlation test between (a) OCT data and clinical data, and (b) OCT data and metabolites concentration data. On these plots, some strong correlation can be observed, and the analysis of the plots generated from data of the others patients will allow us to identify interesting features.

Principal Component Regression analysis didn't allow to spot some useful components for explaining OCT values of PS16. Nevertheless, PLS regression analysis provided some feature with great coefficient of correlation with IRC, SRF and PED values (**Figure 73**). Moreover, fumarate and 3-methyl-2oxovalerate were already identified in analysis conducted on the whole dataset (chapter 4.4.5-**Figure 68**) and by correlation test analysis made on PS16 data. If these features are also depicted using this approach on other patients, it will be of particular interest.
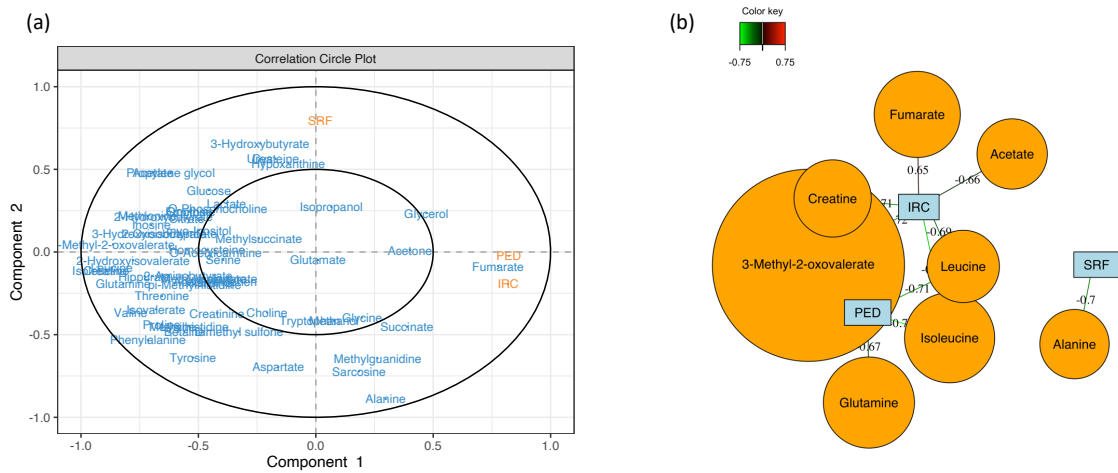
Figure 73. (a) Correlation circle plot from PLS regression analysis between metabolites concentration and OCT data; and (b) the corresponding correlation network representing correlation >0.65.

Analysis of data from patient PS32

Patient PS32 has a similar evolution to PS16, therefore it would be interesting to examine if related outcomes can be observable. As for PS16, none of the correlation test provided interesting results. Regarding the PCA analysis, the evolution of IRC values is well represented along PC1 and represent 37% of the total variation. This observation is reflected what we saw in PCA analysis of PS16 indicating that effect of IRC evolution might impact the metabolism. The loading plot representing the most important variables of PC1 didn't give us the same features as the model made on PS16 data. Thus, when PLS analysis was performed, even if highly correlated feature was identified, none of these was already identified (**Figure 74**). This means that if IRC evolution impact the metabolism of PS32, it does not do it in the same way than it does is PS16 and therefore results lack consistency. In the same line with results coming from PS16 data's, PCR analysis returned no PCs able to predict measures coming from OCT images of PS32.
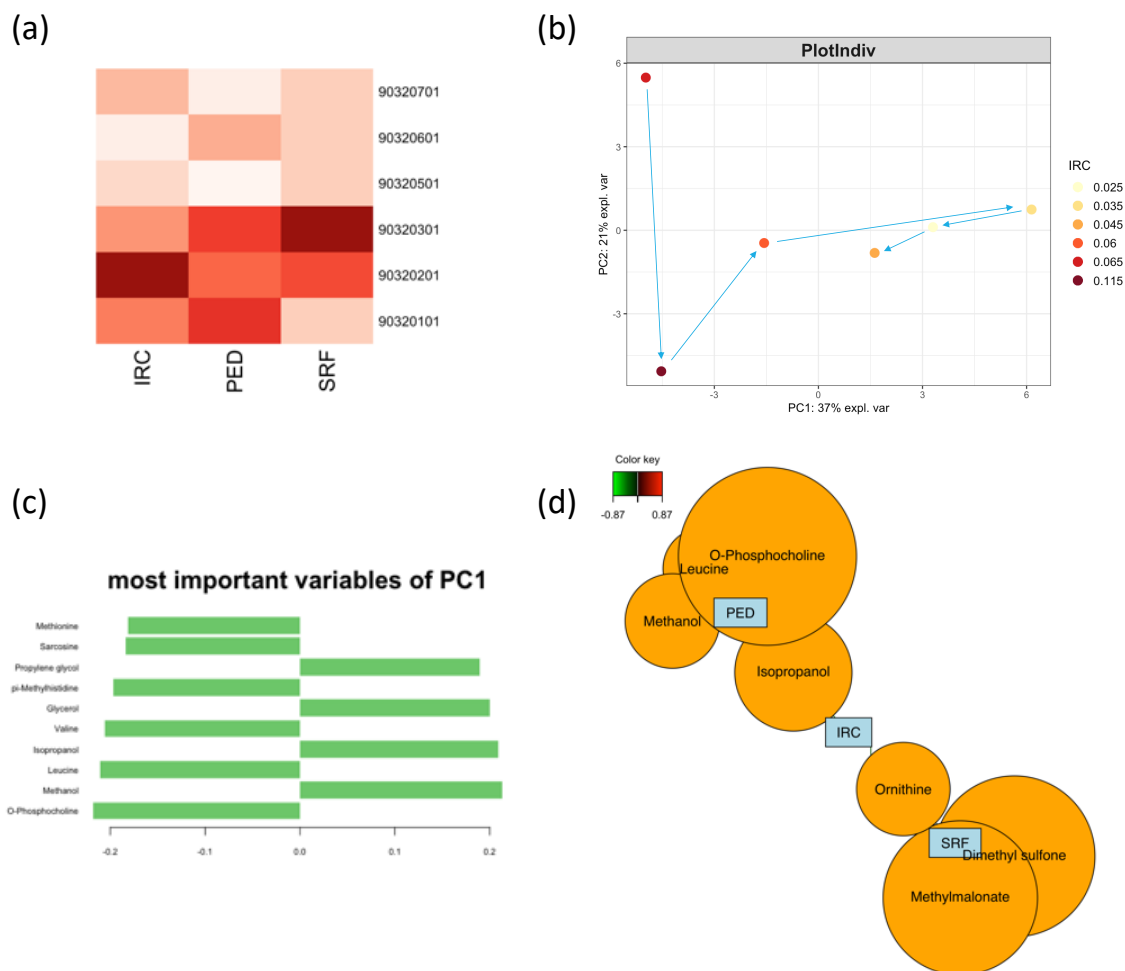
Figure 74. (a) Heatmap representation of OCT data from PS32; (b) score plot of metabolites concentration dataset of PS32 with trajectories representing the evolution of IRC values and (c) the corresponding loading plot showing important variables of PC1; and finally (d) variables highlighted from the PLS regression analysis between OCT dataset and metabolites concentration.

Analysis of data from patient PS31 & PS26

PS31 and PS26 are both patients that exhibited almost no IRC at the beginning of the follow-up. Again, PCA analysis can capture the evolution of IRC status by analyzing the trajectories of the samples along PC1 and PC2 (**Figure 75**). Moreover, as for the precedent patients, effect on the metabolome is different for both patients. Indeed, PLS regression analysis once again returned original features as highly correlated with OCT data while PCR methods provide no relevant PCs.
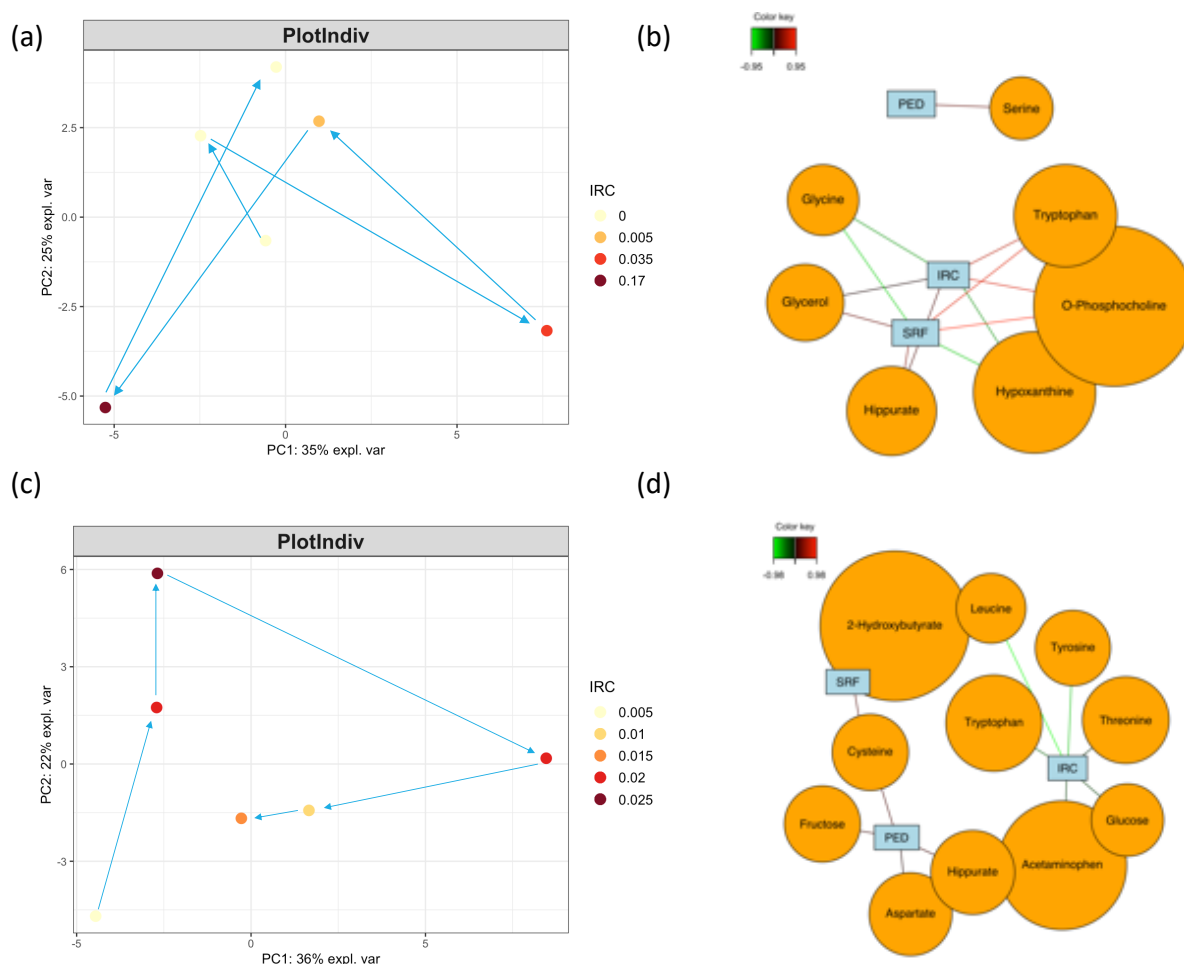
Figure 75. (a) score plot of metabolites concentration dataset of PS31 with trajectories representing the evolution of IRC values and (b) variables highlighted from the PLS regression analysis between OCT dataset and metabolites concentration; (c) score plot of metabolites concentration dataset of PS26 with trajectories representing the evolution of IRC values and (d) variables highlighted from the PLS regression analysis between OCT dataset and metabolites concentration.

When combining the results from all analysis performed on all AMD patients, 98% of variables from NMR metabolites quantifications are identified as relevant by PLS analysis (correlation >0.7) at least one time and 26% 5 times. Furthermore, this conclusion can also be made from all the correlation test made on each individual (**Figure 76**). This shows that our approach is not able to provide consistent results for the whole cohort. No clear information can be extracted from all these tests that could be used for AMD monitoring. If our data allow to highlight metabolites changes that can occur concomitantly with the variation of OCT measures, these changes and the metabolites involve are inconsistent across the cohort. This made these findings poorly relevant and didn't allow us to provide useful information for refining patient's status during AMD management.

Figure 76. Correlogram of Pearson's correlation test between OCT data and metabolites concentration data for patient PS16, 26, 31 and 32. On these plots, some strong correlation can be observed but these are inconsistent from one patient to the other. This observation can be made with all other combination of patients.

4.5 Conclusions

The initial goal of our study was to use the biomarkers previously identified, namely lactate and lipoprotein profile, to monitor the evolution of AMD patients in real-life clinical situation. During a two years patients follow-up, we collected blood samples and recovered relevant information about patient's AMD status. All included patients were diagnosed with advanced and exudative AMD and were under anti-VEGF treatments. They were involved in a treatment regimen adapted by the clinician and based on their evolution. Patients that are not responding well to treatment were followed each month while for some others, visits occurred each 3 months. This regimen is submitted to changes as if a patient involved in a 3-month regimen exhibit bad visit outcomes it can be adapted to a monthly injection schedule. This resulted in an extremely complex and uncontrolled study design driven by the reality of clinical practices. It ends with a cohort of individual in which each patient was followed independently from the others.

The original idea was to compare evolution of our biomarkers with markers coming from OCT images that give the best representation of nAMD evolution. This approach failed as biomarkers' evolution didn't follow the evolution of OCT markers and their values weren't even correlated. This forced us to change our approach by considering our dataset as a chance to find new or more adapted biomarkers that could better fit to our goals. Indeed, if case-control studies are suited to highlight some biomarkers, forced to see that real-life clinical practices don't provide such controlled environment. Therefore, it's not surprising that such markers as lactate or lipoprotein failed when applied to cohort in which people are so heterogenous regarding their lifestyle, ages, health condition (other pathologies than AMD included) or even how they are taken in charges during their treatment. Thus, we hoped that working on samples that are such close to "real" condition could provide more robust and usable information.

Following this goal, we set up a new statistical analysis strategy that suited to the specificity of our dataset. By taking account its longitudinal aspect, we analyzed our dataset using the most relevant approach described in the literature and that could be used on our cohort. Despite all the statistic test performed, by taking all samples together or by individualizing the statistical approach, no consistent result could be found reminding the quote of Thomas Alva Edison *"I have not failed. I've just found 10,000 ways that won't work"*.

At the end, our previously discovered biomarkers were unsuitable for patients monitoring and all data collected during visits were not informative for new biomarker discovery. If the

preliminary study didn't fit to the clinic reality, did our follow-up cohort fitted to statistic reality?

Indeed, by looking at how the study was conducted, and this choice is fully assumed, did we had a chance to extract any relevant information? On the **Figure 77** is represented the timetable of patient's visits. This schedule is fixed by the clinician regarding the evolution of patient's nAMD. Therefore, this schedule is evolving in the way that patient's that are responding well are seen less than the ones that didn't responded to treatments.

Therefore, a patient that was seen each month is not evolving a lot and a patient that was responding well will not come back until his pathology hits back. Thus, we will only have complete dataset for patients that are not presenting evolution regarding their nAMD and we will probably miss the most interesting time points in which the patient that was in a better evolution started to be bad again. At the end the dataset is not as interesting as expected and no relevant information can be extracted.

The interest of this study results in the questions that comes at the end: Does case-control studies give the best chances to identify usable biomarkers in pathologies that need constant follow-up? Are current clinical practices suited to incorporate metabolomics approach to improve patients' health care? Does metabolomics approach suit to such application?

We need to better determine the metabolome baseline of healthy individuals and its time evolution. Knowing the normal variation of metabolites levels through different time points could help to evaluate the usefulness of a discovered biomarkers when applied in real clinical environment. Indeed, for a 10% concentration variation of a given metabolite in a case-control study, the intrinsic and natural variation of this compound over the time must be quite smaller. If not, you might be not able to highlight pathologic events using this biomarker in a real-life context.

Therefore, when the goal is to improve patient's healthcare and monitoring, the study design must include times related variations. This could be directly made by collecting samples at different time points in case-control studies or by already know the intrinsic variation of the studied metabolite under healthy conditions. Anyway, some improvements must be done to produce consistent, pertinent, and useful results that aim to be considered by clinicians as "new tools" for patients' monitoring.

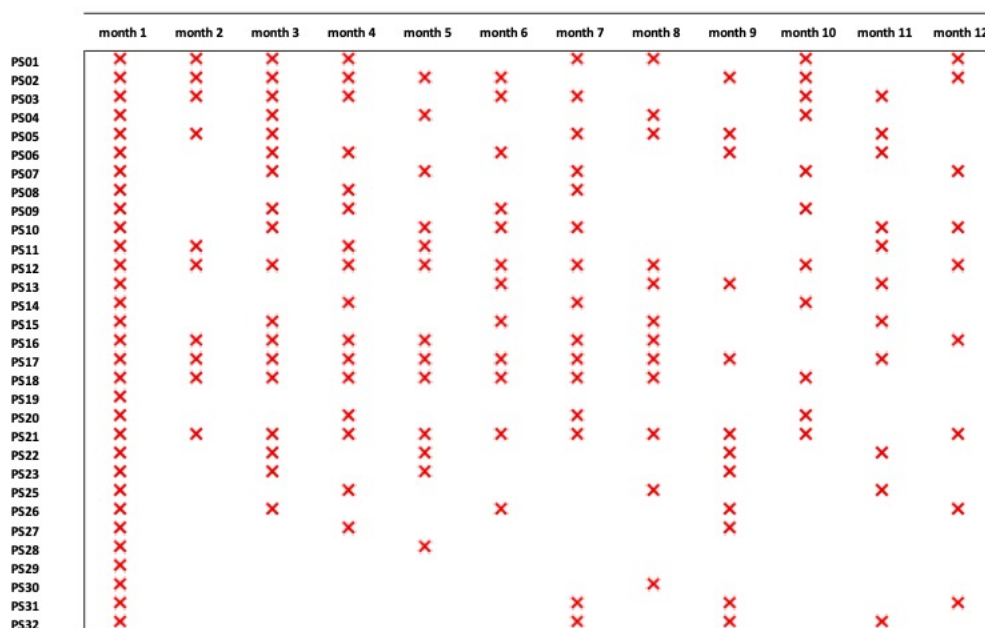|	month 1	month 2	month 3	month 4	month 5	month 6	month 7	month 8	month 9	month 10	month 11	month 12
PS01	×	×	×	×			×	×		×		×
PS02	×	×	×	×	×	×			×	×		×
PS03	×	×	×	×		×	×			×	×	
PS04	×		×		×			×		×		
PS05	×	×	×				×	×	×		×	
PS06	×		×	×		×			×		×	
PS07	×		×		×		×			×		×
PS08	×			×			×					
PS09	×		×	×		×				×		
PS10	×		×		×	×	×				×	×
PS11	×	×		×	×						×	
PS12	×	×	×	×	×	×	×	×		×		×
PS13	×					×	×	×	×		×	
PS14	×			×			×	×		×		
PS15	×		×			×		×			×	
PS16	×	×	×	×	×		×	×			×	×
PS17	×	×	×	×	×	×	×	×	×		×	
PS18	×	×	×	×	×	×	×	×		×		
PS19	×											
PS20	×			×			×			×		
PS21	×	×	×	×	×	×	×	×	×	×		×
PS22	×		×		×				×		×	
PS23	×		×		×				×			
PS25	×			×				×			×	
PS26	×	×				×			×			×
PS27	×			×					×			
PS28	×				×							
PS29	×											
PS30	×							×				
PS31	×						×		×			×
PS32	×						×		×		×	

Figure 77. Representation of patients' follow-up schedule for the twelve first month of the study.

4.6   Materials and Methods

Patients' selection, clinical data, and samples collection

For two years, 14 male and 18 female patients, from 61 to 92 y/o, were followed at each ophthalmologic visit for which blood samples and OCT images of retinas were collected (n total visit= 269). For each sample, lactate concentration, blood pH and HDL/LDL cholesterol were measured in collaboration with the Analytical Chemistry Lab of the CHU. Lactate, glucose levels and HDL/LDL cholesterol values were obtained using dedicated enzymatic dosage kits from Alinity® (Lactic acid, Ultra HDL, Direct LDL and Glucose reagent kit). Measurement of blood pH was obtained using GEM 500 Premier (Werfen) analyzer. Collected plasma samples were conserved at -80°C prior samples preparation and proton NMR metabolomics analysis.

OCT images analysis

Analysis of OCT images (**Figure 78**) has a crucial importance since it'll give access to the best visualization of the pathological event occurring during the follow-up of patients. All data were acquired on Heigelberg HRA+ OCT by the team of the EOL. Analysis of images were made by our team and followed guidelines described in the literature to assess AMD status[65,74]. Based on these, 3 main markers were followed: intra retinal cystoid fluids (IRC), pigment epithelium detachment (PED) and subretinal fluids (SRF).
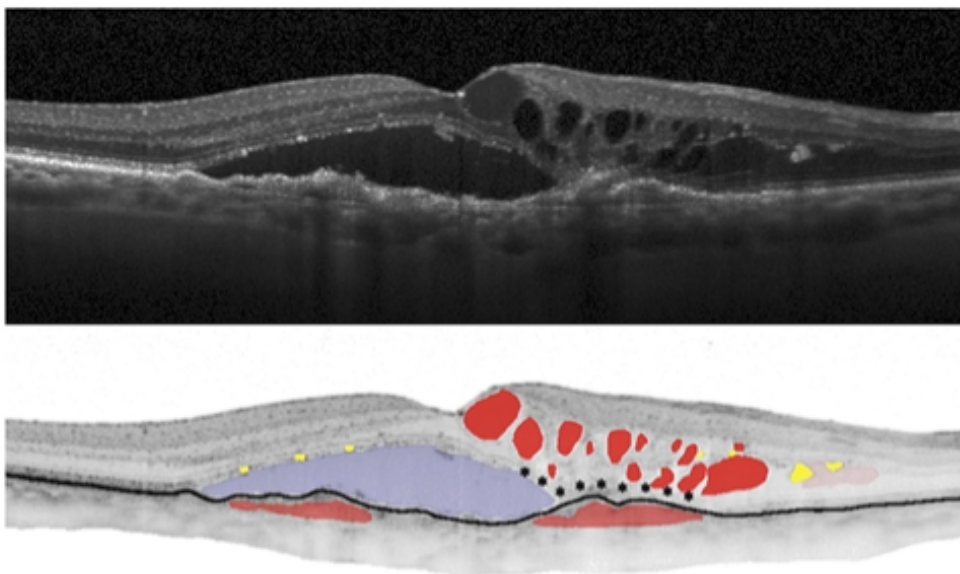

Figure 78. Makers of advanced age-related macular degeneration that can be highlighted from optical coherence tomography images of retinas. In bright red can be found the intra retinal cystoid fluid (IRC), sub retinal fluid (SRF) in blue and pigment epithelium detachment in dark red[74].

From each analysis, 12 scans were taken from each eye and from each scan were measured: the surface of each IRC spot (in mm^2), SRF spot (in mm^2) and the largest pigment epithelium detachment section were calculated from the initial pigment epithelium baseline (in μm). From the 269 visits, a total of 6456 images (12 scan/eye/visits) were generated and 19368 measurements (3 measure/scan/eye/visits) were done to provide a single value for each parameter at single time points. As described in the introduction (*section 1.2.2*), IRC and PED measurement can be used to monitor nAMD progression. Indeed, presence of IRC fluids is linked to CNV processes that is the hallmark of nAMD and PED growth is reported to be associated to long-term vision loss in flexible treatment regimen[74]. On the other hand, the presence of SRF fluids is a sign of healing retinas and is the only measured parameter that is correlated with an improvement of visual function[74].

NMR metabolomic analysis

^{1}H-NMR metabolomic profiles were obtained from plasma samples collected during the two years follow-up and conserved at -80°C. All samples were measured at 298K on a Bruker Avance (or Neo) spectrometer operating at 500,13 MHz for proton detection. The NMR instrument is equipped with TCI 5mm cryoprobe equipped with *Z* gradients. Maleic acid was added to samples as internal standard allowing quantification and trimethylsilyl-3-propanoic acid-d_4 sodium salt (TMSP) for ppm calibration. For sample preparation, 500μL of plasma were mixed with 200μL of deuterated phosphate buffer added of 100μL of 35mM maleic acid solution and 30μL of 10mg/mL TMSP solution in D_2O. Final solution is vortexed and placed in a new 5mm NMR tube before being analyzed in NMR. For the NMR analysis, an edited 1D-CPMG sequence with water presaturation was used: RD-90-(-t-180-t)n with relaxation delay of 4s (RD), spin echo delay of 300ms (t) and 128 loop (n). Pulse for water presaturation is occurring during the relaxation time (RD) and the number of scans fixed at 64. The total acquisition time is of 3.1981568s with 4 dummy scans.

Due to an electronic update (from Avance to Neo), the parameters had to be updated due to signal coupling issues and the spin echo delay was set to 400ms (t). The total acquisition time is of 3.2767999 with 4 dummy scans.

All acquired data were processed using PepsNMR® software, an open-source R package developed by Manon Martin and Bernadette Govaerts from University of Louvain-la-Neuve and our group. After Group Delay Correction, Solvent Suppression, Apodization, Fourier Transform, Zero Phase Order Correction, Internal Referencing, Baseline Correction

(lambda.bc= 10^5), Negative Value Zeroing and Window Selection, the processed spectra were bucketed from 0.5 to 10 ppm with a bucket width of 0.02 ppm. After Region Removal (4.6 to 5.1 ppm for water and 5.85 to 6.3 ppm for AM) and Normalization (mean), the data matrix consisted of 438 spectral zones of 0.02 ppm width containing metabolites information for each sample.

As it was be explained before in this chapter (see section 4.4.1), all samples were analyzed again due to some changes in the experimental NMR sequence that are linked to the upgrade of our analytical platform and led to un-avoided variability between the samples. For this, all NMR samples were filtered through wash-up 10K filter for 60 min at 13000 rpm at 4°C. 270μL of filtrate were added with 30μL of Calcium formate 5mM solution and placed in 3mm NMR tubes. 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.

This step assures the elimination of all lipids and proteins moieties from plasma samples and the NMR sequence used allowed NMR quantification using ChenomX® software.

Lipoproteins profile evaluation

From data recovered using 1D-CPMG NMR analysis of plasma samples, the NMR lipoprotein profile was evaluated using the methods described in the *chapter 3 section 3.5*. Then, we determined the chemical shift corresponding to 4 lipoprotein fractions: F1 = 0.92 ppm (mainly VLDL), F2 = 0.91 ppm, F3 = 0.89 ppm, and F4 = 0.88 ppm (mainly HDL). The signal intensity at these different chemical shifts was measured and then normalized to the total intensities of all fractions to reduce the impact of the global lipoprotein concentrations that could differ between samples. Thus, the obtained values are representing a fraction of the total signal and allow the comparison between the different lipoproteins profiles across all plasma samples.

ChenomX® metabolites quantification

Metabolites concentration was measured by ^{1}H-NMR using spectral data from the analysis of filtered plasma samples. Spectral deconvolution was achieved using the ChenomX® NMR suite software in profiler mode by manually fitting the resonance peak of 61 metabolites. The quantification was based on the signal of the chosen reference, calcium formate at 8.46 ppm. Since samples were not prepared following the guideline of ChenomX® SOP, the concentration values are not absolute but therefore relatives. Anyway, this step allows us to analyze the variations among the different metabolites' concentration among all measured samples.

Statistical analysis

All statistical analysis, tools and codes used in this section can be viewed in the following GitHub repository: *https://github.com/MS28uliege/These_MS_Rmd*. All analysis were made using dedicated R packages. Unsupervised and supervised multivariable models were generated using MixOmics packages as well as PLS1 regression models. Correlation plots were generated using *cor* function and *pls* package used for PCR analysis.

Chapter 5:

Identification of Biomarkers for AMD risk assessment: Prospective cohort from MIRA project

5 Identification of Biomarkers for AMD risk assessment: Prospective cohort from MIRA project

5.1 Introduction:

Over a lifetime, nutritional and environmental factors can alter metabolism and putatively influence the occurrence of AMD and its progression through its last stages[44,250]. As better medical outcomes are obtained with early diagnosed patients[215], identifying new biomarkers able to early predict the occurrence of AMD would be very interesting in view of preventive medicine. If several studies aimed to explore AMD through metabolomics approach, none of them were based on prospective cohorts including patients without any sign of AMD[179–181,251–257]. The objective of the present study, carried out in collaboration with the University of Bordeaux and within the framework of the MIRA project, is to identify metabolic signatures associated with the risk of developing AMD (all stages included) by using untargeted NMR-based metabolomics.

Samples were selected from the ALIENOR[216] cohort and plasma samples were provided by Inserm of Bordeaux Unit U1219. All NMR analysis and data pre-processing were conducted at the University of Liege by the Metabolomics group of the CIRM (Center for Interdisciplinary Research on Medicine) and, after exploratory analysis for outliers' detection, the data were sent to INSERM of Bordeaux (Population Health Research Center Unité INSERM U1219) for statistical analysis.

Once spectral zones were found to be relevant for risk assessment for AMD, we were in charge for the identification of potential biomarkers related to these spectral zones. By this approach, we aimed to identify one or several early biomarkers associated with the risk of developing AMD and to contribute to a better understanding of complex processes that play a role for the development of AMD.

5.2 NMR-based metabolomics:

5.2.1 Dataset

The cohort from MIRA project (**Figure 79**) consist of plasma samples coming from 327 healthy individuals aged of 73 y/o or more at their first ophthalmologic examination (2006-2008). All subjects didn't exhibit any sign of ocular disease at the time of enrolment. They underwent to ophthalmologic exams 4 times with two years gap between each visit during the entire follow-up period (2009-2017). During this period, examinations allowed to assess the progression of healthy individuals through early or advanced AMD. Moreover, information about genetic background, lifestyle, and environmental factors (tobacco uses, sunlight exposure) were collected. More information about the cohort and its origins can be found in the "Materials and Methods section 5.5.1".

Proton NMR metabolomics approach allowed to analyze the metabolic profile of each plasma samples. In house pre-processing tools was used to construct robust dataset in which each individual was described by 411 spectral zone (buckets) containing metabolites information. The former dataset will allow the use of dedicated multivariate and univariate analysis able to capture the metabolic variations that could appear prior the development of AMD.
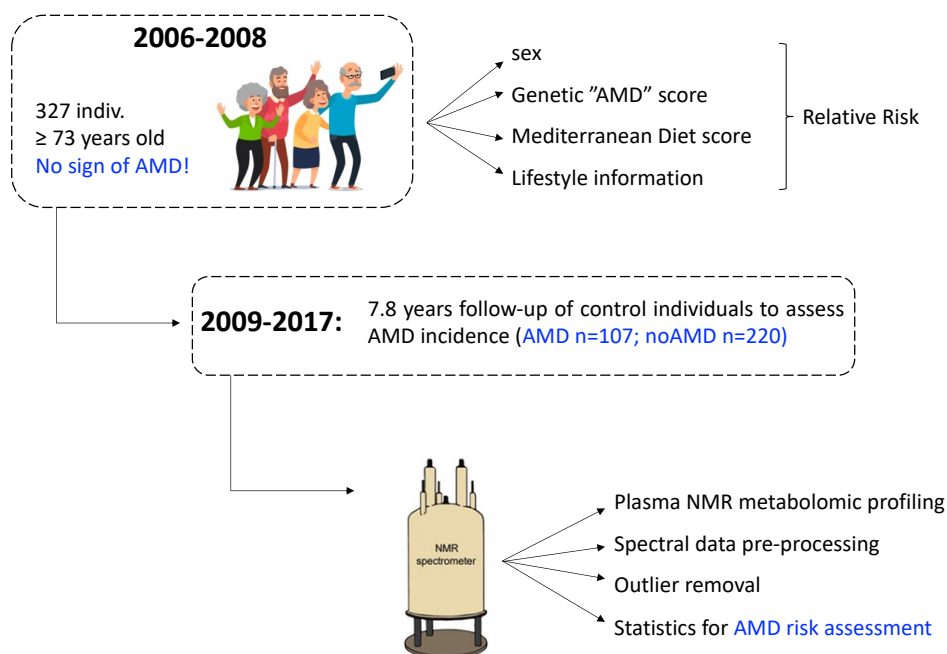

Figure 79. Schematic representation of MIRA project pipeline.

5.2.2 Multivariate exploratory analysis

Principal component analysis (PCA) model was built using the plasma metabolic NMR profiles coming from all patients included in the Mira project to identify source of variations and outliers. From all people included in the study (n=327), 7 exhibited bad spectral data and were not included in the analysis. By analyzing the first PCA model generated, 3 outliers were identified as they exhibited higher contribution for the buckets 1.189, 3.669 and 1.209 (**Figure 80**). These variables are characteristic of the presence of ethanol that is sometimes used during the sampling step to clean the zone of blood taking.
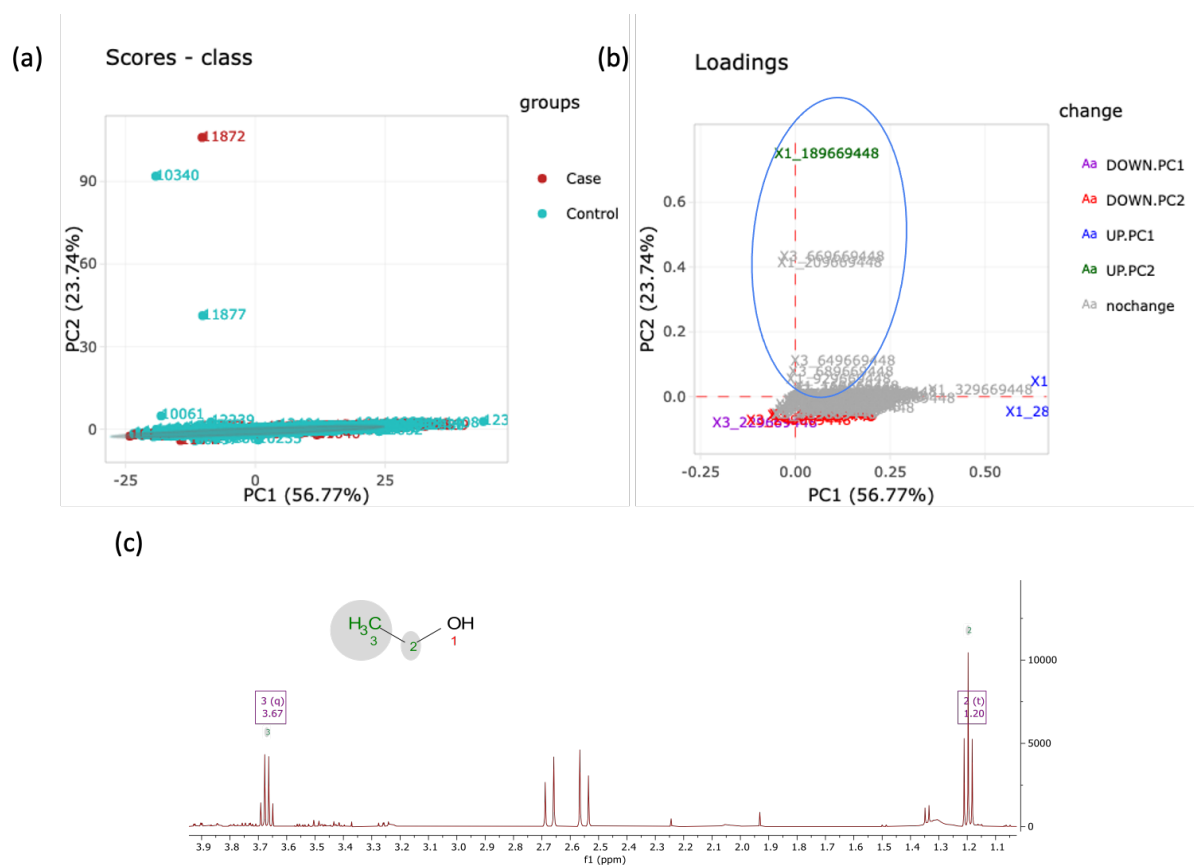
Figure 80. (a) PCA score and (b) loading plot generated with NMR metabolic profiling analysis of patient's plasma samples, control individuals are in blue and case patients, people who had developed AMD during the study, in red. On the score plot the outlier individuals are easily recognized and the buckets involved are spotted on the loading plot. The NMR spectra of the sample coming from the individual 11872 (c) allow the identification of the ethanol as source of contamination. According to this information, these samples were not included in the remaining analyses.

Thus 3 samples were removed from the dataset and a new PCA model was generated (**Figure 81**). From this plot no clear separation can be spotted between control patients and those who had developed AMD during the study. The dataset is well represented, and the loading plot gives information about how samples are different, but no relevant information could arise from this analysis.
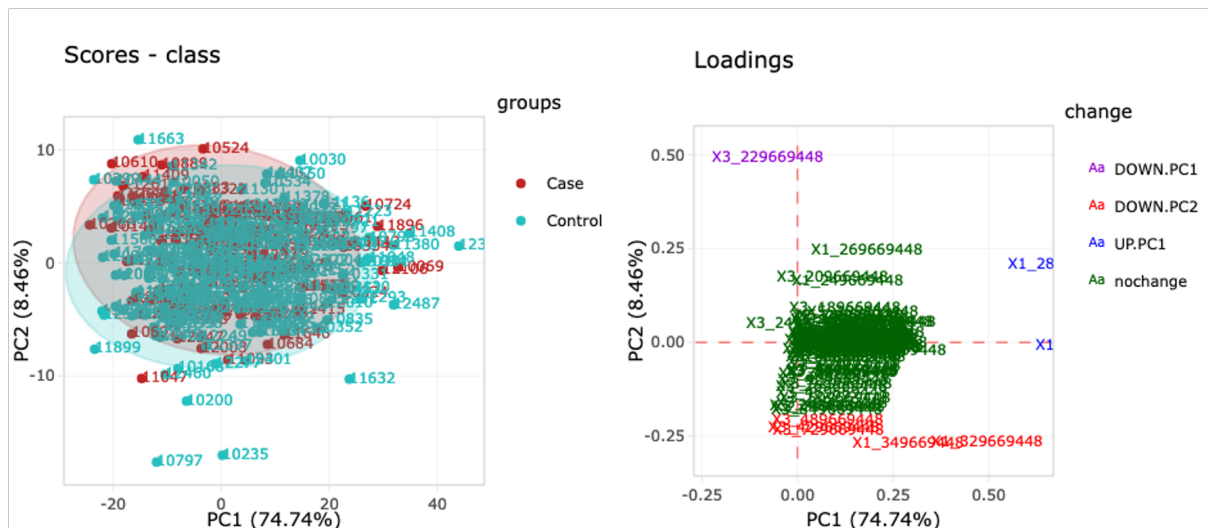
Figure 81. PCA score plot (on the right) and loading plot (on the left) generated from plasma metabolic NMR profiling analysis after outlier removal. No discrimination can be seen between the two groups and as the loading plot shows, only few variables are contributing to the model. This highlights the fact that most of individuals have a close metabolome and only small differences can be found (explained variance of each PCs: PC1 = 74.7, PC2=8.4, PC3=5.3, PC4=2.1).

OPLS-DA models are supervised analysis that allows to investigate the variables able to explain the separation between the two groups. On the score plot a separation is revealed and, even though a big overlap is existing between groups, some variables can be identified (**Figure 82**). Thus, from the loading plot and the VIP score table, four variables are picked out for their ability to discriminate patients with AMD prevalence. These variables are listed on the **Figure 82** and are all part of lipid's related signals indicating that such species could be interesting. Then, a permutation test was made to assess the performance of the model. With a p value of 0.0104 over 100 permutations, the model shows an encouraging performance, and the spotted variables are worth considering.
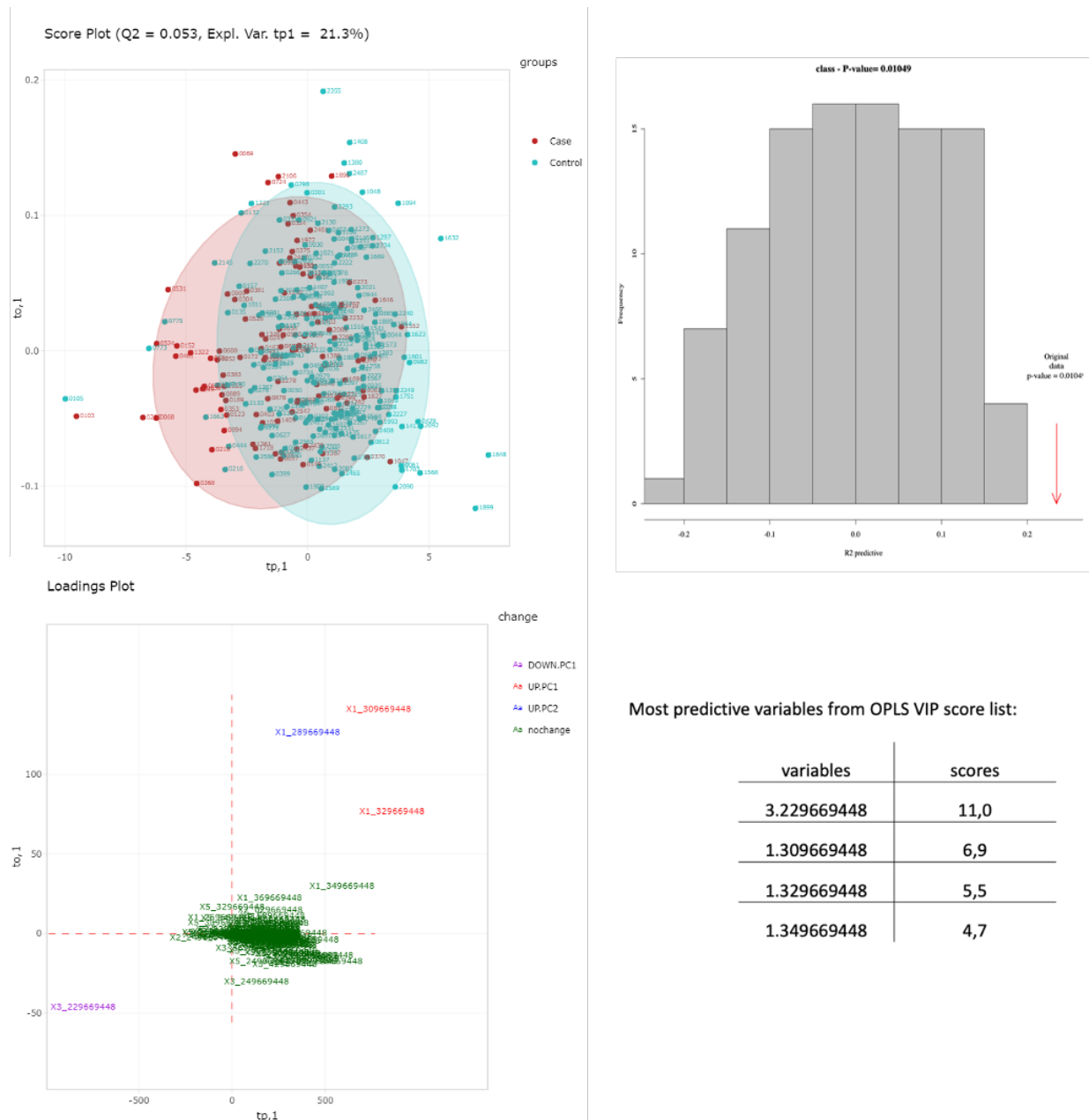
Figure 82. On the top left an (Kernel) OPLS-DA based (Q2= 0.053) on the NMR metabolic profile from plasma samples of patients that have developed AMD during the study (in red) and control patients (in blue). Below on the left the corresponding loading plot with in red the variable positively correlated with tp1 and in purple the variable negatively correlated with tp1 (Expl.var.tp1= 21.3%). On the top right the permutation test (method: K-fold, 100 permutation) with p-value: 0.01049 assessing the performance of the OPLSDA model. Below on the right the most predictive variables from the VIP score list able to predict the AMD outcomes based on the metabolic profile.

When looking at multiple t-test analysis, different spectral zones seem relevant to explain the risk of developing AMD since these are higher in the "case" group (**figure 83**). Interestingly, all these buckets are linked to lipids species and further analysis will confirm their usefulness/ability to predict AMD events. Indeed, by associating this information with data collected from the individuals regarding their ages, sexes, genetics, and lifestyle, we aim to better explain these variations in the metabolic profile and their association with AMD prevalence.

Variables	Case - Control p-value
X5_289669448	0.00017
X5_129669448	0.00026
X5_269669448	0.00026
X5_169669448	0.00037
X5_149669448	0.00047
X5_189669448	0.00118
X5_209669448	0.00124

Figure 83. Most important variables from Multiple t-test (with FDR correction) analysis performed on data from NMR metabolic profiling experiment on plasma samples from p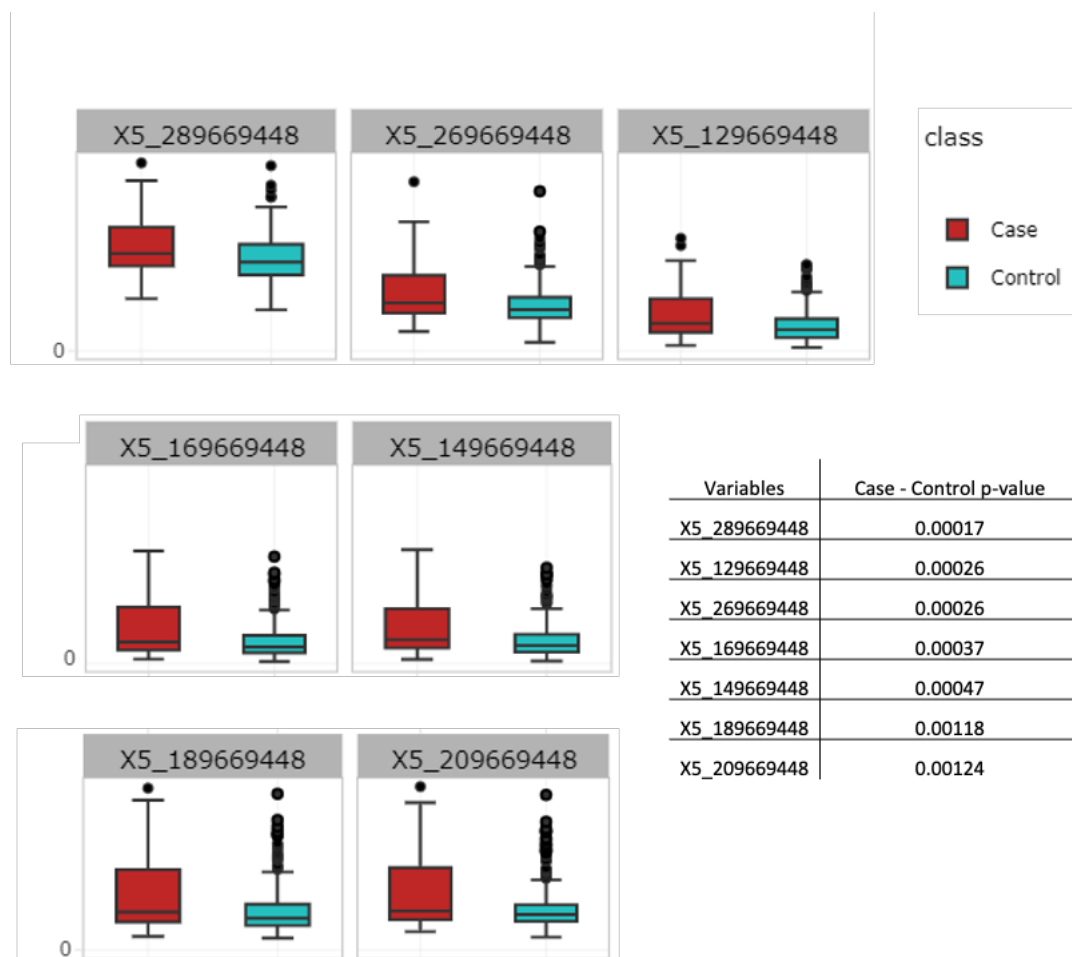atients who had developed AMD in red and control patients in blue. Important variables are listed with their corresponding p-values.

5.2.3 Univariate Analysis for risk assessment (statistical results provided by Soufiane Ajana)[258]

The analysis of the association of each bucket with the risk of early or advanced AMD after adjusting values based on the sex, level of education, tobacco use, the genetic score and the Mediterranean diet score allowed the selection of 7 variables of interest (**Table 7**). Each of the selected buckets had a relative risk (RR) near 2 and higher the value of RR, higher the risk of developing AMD is.

Buckets	RR (95% C.I)	P-value
5,12967	2,23 (1,50 - 3,32)	0,01
5,14967	2,09 (1,44 - 3,02)	0,01
5,16967	1,99 (1,40 - 2,83)	0,01
5,18967	1,92 (1,36 - 2,71)	0,01
5,20967	1,93 (1,36 - 2,75)	0,01
5,26967	1,89 (1,37 - 2,61)	0,01
5,28967	1,93 (1,35 _ 2,73)	0,01

Abbreviation: RR, Relative Risk; C.I, Confidence Interval

Table 7. Selected variables associated with early or advanced AMD incidence[258]. P-value are adjusted for test multiplicity using Benjamini-Hochberg methodology. All models were adjusted on the sex, the education level, tobacco use, genetic score and Mediterranean diet score.

Figure 84 shows the boxplot for each of the selected buckets having higher values for individuals that have progressed through an early or advanced form of AMD compared to healthy individuals. **Figure 85** shows the correlation between buckets and demonstrate a strong positive correlation between buckets 5.12967, 5.14967, 5.16967, 5.18967 and 5.20967 while buckets 5.26967 and 5.28967 ppm are part of their own cluster.

To go further in the work provided by Soufiana Ajana, please refer to: https://theses.hal.science/tel-03415718 or scan the QR code below:
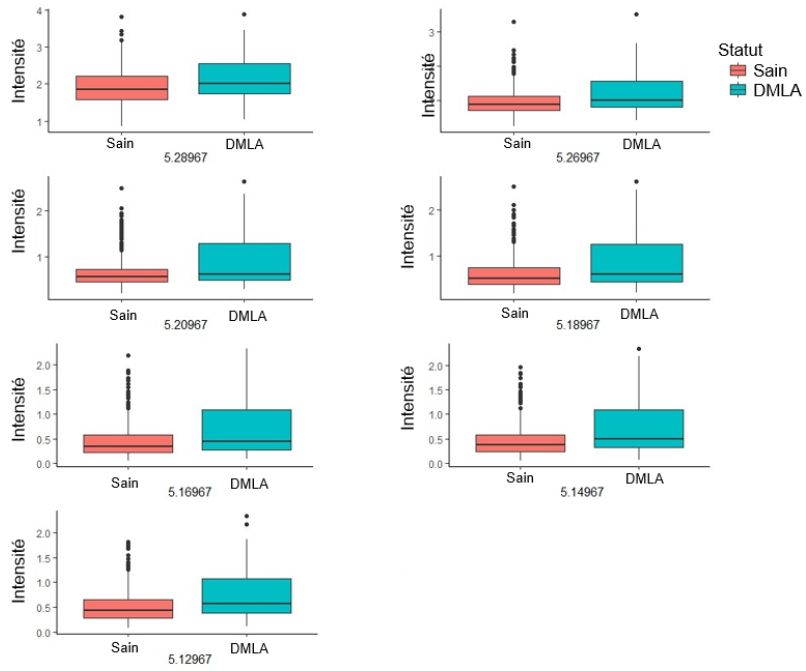
Figure 84. boxplot of selected buckets with increased relative risk for AMD progression[258].
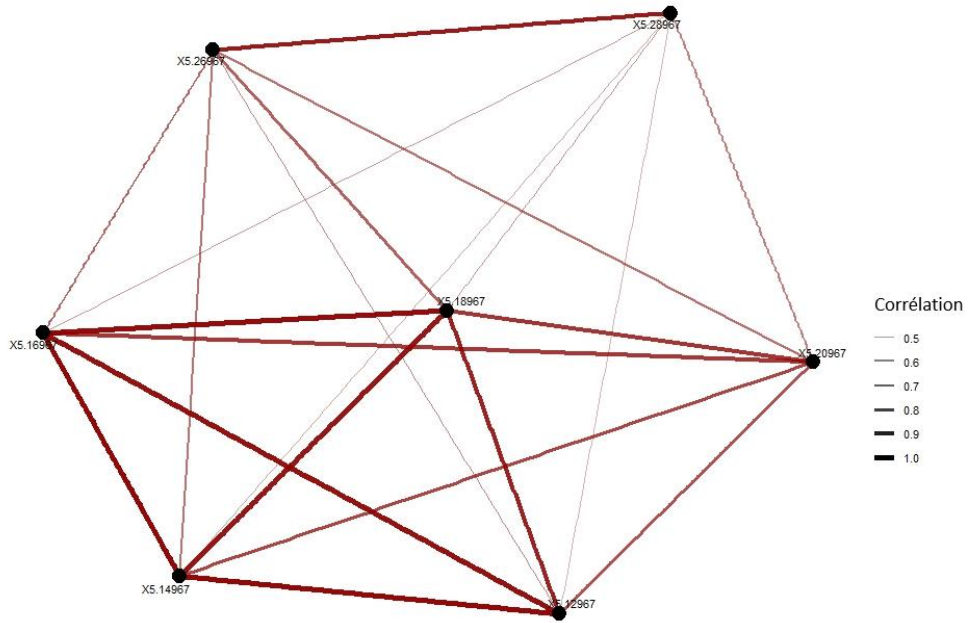

Figure 85. Buckets correlation analysis showing two distinct cluster. Strong correlations (near 1) are showed in dark red while weak correlations (down to 0.5) are in pale red[258].

With this information in hands, we aimed to identify the metabolites behind these spectral zones. All the selected buckets were in the same spectral window, a poorly resolved zone situated near the anomeric peaks of the glucose (proton attached to C2 of α-glucose). Indeed, this spectral zone is mainly composed of signals coming from varieties of lipid moieties, hence the broader and overlapping signals didn't allowed clear identification. However, based on the literature, it's possible to determine which lipids are contributing to these signals: C*H* from Cholesterol/Cholesterol ester; -C*H*=C*H*- from Fatty Acid chains and C*H* from Glycerol backbone C2 (**Figure 86**)
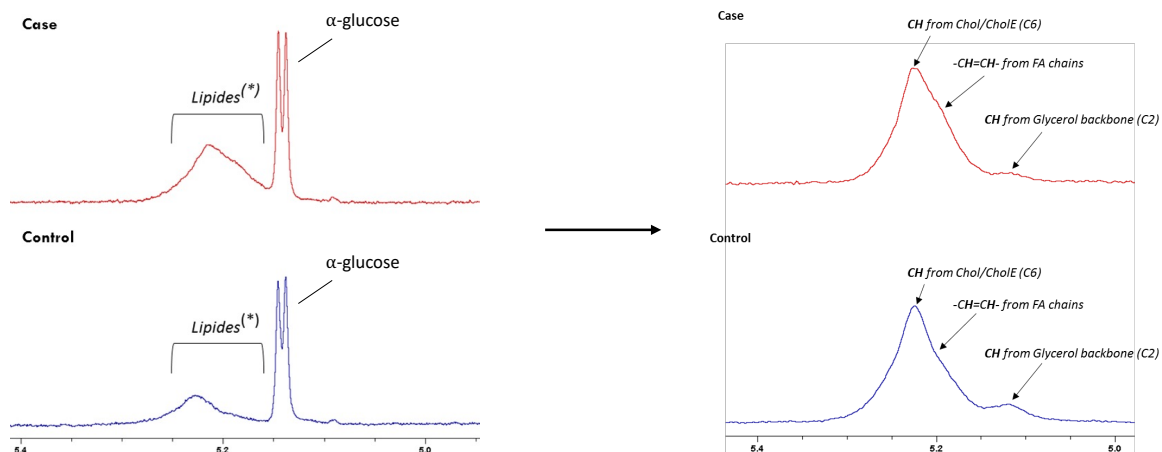
Figure 86. Spectral zones of the selected buckets and the compound having spectral component in these zones. On the left the comparison between 1H-NMR spectra from plasma sample on a healthy individual (control) and an individual that have developed AMD (case). On the right the different lipid moieties having NMR signal in the spectral zone associated with greater risk for AMD progression.

As we demonstrated with our study about lipoprotein profiles changes (see chapter 3)[225], lipid metabolism could play a role in the development and progression of AMD. It's not surprising if change in lipid profile is linked to the diet and it proved that the diet play a role in AMD[1,55,59]. Moreover, several studies already published association between lipid and AMD[180,208,209,213] showing the interest of these feature for risk prevision. However, further analysis must be done to clearly identify the lipids associated with the risk of developing AMD. Indeed, this information could give clues to understand of the complex processes involved in the pathogenesis of AMD. Furthermore, quantification of lipids moieties could be added to predictive models in order to refine them and increase their efficiency to detect individuals with high risk of AMD *(such predictive model already exists and allows to patients to calculate their own risk for developing AMD based on their lifestyle, genetic test, … see www.macutest.net for more information)*.

5.3 NMR-based Lipidomics:

The role of lipids in cell, tissue and organ is crucial since an altered lipid metabolism is associated with several diseases including diabetes, cancer, neurodegenerative and infection disease and in degenerative disease such as AMD. It is therefore not surprising that the concept of "Lipidomics" emerged from related -*omics* sciences. Thus, lipidomics can be defined as "the study of the structure and function of the complete set of lipids (the lipidome) produced in a given cell or organism as well as their interactions with other lipids, proteins, and metabolites"[127].

Even if NMR-based lipidomics methods are less sensitive and suffer of severe peaks overlap, numerous studies reported the use of ^{1}H-NMR for investigating changes among lipid profile[127,128,259–261]. In this work we decided to extract lipids from our plasma samples through solvent-solvent extraction step and analyze the extract by ^{1}H-NMR to perform discriminant statistical analysis on the generated data and identify lipid species responsible for groups separation (**Figure 87**).
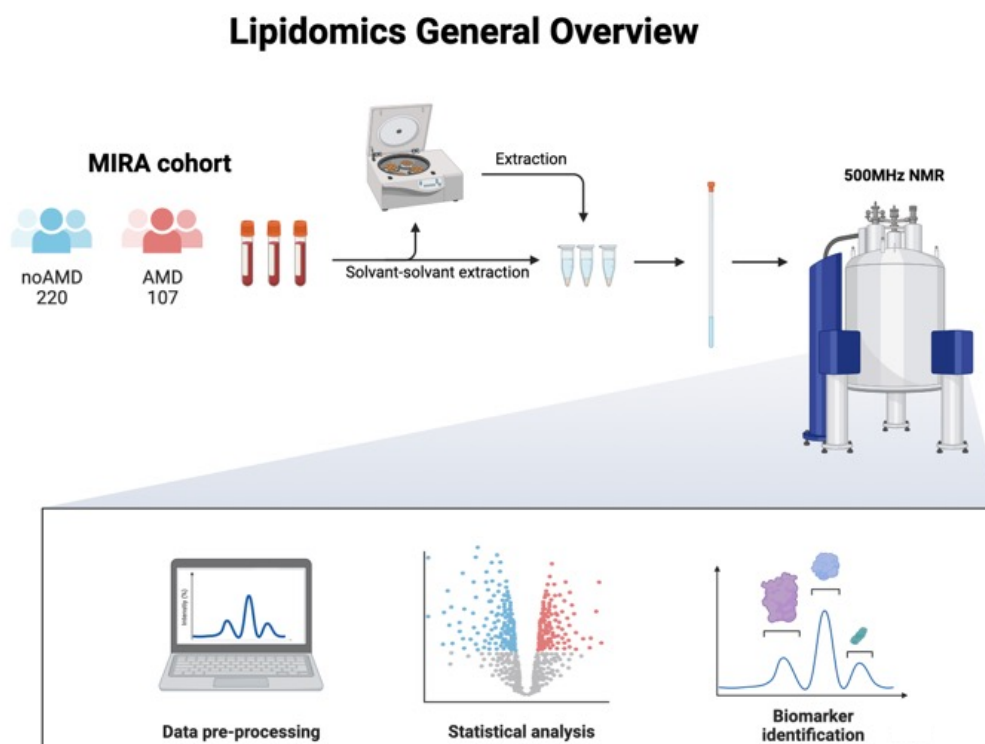

Figure 87. General lipidomic workflow for AMD risk assessment.

5.3.1   Univariate statistical analysis for risk assessment (Soufiane Ajana)[258]

By using the same strategies than for the metabolomics data, none of the generated buckets were found to be relevant by the team of Bordeaux in charge of the statistical analysis. Moreover, when they analyzed the correlation between the new matrix and the previously selected buckets related to the metabolomic profile, only weak correlations were found (**Figure 89**). We explained that lack of correlation and the absence of significative result by an inappropriate processing step.

Indeed, the NMR spectra coming from lipid extract analysis are quite different than those obtained while using classical metabolomics methodology (**Figure 88**). Hence, this analysis results in a poorly resolved spectra where all the information is condensed in a small part of the spectral width (**Figure 90**). Thus, applying a classic bucketing step results in a huge number of useless variables that have no significance. Furthermore, the highly overlapping signals coming from the different lipids' classes lead to non-informative buckets when they are taking at fixed intervals as each bucket that are containing signals can be attributed to different lipids compounds off different lipids' classes. On **Figure 90** we are comparing spectra coming from (88a) classic total plasma analysis and (88b) extracted apolar lipid fraction. As we can see, on the same ppm range, analysis of total plasma sample allows to recover the information about approximatively 7 different polar compounds while the spectra from lipid extract only give information about Glycerol and Glycerophospholipids moieties.
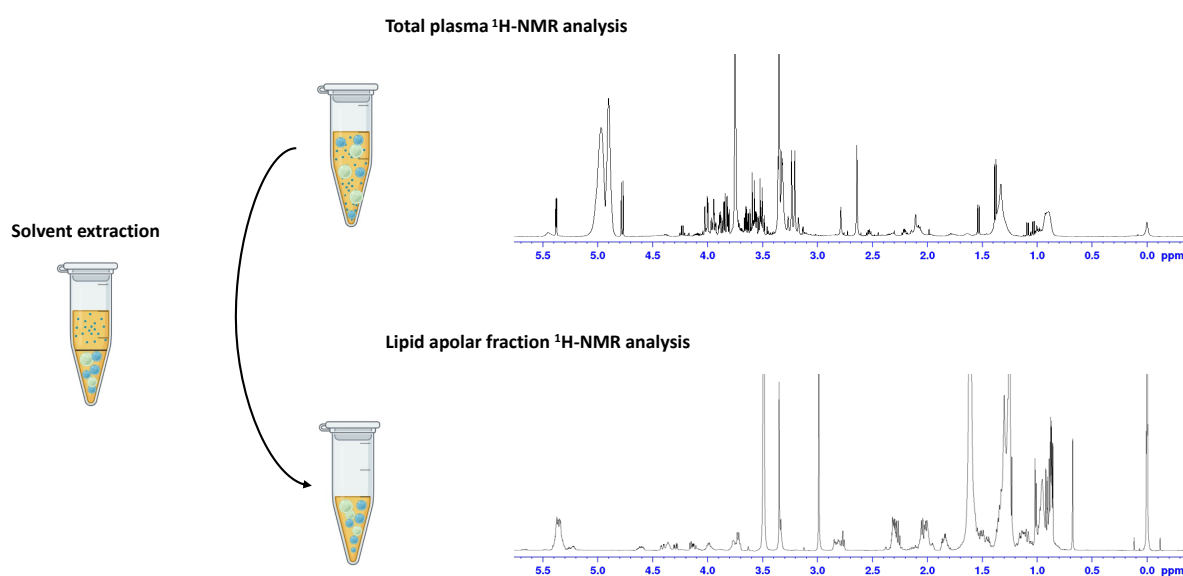
Figure 88. Examples of NMR spectra coming from intact plasma analysis (on the top) and the analysis of lipid apolar fraction recovered from solvent extraction step (on the bottom).

Buckets		
	RR (95% I.C)	P-value*
5,12967	2,23 (1,50 - 3,32)	0,01
5,14967	2,09 (1,44 - 3,02)	0,01
5,16967	1,99 (1,40 - 2,83)	0,01
5,18967	1,92 (1,36 - 2,71)	0,02
5,20967	1,93 (1,36 - 2,75)	0,02
5,26967	1,89 (1,37 - 2,61)	0,01
5,28967	1,93 (1,35 - 2,73)	0,02

Abréviations : RR, Risque Relatif ; I.C, Intervalle de Confiance
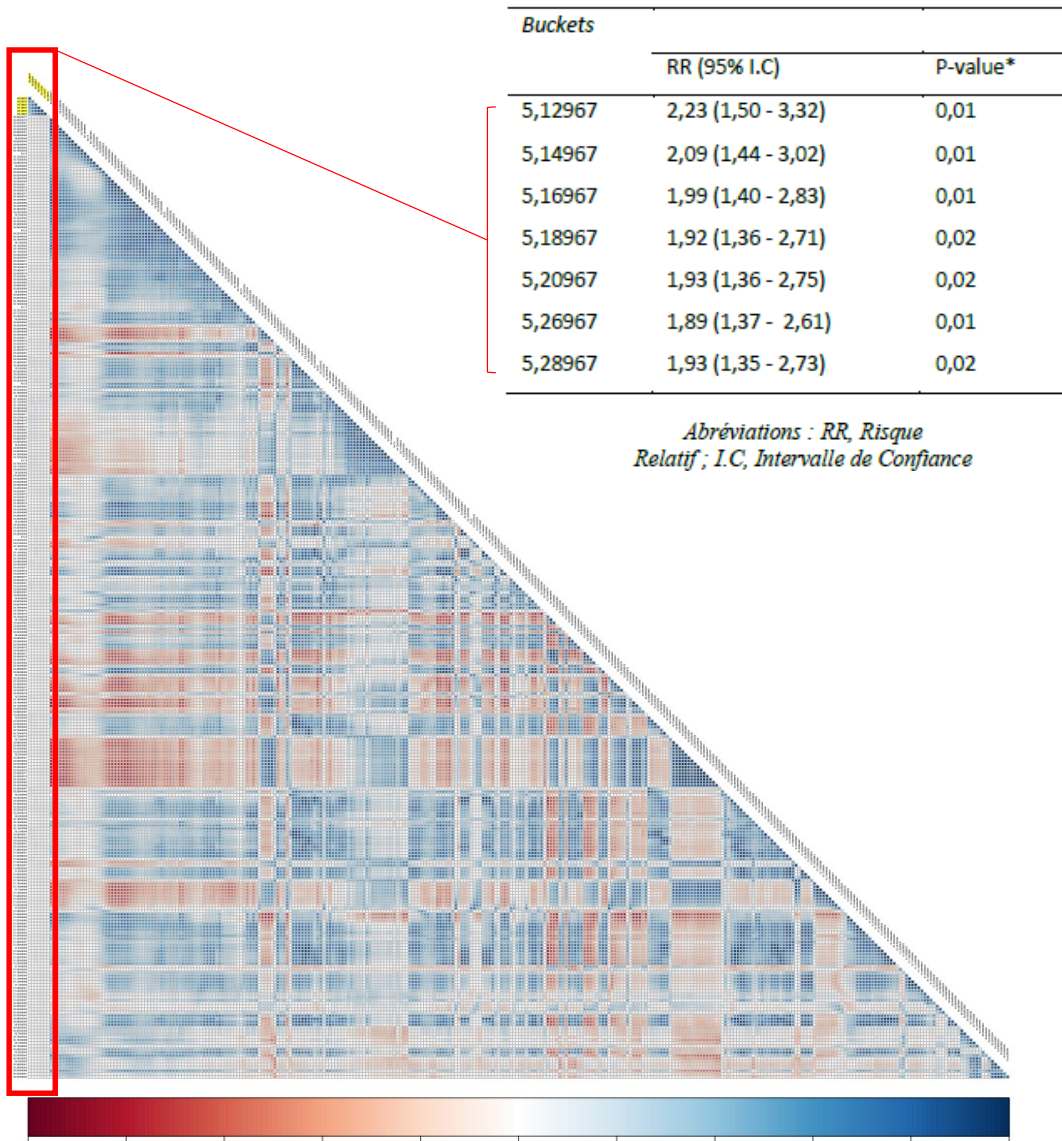
Figure 89. Correlation heatmap: bright and large circles show strong correlations and the color indicate the directions of the correlations (red: positive; blue: negative). As only correlation of 0.6 are worth to be considered, anything that would not appear as evident is not worth seeing[258]. On the top left, the discriminating spectral zone previously identified during the analysis of intact plasma. As shown here, none of the variables from the analysis of extracted lipids are correlated with the variables of interest.

(a)

Bucket ID	Compound
a	/
b	Lactate
c	Lactate
d	Creatinine
e	Serine, Phenylalanine, Asparagine, …
f	Serine, Tyrosine

(b)

Bucket ID	Compound
i	/
j	PL glycerol backbone
k	GPL
l	GPL + TAG glycerol backbone
m	TAG glycerol backbone
n	/
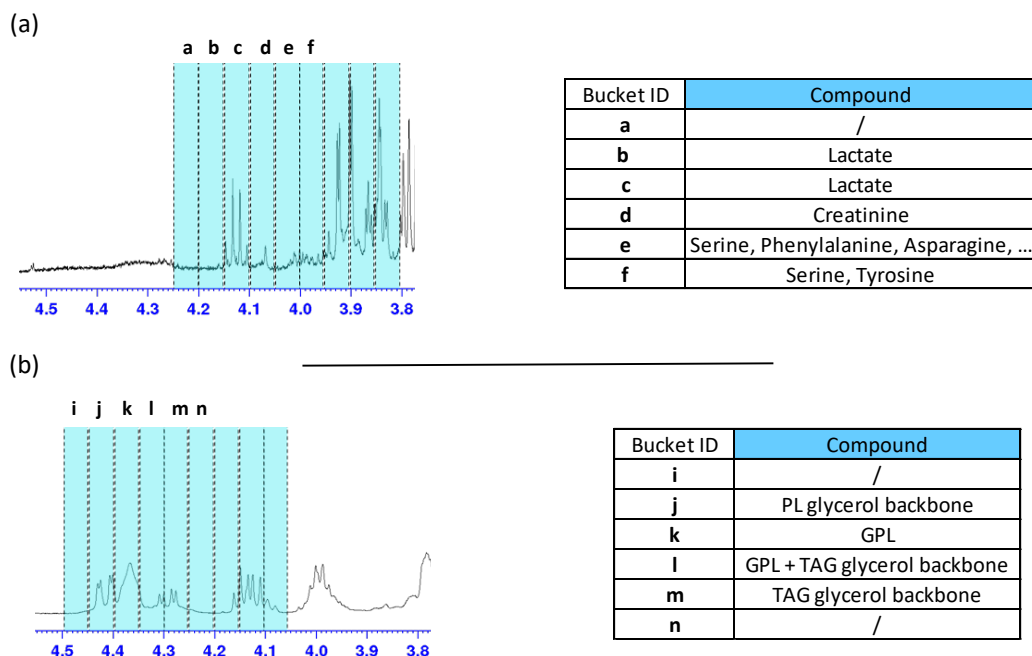
Figure 90. Significance of buckets arising from a traditional bucketing step on (a) plasma sample coming from a healthy donor and (b) the lipid extract coming from the same sample. While analysis of intact plasma samples give access to information on 7 compounds, the lipid extract provides only information about 2 lipid species, Glycerol (PL: phospholipids; TAG: Triacylglycerol) and Glycerophospholipids (GPL).

For these reasons, we decided to change our approach by considering only the interesting part of the spectra and thus reduce redundancy. By using an Adaptative Intelligent-bucketing approach, we aim to build a consistent data matrix in which samples are described mostly by independent variables. Indeed, by reducing the correlation among our buckets table, we hope to better describe the individuals regarding changes in their lipidomes. Moreover, as bucket from all the samples were normalized in respect to the internal standard, the generated buckets will relate to the concentration of the species having NMR component on these spectral zones.

5.3.2 Multivariate analysis for feature identification

In order to identify the lipids behind the zone identified as relevant to prevent AMD prevalence, we changed our bucketing approach to build a more reliable data matrix. In this data matrix, the lipid extracts of each individual are represented by 222 variables extracted from the Adaptive Intelligent-bucketing step. By allowing variable size bucketing and automatic detection of peaks edges, this method provide only relevant buckets that are containing signals, avoid peak splitting and reduce the overlapping of lipid species among the buckets. All the bucket's value were normalized in respect to the value of the internal standard (DMSO$_2$) to have information about the concentrations of lipids species. Hence, we aimed to produce a data matrix that can better individualize the different lipid species among the extract and identify putative lipids for risk assessment. After removal of bad spectral data, our dataset is containing 268 individuals in which 177 have no sign of AMD and 91 had developed AMD during the 7 years of follow-up.

We first performed a PCA analysis to characterize our data set, identify potential outlier and having an overview of how individuals are spread. When we look on the calculated PCA, no clear discrimination between the two groups can be identified and no outliers must be removed. The model is well describing the cohort of samples as the cumulative R2 is of 0,95 and the corresponding loading plot allows to determine how individuals are represented in the variables cloud. PC1 account for 53% of the whole variation among the dataset but, as shown in **Figure 91**, all variables have small weight on the components meaning that their contribution to the model is weak. All individuals seem to have a similar lipidomic profile as they are all located in the middle of the score plot and poorly resolved. A Pearson correlations heatmap plot was generated to investigate how the variables are correlated among the datasets (**Figure 92**). Similarly, with the analysis made on the previous dataset generated with a classical bucketing approach, all variables are strongly correlated together. This strong correlation explains why no clear discrimination between the individuals can be found in PCA analysis. Indeed, only a few samples located down to PC2 (**Figure 91**) seems to be different from the remaining of the dataset. By looking closer to the variables that are contributing negatively to PC2 (red variables spotted on the loading plot), we can observe that these variables are the ones negatively correlated with the other variables. This in fact due to spectra that contain a higher amount of methanol coming from the sample preparation step. This led to a greater distortion of the baseline near the solvent peaks that, even if these signals are removed from prior bucketing step, will give a higher contribution to noise signals. Therefore, as the buckets are normalized

upon the total intensity, a higher contribution of noise background will lead to smaller buckets intensity. Therefore, these variables are negatively correlated with all other variables as a reduction of noise will give higher buckets intensity and vice versa.
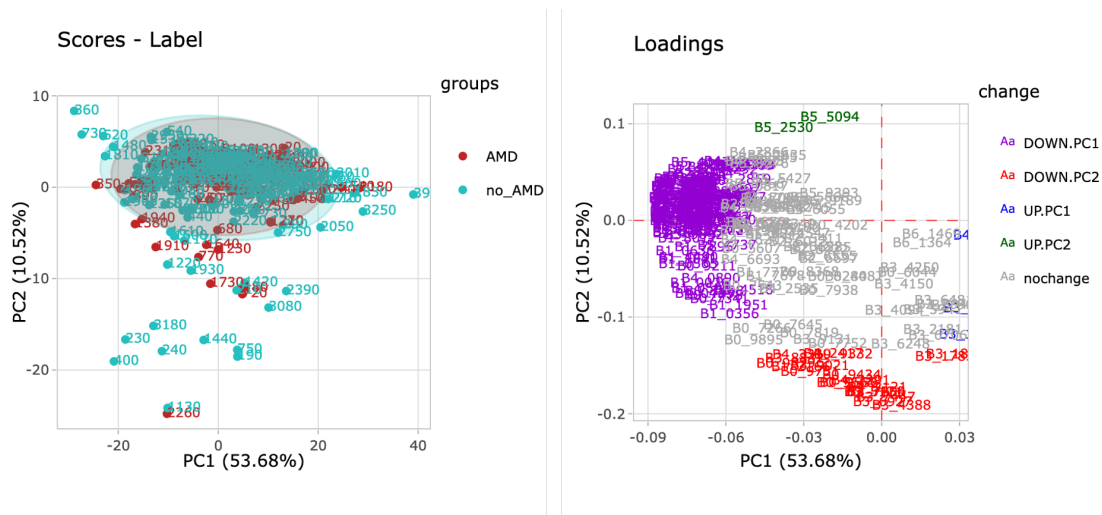
Figure 91. PCA score plot and loading plot of NMR analysis of lipid extract after IA-bucketing step. The score plot on the left allows an overview of the dataset and show how stacked together are the healthy individuals (in pale green) and the subject that has developed AMD (in red). On the right the loading plot identify the variables correlated with the different component of the PCA. Only a small number of variables weighted for the discrimination between the individuals, and most of the variables remain unchanged or only give small contribution.
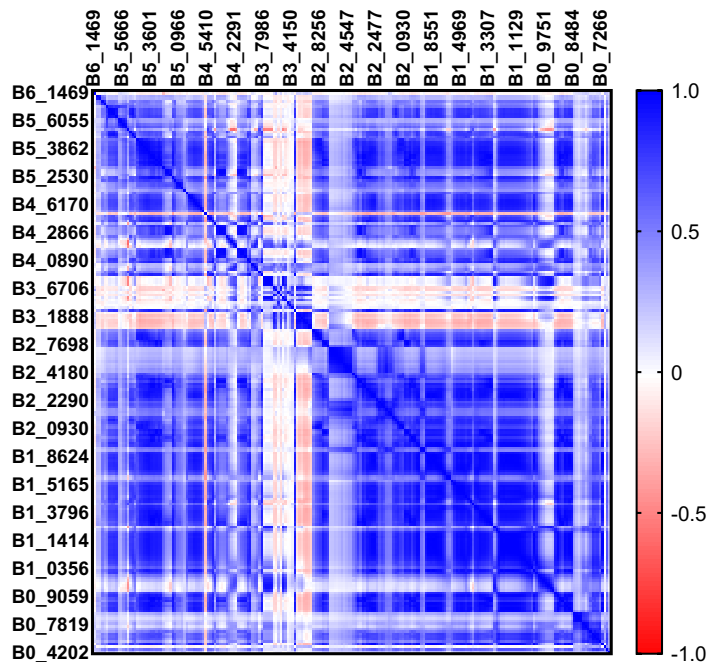
Figure 92. Person's correlations plot showing the correlation between variables (blue positive and red negative).

Our strategy didn't give the expected result and to explain this, several aspects of the dataset/strategies are worth considering:

- Identified spectral zone in intact plasma analysis were not immune from polar compound overlap.
- NMR analysis of lipid extract gives highly poorly resolved spectra.
- The expected changes would be small to fulfill the goal of the research.

First, it is worth mentioning that the spectral zone identified in intact plasma analysis could, aside from lipids, contain hidden signals of polar compounds. These compounds could have been eliminated during extraction step and therefore our samples would have been cleaned from the relevant information. Anyway, NMR analysis of lipid extract gives poor resolved spectra in which all signals coming from the different lipids' species are highly overlapped. Even if several studies were able to identify changes in the lipidome using NMR, most of them were using 2D approach, spectral deconvolution strategies or involved the detection of specific features among samples[128,262,263]. Regarding this information, using classical bucketing approach could only work if strong effect of the pathological condition is impacting the lipidome. If small effects are suspected, more advanced spectral data treatment and acquisition must be made to improve spectral resolution and overcome peaks overlapping.

Indeed, to evaluate how much one lipid species is contributing to the whole data matrix, one lipid extract sample was spiked with 5mM solution of standard Cholesterol and analyzed in NMR. The spectra generated is compared with the original sample and after using an identical pre-processing and bucketing approach than for the whole cohort, the bucket tables of the two spectra were compared. It appeared that 35% (74 over the 206 buckets) of the total amount of buckets were impacted by the addition of Cholesterol (**Figure 93**). Therefore, it's not surprising that the whole data matrix is highly correlated as each lipid classes exhibit similarities in their structural composition. The chance of finding one spectral zone corresponding to one single lipid species is quite small when bucketing even with improved approach is applied.
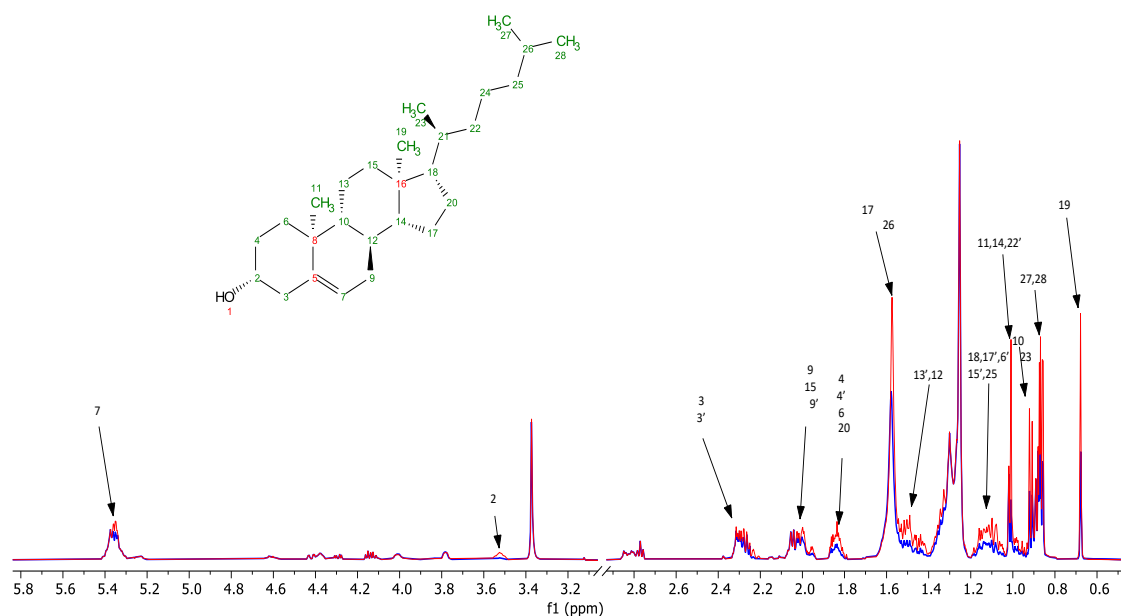
Figure 93. Contribution of cholesterol spiking on lipid extract samples. In blue the original spectra of the lipid extract of plasma sample and in red the corresponding spiking spectra. All signals were assigned using 1D proton NMR and 2D HSQC ^{1}H-^{13}C correlation spectroscopy.

In aging related pathologies, small changes in the metabolome/lipidome are usually the trigger of bigger event. These events are usually diagnosed lately and thus are difficult to prevent. Therefore, we need the most sensitive methods to detect these changes and being able to predict what can occur later. As this study aimed to identify changes in the plasma of an aged population that are linked to the risk of developing AMD over 7 years of follow-up, we need to be able to detect small changes and variability. For these reasons we need to analyze these data using innovative approach that gives access to more relevant information.

To solve these problems, and valorize the dataset, a spectral deconvolution approach could be used to quantify the different lipid species in our plasma lipid extracts. In the literature different tools exist and one retained our attention for its availability, its dedication for lipid analysis and the possibility to work with homemade spectral database. LipSpin is an open access Matbalb package developed by R. Barrilero *et al*[156] (**Figure 94**). This package allows the use of in-house spectral database comprised of lipid standard that will line fitted on the samples spectra. This package uses a constrained line shape fitting (CLS) algorithm based on Voight profile and template from standard lipids spectra to automatically analyze overlapped spectral region coming from lipids signals. Hence, Lipspsin provide the most complete quantification of lipids feature in lipid extract samples by ^{1}H-NMR to date and proved is usefulness in real-case

study[156]. By applying this approach in further investigations, we could access to unique information about changes in the lipidome that could be correlated with AMD prevalence.
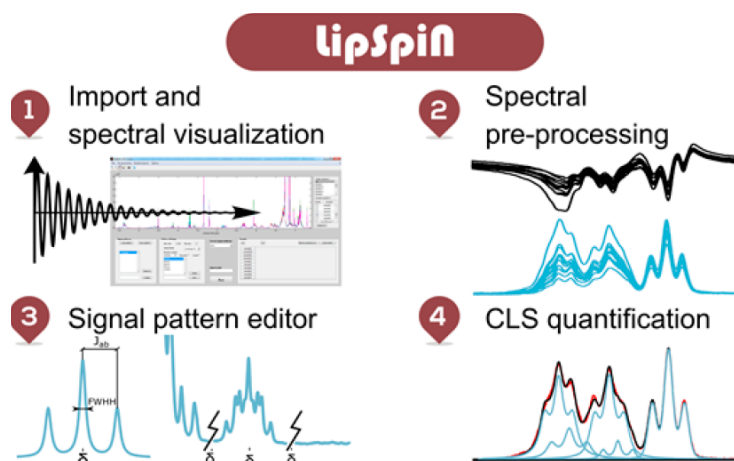
Figure 94. Representation of Lipspin package for lipids quantification in lipid extract samples[156].

5.4 Conclusion and perspective

By collecting plasma samples from healthy donors that were followed over 7 years of ophthalmologic examination follow-up, we aimed to correlate changes in the metabolic profile able to predict AMD issues. This cohort contained plasma samples from people that didn't exhibit any sign of AMD at the date of the enrolment. After 7 years of follow-up, some of them (n= 105) developed sign of AMD compared to the people that remained healthy from all signs of AMD (n= 222). Plasma samples were analyzed in ^{1}H-NMR spectroscopy and, after pre-processing and data reduction through bucketing approach, multivariate and univariate statistical analysis were performed (PCA, OPLS-DA and multiple t-test). Processed data were sent to Inserm of Bordeaux to analyze the association between each bucket and the risk of early or advanced AMD, by using adjusted multivariate Cox models.

From these analyses, a link between lipids metabolism and risk for AMD development was put in light. Indeed, from all analyses, the spectral zone that were able to explain the pathologic condition of patients with AMD were found to contain lipids related signals. However, the lack of resolution of these signals didn't allow us a clear identification of the lipid's species responsible of this effect. This forced us to adapt our strategies and re-analyze our samples to extract the lipid content of the sample and perform new ^{1}H-NMR analysis.

The same analytical strategies were used to process the data coming from the NMR analysis of lipid extracts and to analyze the data. At first, no result was obtained when trying to investigate the association between the generated buckets and the risk of developing AMD. Indeed, the highly correlated data matrix didn't provide useful information and an innovative data-processing strategy giving access to more informative bucket tables had to be done. By these methods we aimed to produce a bucket table containing relevant information about the different lipid's species, the intelligent-bucketing step used here was able to better individualize the different signals coming from highly overlapped regions of the spectra. Unfortunately, the data matrix remained highly correlated and no changes between the individuals were able to be found. The resulting analyses were thus poorly informative as the generated models were not able to differentiate properly the individuals.

Hence, despite all the different strategies used, we were not able to assess which are the lipids related modulation that were able to predict AMD outcomes. To solve this problem, a possibility is the use of an in-house lipid database and spectral deconvolution software that can recover the signal of specific lipid species from the overlapped spectral zone. Once these information in hand, changes among the lipidome of patients that have developed AMD could be spotted and used for risk assessment. Another option is the use of Mass Spectrometry lipidomic approach. Indeed, MS is the most suited tool for lipidomic purpose[264]. The analysis of the remaining samples by such method could provide better lipidome coverage that could be achieved with NMR approach. Indeed, using MS the information about more than hundreds of lipids' species could be recovered using referenced methods. Investigating the change in the lipidome of AMD patients represents a unique opportunity to better characterize AMD patients, allowing a better diagnostic and prevention.

As lipids are representing approximatively 20% of the dry mass of the retina, its thus not surprising that lipid metabolism and lipid oxidation is suspected to play a role in the pathogenesis of AMD. Hence, different studies have investigated the role of lipids in AMD and based on these, several preventive strategies have been tested involving different modification of patient's diet that impact lipid metabolism. For example, recent studies suggested that a Mediterranean-type diet may help to reduce AMD development and progression as these diets are rich in PolyUnsaturated Fatty Acids (PUFAs) and antioxidant. In other studies, the dietary intake of nutrient showed positive effect on AMD defense, but further investigation and clinical trials must be made to better assess and understand the role of nutrition for AMD management.

Photoreceptors require a high quantity of cholesterol to daily replace the shed membrane disk and, even if both RPE and photoreceptors can synthetize their own cholesterol, a large amount of lipids remain to be acquired from the blood circulation. As component of the retina blood-barrier, RPE will achieve this lipid uptake by controlling the efflux of lipoproteins into the retina. Its therefore of particular interest to investigate how the lipoprotein balance and its composition can alter their functional role in the retina. Moreover, determining whether the circulating lipoprotein balance affect the homeostasis within the retina is of great interest. More studies must be made to understand how change in the lipoprotein balance/composition is implicated in the onset and development of AMD.

Lipidomic studies gives a unique opportunity to identify new biomarker related to lipids metabolism and function and understand their role in AMD pathogenesis and development. The identification of such markers could help for risk assessment, patients' stratification across the different stage of AMD and clinical monitoring. Its therefore important to apply such methodologies in order to study cohort of patients that will or have developed AMD and have evolved through the different stage of the pathologies.

Worth considering, in this work we used NMR data from a subset of the 3C cohort. The 327 samples used for the MIRA project were thus part of the 1825 plasma samples received from the University of Bordeaux. Over years they collected information about genetics, lifestyle, and health of all these individuals constituting. Together with our metabolomic NMR data, this constitutes an interesting resource for clinical research. Hence, in the future, data could be reused in a different context and could provide interesting results.

5.5 Materials and Methods

5.5.1 NMR based metabolomics

Study population

All participants were selected from the ALIENOR cohort (**Figure 95**) and included people aged of 73 y/o and more at their first ophthalmologic examination (2006-2008)[216,235]. All these subjects underwent ophthalmologic exams 4 times after the date of enrolment with 2 years gap between each visit (since 2009 to 2017).

The ALIENOR (Antioxydant, Lipides Essentiels, Nutrition et maladies OculaiRes) epidemiologic cohort aimed at finding correlations between nutritional factors and ocular pathologies such as AMD, glaucoma, cataract or dry eye disease[216]. Other factors such as genetic factors, environmental factors (tobacco use, sunlight exposure) and vascular factors were also investigated. This cohort is part of the "étude des 3 Cités (3C)", an epidemiologic study started in 1999 with volunteers from Bordeaux aged of 65 y/o or more. From the 1450 members of the 3C cohort, 963 were enrolled for the ALIENOR study and at the end, 471 people with no sign of AMD at the date of enrolment were selected for our study (MIRA project).

From these 471 individuals, 71 had no NMR data, 50 didn't had genetics information and 23 no nutritional data. At the end, our prospective cohort included 327 subjects with 107 examination events during an average of 7.8 year of follow-up.

Clinic examination

AMD classification was based on retinal images data using an International Classification[12]. The incidence of early or advanced AMD was defined by the progression of healthy individuals that exhibited no sign of AMD in both eyes at the time of enrolment and evolving trough early or advanced AMD during the follow-up in at least one eye.
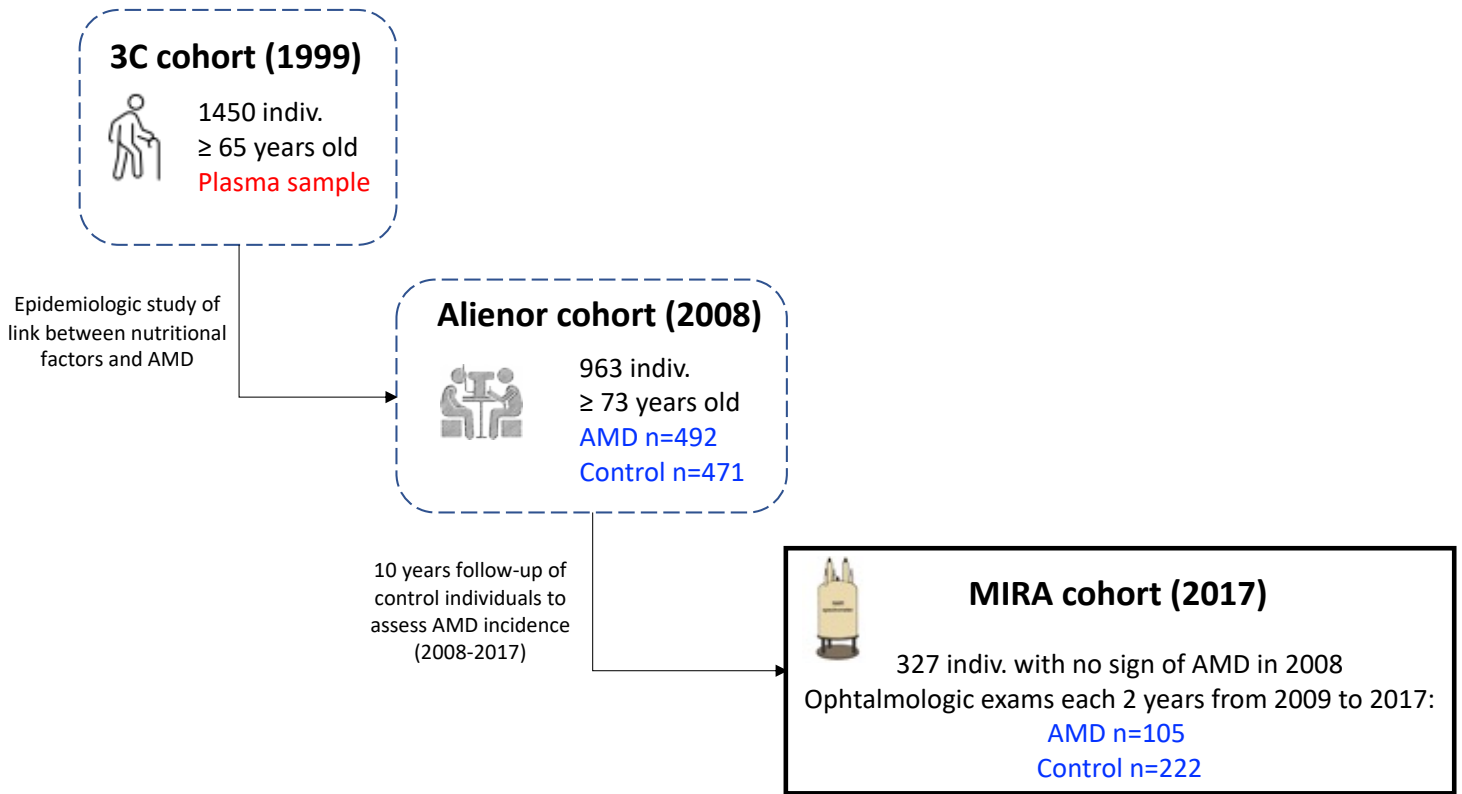
Figure 95. Schematic representation of origins of MIRA cohort.

NMR analysis

^{1}H-NMR metabolomic profiles were obtained from plasma samples collected at the enrolment time of the 3C study (1999-2001) and conserved at -80°C. All samples were measured at 298K on a Bruker Avance spectrometer operating at 500,13 MHz for proton detection. The NMR instrument is equipped with TCI 5mm cryoprobe equipped with Z gradients. Maleic acid was added to samples as internal standard allowing quantification and trimethylsilyl-3-propanoic acid-*d*4 (TMSP) for chemical shift calibration. For sample preparation, 200μL of plasma were mixed with 400μL of deuterated phosphate buffer added of 100μL of 5mM maleic acid solution and 10μL of 10mg/mL TMSP solution in D_2O. Final solution is vortexed and placed in a new 5mm NMR tube before being analyzed in NMR. For the NMR analysis, an edited 1D-CPMG sequence with water presaturation was used: RD-90-(-t-180-t)n with relaxation delay of 4s (RD), spin echo delay of 400ms (t) and 128 loop (n). Pulse for water presaturation is occurring during the relaxation time (RD) and the number of scans fixed at 64. The total acquisition time is of 3.1981568s with 4 dummy scans.

All acquired data were processed using PepsNMR® software[133], an open-source R package developed by Manon Martin and Bernadette Govaerts from University of Louvain-la-Neuve in collaboration with our group. After Group Delay Correction, Solvent Suppression, Apodization, Fourier Transform, Zero Phase Order Correction, Internal Referencing, Baseline Correction, Negative Value Zeroing, Warping (alignment), and Window Selection, the processed spectra were bucketed from 0.5 to 10 ppm with a bucket width of 0.02 ppm. After Region Removal and Normalization (mean), the data matrix consisted of 411 spectral zone of 0.02 ppm width containing metabolites information for each sample coming from the 327 individuals.

Statistical Analysis

At first BioStatFlow (http://biostatflow.org) was used for generating PCA enabling outlier detection and data visualization. BioStatFlow was also used for group analysis and classification case/control was provided by the team of Bordeaux and was based on AMD incidence. PCA, PLS-DA and OPLS-DA model were generated, and important features identified for the discrimination of the two groups. These results were used in complement of results coming from statistical analysis performed by the INSERM team of Bordeaux.

Statistical analysis performed by INSERM of Bordeaux was made as follows: In order to analyze the association between each bucket and the risk of early or advanced AMD, they used multivariate Cox models by adjusting values on the sex, the level of education, tobacco use, a genetic score based on 49 SNPs[42] and a Mediterranean diet score. Correlation between buckets was evaluated with Pearson's coefficient of correlation. The Cox model is a semi-parametric survival analysis that describe the link between the event incidence (AMD outcome) and a set of covariates[265]. This model will assess the Relative Risk (RR) for each variable of interest. Briefly, a RR above 1 indicate a variable that is positively associated with the event probability and therefore negatively associated with the survival (AMD outcomes). All statistics were performed by Soufiane Ajana from INSERM of Bordeaux on R 3.4.4.

5.5.2 NMR based lipidomics

Study population

Plasma samples coming from the 327 individuals included in the precedent study were used for the lipidomics analysis. All these samples were stored at -80°C before analysis.

Lipid's extraction

The extraction procedure is adapted from solvent-solvent extraction methods described in literature. Through the addition of methanol (MeOH), chloroform (CHCl$_3$) and water (H$_2$O) followed by centrifugation step, the organic phase can be extracted and analyzed by proton NMR spectroscopy.

200μL of plasma introduced in a glass tube and 1300μL MeOH, 800μL H$_2$O and 2500μL CHCl$_3$ are added, and the mixture is vortexed for 20s. After centrifugation, 30 min at 3000 rpm, 4°C, 2000μL of the organic phase are collected and placed into a vial. The chloroform is evaporated to dryness under reduced pressure and the dry extract is preserved at -80°C prior NMR analysis.

Sample preparation

The lipidic fraction obtained from plasma is suspended in 660μL of deuterated Chloroform CDCl$_3$ containing 0,03% of trimethylsilane (TMS) as internal reference. To allow quantification, 35μL of a solution 0,3mM of dimethylsulfone (DMSO$_2$) in CDCl$_3$ 0,03%TMS was added to the solution. The final solution is placed in a new 5mm NMR tube before being analyzed in NMR

NMR analysis

All samples were measured at 298K on a Bruker Neo spectrometer operating at 500,13 MHz for proton detection. The NMR instrument is equipped with TCI 5mm cryoprobe equipped with *Z* gradients. For NMR data acquisition, a conventional 1D proton spectrum is recorded using zg30 pulse program (RD-30-acq) with a relaxation delay (RD) of 1s, 64 scans and 2 dummy scans for a total acquisition time of 3,2767999s.

In first intention, all acquired data were processed using PepsNMR® software, an open-source R package developed by Manon Martin and Bernadette Govaerts from University of Louvain-la-Neuve in collaboration with our group. After Group Delay Correction, Solvent Suppression, Apodization, Fourier Transform, Zero Phase Order Correction, Internal Referencing, Baseline

Correction, Negative Value Zeroing, Warping (alignment), and Window Selection, the processed spectra were bucketed from 0.5 to 10ppm with a bucket width of 0.02ppm. After Region Removal and Normalization (mean), the data matrix consisted of 411 spectral zone of 0.02 ppm width containing metabolites information for each sample coming from the 327 individuals.

A second approach was used in order to improve bucketing steps by using NMRprocflow, a free web interface that–allow to perform Adaptative, Intelligent Bucketing (AI-Bucketing) method[147]. This approach will improve the standard, equidistant bucketing, by using a variable buckets sizes technique that will avoid peak splitting and peak alignment modification is minimized. This algorithm determines automatically when to stop further bucket without using references spectra or arbitrary parameters. Thus, this approach aimed to enhance the resolution and the recovery of buckets holding unique information about lipid species. After pre-processing using TopSpin 4, NMR spectra of lipids extract were imported into NMRprocflow web platform. After alignment, selection of the spectral window (0.4-6.2ppm) and suppression on uninformative spectral zone (1.5-7.5ppm, highly misaligned zone containing residual water peaks), Adaptative Intelligent bucketing step was applied, and the corresponding data matrix exported. The data matrix was than imported in Metaboanalyst and samples were normalized upon the peak of the internal reference $DMSO_2$ to provide the final bucket table used for further statistical analysis.

Statistical Analysis

Statistical analysis performed by Inserm of Bordeaux was conducted on the bucket table obtained after the first processing option and was made as follows: In order to analyze the association between each bucket and the risk of early or advanced AMD, they used multivariate Cox models by adjusting values on the sex, the level of education, tobacco use, a genetic score based on 49 SNPs[42] and a Mediterranean diet score. Correlation between buckets was evaluated with Pearson's coefficient of correlation. All statistical were performed by Soufiane Ajana from Inserm of Bordeaux on R 3.4.4.

BioStatFlow was here used to generate all PCA, PLSDA and OPLSDA models created on the data obtained after the second processing option. Variable's correlation analysis was computed on GraphPad Prism9.

Chapter 6:
General conclusions

6   General conclusions

By providing an overview of patient's metabolic status, metabolomics approach is a powerful and interesting tool for personalized medicine. Indeed, this gives access to unique information about patient's health at molecular, physiological, and environmental levels. With current analytical platforms, it's possible to establish a real snapshot of the metabolism based on hundreds of descriptors of the metabolome. This could refine patient's status and help clinicians for clinical decision and build up new therapeutic strategies. Moreover, information about metabolic changes under pathological conditions gives a chance to better understand the biochemical and biological mechanisms that driven the disease evolution or the treatment efficacy.

In the context of AMD, our research group applied NMR metabolomics to find biomarkers that might fill some clinical challenges as early diagnostic of AMD patients, patients' stratification, and evaluation of treatment responses for a better individual follow-up. Identified biomarkers, lactate and lipoprotein levels, could represent an interesting option to improve AMD patient's healthcare and would provide a better understanding of complexes biological mechanisms underlying AMD occurrence and development. Therefore, we decide to continue our study of AMD using the same approach on two different cohorts: a prospective cohort composed on healthy individuals followed over 10 years that focuses on AMD occurrence; and a cohort of AMD patients already diagnosed, engaged in a constant follow-up and under anti-VEGF treatment. Together, these studies represent a unique opportunity to overcome clinical challenges linked to AMD management by providing new tools that could help clinicians to improve diagnostic and individualize the therapeutic approach in a more efficient way.

The first part of this thesis has been done in close collaboration with the Department of Ophthalmology of University Hospital of Liège and aimed to identify key metabolites linked to neovascular Age-related Macular Degeneration (nAMD) process. The objectives were to better characterize patients' status regarding their stage of the pathology (active or non-active nAMD), and their responses to anti-VEGF treatments. Two cohorts were studied for this purpose: the first cohort consisted of a collection of sera samples from healthy volunteers and AMD patients divided on two subgroups "active" and "non-active" depending on the presence of intra ore sub-retinal fluids in OCT analysis; linked to this study, sera samples from mice CNV model that mimic the effect of nAMD were also collected and analyzed; the second cohort involved 32 nAMD patients under anti-VEGF treatment that were followed over 2 years. If first cohort was already constituted and results partially generated before this work, for the second

one the patients had to be recruited and the sampling method, samples storage and analytical workflow had to be optimized.

In the first study conducted either on human or mice CNV model led to the identification of specific lipoprotein signatures of CNV processes. This NMR metabolomics approach identified changes among lipoprotein profile linked to the severity of the pathology. Indeed, we observed a shift toward a VLDL rich lipoprotein profile for patient with active nAMD. These results were confirmed during a kinetics CNV formation NMR metabolomic analysis on mice CNV model. This refined previous observations and identified lipoprotein profile as a new tool for patient's stratification and evaluation of treatment's responses.

For the second study, we applied ^{1}H-NMR metabolomics approach to a two-year follow-up cohort composed of 32 nAMD patients under anti-VEGF treatment. The initial goal was to use the biomarkers previously identified, namely lactate level and lipoprotein profile. The dataset consisted of 269 visits. On each visit blood samples and clinical data were collected as well as OCT images of retinas allowing the measurement of CNV parameters (IRC, SRF, PED). The selected biomarkers were not able to provide good explanation of CNV status and therefore results were not satisfying. This forced us to change our mind and set up new analytical strategies in the way to find new metabolic signatures of CNV events recovered from OCT images of nAMD patients. Analyzing this longitudinal dataset was quite challenging and required to use specific and adapted statistical tools to better describe our samples.

Even if no consistent results were found with this approach, this work pointed out some important questions regarding how metabolomics research and applications should be used in the context of personalized medicine. Indeed, if the lack of consistency of our previously identified biomarkers can be explained by how far from clinical reality are the case-control studies, metabolomics must overcome some challenge to be adapted to the clinic life. Indeed, collection and analysis of samples should take in consideration the reality of daily clinic routine to provide results that could be transposed and used by clinicians. On the other part, this routine does not match with the requirement of current metabolomics and statistics approach and therefore results are difficult to extract from such experimental design. Thus, we need to rethink our way to conduct case-study analysis with the aim to produce results could be used in real clinical situation.

In addition, to take advantage of results that can provide metabolomic-based studies, patient's healthcare routine should be aware of their specific recommendations. By understanding and

merging the exigence of both world, metabolomics, and in particular NMR-based metabolomics, could play a key role for patient's monitoring and development of new therapeutic strategies that put the uniqueness of patients in the center of the playground. This represent a challenging task but is essential to put metabolomics on the center of personalized medicine approach.

The last part of our work aimed to identify metabolite-based biomarkers able to predict AMD occurrence. This work was made in collaboration with Inserm of Bordeaux that provided samples from 471 healthy donors aged of 73 y/o and more and followed over 7 years. Over this period, ophthalmologic examination allowed to identify individuals that developed AMD and snap their evolution through the pathology. Plasma samples were analyzed by ^{1}H-NMR spectroscopy and, after preprocessing and data reduction through bucketing approach, relevant multivariate and univariate analyses were performed. Interesting features were identified and associated with lipids' containing spectral zone. To go deeper, we extracted lipids from our samples and performed new ^{1}H-NMR analysis. Generated lipids profile was analyzed using the same strategies than initially but gave poor results.

Despite all improvement made on our data treatment and analysis strategies, no relevant information was recovered from this highly correlated data matrix. Further studies on this dataset should use in-house lipid databases and spectral deconvolution methods to better characterize lipids samples and identify the lipids correlated with AMD occurrence.

Nevertheless, with our approach we identified interesting spectral zones liked to AMD occurrence and associated with lipids metabolism. The NMR lipidomic study conducted to refine these results wasn't successful as planned but highlighted the weakness of usual NMR metabolomics approach when applied to lipid's extract samples. Indeed, the lack of resolution in 1D ^{1}H-NMR spectra of lipids doesn't allow informative analysis. Highly overlapped spectra generated produce highly correlated dataset that are not suitable for untargeted metabolomics approach. Rapid and robust 2D NMR approach and lipids NMR quantification methods should be developed and could provide interesting tools for NMR metabolomics purpose. Creating a lipid database and performing spectral deconvolution analysis on our NMR data of AMD patients could open the way to new prediction/diagnostic tool for age-related macular degeneration management.

This work was focused on the use of the metabolomic approach for improving patients' healthcare. From early diagnostic using prospective cohort to personalized follow-up in real life clinical routine, we aimed to add metabolomics to the toolbox of clinicians. If none of our attempts gives us satisfying results, it brings to our attention that clinical practices and metabolomics must evolve together to reach the exigence of both fields. Indeed, clinical practices could take massive benefits from the metabolomic approach. For this, we need to optimize the way our research group evolve in this particular environment to be able to move from biomarker discovery to real clinical application.

By standardization of experimental procedure and data analysis through the adoption of standard operating procedure (SOP), we could assess the robustness of our approach and provide to clinician quality assurance of our approach. Moreover, this would set up the exigence of metabolomics and pave the way to a better communication with the hospital environment. In this way, we could encourage medical unit to take an active part of studies by providing quality samples, improve study design and set up achievable goals in concertation with the research team.

In addition, we need to better define individual's metabolome variability among healthy and unhealthy individuals coming from any class of the population. This could help to better determine if a given variation of a biomarker, identified case-control study, could be truly relevant for daily clinical practices. By having an idea of how variations of the metabolome are occurring for a given individuals during time course, the usefulness of discovered feature could be assessed. With this information we could therefore adapt the way to conduct metabolomics studies. Indeed, especially for degenerative pathologies and pathologies that are in continuous evolution, this time course evolution of the metabolome is mandatory.

By taking a step back and looking at all our work, having such information would help us to be more critical regarding our results and force us to conduct more controlled studies regarding the patient follow-up cohort. With a better communication with the medical unit of the hospital, maybe we would have realized that the goals of the study weren't achievable in a small-time gap regarding the evolution of the pathology. At the end, even if a lot of efforts were made to add value to the generated datasets, information collected, and the developed analytical strategies could be the starting point of a more controlled, optimized, and rational study that could make NMR-based metabolomics the new Amsler chart for ophthalmologist.

Chapter 7:
Bibliography

7   Bibliography

(1)   Colijn, J. M.; Buitendijk, G. H. S.; Prokofyeva, E.; Alves, D.; Cachulo, M. L.; Khawaja, A. P.; Cougnard-Gregoire, A.; Merle, B. M. J.; Korb, C.; Erke, M. G.; Bron, A.; Anastasopoulos, E.; Meester-Smoor, M. A.; Segato, T.; Piermarocchi, S.; de Jong, P. T. V. M.; Vingerling, J. R.; Topouzis, F.; Creuzot-Garcher, C.; Bertelsen, G.; Pfeiffer, N.; Fletcher, A. E.; Foster, P. J.; Silva, R.; Korobelnik, J.-F.; Delcourt, C.; Klaver, C. C. W.; Ajana, S.; Arango-Gonzalez, B.; Arndt, V.; Bhatia, V.; Bhattacharya, S. S.; Biarnés, M.; Borrell, A.; Bühren, S.; Calado, S. M.; Colijn, J. M.; Cougnard-Grégoire, A.; Dammeier, S.; de Jong, E. K.; De la Cerda, B.; Delcourt, C.; den Hollander, A. I.; Diaz-Corrales, F. J.; Diether, S.; Emri, E.; Endermann, T.; Ferraro, L. L.; Garcia, M.; Heesterbeek, T. J.; Honisch, S.; Hoyng, C. B.; Kersten, E.; Kilger, E.; Klaver, C. C. W.; Langen, H.; Lengyel, I.; Luthert, P.; Maugeais, C.; Meester-Smoor, M.; Merle, B. M. J.; Monés, J.; Nogoceke, E.; Peto, T.; Pool, F. M.; Rodríguez, E.; Ueffing, M.; Ulrich Bartz-Schmidt, K. U.; van Leeuwen, E. M.; Verzijden, T.; Zumbansen, M.; Acar, N.; Anastosopoulos, E.; Azuara-Blanco, A.; Bergen, A.; Bertelsen, G.; Binquet, C.; Bird, A.; Brétillon, L.; Bron, A.; Buitendijk, G.; Cachulo, M. L.; Chakravarthy, U.; Chan, M.; Chang, P.; Colijn, J.; Cougnard-Grégoire, A.; Creuzot-Garcher, C.; Cumberland, P.; Cunha-Vaz, J.; Daien, V.; Deak, G.; Delcourt, C.; Delyfer, M.-N.; den Hollander, A.; Dietzel, M.; Erke, M. G.; Fauser, S.; Finger, R.; Fletcher, A.; Foster, P.; Founti, P.; Göbel, A.; Gorgels, T.; Grauslund, J.; Grus, F.; Hammond, C.; Helmer, C.; Hense, H.-W.; Hermann, M.; Hoehn, R.; Hogg, R.; Holz, F.; Hoyng, C.; Jansonius, N.; Janssen, S.; Khawaja, A.; Klaver, C.; Korobelnik, J.-F.; Lamparter, J.; Le Goff, M.; Leal, S.; Lechanteur, Y.; Lehtimäki, T.; Lotery, A.; Leung, I.; Mauschitz, M.; Merle, B.; Meyer zu Westrup, V.; Midena, E.; Miotto, S.; Mirshahi, A.; Mohan-Saïd, S.; Mueller, M.; Muldrew, A.; Nunes, S.; Oexle, K.; Peto, T.; Piermarocchi, S.; Prokofyeva, E.; Rahi, J.; Raitakari, O.; Ribeiro, L.; Rougier, M.-B.; Sahel, J.; Salonikiou, A.; Sanchez, C.; Schmitz-Valckenberg, S.; Schweitzer, C.; Segato, T.; Shehata, J.; Silva, R.; Silvestri, G.; Simader, C.; Souied, E.; Springelkamp, H.; Tapp, R.; Topouzis, F.; Verhoeven, V.; Von Hanno, T.; Vujosevic, S.; Williams, K.; Wolfram, C.; Yip, J.; Zerbib, J.; Zwiener, I. Prevalence of Age-Related Macular Degeneration in Europe. *Ophthalmology* **2017**, *124* (12), 1753–1763. https://doi.org/10.1016/j.ophtha.2017.05.035.

(2)   Wong, W. L.; Su, X.; Li, X.; Cheung, C. M. G.; Klein, R.; Cheng, C.-Y.; Wong, T. Y. Global Prevalence of Age-Related Macular Degeneration and Disease Burden Projection for 2020 and 2040: A Systematic Review and Meta-Analysis. *Lancet Glob. Health* **2014**, *2* (2), e106–e116. https://doi.org/10.1016/S2214-109X(13)70145-1.

(3)   Prevalence of Age-Related Macular Degeneration in the United States. *Arch. Ophthalmol.* **2004**, *122* (4), 564. https://doi.org/10.1001/archopht.122.4.564.

(4)   Ebrahimi, K. B.; Handa, J. T. Lipids, Lipoproteins, and Age-Related Macular Degeneration. *J. Lipids* **2011**, *2011*, 1–14. https://doi.org/10.1155/2011/802059.

(5)   Mitchell, J.; Bradley, C. Quality of Life in Age-Related Macular Degeneration: A Review of the Literature. *Health Qual. Life Outcomes* **2006**, *4* (1), 97. https://doi.org/10.1186/1477-7525-4-97.

(6)   van Lookeren Campagne, M.; LeCouter, J.; Yaspan, B. L.; Ye, W. Mechanisms of Age-Related Macular Degeneration and Therapeutic Opportunities: Pathology, Genetics, Animal Models, and Therapeutic Rationale of AMD. *J. Pathol.* **2014**, *232* (2), 151–164. https://doi.org/10.1002/path.4266.

(7)   Bressler, N. M.; Doan, Q. V.; Varma, R.; Lee, P. P.; Suñer, I. J.; Dolan, C.; Danese, M. D.; Yu, E.; Tran, I.; Colman, S. Estimated Cases of Legal Blindness and Visual Impairment Avoided Using Ranibizumab for Choroidal Neovascularization: Non-Hispanic White Population in the United States With Age-Related Macular Degeneration. *Arch. Ophthalmol.* **2011**, *129* (6).

https://doi.org/10.1001/archophthalmol.2011.140.

(8)   Curcio, C. A.; Messinger, J. D.; Sloan, K. R.; McGwin, G.; Medeiros, N. E.; Spaide, R. F. SUBRETINAL DRUSENOID DEPOSITS IN NON-NEOVASCULAR AGE-RELATED MACULAR DEGENERATION: Morphology, Prevalence, Topography, and Biogenesis Model. *Retina* **2013**, *33* (2), 265–276. https://doi.org/10.1097/IAE.0b013e31827e25e0.

(9)   Mitchell et al. - 2018 - Age-Related Macular Degeneration.Pdf.

(10)   Zarubina, A. V.; Neely, D. C.; Clark, M. E.; Huisingh, C. E.; Samuels, B. C.; Zhang, Y.; McGwin, G.; Owsley, C.; Curcio, C. A. Prevalence of Subretinal Drusenoid Deposits in Older Persons with and without Age-Related Macular Degeneration, by Multimodal Imaging. *Ophthalmology* **2016**, *123* (5), 1090–1100. https://doi.org/10.1016/j.ophtha.2015.12.034.

(11)   Mitchell, P.; Liew, G.; Gopinath, B.; Wong, T. Y. Age-Related Macular Degeneration. *The Lancet* **2018**, *392* (10153), 1147–1159. https://doi.org/10.1016/S0140-6736(18)31550-2.

(12)   Bird, A. C.; Bressler, N. M.; Bressler, S. B.; Chisholm, I. H.; Coscas, G.; Davis, M. D.; de Jong, P. T. V. M.; Klaver, C. C. W.; Klein, B. E. K.; Klein, R.; Mitchell, P.; Sarks, J. P.; Sarks, S. H.; Soubrane, G.; Taylor, H. R.; Vingerling, J. R. An International Classification and Grading System for Age-Related Maculopathy and Age-Related Macular Degeneration. *Surv. Ophthalmol.* **1995**, *39* (5), 367–374. https://doi.org/10.1016/S0039-6257(05)80092-X.

(13)   Joachim, N.; Mitchell, P.; Kifley, A.; Rochtchina, E.; Hong, T.; Wang, J. J. Incidence and Progression of Geographic Atrophy. *Ophthalmology* **2013**, *120* (10), 2042–2050. https://doi.org/10.1016/j.ophtha.2013.03.029.

(14)   Gheorghe, A.; Mahdi, L.; Musat, O. AGE-RELATED MACULAR DEGENERATION. *Romanian J. Ophthalmol.* **2015**, 4.

(15)   Klein, R.; Meuer, S. M.; Knudtson, M. D.; Klein, B. E. K. The Epidemiology of Progression of Pure Geographic Atrophy: The Beaver Dam Eye Study. *Am. J. Ophthalmol.* **2008**, *146* (5), 692-699.e1. https://doi.org/10.1016/j.ajo.2008.05.050.

(16)   Miceli, M. V.; Liles, M. R.; Newsome, D. A. Evaluation of Oxidative Processes in Human Pigment Epithelial Cells Associated with Retinal Outer Segment Phagocytosis. *Exp. Cell Res.* **1994**, *214* (1), 242–249. https://doi.org/10.1006/excr.1994.1254.

(17)   Anderson, D. H.; Mullins, R. F.; Hageman, G. S.; Johnson, L. V. A Role for Local Inflammation in the Formation of Drusen in the Aging Eye. *Am. J. Ophthalmol.* **2002**, *134* (3), 411–431. https://doi.org/10.1016/S0002-9394(02)01624-0.

(18)   Fine, S. L.; Berger, J. W.; Maguire, M. G.; Ho, A. C. Age-Related Macular Degeneration. *N. Engl. J. Med.* **2000**, *342* (7), 483–492. https://doi.org/10.1056/NEJM200002173420707.

(19)   Campochiaro, P. A. Ocular Neovascularisation and Excessive Vascular Permeability. *Expert Opin. Biol. Ther.* **2004**, *4* (9), 1395–1402. https://doi.org/10.1517/14712598.4.9.1395.

(20)   Kijlstra, A.; La Heij, E. C.; Hendrikse, F. REVIEW ARTICLE, Immunological Factors in the Pathogenesis and Treatment of Age-Related Macular Degeneration. *Ocul. Immunol. Inflamm.* **2005**, *13* (1), 3–11. https://doi.org/10.1080/09273940590909185.

(21)   Nowak, J. Z. Role of Lipofuscin in Pathogenesis of Age-Related Macular Degeneration (AMD). *Mag Okul* **2005**, *2*, 103–114.

(22)   Sparrow, J. R.; Boulton, M. RPE Lipofuscin and Its Role in Retinal Pathobiology. *Exp. Eye Res.* **2005**, *80* (5), 595–606. https://doi.org/10.1016/j.exer.2005.01.007.

(23)   Klein, R.; Peto, T.; Bird, A.; Vannewkirk, M. R. The Epidemiology of Age-Related Macular Degeneration. *Am. J. Ophthalmol.* **2004**, *137* (3), 486–495. https://doi.org/10.1016/j.ajo.2003.11.069.

(24)   Nowak, J. Z. Age-Related Macular Degeneration (AMD): Pathogenesis and Therapy. *Pharmacol. Rep.* **2006**, 11.

(25)   Strauss, O. The Retinal Pigment Epithelium in Visual Function. *Physiol. Rev.* **2005**, *85* (3), 845–881. https://doi.org/10.1152/physrev.00021.2004.

(26)   Warburton, S.; Southwick, K.; Hardman, R. M.; Secrest, A. M.; Grow, R. K.; Xin, H.; Woolley, A. T.; Burton, G. F.; Thulin, C. D. Examining the Proteins of Functional Retinal Lipofuscin Using Proteomic Analysis as a Guide for Understanding Its Origin. *Mol. Vis.* 13.

(27)   Kumar-Singh, R. The Role of Complement Membrane Attack Complex in Dry and Wet AMD - From Hypothesis to Clinical Trials. *Exp. Eye Res.* **2019**, *184*, 266–277. https://doi.org/10.1016/j.exer.2019.05.006.

(28)   Walport, M. J. Complement. First of Two Parts. *N. Engl. J. Med.* **2001**, *344* (14), 1058–1066. https://doi.org/10.1056/NEJM200104053441406.

(29)   Walport, M. J. Complement. Second of Two Parts. *N. Engl. J. Med.* **2001**, *344* (15), 1140–1144. https://doi.org/10.1056/NEJM200104123441506.

(30)   Klos, A.; Tenner, A. J.; Johswich, K.-O.; Ager, R. R.; Reis, E. S.; Köhl, J. The Role of the Anaphylatoxins in Health and Disease. *Mol. Immunol.* **2009**, *46* (14), 2753–2766. https://doi.org/10.1016/j.molimm.2009.04.027.

(31)   Cao, X.; Shen, D.; Patel, M. M.; Tuo, J.; Johnson, T. M.; Olsen, T. W.; Chan, C.-C. Macrophage Polarization in the Maculae of Age-Related Macular Degeneration: A Pilot Study: Macrophage Polarization in AMD. *Pathol. Int.* **2011**, *61* (9), 528–535. https://doi.org/10.1111/j.1440-1827.2011.02695.x.

(32)   Cherepanoff, S.; McMenamin, P.; Gillies, M. C.; Kettle, E.; Sarks, S. H. Bruch's Membrane and Choroidal Macrophages in Early and Advanced Age-Related Macular Degeneration. *Br. J. Ophthalmol.* **2010**, *94* (7), 918–925. https://doi.org/10.1136/bjo.2009.165563.

(33)   Campa, C.; Costagliola, C.; Incorvaia, C.; Sheridan, C.; Semeraro, F.; De Nadai, K.; Sebastiani, A.; Parmeggiani, F. Inflammatory Mediators and Angiogenic Factors in Choroidal Neovascularization: Pathogenetic Interactions and Therapeutic Implications. *Mediators Inflamm.* **2010**, *2010*, 1–14. https://doi.org/10.1155/2010/546826.

(34)   Campa, C.; Kasman, I.; Ye, W.; Lee, W. P.; Fuh, G.; Ferrara, N. Effects of an Anti–VEGF-A Monoclonal Antibody on Laser-Induced Choroidal Neovascularization in Mice: Optimizing Methods to Quantify Vascular Changes. *Investig. Opthalmology Vis. Sci.* **2008**, *49* (3), 1178. https://doi.org/10.1167/iovs.07-1194.

(35)   Bhutto, I. A.; McLeod, D. S.; Hasegawa, T.; Kim, S. Y.; Merges, C.; Tong, P.; Lutty, G. A. Pigment Epithelium-Derived Factor (PEDF) and Vascular Endothelial Growth Factor (VEGF) in Aged Human Choroid and Eyes with Age-Related Macular Degeneration. *Exp. Eye Res.* **2006**, *82* (1), 99–110. https://doi.org/10.1016/j.exer.2005.05.007.

(36)   Kwak, N.; Okamoto, N.; Wood, J. M.; Campochiaro, P. A. VEGF Is Major Stimulator in Model of Choroidal Neovascularization. **2000**, *41* (10), 7.

(37)   Saishin, Y.; Saishin, Y.; Takahashi, K.; Silva, R. L. e; Hylton, D.; Rudge, J. S.; Wiegand, S. J.; Campochiaro, P. A. VEGF-TRAPR1R2 Suppresses Choroidal Neovascularization and VEGF-Induced Breakdown of the Blood-Retinal Barrier. *J. Cell. Physiol.* **2003**, *195* (2), 241–248.

https://doi.org/10.1002/jcp.10246.

(38)   Heier, J. S.; Brown, D. M.; Chong, V.; Korobelnik, J.-F.; Kaiser, P. K.; Nguyen, Q. D.; Kirchhof, B.; Ho, A.; Ogura, Y.; Yancopoulos, G. D.; Stahl, N.; Vitti, R.; Berliner, A. J.; Soo, Y.; Anderesi, M.; Groetzbach, G.; Sommerauer, B.; Sandbrink, R.; Simader, C.; Schmidt-Erfurth, U. Intravitreal Aflibercept (VEGF Trap-Eye) in Wet Age-Related Macular Degeneration. *Ophthalmology* **2012**, *119* (12), 2537–2548. https://doi.org/10.1016/j.ophtha.2012.09.006.

(39)   Sheridan, C. M.; Pate, S.; Hiscott, P.; Wong, D.; Pattwell, D. M.; Kent, D. Expression of Hypoxia-Inducible Factor−1α and −2α in Human Choroidal Neovascular Membranes. *Graefes Arch. Clin. Exp. Ophthalmol.* **2009**, *247* (10), 1361–1367. https://doi.org/10.1007/s00417-009-1133-3.

(40)   Kijlstra, A.; Berendschot, T. T. J. M. Age-Related Macular Degeneration: A Complementopathy? *Ophthalmic Res.* **2015**, *54* (2), 64–73. https://doi.org/10.1159/000432401.

(41)   Seddon, J. M. The US Twin Study of Age-Related Macular Degeneration: Relative Roles of Genetic and Environmental Influences. *Arch. Ophthalmol.* **2005**, *123* (3), 321. https://doi.org/10.1001/archopht.123.3.321.

(42)   Fritsche, L. G.; Igl, W.; Bailey, J. N. C.; Grassmann, F.; Sengupta, S.; Bragg-Gresham, J. L.; Burdon, K. P.; Hebbring, S. J.; Wen, C.; Gorski, M.; Kim, I. K.; Cho, D.; Zack, D.; Souied, E.; Scholl, H. P. N.; Bala, E.; Lee, K. E.; Hunter, D. J.; Sardell, R. J.; Mitchell, P.; Merriam, J. E.; Cipriani, V.; Hoffman, J. D.; Schick, T.; Lechanteur, Y. T. E.; Guymer, R. H.; Johnson, M. P.; Jiang, Y.; Stanton, C. M.; Buitendijk, G. H. S.; Zhan, X.; Kwong, A. M.; Boleda, A.; Brooks, M.; Gieser, L.; Ratnapriya, R.; Branham, K. E.; Foerster, J. R.; Heckenlively, J. R.; Othman, M. I.; Vote, B. J.; Liang, H. H.; Souzeau, E.; McAllister, I. L.; Isaacs, T.; Hall, J.; Lake, S.; Mackey, D. A.; Constable, I. J.; Craig, J. E.; Kitchner, T. E.; Yang, Z.; Su, Z.; Luo, H.; Chen, D.; Ouyang, H.; Flagg, K.; Lin, D.; Mao, G.; Ferreyra, H.; Stark, K.; von Strachwitz, C. N.; Wolf, A.; Brandl, C.; Rudolph, G.; Olden, M.; Morrison, M. A.; Morgan, D. J.; Schu, M.; Ahn, J.; Silvestri, G.; Tsironi, E. E.; Park, K. H.; Farrer, L. A.; Orlin, A.; Brucker, A.; Li, M.; Curcio, C. A.; Mohand-Saïd, S.; Sahel, J.-A.; Audo, I.; Benchaboune, M.; Cree, A. J.; Rennie, C. A.; Goverdhan, S. V.; Grunin, M.; Hagbi-Levi, S.; Campochiaro, P.; Katsanis, N.; Holz, F. G.; Blond, F.; Blanché, H.; Deleuze, J.-F.; Igo, R. P.; Truitt, B.; Peachey, N. S.; Meuer, S. M.; Myers, C. E.; Moore, E. L.; Klein, R.; Hauser, M. A.; Postel, E. A.; Courtenay, M. D.; Schwartz, S. G.; Kovach, J. L.; Scott, W. K.; Liew, G.; Tan, A. G.; Gopinath, B.; Merriam, J. C.; Smith, R. T.; Khan, J. C.; Shahid, H.; Moore, A. T.; McGrath, J. A.; Laux, R.; Brantley, M. A.; Agarwal, A.; Ersoy, L.; Caramoy, A.; Langmann, T.; Saksens, N. T. M.; de Jong, E. K.; Hoyng, C. B.; Cain, M. S.; Richardson, A. J.; Martin, T. M.; Blangero, J.; Weeks, D. E.; Dhillon, B.; van Duijn, C. M.; Doheny, K. F.; Romm, J.; Klaver, C. C. W.; Hayward, C.; Gorin, M. B.; Klein, M. L.; Baird, P. N.; den Hollander, A. I.; Fauser, S.; Yates, J. R. W.; Allikmets, R.; Wang, J. J.; Schaumberg, D. A.; Klein, B. E. K.; Hagstrom, S. A.; Chowers, I.; Lotery, A. J.; Léveillard, T.; Zhang, K.; Brilliant, M. H.; Hewitt, A. W.; Swaroop, A.; Chew, E. Y.; Pericak-Vance, M. A.; DeAngelis, M.; Stambolian, D.; Haines, J. L.; Iyengar, S. K.; Weber, B. H. F.; Abecasis, G. R.; Heid, I. M. A Large Genome-Wide Association Study of Age-Related Macular Degeneration Highlights Contributions of Rare and Common Variants. *Nat. Genet.* **2016**, *48* (2), 134–143. https://doi.org/10.1038/ng.3448.

(43)   Al-Zamil, W.; Yassin, S. Recent Developments in Age-Related Macular Degeneration: A Review. *Clin. Interv. Aging* **2017**, *Volume 12*, 1313–1330. https://doi.org/10.2147/CIA.S143508.

(44)   Sobrin, L.; Seddon, J. M. Nature and Nurture- Genes and Environment- Predict Onset and Progression of Macular Degeneration. *Prog. Retin. Eye Res.* **2014**, *40*, 1–15. https://doi.org/10.1016/j.preteyeres.2013.12.004.

(45)   Risk Factors for Age-Related Maculopathy: The Visual Impairment Project. *ARCH*

OPHTHALMOL **2001**, *119*, 8.

(46)   Delcourt, C.; Michel, F.; Colvez, A.; Lacroux, A.; Delage, M.; Vernet, M.-H. Associations of Cardiovascular Disease and Its Risk Factors with Age-Related Macular Degeneration: The POLA Study. *Ophthalmic Epidemiol.* **2001**, *8* (4), 237–249. https://doi.org/10.1076/opep.8.4.237.1613.

(47)   Seddon, J. M. Progression of Age-Related Macular Degeneration: Association With Body Mass Index, Waist Circumference, and Waist-Hip Ratio. *Arch. Ophthalmol.* **2003**, *121* (6), 785. https://doi.org/10.1001/archopht.121.6.785.

(48)   Delcourt, C.; Carrie`re, I.; Delage, M.; Barberger-Gateau, P.; Schalch, W. Plasma Lutein and Zeaxanthin and Other Carotenoids as Modifiable Risk Factors for Age-Related Maculopathy and Cataract: The POLA Study. *Investig. Opthalmology Vis. Sci.* **2006**, *47* (6), 2329. https://doi.org/10.1167/iovs.05-1235.

(49)   Gale, C. R.; Hall, N. F.; Phillips, D. I. W.; Martyn, C. N. Lutein and Zeaxanthin Status and Risk of Age-Related Macular Degeneration. *Investig. Opthalmology Vis. Sci.* **2003**, *44* (6), 2461. https://doi.org/10.1167/iovs.02-0929.

(50)   Cho, E. Prospective Study of Intake of Fruits, Vegetables, Vitamins, and Carotenoidsand Risk of Age-Related Maculopathy. *Arch. Ophthalmol.* **2004**, *122* (6), 883. https://doi.org/10.1001/archopht.122.6.883.

(51)   Krinsky, N. I.; Landrum, J. T.; Bone, R. A. B IOLOGIC M ECHANISMS OF THE P ROTECTIVE R OLE OF L UTEIN AND Z EAXANTHIN IN THE E YE. *Annu. Rev. Nutr.* **2003**, *23* (1), 171–201. https://doi.org/10.1146/annurev.nutr.23.011702.073307.

(52)   Age-Related Eye Disease Study Research Group. A Randomized, Placebo-Controlled, Clinical Trial of High-Dose Supplementation With Vitamins C and E, Beta Carotene, and Zinc for Age-Related Macular Degeneration and Vision Loss: AREDS Report No. 8. *Arch. Ophthalmol.* **2001**, *119* (10), 1417–1436. https://doi.org/10.1001/archopht.119.10.1417.

(53)   Seddon, J. M. Dietary Fat and Risk for Advanced Age-Related Macular Degeneration. *Arch. Ophthalmol.* **2001**, *119* (8), 1191. https://doi.org/10.1001/archopht.119.8.1191.

(54)   the POLANUT Study Group; Delcourt, C.; Carrière, I.; Cristol, J.-P.; Lacroux, A.; Gerber, M. Dietary Fat and the Risk of Age-Related Maculopathy: The POLANUT Study. *Eur. J. Clin. Nutr.* **2007**, *61* (11), 1341–1344. https://doi.org/10.1038/sj.ejcn.1602685.

(55)   Mares-Perlman, J. A.; Brady, W. E.; Klein, R.; VandenLangenberg, G. M.; Klein, B. E. K.; Palta, M. Dietary Fat and Age-Related Maculopathy. *Arch. Ophthalmol.* **1995**, *113* (6), 743–748. https://doi.org/10.1001/archopht.1995.01100060069034.

(56)   Calder, P. C. Polyunsaturated Fatty Acids, Inflammation, and Immunity. *Lipids* **2001**, *36* (9), 1007–1024. https://doi.org/10.1007/s11745-001-0812-7.

(57)   Mukutmoni-Norris, M.; Hubbard, N. E.; Erickson, K. L. Modulation of Murine Mammary Tumor Vasculature by Dietary N-3 Fatty Acids in Fish Oil. *Cancer Lett.* **2000**, *150* (1), 101–109. https://doi.org/10.1016/S0304-3835(99)00380-8.

(58)   The Age-Related Eye Disease Study 2 (AREDS2) Research Group*. Lutein + Zeaxanthin and Omega-3 Fatty Acids for Age-Related Macular Degeneration: The Age-Related Eye Disease Study 2 (AREDS2) Randomized Clinical Trial. *JAMA* **2013**, *309* (19), 2005–2015. https://doi.org/10.1001/jama.2013.4997.

(59)   Merle, B. M. J. Mediterranean Diet and Incidence of Advanced Age-Related Macular Degeneration. **2019**, *126* (3), 10.

(60)   Chew, E. Y. Nutrition, Genes and Age-Related Macular Degeneration: What Have We Learned from the Trials? **2018**, 8.

(61)   Chakravarthy, U.; Wong, T. Y.; Fletcher, A.; Piault, E.; Evans, C.; Zlateva, G.; Buggage, R.; Pleil, A.; Mitchell, P. Clinical Risk Factors for Age-Related Macular Degeneration: A Systematic Review and Meta-Analysis. *BMC Ophthalmol.* **2010**, *10* (1), 31. https://doi.org/10.1186/1471-2415-10-31.

(62)   Wong, I. Y. H.; Koo, S. C. Y.; Chan, C. W. N. Prevention of Age-Related Macular Degeneration. *Int Ophthalmol* **2011**, 10.

(63)   Brown, D. M.; Kaiser, P. K.; Michels, M.; Soubrane, G.; Heier, J. S.; Kim, R. Y.; Sy, J. P.; Schneider, S.; ANCHOR Study Group. Ranibizumab versus Verteporfin for Neovascular Age-Related Macular Degeneration. *N. Engl. J. Med.* **2006**, *355* (14), 1432–1444. https://doi.org/10.1056/NEJMoa062655.

(64)   Rosenfeld, P. J.; Brown, D. M.; Heier, J. S.; Boyer, D. S.; Kaiser, P. K.; Chung, C. Y.; Kim, R. Y.; MARINA Study Group. Ranibizumab for Neovascular Age-Related Macular Degeneration. *N. Engl. J. Med.* **2006**, *355* (14), 1419–1431. https://doi.org/10.1056/NEJMoa054481.

(65)   Schmidt-Erfurth, U.; Chong, V.; Loewenstein, A.; Larsen, M.; Souied, E.; Schlingemann, R.; Eldem, B.; Monés, J.; Richard, G.; Bandello, F. Guidelines for the Management of Neovascular Age-Related Macular Degeneration by the European Society of Retina Specialists (EURETINA). 24.

(66)   Meyer, C. H.; Michels, S.; Rodrigues, E. B.; Hager, A.; Mennel, S.; Schmidt, J. C.; Helb, H.-M.; Farah, M. E. Incidence of Rhegmatogenous Retinal Detachments after Intravitreal Antivascular Endothelial Factor Injections. *Acta Ophthalmol. (Copenh.)* **2011**, *89* (1), 70–75. https://doi.org/10.1111/j.1755-3768.2010.02064.x.

(67)   Grunwald, J. E.; Daniel, E.; Huang, J.; Ying, G.; Maguire, M. G.; Toth, C. A.; Jaffe, G. J.; Fine, S. L.; Blodi, B.; Klein, M. L.; Martin, A. A.; Hagstrom, S. A.; Martin, D. F. Risk of Geographic Atrophy in the Comparison of Age-Related Macular Degeneration Treatments Trials. *Ophthalmology* **2014**, *121* (1), 150–161. https://doi.org/10.1016/j.ophtha.2013.08.015.

(68)   Modi, Y. S.; Tanchon, C.; Ehlers, J. P. Comparative Safety and Tolerability of Anti-VEGF Therapy in Age-Related Macular Degeneration. *Drug Saf.* **2015**, *38* (3), 279–293. https://doi.org/10.1007/s40264-015-0273-0.

(69)   Lalwani, G. A.; Rosenfeld, P. J.; Fung, A. E.; Dubovy, S. R.; Michels, S.; Feuer, W.; Davis, J. L.; Flynn, H. W.; Esquiabro, M. A Variable-Dosing Regimen with Intravitreal Ranibizumab for Neovascular Age-Related Macular Degeneration: Year 2 of the PrONTO Study. *Am. J. Ophthalmol.* **2009**, *148* (1), 43-58.e1. https://doi.org/10.1016/j.ajo.2009.01.024.

(70)   Holz, F. G.; Amoaku, W.; Donate, J.; Guymer, R. H.; Kellner, U.; Schlingemann, R. O.; Weichselberger, A.; Staurenghi, G. Safety and Efficacy of a Flexible Dosing Regimen of Ranibizumab in Neovascular Age-Related Macular Degeneration: The SUSTAIN Study. *Ophthalmology* **2011**, *118* (4), 663–671. https://doi.org/10.1016/j.ophtha.2010.12.019.

(71)   Writing Committee for the UK Age-Related Macular Degeneration EMR Users Group. The Neovascular Age-Related Macular Degeneration Database: Multicenter Study of 92 976 Ranibizumab Injections: Report 1: Visual Acuity. *Ophthalmology* **2014**, *121* (5), 1092–1101. https://doi.org/10.1016/j.ophtha.2013.11.031.

(72)   Chakravarthy, U.; Harding, S. P.; Rogers, C. A.; Downes, S. M.; Lotery, A. J.; Culliford, L. A.; Reeves, B. C. Alternative Treatments to Inhibit VEGF in Age-Related Choroidal Neovascularisation: 2-Year Findings of the IVAN Randomised Controlled Trial. *The Lancet* **2013**, *382*

(9900), 1258–1267. https://doi.org/10.1016/S0140-6736(13)61501-9.

(73)   Fercher, A. F.; Hitzenberger, C. K.; Drexler, W.; Kamp, G.; Sattmann, H. In Vivo Optical Coherence Tomography. *Am. J. Ophthalmol.* **1993**, *116* (1), 113–114. https://doi.org/10.1016/S0002-9394(14)71762-3.

(74)   Schmidt-Erfurth, U.; Waldstein, S. M. A Paradigm Shift in Imaging Biomarkers in Neovascular Age-Related Macular Degeneration. *Prog. Retin. Eye Res.* **2016**, *50*, 1–24. https://doi.org/10.1016/j.preteyeres.2015.07.007.

(75)   Bhende, M.; Shetty, S.; Parthasarathy, M.; Ramya, S. Optical Coherence Tomography: A Guide to Interpretation of Common Macular Diseases. *Indian J. Ophthalmol.* **2018**, *66* (1), 20. https://doi.org/10.4103/ijo.IJO_902_17.

(76)   Hee, M. R. Quantitative Assessment of Macular Edema With Optical Coherence Tomography. *Arch. Ophthalmol.* **1995**, *113* (8), 1019. https://doi.org/10.1001/archopht.1995.01100080071031.

(77)   Krebs, I.; Falkner-Radler, C.; Hagen, S.; Haas, P.; Brannath, W.; Lie, S.; Ansari-Shahrezaei, S.; Binder, S. Quality of the Threshold Algorithm in Age-Related Macular Degeneration: Stratus versus Cirrus OCT. *Investig. Opthalmology Vis. Sci.* **2009**, *50* (3), 995. https://doi.org/10.1167/iovs.08-2617.

(78)   Jaffe, G. J.; Martin, D. F.; Toth, C. A.; Daniel, E.; Maguire, M. G.; Ying, G.-S.; Grunwald, J. E.; Huang, J. Macular Morphology and Visual Acuity in the Comparison of Age-Related Macular Degeneration Treatments Trials. *Ophthalmology* **2013**, *120* (9), 1860–1870. https://doi.org/10.1016/j.ophtha.2013.01.073.

(79)   Keane, P. A.; Liakopoulos, S.; Chang, K. T.; Wang, M.; Dustin, L.; Walsh, A. C.; Sadda, S. R. Relationship Between Optical Coherence Tomography Retinal Parameters and Visual Acuity in Neovascular Age-Related Macular Degeneration. *Ophthalmology* **2008**, *115* (12), 2206–2214. https://doi.org/10.1016/j.ophtha.2008.08.016.

(80)   Keane, P. A.; Patel, P. J.; Ouyang, Y.; Chen, F. K.; Ikeji, F.; Walsh, A. C.; Tufail, A.; Sadda, S. R. Effects of Retinal Morphology on Contrast Sensitivity and Reading Ability in Neovascular Age-Related Macular Degeneration. *Investig. Opthalmology Vis. Sci.* **2010**, *51* (11), 5431. https://doi.org/10.1167/iovs.09-4846.

(81)   Moutray, T.; Alarbi, M.; Mahon, G.; Stevenson, M.; Chakravarthy, U. Relationships between Clinical Measures of Visual Function, Fluorescein Angiographic and Optical Coherence Tomography Features in Patients with Subfoveal Choroidal Neovascularisation. *Br. J. Ophthalmol.* **2008**, *92* (3), 361–364. https://doi.org/10.1136/bjo.2007.123976.

(82)   Simader, C.; Ritter, M.; Bolz, M.; Deák, G. G.; Mayr-Sponer, U.; Golbaz, I.; Kundi, M.; Schmidt-Erfurth, U. M. Morphologic Parameters Relevant for Visual Outcome During Anti-Angiogenic Therapy of Neovascular Age-Related Macular Degeneration. *Ophthalmology* **2014**, *121* (6), 1237–1245. https://doi.org/10.1016/j.ophtha.2013.12.029.

(83)   Penha, F. M.; Rosenfeld, P. J.; Gregori, G.; Falcão, M.; Yehoshua, Z.; Wang, F.; Feuer, W. J. Quantitative Imaging of Retinal Pigment Epithelial Detachments Using Spectral-Domain Optical Coherence Tomography. *Am. J. Ophthalmol.* **2012**, *153* (3), 515–523. https://doi.org/10.1016/j.ajo.2011.08.031.

(84)   Keane, P. A.; Patel, P. J.; Liakopoulos, S.; Heussen, F. M.; Sadda, S. R.; Tufail, A. Evaluation of Age-Related Macular Degeneration With Optical Coherence Tomography. *Surv. Ophthalmol.* **2012**, *57* (5), 389–414. https://doi.org/10.1016/j.survophthal.2012.01.006.

(85)   Akagi-Kurashige, Y.; Tsujikawa, A.; Oishi, A.; Ooto, S.; Yamashiro, K.; Tamura, H.; Nakata, I.; Ueda-Arakawa, N.; Yoshimura, N. Relationship between Retinal Morphological Findings and Visual Function in Age-Related Macular Degeneration. *Graefes Arch. Clin. Exp. Ophthalmol.* **2012**, *250* (8), 1129–1136. https://doi.org/10.1007/s00417-012-1928-5.

(86)   Coscas, G.; De Benedetto, U.; Coscas, F.; Li Calzi, C. I.; Vismara, S.; Roudot-Thoraval, F.; Bandello, F.; Souied, E. Hyperreflective Dots: A New Spectral-Domain Optical Coherence Tomography Entity for Follow-up and Prognosis in Exudative Age-Related Macular Degeneration. *Ophthalmol. J. Int. Ophtalmol. Int. J. Ophthalmol. Z. Augenheilkd.* **2013**, *229* (1), 32–37. https://doi.org/10.1159/000342159.

(87)   Schmidt-Erfurth, U.; Waldstein, S. M.; Deak, G.-G.; Kundi, M.; Simader, C. Pigment Epithelial Detachment Followed by Retinal Cystoid Degeneration Leads to Vision Loss in Treatment of Neovascular Age-Related Macular Degeneration. **2015**, *122* (4), 11.

(88)   Dirani, A.; Ambresin, A.; Marchionno, L.; Decugis, D.; Mantel, I. Factors Influencing the Treatment Response of Pigment Epithelium Detachment in Age-Related Macular Degeneration. *Am. J. Ophthalmol.* **2015**, *160* (4), 732-738.e2. https://doi.org/10.1016/j.ajo.2015.06.025.

(89)   Waldstein, S. M.; Philip, A.-M.; Leitner, R.; Simader, C.; Langs, G.; Gerendas, B. S.; Schmidt-Erfurth, U. Correlation of 3-Dimensionally Quantified Intraretinal and Subretinal Fluid With Visual Acuity in Neovascular Age-Related Macular Degeneration. *JAMA Ophthalmol.* **2016**, *134* (2), 182. https://doi.org/10.1001/jamaophthalmol.2015.4948.

(90)   Sun, X.; Yang, S.; Zhao, J. Resistance to Anti-VEGF Therapy in Neovascular Age-Related Macular Degeneration: A Comprehensive Review. *Drug Des. Devel. Ther.* **2016**, 1857. https://doi.org/10.2147/DDDT.S97653.

(91)   Grewal, D. S.; Gill, M. K.; Sarezky, D.; Lyon, A. T.; Mirza, R. G. Visual and Anatomical Outcomes Following Intravitreal Aflibercept in Eyes with Recalcitrant Neovascular Age-Related Macular Degeneration: 12-Month Results. *Eye* **2014**, *28* (7), 895–899. https://doi.org/10.1038/eye.2014.101.

(92)   Broadhead, G. K.; Hong, T.; Chang, A. A. Treating the Untreatable Patient: Current Options for the Management of Treatment-Resistant Neovascular Age-Related Macular Degeneration. *Acta Ophthalmol. (Copenh.)* **2014**, *92* (8), 713–723. https://doi.org/10.1111/aos.12463.

(93)   Fung, A. T.; Kumar, N.; Vance, S. K.; Slakter, J. S.; Klancnik, J. M.; Spaide, R. S.; Freund, K. B. Pilot Study to EvaLuate the Role of High-Dose RAnibizumab 2.0 Mg in the Management of Neovascular Age-Related Macular Degeneration in Patients with PerSistent/RecurrenT Macular Fluid <30 Days Following Treatment with Intravitreal Anti-VEGF Therapy (the LAST Study). *Eye* **2012**, *26* (9), 1181–1187. https://doi.org/10.1038/eye.2012.174.

(94)   Kuroda, Y.; Yamashiro, K.; Miyake, M.; Yoshikawa, M.; Nakanishi, H.; Oishi, A.; Tamura, H.; Ooto, S.; Tsujikawa, A.; Yoshimura, N. Factors Associated with Recurrence of Age-Related Macular Degeneration after Anti-Vascular Endothelial Growth Factor Treatment. *Ophthalmology* **2015**, *122* (11), 2303–2310. https://doi.org/10.1016/j.ophtha.2015.06.053.

(95)   Arcinue, C. A.; Ma, F.; Barteselli, G.; Sharpsten, L.; Gomez, M. L.; Freeman, W. R. One-Year Outcomes of Aflibercept in Recurrent or Persistent Neovascular Age-Related Macular Degeneration. *Am. J. Ophthalmol.* **2015**, *159* (3), 426-436.e2. https://doi.org/10.1016/j.ajo.2014.11.022.

(96)   Leenders, J. APPLICATION OF METABOLOMICS IN TRANSLATIONAL BIOMEDICAL RESEARCH.

(97)   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* **1999**, *29* (11), 1181–1189. https://doi.org/10.1080/004982599238047.

(98)   Dunn, W. B.; Bailey, N. J. C.; Johnson, H. E. Measuring the Metabolome: Current Analytical Technologies. *The Analyst* **2005**, *130* (5), 606. https://doi.org/10.1039/b418288j.

(99)   Wishart, D. S. Quantitative Metabolomics Using NMR. *TrAC Trends Anal. Chem.* **2008**, *27* (3), 228–237. https://doi.org/10.1016/j.trac.2007.12.001.

(100)   Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. Modern Analytical Techniques in Metabolomics Analysis. *The Analyst* **2012**, *137* (2), 293–300. https://doi.org/10.1039/C1AN15605E.

(101)   Mussap, M.; Antonucci, R.; Noto, A.; Fanos, V. The Role of Metabolomics in Neonatal and Pediatric Laboratory Medicine. *Clin. Chim. Acta* **2013**, *426*, 127–138. https://doi.org/10.1016/j.cca.2013.08.020.

(102)   Psychogios, N.; Hau, D. D.; Peng, J.; Guo, A. C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.; Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Xia, J.; Knox, C.; Dong, E.; Huang, P.; Hollander, Z.; Pedersen, T. L.; Smith, S. R.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J. W.; Goodfriend, T.; Wishart, D. S. The Human Serum Metabolome. *PLoS ONE* **2011**, *6* (2), e16957. https://doi.org/10.1371/journal.pone.0016957.

(103)   Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z. T.; Poelzer, J.; Huynh, J.; Yallou, F. S.; Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D. S. The Human Urine Metabolome. *PLOS ONE* **2013**, *8* (9), 29.

(104)   Kaluarachchi, M.; Boulangé, C. L.; Karaman, I.; Lindon, J. C.; Ebbels, T. M. D.; Elliott, P.; Tracy, R. P.; Olson, N. C. A Comparison of Human Serum and Plasma Metabolites Using Untargeted 1H NMR Spectroscopy and UPLC-MS. *Metabolomics* **2018**, *14* (3), 32. https://doi.org/10.1007/s11306-018-1332-1.

(105)   Frédérich, M.; Pirotte, B.; Fillet, M.; de Tullio, P. Metabolomics as a Challenging Approach for Medicinal Chemistry and Personalized Medicine. *J. Med. Chem.* **2016**, *59* (19), 8649–8666. https://doi.org/10.1021/acs.jmedchem.5b01335.

(106)   Schattka, B.; Alexander, M.; Ying, S. L.; Man, A.; Shaw, R. A. Metabolic Fingerprinting of Biofluids by Infrared Spectroscopy: Modeling and Optimization of Flow Rates for Laminar Fluid Diffusion Interface Sample Preconditioning. *Anal. Chem.* **2011**, *83* (2), 555–562. https://doi.org/10.1021/ac102338n.

(107)   Markley, J. L.; Brüschweiler, R.; Edison, A. S.; Eghbalnia, H. R.; Powers, R.; Raftery, D.; Wishart, D. S. The Future of NMR-Based Metabolomics. *Curr. Opin. Biotechnol.* **2017**, *43*, 34–40. https://doi.org/10.1016/j.copbio.2016.08.001.

(108)   Fan, T. W.-M.; Lane, A. N. Applications of NMR Spectroscopy to Systems Biochemistry. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *92–93*, 18–53. https://doi.org/10.1016/j.pnmrs.2016.01.005.

(109)   Nagana Gowda, G. A.; Raftery, D. Can NMR Solve Some Significant Challenges in Metabolomics? *J. Magn. Reson.* **2015**, *260*, 144–160. https://doi.org/10.1016/j.jmr.2015.07.014.

(110)   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.* **2021**, *8*, 698337. https://doi.org/10.3389/fmolb.2021.698337.

(111)   Fang, Z.-Z.; Gonzalez, F. J. LC–MS-Based Metabolomics: An Update. *Arch. Toxicol.* **2014**, *88* (8), 1491–1502. https://doi.org/10.1007/s00204-014-1234-6.

(112)   Gika, H. G.; Theodoridis, G. A.; Plumb, R. S.; Wilson, I. D. Current Practice of Liquid Chromatography–Mass Spectrometry in Metabolomics and Metabonomics. *J. Pharm. Biomed. Anal.* **2014**, *87*, 12–25. https://doi.org/10.1016/j.jpba.2013.06.032.

(113)   Schrimpe-Rutledge, A. C.; Codreanu, S. G.; Sherrod, S. D.; McLean, J. A. Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (12), 1897–1905. https://doi.org/10.1007/s13361-016-1469-y.

(114)   Jones, O. A. H. Illuminating the Dark Metabolome to Advance the Molecular Characterisation of Biological Systems. *Metabolomics* **2018**, *14* (8), 101. https://doi.org/10.1007/s11306-018-1396-y.

(115)   Ardenkjaer-Larsen, J.-H.; Boebinger, G. S.; Comment, A.; Duckett, S.; Edison, A. S.; Engelke, F.; Griesinger, C.; Griffin, R. G.; Hilty, C.; Maeda, H.; Parigi, G.; Prisner, T.; Ravera, E.; van Bentum, J.; Vega, S.; Webb, A.; Luchinat, C.; Schwalbe, H.; Frydman, L. Facing and Overcoming Sensitivity Challenges in Biomolecular NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2015**, *54* (32), 9162–9185. https://doi.org/10.1002/anie.201410653.

(116)   Mili, M. Metabolomics Investigation of Cancer Cells by High Field NMR, 2019.

(117)   Emwas, A.-H.; Roy, R.; McKay, R. T.; Tenori, L.; Saccenti, E.; Gowda, G. A. N.; Raftery, D.; Alahmari, F.; Jaremko, L.; Jaremko, M.; Wishart, D. S. NMR Spectroscopy for Metabolomics Research. *Metabolites* **2019**, *9* (7), 123. https://doi.org/10.3390/metabo9070123.

(118)   Kovacs, H.; Moskau, D.; Spraul, M. Cryogenically Cooled Probes—a Leap in NMR Technology. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46* (2–3), 131–155. https://doi.org/10.1016/j.pnmrs.2005.03.001.

(119)   Moskau, D. Application of Real Time Digital Filters in NMR Spectroscopy. *Concepts Magn. Reson.* **2002**, *15* (2), 164–176. https://doi.org/10.1002/cmr.10031.

(120)   Keun, H. C. CHAPTER 4. NMR Spectroscopy of Serum and Plasma. In *New Developments in NMR*; Keun, H. C., Ed.; Royal Society of Chemistry: Cambridge, 2018; pp 85–132. https://doi.org/10.1039/9781782627937-00085.

(121)   Nagana Gowda, G. A.; Gowda, Y. N.; Raftery, D. Expanding the Limits of Human Blood Metabolite Quantitation Using NMR Spectroscopy. *Anal. Chem.* **2015**, *87* (1), 706–715. https://doi.org/10.1021/ac503651e.

(122)   Wevers, R. A.; Engelke, U.; Heerschap, A. High-Resolution 1H-NMR Spectroscopy of Blood Plasma for Metabolic Studies. *Clin. Chem.* **1994**, *40* (7), 1245–1250. https://doi.org/10.1093/clinchem/40.7.1245.

(123)   McHugh, C.; Flott, T.; Schooff, C.; Smiley, Z.; Puskarich, M.; Myers, D.; Younger, J.; Jones, A.; Stringer, K. Rapid, Reproducible, Quantifiable NMR Metabolomics: Methanol and Methanol: Chloroform Precipitation for Removal of Macromolecules in Serum and Whole Blood. *Metabolites* **2018**, *8* (4), 93. https://doi.org/10.3390/metabo8040093.

(124)   Tiziani, S.; Emwas, A.-H.; Lodi, A.; Ludwig, C.; Bunce, C. M.; Viant, M. R.; Günther, U. L. Optimized Metabolite Extraction from Blood Serum for 1H Nuclear Magnetic Resonance Spectroscopy. *Anal. Biochem.* **2008**, *377* (1), 16–23. https://doi.org/10.1016/j.ab.2008.01.037.

(125)   Daykin, C. A.; Foxall, P. J. D.; Connor, S. C.; Lindon, J. C.; Nicholson, J. K. The Comparison of Plasma Deproteinization Methods for the Detection of Low-Molecular-Weight Metabolites by 1H Nuclear Magnetic Resonance Spectroscopy. *Anal. Biochem.* **2002**, *304* (2), 220–230.

https://doi.org/10.1006/abio.2002.5637.

(126)   Gowda, G. A. N.; Raftery, D. Quantitating Metabolites in Protein Precipitated Serum Using NMR Spectroscopy. *Anal. Chem.* **2014**, *86* (11), 5433–5440. https://doi.org/10.1021/ac5005103.

(127)   Li, J.; Vosegaard, T.; Guo, Z. Applications of Nuclear Magnetic Resonance in Lipid Analyses: An Emerging Powerful Tool for Lipidomics Studies. *Prog. Lipid Res.* **2017**, *68*, 37–56. https://doi.org/10.1016/j.plipres.2017.09.003.

(128)   Marchand, J.; Martineau, E.; Guitton, Y.; Le Bizec, B.; Dervilly-Pinel, G.; Giraudeau, P. A Multidimensional 1H NMR Lipidomics Workflow to Address Chemical Food Safety Issues. *Metabolomics* **2018**, *14* (5), 60. https://doi.org/10.1007/s11306-018-1360-x.

(129)   Giraudeau, P. NMR-Based Metabolomics and Fluxomics: Developments and Future Prospects. *The Analyst* **2020**, *145* (7), 2457–2472. https://doi.org/10.1039/D0AN00142B.

(130)   Grzesiek, S.; Bax, A. The Importance of Not Saturating Water in Protein NMR. Application to Sensitivity Enhancement and NOE Measurements. *J. Am. Chem. Soc.* **1993**, *115* (26), 12593–12594. https://doi.org/10.1021/ja00079a052.

(131)   Mckay, R. T. How the 1D-NOESY Suppresses Solvent Signal in Metabonomics NMR Spectroscopy: An Examination of the Pulse Sequence Components and Evolution. *Concepts Magn. Reson. Part A* **2011**, *38A* (5), 197–220. https://doi.org/10.1002/cmr.a.20223.

(132)   Liland, K. H. Multivariate Methods in Metabolomics – from Pre-Processing to Dimension Reduction and Statistical Analysis. *TrAC Trends Anal. Chem.* **2011**, *30* (6), 827–841. https://doi.org/10.1016/j.trac.2011.02.007.

(133)   Martin, M.; Legat, B.; Leenders, J.; Vanwinsberghe, J.; Rousseau, R.; Boulanger, B.; Eilers, P. H. C.; De Tullio, P.; Govaerts, B. PepsNMR for 1H NMR Metabolomic Data Pre-Processing. *Anal. Chim. Acta* **2018**, *1019*, 1–13. https://doi.org/10.1016/j.aca.2018.02.067.

(134)   Eilers, P. H. C. A Perfect Smoother. *Anal. Chem.* **2003**, *75* (14), 3631–3636. https://doi.org/10.1021/ac034173t.

(135)   Whittaker, P. E. T. On a New Method of Graduation. 13.

(136)   Frasso, G.; Eilers, P. H. L- and V-Curves for Optimal Smoothing. *Stat. Model.* **2015**, *15* (1), 91–111. https://doi.org/10.1177/1471082X14549288.

(137)   Eilers, P. H.; Boelens, H. F. Baseline Correction with Asymmetric Least Squares Smoothing. *Leiden Univ. Med. Cent. Rep.* **2005**, *1* (1), 5.

(138)   Vu, T.; Laukens, K. Getting Your Peaks in Line: A Review of Alignment Methods for NMR Spectral Data. *Metabolites* **2013**, *3* (2), 259–276. https://doi.org/10.3390/metabo3020259.

(139)   Bloemberg, T. G.; Gerretzen, J.; Lunshof, A.; Wehrens, R.; Buydens, L. M. C. Warping Methods for Spectroscopic and Chromatographic Signal Alignment: A Tutorial. *Anal. Chim. Acta* **2013**, *781*, 14–32. https://doi.org/10.1016/j.aca.2013.03.048.

(140)   Savorani, F.; Tomasi, G.; Engelsen, S. B. Icoshift: A Versatile Tool for the Rapid Alignment of 1D NMR Spectra. *J. Magn. Reson.* **2010**, *202* (2), 190–202. https://doi.org/10.1016/j.jmr.2009.11.012.

(141)   Eilers, P. H. C. Parametric Time Warping. *Anal. Chem.* **2004**, *76* (2), 404–411. https://doi.org/10.1021/ac034800e.

(142)   van Nederkassel, A. M.; Daszykowski, M.; Eilers, P. H. C.; Heyden, Y. V. A Comparison of

Three Algorithms for Chromatograms Alignment. *J. Chromatogr. A* **2006**, *1118* (2), 199–210. https://doi.org/10.1016/j.chroma.2006.03.114.

(143)   Giraudeau, P.; Tea, I.; Remaud, G. S.; Akoka, S. Reference and Normalization Methods: Essential Tools for the Intercomparison of NMR Spectra. *J. Pharm. Biomed. Anal.* **2014**, *93*, 3–16. https://doi.org/10.1016/j.jpba.2013.07.020.

(144)   Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in [1] H NMR Metabonomics. *Anal. Chem.* **2006**, *78* (13), 4281–4290. https://doi.org/10.1021/ac051632c.

(145)   Craig, A.; Cloarec, O.; Holmes, E.; Nicholson, J. K.; Lindon, J. C. Scaling and Normalization Effects in NMR Spectroscopic Metabonomic Data Sets. *Anal. Chem.* **2006**, *78* (7), 2262–2267. https://doi.org/10.1021/ac0519312.

(146)   Vu, T.; Riekeberg, E.; Qiu, Y.; Powers, R. Comparing Normalization Methods and the Impact of Noise. *Metabolomics* **2018**, *14* (8), 108. https://doi.org/10.1007/s11306-018-1400-6.

(147)   De Meyer, T.; Sinnaeve, D.; Van Gasse, B.; Tsiporkova, E.; Rietzschel, E. R.; De Buyzere, M. L.; Gillebert, T. C.; Bekaert, S.; Martins, J. C.; Van Criekinge, W. NMR-Based Characterization of Metabolic Alterations in Hypertension Using an Adaptive, Intelligent Binning Algorithm. *Anal. Chem.* **2008**, *80* (10), 3783–3790. https://doi.org/10.1021/ac7025964.

(148)   Beirnaert, C.; Meysman, P.; Vu, T. N.; Hermans, N.; Apers, S.; Pieters, L.; Covaci, A.; Laukens, K. Speaq 2.0: A Complete Workflow for High-Throughput 1D NMR Spectra Processing and Quantification. *PLOS Comput. Biol.* **2018**, *14* (3), e1006018. https://doi.org/10.1371/journal.pcbi.1006018.

(149)   Weljie, A. M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C. M. Targeted Profiling: Quantitative Analysis of [1] H NMR Metabolomics Data. *Anal. Chem.* **2006**, *78* (13), 4430–4442. https://doi.org/10.1021/ac060209g.

(150)   Bingol, K. Recent Advances in Targeted and Untargeted Metabolomics by NMR and MS/NMR Methods. *High-Throughput* **2018**, *7* (2), 9. https://doi.org/10.3390/ht7020009.

(151)   Hao, J.; Astle, W.; De Iorio, M.; Ebbels, T. M. D. BATMAN--an R Package for the Automated Quantification of Metabolites from Nuclear Magnetic Resonance Spectra Using a Bayesian Model. *Bioinformatics* **2012**, *28* (15), 2088–2090. https://doi.org/10.1093/bioinformatics/bts308.

(152)   Ravanbakhsh, S.; Liu, P.; Bjordahl, T. C.; Mandal, R.; Grant, J. R.; Wilson, M.; Eisner, R.; Sinelnikov, I.; Hu, X.; Luchinat, C.; Greiner, R.; Wishart, D. S. Accurate, Fully-Automated NMR Spectral Profiling for Metabolomics. *PLOS ONE* **2015**, *10* (5), e0124219. https://doi.org/10.1371/journal.pone.0124219.

(153)   Tardivel, P. J. C.; Canlet, C.; Lefort, G.; Tremblay-Franco, M.; Debrauwer, L.; Concordet, D.; Servien, R. ASICS: An Automatic Method for Identification and Quantification of Metabolites in Complex 1D 1H NMR Spectra. *Metabolomics* **2017**, *13* (10), 109. https://doi.org/10.1007/s11306-017-1244-5.

(154)   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* **2018**, *14* (3), 24. https://doi.org/10.1007/s11306-018-1319-y.

(155)   Röhnisch, H. E.; Eriksson, J.; Müllner, E.; Agback, P.; Sandström, C.; Moazzami, A. A. AQuA: An Automated Quantification Algorithm for High-Throughput NMR-Based Metabolomics

and Its Application in Human Plasma. *Anal. Chem.* **2018**, *90* (3), 2095–2102. https://doi.org/10.1021/acs.analchem.7b04324.

(156)   Barrilero, R.; Gil, M.; Amigó, N.; Dias, C. B.; Wood, L. G.; Garg, M. L.; Ribalta, J.; Heras, M.; Vinaixa, M.; Correig, X. LipSpin: A New Bioinformatics Tool for Quantitative [1] H NMR Lipid Profiling. *Anal. Chem.* **2018**, *90* (3), 2031–2040. https://doi.org/10.1021/acs.analchem.7b04148.

(157)   Emwas, A.-H.; Saccenti, E.; Gao, X.; McKay, R. T.; dos Santos, V. A. P. M.; Roy, R.; Wishart, D. S. Recommended Strategies for Spectral Processing and Post-Processing of 1D 1H-NMR Data of Biofluids with a Particular Focus on Urine. *Metabolomics* **2018**, *14* (3), 31. https://doi.org/10.1007/s11306-018-1321-4.

(158)   Madsen, R.; Lundstedt, T.; Trygg, J. Chemometrics in Metabolomics—A Review in Human Disease Diagnosis. *Anal. Chim. Acta* **2010**, *659* (1–2), 23–33. https://doi.org/10.1016/j.aca.2009.11.042.

(159)   Alonso, A.; Marsal, S.; Juliã, A. Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. *Front. Bioeng. Biotechnol.* **2015**, *3*. https://doi.org/10.3389/fbioe.2015.00023.

(160)   Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2* (1–3), 37–52. https://doi.org/10.1016/0169-7439(87)80084-9.

(161)   Gika, H. G.; Theodoridis, G. A.; Wilson, I. D. Liquid Chromatography and Ultra-Performance Liquid Chromatography–Mass Spectrometry Fingerprinting of Human Urine. *J. Chromatogr. A* **2008**, *1189* (1–2), 314–322. https://doi.org/10.1016/j.chroma.2007.10.066.

(162)   Rasmussen, L. G.; Savorani, F.; Larsen, T. M.; Dragsted, L. O.; Astrup, A.; Engelsen, S. B. Standardization of Factors That Influence Human Urine Metabolomics. *Metabolomics* **2011**, *7* (1), 71–83. https://doi.org/10.1007/s11306-010-0234-7.

(163)   Fonville, J. M.; Richards, S. E.; Barton, R. H.; Boulange, C. L.; Ebbels, T. M. D.; Nicholson, J. K.; Holmes, E.; Dumas, M.-E. The Evolution of Partial Least Squares Models and Related Chemometric Approaches in Metabonomics and Metabolic Phenotyping. *J. Chemom.* **2010**, *24* (11–12), 636–649. https://doi.org/10.1002/cem.1359.

(164)   Trygg, J.; Wold, S. Orthogonal Projections to Latent Structures (O-PLS). *J. Chemom.* **2002**, *16* (3), 119–128. https://doi.org/10.1002/cem.695.

(165)   Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S. Translational Biomarker Discovery in Clinical Metabolomics: An Introductory Tutorial. *Metabolomics* **2013**, *9* (2), 280–299. https://doi.org/10.1007/s11306-012-0482-9.

(166)   Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; van Velzen, E. J. J.; van Duijnhoven, J. P. M.; van Dorsten, F. A. Assessment of PLSDA Cross Validation. *Metabolomics* **2008**, *4* (1), 81–89. https://doi.org/10.1007/s11306-007-0099-6.

(167)   Chong, J.; Wishart, D. S.; Xia, J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Curr. Protoc. Bioinforma.* **2019**, *68* (1). https://doi.org/10.1002/cpbi.86.

(168)   Kohl, M.; Wiese, S.; Warscheid, B. Cytoscape: Software for Visualization and Analysis of Biological Networks. In *Data Mining in Proteomics*; Hamacher, M., Eisenacher, M., Stephan, C., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2011; Vol. 696, pp 291–303. https://doi.org/10.1007/978-1-60761-987-1_18.

(169)   Cottret, L.; Wildridge, D.; Vinson, F.; Barrett, M. P.; Charles, H.; Sagot, M.-F.; Jourdan, F. MetExplore: A Web Server to Link Metabolomic Experiments and Genome-Scale Metabolic Networks. *Nucleic Acids Res.* **2010**, *38* (Web Server), W132–W137.

https://doi.org/10.1093/nar/gkq312.

(170)   Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. Proposed Minimum Reporting Standards for Chemical Analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3* (3), 211–221. https://doi.org/10.1007/s11306-007-0082-2.

(171)   Kale, N. S.; Haug, K.; Conesa, P.; Jayseelan, K.; Moreno, P.; Rocca-Serra, P.; Nainala, V. C.; Spicer, R. A.; Williams, M.; Li, X.; Salek, R. M.; Griffin, J. L.; Steinbeck, C. MetaboLights: An Open-Access Database Repository for Metabolomics Data. *Curr. Protoc. Bioinforma.* **2016**, *53* (1). https://doi.org/10.1002/0471250953.bi1413s53.

(172)   Clish, C. B. Metabolomics: An Emerging but Powerful Tool for Precision Medicine. *Mol. Case Stud.* **2015**, *1* (1), a000588. https://doi.org/10.1101/mcs.a000588.

(173)   Frédérich, M.; Pirotte, B.; Fillet, M.; de Tullio, P. Metabolomics as a Challenging Approach for Medicinal Chemistry and Personalized Medicine. *J. Med. Chem.* **2016**, *59* (19), 8649–8666. https://doi.org/10.1021/acs.jmedchem.5b01335.

(174)   Strimbu, K.; Tavel, J. A. What Are Biomarkers?: *Curr. Opin. HIV AIDS* **2010**, *5* (6), 463–466. https://doi.org/10.1097/COH.0b013e32833ed177.

(175)   K. Trivedi, D.; A. Hollywood, K.; Goodacre, R. Metabolomics for the Masses: The Future of Metabolomics in a Personalized World. *Eur. J. Mol. Clin. Med.* **2017**, *3* (6), 294. https://doi.org/10.1016/j.nhtm.2017.06.001.

(176)   Dhanasekaran, S. M.; Barrette, T. R.; Ghosh, D.; Shah, R.; Varambally, S.; Kurachi, K.; Pienta, K. J.; Rubin, M. A.; Chinnaiyan, A. M. Delineation of Prognostic Biomarkers in Prostate Cancer. *Nature* **2001**, *412* (6849), 822–826. https://doi.org/10.1038/35090585.

(177)   for "Precision Medicine and Pharmacometabolomics Task Group"-Metabolomics Society Initiative; Beger, R. D.; Dunn, W.; Schmidt, M. A.; Gross, S. S.; Kirwan, J. A.; Cascante, M.; Brennan, L.; Wishart, D. S.; Oresic, M.; Hankemeier, T.; Broadhurst, D. I.; Lane, A. N.; Suhre, K.; Kastenmüller, G.; Sumner, S. J.; Thiele, I.; Fiehn, O.; Kaddurah-Daouk, R. Metabolomics Enables Precision Medicine: "A White Paper, Community Perspective." *Metabolomics* **2016**, *12* (9), 149. https://doi.org/10.1007/s11306-016-1094-6.

(178)   Brown, C.; Green, B.; Thompson, R.; den Hollander, A.; Lengyel, I.; on behalf of the EYE-RISK consortium. Metabolomics and Age-Related Macular Degeneration. *Metabolites* **2018**, *9* (1), 4. https://doi.org/10.3390/metabo9010004.

(179)   Osborn, M. P.; Park, Y.; Parks, M. B.; Burgess, L. G.; Uppal, K.; Lee, K.; Jones, D. P.; Brantley, M. A. Metabolome-Wide Association Study of Neovascular Age-Related Macular Degeneration. *PLoS ONE* **2013**, *8* (8), e72737. https://doi.org/10.1371/journal.pone.0072737.

(180)   Laíns, I.; Kelly, R. S.; Miller, J. B.; Silva, R.; Vavvas, D. G.; Kim, I. K.; Murta, J. N.; Lasky-Su, J.; Miller, J. W.; Husain, D. Human Plasma Metabolomics Study across All Stages of Age-Related Macular Degeneration Identifies Potential Lipid Biomarkers. *Ophthalmology* **2018**, *125* (2), 245–254. https://doi.org/10.1016/j.ophtha.2017.08.008.

(181)   Mitchell, S. L.; Uppal, K.; Williamson, S. M.; Liu, K.; Burgess, L. G.; Tran, V.; Umfress, A. C.; Jarrell, K. L.; Cooke Bailey, J. N.; Agarwal, A.; Pericak-Vance, M.; Haines, J. L.; Scott, W. K.; Jones, D. P.; Brantley, M. A. The Carnitine Shuttle Pathway Is Altered in Patients With Neovascular Age-Related Macular Degeneration. *Investig. Opthalmology Vis. Sci.* **2018**, *59* (12), 4978.

https://doi.org/10.1167/iovs.18-25137.

(182)   Kersten, E.; Paun, C. C.; Schellevis, R. L.; Hoyng, Carel. B.; Delcourt, C.; Lengyel, I.; Peto, T.; Ueffing, M.; Klaver, C. C. W.; Dammeier, S.; den Hollander, A. I.; de Jong, E. K. Systemic and Ocular Fluid Compounds as Potential Biomarkers in Age-Related Macular Degeneration. *Surv. Ophthalmol.* **2018**, *63* (1), 9–39. https://doi.org/10.1016/j.survophthal.2017.05.003.

(183)   Cai, J.; Nelson, K. C.; Wu, M.; Sternberg, P.; Jones, D. P. Oxidative Damage and Protection of the RPE. *Prog. Retin. Eye Res.* **2000**, *19* (2), 205–221. https://doi.org/10.1016/S1350-9462(99)00009-9.

(184)   Uğurlu, N.; Aşık, M. D.; Yülek, F.; Neselioglu, S.; Cagil, N. Oxidative Stress and Anti-Oxidative Defence in Patients with Age-Related Macular Degeneration. *Curr. Eye Res.* **2013**, *38* (4), 497–502. https://doi.org/10.3109/02713683.2013.774023.

(185)   Totan, Y. Plasma Malondialdehyde and Nitric Oxide Levels in Age Related Macular Degeneration. *Br. J. Ophthalmol.* **2001**, *85* (12), 1426–1428. https://doi.org/10.1136/bjo.85.12.1426.

(186)   Totan, Y.; Yağcı, R.; Bardak, Y.; Özyurt, H.; Kendir, F.; Yılmaz, G.; Şahin, Ş.; Şahin Tığ, U. Oxidative Macromolecular Damage in Age-Related Macular Degeneration. *Curr. Eye Res.* **2009**, *34* (12), 1089–1093. https://doi.org/10.3109/02713680903353772.

(187)   Jia, L.; Dong, Y.; Yang, H.; Pan, X.; Fan, R.; Zhai, L. Serum Superoxide Dismutase and Malondialdehyde Levels in a Group of Chinese Patients with Age-Related Macular Degeneration. *Aging Clin. Exp. Res.* **2011**, *23* (4), 264–267. https://doi.org/10.1007/BF03324965.

(188)   Shen, X.; Jia, L. H.; Zhao, P.; Fan, R.; Pan, X. Y.; Yang, H. M.; Liu, L. Changes in Blood Oxidative and Antioxidant Parameters in a Group of Chinese Patients with Age-Related Macular Degeneration. *J. Nutr. Health Aging* **2012**, *16* (3), 201–204. https://doi.org/10.1007/s12603-011-0350-8.

(189)   Ikeda, T.; Obayashi, H.; Hasegawa, G.; Nakamura, N.; Yoshikawa, T.; Imamura, Y.; Koizumi, K.; Kinoshita, S. Paraoxonase Gene Polymorphisms and Plasma Oxidized Low-Density Lipoprotein Level as Possible Risk Factors for Exudative Age-Related Macular Degeneration. *Am. J. Ophthalmol.* **2001**, *132* (2), 191–195. https://doi.org/10.1016/S0002-9394(01)00975-8.

(190)   Javadzadeh, A.; Ghorbanihaghjo, A.; Bahreini, E.; Rashtchizadeh, N.; Argani, H.; Alizadeh, S. Plasma Oxidized LDL and Thiol-Containing Molecules in Patients with Exudative Age-Related Macular Degeneration. *Mol. Vis.* **2010**, 7.

(191)   Javadzadeh, A.; Ghorbanihaghjo, A.; Bahreini, E.; Rashtchizadeh, N.; Argani, H.; Alizadeh, S. SERUM PARAOXONASE PHENOTYPE DISTRIBUTION IN EXUDATIVE AGE-RELATED MACULAR DEGENERATION AND ITS RELATIONSHIP TO HOMOCYSTEINE AND OXIDIZED LOW-DENSITY LIPOPROTEIN. *Retina* **2012**, *32* (4), 658–666. https://doi.org/10.1097/IAE.0b013e31822529b1.

(192)   Tsai, D.-C.; Charng, M.-J.; Lee, F.-L.; Hsu, W.-M.; Chen, S.-J. Different Plasma Levels of Vascular Endothelial Growth Factor and Nitric Oxide between Patients with Choroidal and Retinal Neovascularization. *Ophthalmologica* **2006**, *220* (4), 246–251. https://doi.org/10.1159/000093079.

(193)   Evereklioglu, C.; Er, H.; Doganay, S.; Cekmen, M.; Turkoz, Y.; Otlu, B.; Ozerol, E. [No Title Found]. *Doc. Ophthalmol.* **2003**, *106* (2), 129–136. https://doi.org/10.1023/A:1022512402811.

(194)   Heuberger, R. A.; Fisher, A. I.; Jacques, P. F.; Klein, R.; Klein, B. E.; Palta, M.; Mares-Perlman, J. A. Relation of Blood Homocysteine and Its Nutritional Determinants to Age-Related Maculopathy in the Third National Health and Nutrition Examination Survey. *Am. J. Clin. Nutr.* **2002**,

76 (4), 897–902. https://doi.org/10.1093/ajcn/76.4.897.

(195)   Axer-Siegel, R.; Bourla, D.; Ehrlich, R.; Dotan, G.; Benjamini, Y.; Gavendo, S.; Weinberger, D.; Sela, B.-A. Association of Neovascular Age-Related Macular Degeneration and Hyperhomocysteinemia. *Am. J. Ophthalmol.* **2004**, *137* (1), 84–89. https://doi.org/10.1016/S0002-9394(03)00864-X.

(196)   Ghosh, S.; Saha, M.; Das, D. A Study on Plasma Homocysteine Level in Age-Related Macular Degeneration. *Nepal. J. Ophthalmol.* **2013**, *5* (2), 195–200. https://doi.org/10.3126/nepjoph.v5i2.8728.

(197)   Manresa, N.; Mulero, J.; Losada, M.; Zafrilla, P. EFFECT OF PEGAPTANIB AND RANIBIZUMAB ON PLASMA AND VITREOUS HOMOCYSTEINE IN PATIENTS WITH EXUDATIVE AGE-RELATED MACULAR DEGENERATION. *Retina* **2015**, *35* (9), 1765–1771. https://doi.org/10.1097/IAE.0000000000000552.

(198)   Christen, W. G.; Cook, N. R.; Ridker, P. M.; Buring, J. E. Prospective Study of Plasma Homocysteine Level and Risk of Age-Related Macular Degeneration in Women. *Ophthalmic Epidemiol.* **2015**, *22* (2), 85–93. https://doi.org/10.3109/09286586.2015.1012272.

(199)   Delcourt, C. Age-Related Macular Degeneration and Antioxidant Status in the POLA Study. *Arch. Ophthalmol.* **1999**, *117* (10), 1384. https://doi.org/10.1001/archopht.117.10.1384.

(200)   Coral, K.; Raman, R.; Rathi, S.; Rajesh, M.; Sulochana, K. N.; Angayarkanni, N.; Paul, P. G.; Ramakrishnan, S. Plasma Homocysteine and Total Thiol Content in Patients with Exudative Age-Related Macular Degeneration. *Eye* **2006**, *20* (2), 203–207. https://doi.org/10.1038/sj.eye.6701853.

(201)   Michikawa, T.; Ishida, S.; Nishiwaki, Y.; Kikuchi, Y.; Tsuboi, T.; Hosoda, K.; Ishigami, A.; Iwasawa, S.; Nakano, M.; Takebayashi, T. Serum Antioxidants and Age-Related Macular Degeneration among Older Japanese. *Asia Pac. J. Clin. Nutr.* **2009**, *18* (1), 1–7.

(202)   Zhou, H.; Zhao, X.; Johnson, E. J.; Lim, A.; Sun, E.; Yu, J.; Zhang, Y.; Liu, X.; Snellingen, T.; Shang, F.; Liu, N. Serum Carotenoids and Risk of Age-Related Macular Degeneration in a Chinese Population Sample. *Investig. Opthalmology Vis. Sci.* **2011**, *52* (7), 4338. https://doi.org/10.1167/iovs.10-6519.

(203)   Tan, W.; Zou, J.; Yoshida, S.; Jiang, B.; Zhou, Y. The Role of Inflammation in Age-Related Macular Degeneration. *Int. J. Biol. Sci.* **2020**, *16* (15), 2989–3001. https://doi.org/10.7150/ijbs.49890.

(204)   Diakos, C. I.; Charles, K. A.; McMillan, D. C.; Clarke, S. J. Cancer-Related Inflammation and Treatment Effectiveness. *Lancet Oncol.* **2014**, *15* (11), e493–e503. https://doi.org/10.1016/S1470-2045(14)70263-3.

(205)   Heppner, F. L.; Ransohoff, R. M.; Becher, B. Immune Attack: The Role of Inflammation in Alzheimer Disease. *Nat. Rev. Neurosci.* **2015**, *16* (6), 358–372. https://doi.org/10.1038/nrn3880.

(206)   Hageman, G. An Integrated Hypothesis That Considers Drusen as Biomarkers of Immune-Mediated Processes at the RPE-Bruch's Membrane Interface in Aging and Age-Related Macular Degeneration. *Prog. Retin. Eye Res.* **2001**, *20* (6), 705–732. https://doi.org/10.1016/S1350-9462(01)00010-6.

(207)   Li, C.-M.; Clark, M. E.; Rudolf, M.; Curcio, C. A. Distribution and Composition of Esterified and Unesterified Cholesterol in Extra-Macular Drusen. *Exp. Eye Res.* **2007**, *85* (2), 192–201. https://doi.org/10.1016/j.exer.2007.04.002.

(208)   Wang, L.; Clark, M. E.; Crossman, D. K.; Kojima, K.; Messinger, J. D.; Mobley, J. A.; Curcio, C. A. Abundant Lipid and Protein Components of Drusen. *PLoS ONE* **2010**, *5* (4), e10329. https://doi.org/10.1371/journal.pone.0010329.

(209)   van Leeuwen, E. M.; Emri, E.; Merle, B. M. J.; Colijn, J. M.; Kersten, E.; Cougnard-Gregoire, A.; Dammeier, S.; Meester-Smoor, M.; Pool, F. M.; de Jong, E. K.; Delcourt, C.; Rodrigez-Bocanegra, E.; Biarnés, M.; Luthert, P. J.; Ueffing, M.; Klaver, C. C. W.; Nogoceke, E.; den Hollander, A. I.; Lengyel, I. A New Perspective on Lipid Research in Age-Related Macular Degeneration. *Prog. Retin. Eye Res.* **2018**, *67*, 56–86. https://doi.org/10.1016/j.preteyeres.2018.04.006.

(210)   Ouchi, M.; Ikeda, T.; Nakamura, K.; Harino, S.; Kinoshita, S. A Novel Relation of Fatty Acid with Age-Related Macular Degeneration. *Ophthalmologica* **2002**, *216* (5), 363–367. https://doi.org/10.1159/000066178.

(211)   Kabasawa, S.; Mori, K.; Horie-Inoue, K.; Gehlbach, P. L.; Inoue, S.; Awata, T.; Katayama, S.; Yoneya, S. Associations of Cigarette Smoking But Not Serum Fatty Acids with Age-Related Macular Degeneration in a Japanese Population. *Ophthalmology* **2011**, *118* (6), 1082–1088. https://doi.org/10.1016/j.ophtha.2010.10.012.

(212)   Merle, B. M. J.; Delyfer, M.-N.; Korobelnik, J.-F.; Rougier, M.-B.; Malet, F.; Féart, C.; Le Goff, M.; Peuchant, E.; Letenneur, L.; Dartigues, J.-F.; Colin, J.; Barberger-Gateau, P.; Delcourt, C. High Concentrations of Plasma N3 Fatty Acids Are Associated with Decreased Risk for Late Age-Related Macular Degeneration. *J. Nutr.* **2013**, *143* (4), 505–511. https://doi.org/10.3945/jn.112.171033.

(213)   Orban, T.; Johnson, W. M.; Dong, Z.; Maeda, T.; Maeda, A.; Sakai, T.; Tsuneoka, H.; Mieyal, J. J.; Palczewski, K. Serum Levels of Lipid Metabolites in Age-related Macular Degeneration. *FASEB J.* **2015**, *29* (11), 4579–4588. https://doi.org/10.1096/fj.15-275289.

(214)   Merle, B. M. J.; Benlian, P.; Puche, N.; Bassols, A.; Delcourt, C.; Souied, E. H. Circulating Omega-3 Fatty Acids and Neovascular Age-Related Macular Degeneration. *Investig. Opthalmology Vis. Sci.* **2014**, *55* (3), 2010. https://doi.org/10.1167/iovs.14-13916.

(215)   Piepoli, M. F.; Hoes, A. W.; Agewall, S.; Albus, C.; Brotons, C.; Catapano, A. L.; Cooney, M.-T.; Corrà, U.; Cosyns, B.; Deaton, C.; Graham, I.; Hall, M. S.; Hobbs, F. D. R.; Løchen, M.-L.; Löllgen, H.; Marques-Vidal, P.; Perk, J.; Prescott, E.; Redon, J.; Richter, D. J.; Sattar, N.; Smulders, Y.; Tiberi, M.; van der Worp, H. B.; van Dis, I.; Verschuren, W. M. M.; Binno, S.; ESC Scientific Document Group. 2016 European Guidelines on Cardiovascular Disease Prevention in Clinical Practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (Constituted by Representatives of 10 Societies and by Invited Experts)Developed with the Special Contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur. Heart J.* **2016**, *37* (29), 2315–2381. https://doi.org/10.1093/eurheartj/ehw106.

(216)   Cougnard-Grégoire, A.; Delyfer, M.-N.; Korobelnik, J.-F.; Rougier, M.-B.; Le Goff, M.; Dartigues, J.-F.; Barberger-Gateau, P.; Delcourt, C. Elevated High-Density Lipoprotein Cholesterol and Age-Related Macular Degeneration: The Alienor Study. *PLoS ONE* **2014**, *9* (3), e90973. https://doi.org/10.1371/journal.pone.0090973.

(217)   G, H. B.; Rao, V. S.; Kakkar, V. V. Friend Turns Foe: Transformation of Anti-Inflammatory HDL to Proinflammatory HDL during Acute-Phase Response. *Cholesterol* **2011**, *2011*, 1–7. https://doi.org/10.1155/2011/274629.

(218)   Eren, E.; Yilmaz, N.; Aydin, O. High Density Lipoprotein and It's Dysfunction. *Open Biochem. J.* **2012**, *6*, 78–93. https://doi.org/10.2174/1874091X01206010078.

(219)   Parmeggiani, F.; Romano, M. R.; Costagliola, C.; Semeraro, F.; Incorvaia, C.; D'Angelo, S.; Perri, P.; De Palma, P.; De Nadai, K.; Sebastiani, A. Mechanism of Inflammation in Age-Related Macular Degeneration. *Mediators Inflamm.* **2012**, *2012*, 1–16. https://doi.org/10.1155/2012/546786.

(220)   Javadzadeh, A.; Ghorbanihaghjo, A.; Rashtchizadeh, N.; Rafeey, M.; Rahimi-Ardabili, B. Enhanced Susceptibility of Low-Density Lipoprotein to Oxidation in Wet Type Age-Related Macular Degeneration in Male Patients. *Saudi Med. J.* **2007**, *28* (2), 221–224.

(221)   Reynolds, R.; Rosner, B.; Seddon, J. M. Serum Lipid Biomarkers and Hepatic Lipase Gene Associations with Age-Related Macular Degeneration. *Ophthalmology* **2010**, *117* (10), 1989–1995. https://doi.org/10.1016/j.ophtha.2010.07.009.

(222)   Superko, H. R. Advanced Lipoprotein Testing and Subfractionation Are Clinically Useful. *Circulation* **2009**, *119* (17), 2383–2395. https://doi.org/10.1161/CIRCULATIONAHA.108.809582.

(223)   Friedewald, W. T.; Levy, R. I.; Fredrickson, D. S. Estimation of the Concentration of Low-Density Lipoprotein Cholesterol in Plasma, Without Use of the Preparative Ultracentrifuge. *Clin. Chem.* **1972**, *18* (6), 499–502. https://doi.org/10.1093/clinchem/18.6.499.

(224)   Superko, H. R. High-Density Lipoprotein Cholesterol Measurements: A Help or Hindrance in Practical Clinical Medicine? *JAMA* **1986**, *256* (19), 2714. https://doi.org/10.1001/jama.1986.03380190084030.

(225)   Lambert, V.; Hansen, S.; Schoumacher, M.; Lecomte, J.; Leenders, J.; Hubert, P.; Herfs, M.; Blacher, S.; Carnet, O.; Yip, C.; Blaise, P.; Duchateau, E.; Locht, B.; Thys, M.; Cavalier, E.; Gothot, A.; Govaerts, B.; Rakic, J.-M.; Noel, A.; de Tullio, P. Pyruvate Dehydrogenase Kinase/Lactate Axis: A Therapeutic Target for Neovascular Age-Related Macular Degeneration Identified by Metabolomics. *J. Mol. Med.* **2020**, *98* (12), 1737–1751. https://doi.org/10.1007/s00109-020-01994-9.

(226)   Lambert, V.; Lecomte, J.; Hansen, S.; Blacher, S.; Gonzalez, M.-L. A.; Struman, I.; Sounni, N. E.; Rozet, E.; de Tullio, P.; Foidart, J. M.; Rakic, J.-M.; Noel, A. Laser-Induced Choroidal Neovascularization Model to Study Age-Related Macular Degeneration in Mice. **2013**, 15.

(227)   Beckert, S.; Farrahi, F.; Aslam, R. S.; Scheuenstuhl, H.; Königsrainer, A.; Hussain, M. Z.; Hunt, T. K. Lactate Stimulates Endothelial Cell Migration: Lactate Stimulates Endothelial Cell Migration. *Wound Repair Regen.* **2006**, *14* (3), 321–324. https://doi.org/10.1111/j.1743-6109.2006.00127.x.

(228)   Porporato, P. E.; Payen, V. L.; De Saedeleer, C. J.; Préat, V.; Thissen, J.-P.; Feron, O.; Sonveaux, P. Lactate Stimulates Angiogenesis and Accelerates the Healing of Superficial and Ischemic Wounds in Mice. *Angiogenesis* **2012**, *15* (4), 581–592. https://doi.org/10.1007/s10456-012-9282-0.

(229)   Skeie, J. M.; Mullins, R. F. Macrophages in Neovascular Age-Related Macular Degeneration: Friends or Foes? *Eye* **2009**, *23* (4), 747–755. https://doi.org/10.1038/eye.2008.206.

(230)   Sakurai, E.; Anand, A.; Ambati, B. K.; van Rooijen, N.; Ambati, J. Macrophage Depletion Inhibits Experimental Choroidal Neovascularization. *Investig. Opthalmology Vis. Sci.* **2003**, *44* (8), 3578. https://doi.org/10.1167/iovs.03-0097.

(231)   Zandi, S.; Nakao, S.; Chun, K.-H.; Fiorina, P.; Sun, D.; Arita, R.; Zhao, M.; Kim, E.; Schueller, O.; Campbell, S.; Taher, M.; Melhorn, M. I.; Schering, A.; Gatti, F.; Tezza, S.; Xie, F.; Vergani, A.; Yoshida, S.; Ishikawa, K.; Yamaguchi, M.; Sasaki, F.; Schmidt-Ullrich, R.; Hata, Y.; Enaida, H.; Yuzawa, M.; Yokomizo, T.; Kim, Y.-B.; Sweetnam, P.; Ishibashi, T.; Hafezi-Moghadam, A. ROCK-Isoform-Specific Polarization of Macrophages Associated with Age-Related Macular Degeneration. *Cell Rep.* **2015**, *10* (7), 1173–1186. https://doi.org/10.1016/j.celrep.2015.01.050.

(232)   Jetten, N.; Verbruggen, S.; Gijbels, M. J.; Post, M. J.; De Winther, M. P. J.; Donners, M. M. P. C. Anti-Inflammatory M2, but Not pro-Inflammatory M1 Macrophages Promote Angiogenesis in Vivo. *Angiogenesis* **2014**, *17* (1), 109–118. https://doi.org/10.1007/s10456-013-9381-6.

(233)   Colegio, O. R.; Chu, N.-Q.; Szabo, A. L.; Chu, T.; Rhebergen, A. M.; Jairam, V.; Cyrus, N.; Brokowski, C. E.; Eisenbarth, S. C.; Phillips, G. M.; Cline, G. W.; Phillips, A. J.; Medzhitov, R. Functional Polarization of Tumour-Associated Macrophages by Tumour-Derived Lactic Acid. *Nature* **2014**, *513* (7519), 559–563. https://doi.org/10.1038/nature13490.

(234)   Dhup, S.; Kumar Dadhich, R.; Ettore Porporato, P.; Sonveaux, P. Multiple Biological Activities of Lactic Acid in Cancer: Influences on Tumor Growth,Angiogenesis and Metastasis. *Curr. Pharm. Des.* **2012**, *18* (10), 1319–1330. https://doi.org/10.2174/138161212799504902.

(235)   Saunier, V.; Merle, B. M. J.; Delyfer, M.-N.; Cougnard-Grégoire, A.; Rougier, M.-B.; Amouyel, P.; Lambert, J.-C.; Dartigues, J.-F.; Korobelnik, J.-F.; Delcourt, C. Incidence of and Risk Factors Associated With Age-Related Macular Degeneration: Four-Year Follow-up From the ALIENOR Study. *JAMA Ophthalmol.* **2018**, *136* (5), 473. https://doi.org/10.1001/jamaophthalmol.2018.0504.

(236)   Rodríguez, I. R.; Larrayoz, I. M. Cholesterol Oxidation in the Retina: Implications of 7KCh Formation in Chronic Inflammation and Age-Related Macular Degeneration. *J. Lipid Res.* **2010**, *51* (10), 2847–2862. https://doi.org/10.1194/jlr.R004820.

(237)   Mallol, R.; Rodriguez, M. A.; Brezmes, J.; Masana, L.; Correig, X. Human Serum/Plasma Lipoprotein Analysis by NMR: Application to the Study of Diabetic Dyslipidemia. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *70*, 1–24. https://doi.org/10.1016/j.pnmrs.2012.09.001.

(238)   Lounila, J.; Ala-Korpela, M.; Jokisaari, J.; Savolainen, M. J.; Kesäniemi, Y. A. Effects of Orientational Order and Particle Size on the NMR Line Positions of Lipoproteins. *Phys. Rev. Lett.* **1994**, *72* (25), 4049–4052. https://doi.org/10.1103/PhysRevLett.72.4049.

(239)   Otvos, J. D.; Jeyarajah, E. J.; Hayes, L. W.; Freedman, D. S.; Janjan, N. A.; Anderson, T. Relationships between the Proton Nuclear Magnetic Resonance Properties of Plasma Lipoproteins and Cancer. *Clin. Chem.* **1991**, *37* (3), 369–376. https://doi.org/10.1093/clinchem/37.3.369.

(240)   Mallol, R.; Rodriguez, M. A.; Brezmes, J.; Masana, L.; Correig, X. Human Serum/Plasma Lipoprotein Analysis by NMR: Application to the Study of Diabetic Dyslipidemia. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *70*, 1–24. https://doi.org/10.1016/j.pnmrs.2012.09.001.

(241)   Mallol, R.; Rodríguez, M. A.; Heras, M.; Vinaixa, M.; Plana, N.; Masana, L.; Morris, G. A.; Correig, X. Particle Size Measurement of Lipoprotein Fractions Using Diffusion-Ordered NMR Spectroscopy. *Anal. Bioanal. Chem.* **2012**, *402* (7), 2407–2415. https://doi.org/10.1007/s00216-011-5705-9.

(242)   Otvos, J. D.; Jeyarajah, E. J.; Bennett, D. W. Quantification of Plasma Lipoproteins by Proton Nuclear Magnetic Resonance Spectroscopy. *Clin. Chem.* **1991**, *37* (3), 377–386. https://doi.org/10.1093/clinchem/37.3.377.

(243)   Jeyarajah, E. J.; Cromwell, W. C.; Otvos, J. D. Lipoprotein Particle Analysis by Nuclear Magnetic Resonance Spectroscopy. *Clin. Lab. Med.* **2006**, *26* (4), 847–870. https://doi.org/10.1016/j.cll.2006.07.006.

(244)   Mallol, R.; Amigó, N.; Rodríguez, M. A.; Heras, M.; Vinaixa, M.; Plana, N.; Rock, E.; Ribalta, J.; Yanes, O.; Masana, L.; Correig, X. Liposcale: A Novel Advanced Lipoprotein Test Based on 2D Diffusion-Ordered 1H NMR Spectroscopy. *J. Lipid Res.* **2015**, *56* (3), 737–746. https://doi.org/10.1194/jlr.D050120.

(245)   Aguilar, J. A.; Nilsson, M.; Bodenhausen, G.; Morris, G. A. Spin Echo NMR Spectra without J Modulation. *Chem Commun* **2012**, *48* (6), 811–813. https://doi.org/10.1039/C1CC16699A.

(246)   Topgaard, D.; Sakellariou, D.; Pines, A. NMR Spectroscopy in Inhomogeneous B0 and B1 Fields with Non-Linear Correlation. *J. Magn. Reson.* **2005**, *175* (1), 1–10. https://doi.org/10.1016/j.jmr.2005.03.006.

(247)   Liquet, B.; Cao, K.-A. L.; Hocini, H.; Thiébaut, R. A Novel Approach for Biomarker Selection and the Integration of Repeated Measures Experiments from Two Assays. *BMC Bioinformatics* **2012**, *13* (1), 325. https://doi.org/10.1186/1471-2105-13-325.

(248)   Westerhuis, J. A.; van Velzen, E. J. J.; Hoefsloot, H. C. J.; Smilde, A. K. Multivariate Paired Data Analysis: Multilevel PLSDA versus OPLSDA. *Metabolomics* **2010**, *6* (1), 119–128. https://doi.org/10.1007/s11306-009-0185-z.

(249)   Lê Cao, K.-A.; Boitard, S.; Besse, P. Sparse PLS Discriminant Analysis: Biologically Relevant Feature Selection and Graphical Displays for Multiclass Problems. *BMC Bioinformatics* **2011**, *12* (1), 253. https://doi.org/10.1186/1471-2105-12-253.

(250)   Chew, E. Y. Nutrition, Genes, and Age-Related Macular Degeneration: What Have We Learned from the Trials? *Ophthalmologica* **2017**, *238* (1–2), 1–5. https://doi.org/10.1159/000473865.

(251)   Chao de la Barca, J. M.; Rondet-Courbis, B.; Ferré, M.; Muller, J.; Buisset, A.; Leruez, S.; Plubeau, G.; Macé, T.; Moureauzeau, L.; Chupin, S.; Tessier, L.; Blanchet, O.; Lenaers, G.; Procaccio, V.; Mirebeau-Prunier, D.; Simard, G.; Gohier, P.; Miléa, D.; Reynier, P. A Plasma Metabolomic Profiling of Exudative Age-Related Macular Degeneration Showing Carnosine and Mitochondrial Deficiencies. *J. Clin. Med.* **2020**, *9* (3), 631. https://doi.org/10.3390/jcm9030631.

(252)   Laíns, I.; Duarte, D.; Barros, A. S.; Martins, A. S.; Gil, J.; Miller, J. B.; Marques, M.; Mesquita, T.; Kim, I. K.; Cachulo, M. da L.; Vavvas, D.; Carreira, I. M.; Murta, J. N.; Silva, R.; Miller, J. W.; Husain, D.; Gil, A. M. Human Plasma Metabolomics in Age-Related Macular Degeneration (AMD) Using Nuclear Magnetic Resonance Spectroscopy. *PLOS ONE* **2017**, *12* (5), e0177749. https://doi.org/10.1371/journal.pone.0177749.

(253)   Han, G.; Wei, P.; He, M.; Teng, H.; Chu, Y. Metabolomic Profiling of the Aqueous Humor in Patients with Wet Age-Related Macular Degeneration Using UHPLC–MS/MS. *J. Proteome Res.* **2020**, *19* (6), 2358–2366. https://doi.org/10.1021/acs.jproteome.0c00036.

(254)   Kersten, E.; Dammeier, S.; Ajana, S.; Groenewoud, J. M. M.; Codrea, M.; Klose, F.; Lechanteur, Y. T.; Fauser, S.; Ueffing, M.; Delcourt, C.; Hoyng, C. B.; de Jong, E. K.; den Hollander, A. I.; EYE-RISK Consortium. Metabolomics in Serum of Patients with Non-Advanced Age-Related Macular Degeneration Reveals Aberrations in the Glutamine Pathway. *PLOS ONE* **2019**, *14* (6), e0218457. https://doi.org/10.1371/journal.pone.0218457.

(255)   Luo, D.; Deng, T.; Yuan, W.; Deng, H.; Jin, M. Plasma Metabolomic Study in Chinese Patients with Wet Age-Related Macular Degeneration. *BMC Ophthalmol.* **2017**, *17* (1), 165. https://doi.org/10.1186/s12886-017-0555-7.

(256)   Laíns, I.; Duarte, D.; Barros, A. S.; Martins, A. S.; Carneiro, T. J.; Gil, J. Q.; Miller, J. B.; Marques, M.; Mesquita, T. S.; Barreto, P.; Kim, I. K.; da Luz Cachulo, M.; Vavvas, D. G.; Carreira, I. M.; Murta, J. N.; Silva, R.; Miller, J. W.; Husain, D.; Gil, A. M. Urine Nuclear Magnetic Resonance (NMR) Metabolomics in Age-Related Macular Degeneration. *J. Proteome Res.* **2019**, *18* (3), 1278–1288. https://doi.org/10.1021/acs.jproteome.8b00877.

(257)   Deng, Y.; Shuai, P.; Wang, H.; Zhang, S.; Li, J.; Du, M.; Huang, P.; Qu, C.; Huang, L. Untargeted Metabolomics for Uncovering Plasma Biological Markers of Wet Age-Related Macular Degeneration. *Aging* **2021**, *13* (10), 13968–14000. https://doi.org/10.18632/aging.203006.

(258)   Ajana, S. Prédiction du risque de DMLA : identification de nouveaux biomarqueurs et

modélisation du risque. phdthesis, Université de Bordeaux, 2019. https://tel.archives-ouvertes.fr/tel-03415718 (accessed 2022-08-16).

(259)   Beger, R. D.; Schnackenberg, L. K.; Holland, R. D.; Li, D.; Dragan, Y. Metabonomic Models of Human Pancreatic Cancer Using 1D Proton NMR Spectra of Lipids in Plasma. *Metabolomics* **2006**, *2* (3), 125–134. https://doi.org/10.1007/s11306-006-0026-2.

(260)   Ekman, D. R.; Teng, Q.; Villeneuve, D. L.; Kahl, M. D.; Jensen, K. M.; Durhan, E. J.; Ankley, G. T.; Collette, T. W. Profiling Lipid Metabolites Yields Unique Information on Sex- and Time-Dependent Responses of Fathead Minnows (Pimephales Promelas) Exposed to 17α-Ethynylestradiol. *Metabolomics* **2009**, *5* (1), 22–32. https://doi.org/10.1007/s11306-008-0138-y.

(261)   Fernando, H.; Bhopale, K. K.; Kondraganti, S.; Kaphalia, B. S.; Shakeel Ansari, G. A. Lipidomic Changes in Rat Liver after Long-Term Exposure to Ethanol. *Toxicol. Appl. Pharmacol.* **2011**, *255* (2), 127–137. https://doi.org/10.1016/j.taap.2011.05.022.

(262)   Tynkkynen, T.; Mursu, J.; Nurmi, T.; Tuppurainen, K.; Laatikainen, R.; Soininen, P. NMR Protocol for Determination of Oxidation Susceptibility of Serum Lipids and Application of the Protocol to a Chocolate Study. *Metabolomics* **2012**, *8* (3), 386–398. https://doi.org/10.1007/s11306-011-0323-2.

(263)   Martineau, E.; Dumez, J.-N.; Giraudeau, P. Fast Quantitative 2D NMR for Metabolomics and Lipidomics: A Tutorial. *Magn. Reson. Chem.* **2020**, *58* (5), 390–403. https://doi.org/10.1002/mrc.4899.

(264)   Jurowski, K.; Kochan, K.; Walczak, J.; Barańska, M.; Piekoszewski, W.; Buszewski, B. Analytical Techniques in Lipidomics: State of the Art. *Crit. Rev. Anal. Chem.* **2017**, *47* (5), 418–437. https://doi.org/10.1080/10408347.2017.1310613.

(265)   Bradburn, M. J.; Clark, T. G.; Love, S. B.; Altman, D. G. Survival Analysis Part II: Multivariate Data Analysis – an Introduction to Concepts and Methods. *Br. J. Cancer* **2003**, *89* (3), 431–436. https://doi.org/10.1038/sj.bjc.6601119.

Annexes

<h1 style="text-align:center"><u>Publication – Communication – Awards</u></h1>

8.1   Publication linked to the thesis

Lambert, V.* Hansen, S.* , Schoumacher, M.* , Lecomte, J., Leenders, J., Hubert, P., Herfs, M., Blacher, S., Carnet, O., Yip, C., BLAISE, P., DUCHATEAU, E., LOCHT, B., THYS, M., CAVALIER, E., GOTHOT, A., Govaerts, B., RAKIC, J.-M., Noël, A.* , & De Tullio, P.*. (2020). Pyruvate dehydrogenase kinase/lactate axis: a therapeutic target for neovascular age-related macular degeneration identified by metabolomics. *Journal of Molecular Medicine*. Doi:10.1007/s00109-020-01994-9 https://hdl.handle.net/2268/251788 * *These authors have contributed equally to this work.*

8.2   Other publication

Arslan, D., Schoumacher, M., Dilly, S., Elmoualij, B., Zorzi, D., Quatresooz, P., Lambert, V., Noël, A., Pirotte, B., & De Tullio, P. (19 August 2022). Design, Synthesis, and Evaluation of Novel Pyruvate Dehydrogenase Kinase Inhibitors. *Medicinal Chemistry, 19* (3), 276-296. doi:10.2174/1573406418666220819102627 https://hdl.handle.net/2268/294183

Coelho Cristino Mamede, L. C., Fall, F., Schoumacher, M., Ledoux, A., De Tullio, P., Quetin-Leclercq, J., & Frederich, M. (2022). Recent metabolomic developments for antimalarial drug discovery. *Parasitology Research*. doi:10.1007/s00436-022-07673-7 https://hdl.handle.net/2268/295855

8.3   Oral and poster communication

Schoumacher, M., LAMBERT, V., Hansen, S., Leenders, J., Govaerts, B., Rakic, J.-M., Noël, A., & De Tullio, P. (11 December 2019). *New Insight into exudative Age-related Macular Degeneration (AMD): A Metabolomics Approach.* Paper presented at nd CIRM Day, Liège, Belgium. https://hdl.handle.net/2268/247404

Schoumacher, M., LAMBERT, V., Hansen, S., leenders, J., govaerts, B., Pirotte, B., Rakic, J.-M., Noël, A., & De Tullio, P. (23 June 2019). *New insight in Metabolomics based study of Age Related Macular Degeneration (AMD): Lipoprotein profile and patients follow-up.* Poster session presented at Metabolomics 2019, The Hague, Netherlands. https://hdl.handle.net/2268/247402

Schoumacher, M., LAMBERT, V., Hansen, S., Leenders, J., Govaerts, B., Rakic, J.-M., Noël, A., & De Tullio, P. (23 May 2019). *New insight in Metabolomics based study of Age Related Macular Degeneration (AMD): Lipoprotein profile and patients follow-up.* Paper presented at 12èmes journées RFMF, Clermont-Ferrand, France. https://hdl.handle.net/2268/247401

Schoumacher, M., LAMBERT, V., Hansen, S., Leenders, J., Govaerts, B., Rakic, J.-M., Noël, A., & De Tullio, P. (19 September 2018). *NMR-based metabolomics study of Age-related Macular Degeneration (AMD): patient follow-up and new target discovery.* Poster session presented at Benelux metabolomics Days, Netherlands. https://hdl.handle.net/2268/247399

Schoumacher, M., LAMBERT, V., Leenders, J., Roblain, Q., govaerts, B., Rakic, J.-M., Noël, A., & De Tullio, P. (05 July 2018). *NMR-based Metabolomics for New Target Discovery and Personalized Medicine: Application to Age-Related Macular Degeneration (AMD).* Paper presented at EUROMAR (European Magnetic Resonance Meeting), Nantes, France. https://hdl.handle.net/2268/227306

Schoumacher, M., LAMBERT, V., Hansen, S., Leenders, J., Govaerts, B., Pirotte, B., Rakic, J.-M., Noël, A., & De Tullio, P. (24 June 2018). *NMR-Based metabolomics study of Age-related Macular Degeneration (AMD): from pre-clinic model to new target discovery.* Poster session presented at Metabolomics 2017, Seattle, United States. https://hdl.handle.net/2268/247398

Schoumacher, M., LAMBERT, V., Govaerts, B., Rakic, J.-M., Noël, A., & De Tullio, P. (23 May 2018). *New insight in Metabolomics based study of Age Related Macular Degeneration (AMD): Lipoprotein profile and subclass analysis.* Poster session presented at 11e JS du RFMF, Liege, Belgium. https://hdl.handle.net/2268/247392

Schoumacher, M., LAMBERT, V., Hansen, S., Leenders, J., Govaerts, B., Pirotte, B., Rakic, J.-M., Noël, A., & De Tullio, P. (02 October 2017). *New Translational Metabolomics Exploration of Age-Related Macular Degeneration: From Patients to Functional Implication of Lactate.* Poster session presented at SMMAP 2017, Paris, France. https://hdl.handle.net/2268/247372

Schoumacher, M., LAMBERT, V., Hansen, S., Leenders, J., Govaerts, B., Pirotte, B., Rakic, J.-M., Noël, A., & De Tullio, P. (02 October 2017). *New translational metabolomics exploration of age-related macular degeneration: from patients to functional implication of lactate.* Paper presented at SMMAP 2017, Paris, France. https://hdl.handle.net/2268/247370

Schoumacher, M., LAMBERT, V., Hansen, S., Leenders, J., Govaerts, B., Pirotte, B., Rakic, J.-M., Noël, A., & De Tullio, P. (27 June 2016). *New Insight into exudative Age-related Macular Degeneration (AMD): A Metabolomics Approach*. Poster session presented at Metabolomics 2016, Dublin, Ireland. https://hdl.handle.net/2268/247393

8.4   Awards

Poster price CIRMday 2019: "*New Insight into exudative Age-related Macular Degeneration (AMD): A Metabolomics Approach.*"

Best Flash presentation award 12ème JS du RFMF 2019; "*New insight in Metabolomics based study of Age Related Macular Degeneration (AMD): Lipoprotein profile and patients follow-up.*"

8.5   Master Thesis

Maïlis Bianu (2017): « La dégénérescence maculaire liée à l'âge : contribution à la mise en place de nouveaux outils diagnostiques et à l'évaluation de l'efficacité de nouveaux axes thérapeutiques. »

Arianna Cirillo (2018): "study of dried blood spot sampling for NMR based metabolomics"

Manon Campas (2019): "Establishment of a metabolomics and lipidomics analysis workflow for pre-clinical and clinical studies"

JMolMed

ORIGINAL ARTICLE

Pyruvate dehydrogenase kinase/lactate axis: a therapeutic target for neovascular age-related macular degeneration identified by metabolomics

Vincent Lambert[1,2] · Sylvain Hansen[2] · Matthieu Schoumacher[3] · Julie Lecomte[2] · Justine Leenders[3] · Pascale Hubert[4] · Michael Herfs[4] · Silvia Blacher[2] · Oriane Carnet[2] · Cassandre Yip[2] · Pierre Blaise[1] · Edouard Duchateau[1] · Bénédicte Locht[1] · Michèle Thys[1] · Etienne Cavalier[5] · André Gothot[6] · Bernadette Govaerts[7] · Jean-Marie Rakic[1] · Agnès Noel[2] · Pascal de Tullio[3]

Abstract
Neovascular age-related macular degeneration (nAMD) is the leading cause of blindness in aging populations. Here, we applied metabolomics to human sera of patients with nAMD during an active (exudative) phase of the pathology and found higher lactate levels and a shift in the lipoprotein profile (increased VLDL-LDL/HDL ratio). Similar metabolomics changes were detected in the sera of mice subjected to laser-induced choroidal neovascularization (CNV). In this experimental model, we provide evidence for two sites of lactate production: first, a local one in the injured eye, and second a systemic site associated with the recruitment of bone marrow–derived inflammatory cells. Mechanistically, lactate promotes the angiogenic response and M2-like macrophage accumulation in the eyes. The therapeutic potential of our findings is demonstrated by the pharmacological control of lactate levels through pyruvate dehydrogenase kinase (PDK) inhibition by dichloroacetic acid (DCA). Mice treated with DCA exhibited normalized lactate levels and lipoprotein profiles, and inhibited CNV formation. Collectively, our findings implicate the key role of the PDK/lactate axis in AMD pathogenesis and reveal that the regulation of PDK activity has potential therapeutic value in this ocular disease. The results indicate that the lipoprotein profile is a traceable pattern that is worth considering for patient follow-up.

Vincent Lambert, Sylvain Hansen and Matthieu Schoumacher contributed equally to this work.

Agnès Noel and Pascal de Tullio equally supervised this paper

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00109-020-01994-9) contains supplementary material, which is available to authorized users.

✉ Pascal de Tullio
  P.deTullio@uLiege.be

1 Department of Ophthalmology, University Hospital of Liège, Liège, Belgium

2 Laboratory of Tumor and Development Biology, GIGA, Université de Liège, Liège, Belgium

3 Center for Interdisciplinary Research on Medicines, Metabolomics Group, Université de Liège, Liège, Belgium

4 Laboratory of Experimental Pathology, GIGA, Université de Liège, avenue Hippocrate, Liège, Belgium

5 Department of Medical Chemistry, University Hospital of Liège, Liège, Belgium

6 Department of Hematology and Immuno-Hematology, University Hospital of Liège, Liège, Belgium

7 Institute of Statistics Biostatistics and Actuarial Sciences, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

✿ Springer

Key messages

- Lactate and lipoprotein profile are associated with the active phase of AMD and CNV development.
- Lactate is a relevant and functional metabolite correlated with AMD progression.
- Modulating lactate through pyruvate dehydrogenase kinase led to a decrease of CNV progression.
- Pyruvate dehydrogenase kinase is a new therapeutic target for neovascular AMD.

Introduction

Age-related macular degeneration (AMD) is the leading cause of vision impairment in the elderly, in the Western world [1]. Clinically, AMD is classified into early asymptomatic retinal abnormalities, an intermediate form, and two advanced forms associated with severe visual impairment: "dry" AMD (or geographic atrophic form) and neovascular/exudative AMD (nAMD or "wet" AMD). The nAMD is the most severe form of this pathology and the major cause of blindness in people over 50 years old. Choroidal neovascularization (CNV), the hallmark of nAMD, results in the formation of immature and fragile blood vessels causing exudates in the sub-retinal spaces. In the clinics, optical coherence tomography (OCT) examination can monitor disease activity (i.e., the presence of intraretinal and sub-retinal fluids) [2, 3]. Numerous studies have established age, lifestyle, and genetic predispositions as key risk factors for AMD [4–6]. The disease is characterized by a low-grade/subclinical degree of inflammation termed "para-inflammation" [7]. Several cytokines, such as IL-1b [8], IL-18 [9], and IL-33 [10], have been implicated in immune and vascular responses. Moreover, altered cholesterol homeostasis and higher concentrations of high-density lipoproteins have been linked to AMD [11]. Despite these advances, a comprehensive understanding of the pathogenesis and the evolution of this complex multifactorial disease remains incomplete.

The intravitreal administration of anti-vascular endothelial growth factor (anti-VEGF) drugs such as bevacizumab (Avastin®, Genentech), ranibizumab (Lucentis®, Novartis), and aflibercept (Eylea®, Regeneron Pharmaceuticals) represents the cornerstone of the current treatment for patients with nAMD [12, 13]. Non-response to those drugs occurs in up to 21% of eyes [14, 15]. Non-responders to a specific anti-VEGF compound might benefit from a switch to an alternative anti-VEGF drug [14, 15]. Currently, patient follow-up mainly relies on ophthalmologic exams [16] and crucially lacks powerful tools for unbiased biological analyses. Furthermore, the identification of novel therapeutic targets is mandatory to overcome resistance to anti-VEGF treatments.

Metabolomics, defined as the comprehensive identification and quantification of low molecular weight endogenous metabolites in biological samples, offers new opportunities to obtain innovative insights into pathologies and to identify biomarkers/patterns that could be helpful for personalized medicine and improve patient care [17, 18]. Directly linked to the phenotype, this approach provides unique challenging opportunities to correlate dynamic variations in the metabolome with pathological disease status (occurrence and progression) and/or to identify metabolites that are markers and/or key players of the disease [17, 19]. Notably, recent studies have highlighted the implication of metabolism and some energetic metabolites in inflammation and angiogenesis, two identified processes involved in CNV development [20, 21]. In this context, metabolomics has emerged as a relevant tool to obtain new insights into nAMD development through experimental and clinical approaches. Nevertheless, in comparison with other omics sciences, only a small number of metabolomics studies applied to different stages of AMD have been reported [22]. Most of them, based on a mass-spectrometry approach, identified relevant markers and putative biochemical pathways that could be implied in this pathology, mainly in lipid and energetic metabolism [23–25]. Studies using nuclear magnetic resonance (NMR)–based metabolomics are scarce and report small differences between AMD stages [26]. None of those studies has combined a metabolomics screening with an experimental study with the aim of establishing the functional impact of an identified metabolite/marker that could lay down the foundation for novel therapeutic approaches.

Here, we report a NMR-based metabolomics analysis applied to the sera of mice subjected to laser-induced CNV and to serum samples of patients with nAMD. In both experimental and clinical settings, alterations in lactate levels and lipoprotein profiles were detected. Laser-induced CNV [27] was used to mechanistically explore the functional implication of the pyruvate dehydrogenase kinase (PDK)/lactate axis, which is a targetable pathway to modulate the lactate levels [28, 29]. The exciting discovery of lactate as a functional and targetable molecule and the lipoprotein profile as a traceable pattern opens new perspectives for optimizing AMD treatment and patient follow-up.

Methods

Patient selection

The study population consisted of unrelated European-Caucasian individuals (> 59 years old) affected with nAMD ($n = 72$) and healthy volunteers without signs of macular disease or a known family history of AMD ($n = 50$). Patients with AMD and volunteers were not matched for age or sex. Trained ophthalmologists examined all patients with nAMD and divided them into clinically active or inactive subgroups depending on the presence or absence of intraretinal or sub-retinal exudative fluids as assessed by OCT, respectively. A complete medical history of each patient was obtained by using a standardized questionnaire (i.e., lifestyle, pathologies, treatments, BMI, etc.). Patients and volunteers with diabetes mellitus were excluded from the cohort, while those under treatment for hypercholesterolemia and hypertension were accepted.

Blood sample collection

For all the participants, peripheral blood was collected after ophthalmological exams in: (1) K_2 EDTA blood collection tubes for blood cell counts, leucocyte differential, and C-reactive protein (CRP) measurement using CRPLX kit (Cobas®, Roche/Hitachi); (2) serum-separating tubes (Greiner); and (3) sodium fluoride/oxalate tubes (Greiner) for the lactate quantification. Serum samples were routinely taken in the morning with a fasting period of at least 2 h. Samples were treated according to clinical standard processes and stored at -80 °C after sampling until the metabolomics analysis. Clinical biology analyses (lactic acid, CRP, red and white blood cells) were also performed on the sera in order to evaluate patient inflammatory status.

Laser-induced choroidal neovascularization in mice

CNV was induced in 8-week-old C57BL/6J mice by laser burns as previously described ($n = 4$ to 6 mice per experimental group, minimum 4 impacts per eye) [27]. Mice developing cataracts were excluded from further analyses [27]. Dichloroacetic acid (DCA) was provided in the drinking water according to a described protocol [30]. After measuring the average volume of water consumed by mice every day, the DCA concentration in the drinking water was determined to achieve a daily dose of 3 mg/mouse per day (150 mg/kg/day). At mouse sacrifice, the posterior segments of enucleated eyes were flat-mounted for immunostaining or treated with collagenase for FACS analyses. For CNV quantification, we used FITC-dextran-labeled flat-mounted choroids, as described [27]. In some assays, intravitreous injection of saline or Avastin solution (2 μl −50 μg Avastin/eye) was performed

immediately after laser induction, during anesthesia. All animal experiments were approved by the Animal Ethics Committee of the University of Liège.

^{1}H-NMR spectroscopy and post-treatment of spectra

All samples were recorded at 298 K on a Bruker Avance spectrometer operating at 500.13 MHz for the proton signal acquisition. The instrument was equipped with a 5-mm TCI cryoprobe with a Z-gradient. Maleic acid was used as an internal standard for quantification and trimethylsilyl-3-propionic acid-d4 (TMSP) was used for the zero calibration. Human sera (500 μl) were mixed with D_2O phosphate buffer (100 μl) (0.1 M, pH 7.4), a 35 mM solution of maleic acid (100 μl) (Aldrich, Germany), and TMSP (30 μl) (sodium trimethylsilyl[2,2,3,3-D$_4$]propionate) in D_2O (10 mg/ml). Mouse sera (200 μl) were mixed with D_2O phosphate buffer (400 μl) (0.1 M, pH 7.4), 35 mM solution of maleic acid (100 μl), and TMSP (30 μl) in D_2O (10 mg/ml). For analyses on other organs, samples were weighed (liver, spleen) or cells were counted (bone marrow) to normalize data (per mg or cell). Samples were prepared as previously described [31].

^{1}H-NMR spectra were acquired using a 1D-CPMG (Carr-Purcell-Meiboom-Gill) relaxation-editing sequence with presaturation for serum samples. The CPMG experiment used a RD-90-(t-180-t)n-sequence with a relaxation delay (RD) of 2 s, a spin echo delay (t) of 400 ms, and the number of loops (n) equal to 80. The water suppression pulse was placed during the RD. The number of transients was typically 32. The acquisition time was set to 3.982555 s, and a quantity of four dummy scans was chosen. Data were processed with the Bruker Topspin 3.1 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.

Metabolomics analyses

Optimized ^{1}H-NMR spectra were automatically baseline-corrected and reduced to ASCII files using AMIX software (version 3.9.14; Bruker). The spectral intensities were normalized to total intensities and reduced to integrated regions of equal width (0.04 ppm) corresponding to the 0.5–10.00 ppm region. Because of the residual signals of water and maleic acid, regions between 4.7 and 5 ppm (water signal) and 5.6 and 6.2 ppm (maleic acid signal) were removed before analysis. The reduced and normalized NMR spectra were imported to SIMCA (version 14.1, Umetrics AB, Umea Sweden). Pareto scaling was applied to bucket tables, and discriminant analysis (DA), such as principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA), and orthogonal partial least squares discriminant analysis (OPLS-DA), was

performed. SIMCA was used to generate all PCA, PLS-DA, and OPLS-DA models and plots. Unsupervised multivariate analysis (PCA) was first used to determine intrinsic clusters within the data set, while PLS-DA maximized the separation and OPLS-DA facilitated the graphic visualization of differences and similarities between groups. The quality of OPLS-DA models was determined by the goodness of fit (R^2), and the predictability was calculated on the basis of the fraction correctly predicted in one-seventh cross-validation (Q^2). This workflow was applied for both human and mouse serum samples.

Lipoprotein profile ^{1}H-NMR analysis

Estimation of lipoprotein profile modification among spectra collected from patients and mice was performed by using peak picking methods that compare intensities between the different fractions. In human and mouse blood NMR spectra, the global signal of lipoproteins between 0.80 and 0.95 ppm is due to an overlap of several peaks that could be linked to the main classes of lipoproteins: very low-density lipoproteins (VLDL), low-density lipoproteins (LDL), intermediary density lipoproteins (IDL), high-density lipoproteins (HDL), and chylomicron. Then, it could be decomposed into distinguishable signals corresponding to these different classes or to a mixture of 2 classes (4 in humans and 5 in mice). To evaluate the proportion of each lipoprotein fraction in the samples, a method based on normalized peak intensity calculation was developed. Then, for each class, we determined the chemical shift corresponding to the peak of signal intensity. For human samples, 4 fractions are selected (F1 = 0.92 ppm (mainly VLDL), F2 = 0.91 ppm, F3 = 0.89 ppm, and F4 = 0.88 ppm (mainly HDL). For mouse samples, 5 fractions are selected (F1 = 0.93 ppm (mainly VLDL), F2 = 0.92 ppm, F3 = 0.90 ppm, F4 = 0.89 ppm, and F5 = 0.88 (mainly HDL)). For each sample, the signal intensity at these different chemical shifts was measured and then normalized to the total intensities of all fractions to reduce the impact of the global lipoprotein concentrations that could differ between samples. Therefore, the obtained values represent a fraction of the total signal. This method allows the comparison between lipoprotein profiles issued from the spectra of blood samples collected from patients with AMD, control subjects, and induced/non-induced mice.

Enzymatic and NMR lactate dosages

Enzymatic lactate measurements in whole blood were performed using lactate kit (Cobas®, Roche/Hitachi). NMR lactate quantification was conducted in sera following a described procedure [32].

Immunostaining

Flat-mounted posterior ocular poles of mouse eyes were incubated with blocking buffer (PBS containing 10% bovine serum albumin) for 30 min, followed by an overnight incubation, at room temperature with primary antibodies: rat anti-mouse Ly-6B.2 (dilution 1:100) (AbDSerotec, MCA771G) or rat anti-mouse Alexa Fluor 647-conjugated F4/80 (dilution 1:100) (Invitrogen, MF48021). To reveal Ly-6B.2 staining, flat-mounted samples were incubated for 1.5 h with a goat anti-rat Alexa Fluor 555-conjugated antibody (dilution 1:1000) (Invitrogen, A-21434). Samples were mounted in Vectashield mounting medium (Vector Laboratories, Burlingame, CA) after washing with PBS. Images were captured with a confocal Leica TCS SP5 microscope using a × 20 objective lens.

Bone marrow isolation

In some assays, at sacrifice, bone marrow (BM) cells were isolated, as previously described [22], from the tibia and femur of mice subjected or not to laser-induced CNV.

Macrophage isolation in mouse eye and staining for FACS analysis

After excision, posterior ocular segments were immediately incubated with 2.5 mg/ml collagenase (Sigma) for 1.5 h at 37 °C. The mixture was subjected to mechanical dissociation using a 23G syringe needle and filtered through a 50-μm cell strainer. After FcR blockade with a control isotype (anti-CD16/32, clone R35-95 from BD Biosciences, Erembodegem, Belgium), the immunostaining was performed with the following antibodies: anti-F4/80 (PE, clone BM8, eBioscience, San Diego, CA, USA), anti-CD45 (APC-Cy7, clone 30-F11, eBioscience), and anti-CD206 (APC, clone FAB2535A, R&D systems, Minneapolis, MN, USA). Cell percentages were measured using a FACS Canto II flow cytometer (BD Biosciences) and FACS Diva V 6.1.2 software. F4/80$^+$/CD206$^-$ cells were considered M1-like macrophages and F4/80$^+$/CD206$^+$ cells were considered M2-like macrophages.

Generation of human M1/M2 monocyte-derived macrophages

Human peripheral blood mononuclear cells (PBMCs) were plated overnight in 6-well flat bottom plates (10^7 cells/well) (Nunc, Roskilde, Denmark). The adherent PBMC fraction was grown in complete Roswell Park Memorial Institute (RPMI) medium 1640 containing 50 mM 2-mercaptoethanol and antibiotics (all from Gibco, Merelbeke, Belgium), at 37 °C and 5% CO_2. For macrophage differentiation, monocytes were incubated with M-CSF (100 ng/ml), for 6 days. The M1-like or M2-like polarized macrophages were generated by stimulating

cells for 3 days with lipopolysaccharide (LPS) (100 ng/ml) (Sigma-Aldrich, St Louis, MO, USA) and IFN-γ (20 ng/ml) (PeproTech) or IL-4 (20 ng/ml) (ImmunoTools, Friesoythe, Germany), respectively [33, 34]. In some experiments, lactate or lactic acid (10 mM) was added to M1-like macrophages, for 2 days. Macrophage phenotypes were analyzed through labelling with anti-CD68 (APC, clone Y1/82A, Miltenyi), anti-CD80 (FITC, clone L307, BD Biosciences), and anti-CD206 (PE, clone 19.21, BD Biosciences). CD68$^+$/CD206$^-$/CD80$^-$ cells were considered M0, CD68$^+$/CD206$^-$/CD80+ were considered M1-like, and CD68+/CD206+/CD80$^-$ were considered M2-like macrophages. Positive cell percentages were analyzed on a FACS Canto II flow cytometer using FACS Diva software V 6.1.2 (BD Biosciences).

Boyden chamber migration assay

The chemotactic migration of macrophages was evaluated using Boyden chambers (48-well Boyden microchamber; Neuroprobe, Cabin John, MD, USA) and gelatine-coated filters in the presence of serum-containing RPMI [35]. Briefly, lower wells were filled with 27 μl of RPMI medium containing 0.1% bovine serum albumins (BSA) (as a control for random migration) or 1% BSA (as a positive control). Lactate was added at different concentrations (0.1 to 20 mM) to RPMI in the presence of 0.1% BSA. Cell suspension (15 μl) (1×10^6 cells/ml of RPMI + 0.1% BSA) was added to the upper wells for 5 h at 37 °C. Six wells were used for each experimental condition. One random field/well was counted using an eyepiece with a calibrated grid to evaluate the number of migrating cells.

Statistical analyses

For the in vivo assays, the results related to CNV formation, cell densities, and cell percentages were compared by using a linear mixed model as described [27]. Lactate measurements and Boyden chamber assays were analyzed by using a one-way ANOVA (with Tukey's multiple comparison test), while inflammatory cell quantification was compared by using a two-way ANOVA. For metabolomics studies, unsupervised PCAs and supervised multivariate PLS-DAs and OPLS-DAs were used to discriminate between groups and to identify the main biomarkers (see above in "Metabolomics analyses"). For human and mouse samples, computations were conducted with SIMCA 14.1 software with Pareto scaling of bucket variables.

Univariate statistical analyses were performed on each lipoprotein fraction (see above "Lipoprotein profile ^{1}H-NMR analysis") of both mouse and human serum spectra using GraphPad Prism version 7.0. A nonparametric Kruskal-Wallis test with Dunn's multiple comparison was used to compare Fx controls with Fx AMD patients and Fx mice control versus CNV mice (J5 and J7 post-laser) with Fx being

the lipoprotein fraction number 1 to 4 for humans and 1 to 5 for mice.

Hypothesis tests were used to compare AMD cases and controls. P values of homogeneity (chi-square or Fisher exact tests) are reported for qualitative variables. For quantitative variables, the P values result from one-way ANOVA (parametric or nonparametric (Kruskal-Wallis test) according to D'Agostino and Pearson normality test).

Study approval

The human study was conducted under protocols approved by the Ethical Committee of the University Hospital of Liège, B7072006295 (Belgium). Informed consents were obtained from all study subjects before participation. All animal experiments were approved by the Animal Ethics Committee of the University of Liège.

Results

Metabolomics discriminates patients with AMD from volunteers and active patients from inactive patients

The prospective cohort of patients with and healthy volunteers (Supplemental Table 1) was sub-divided into 3 groups: active AMD (patients with AMD associated with CNV development and exudate), inactive AMD (patients with stabilized AMD without exudate), and healthy donors. Supplemental Table 1 describes patient parameters (sex, age, BMI, clinical biology parameters, treatments). Blood cell counts and C-reactive protein (CRP) measurements were similar in all groups (Supplemental Table 1). However, a slight subclinical increase in monocyte and basophil percentages was noted in patients with active AMD and in patients with inactive AMD compared to healthy volunteers, respectively (Supplemental Table 1). Partial discrimination between all AMD patients ($n = 72$) and healthy volunteers ($n = 50$) was highlighted by supervised PLS-DAs and O-PLS-DAs of the NMR-based metabolomics spectral data (Fig. 1a). We next focused on the 2 subgroups of AMD patients and noticed also discrimination between patients who are in an active phase (active) and those in a non-active phase (inactive) (Fig. 1b). Importantly, some specific spectral zones and metabolites were responsible for the discrimination of these groups, as shown in the loading plot (Fig. 1c). Increased lactate levels and changes in lipoprotein profiles were linked to the active phase of the pathology (Fig. 1c). Both ^{1}H-NMR (Fig. 1d) and enzymatic (Fig. 1e) dosages in blood samples of patients confirmed increasing lactate concentrations in patients with active AMD, as compared to patients with inactive AMD and healthy volunteers. The NMR signal corresponding to the main different classes of lipoproteins (Fig. 1f) was separated into 4 fractions, namely, F1 to F4, which progressively shifted from a fraction rich in VLDL (F1), to LDL
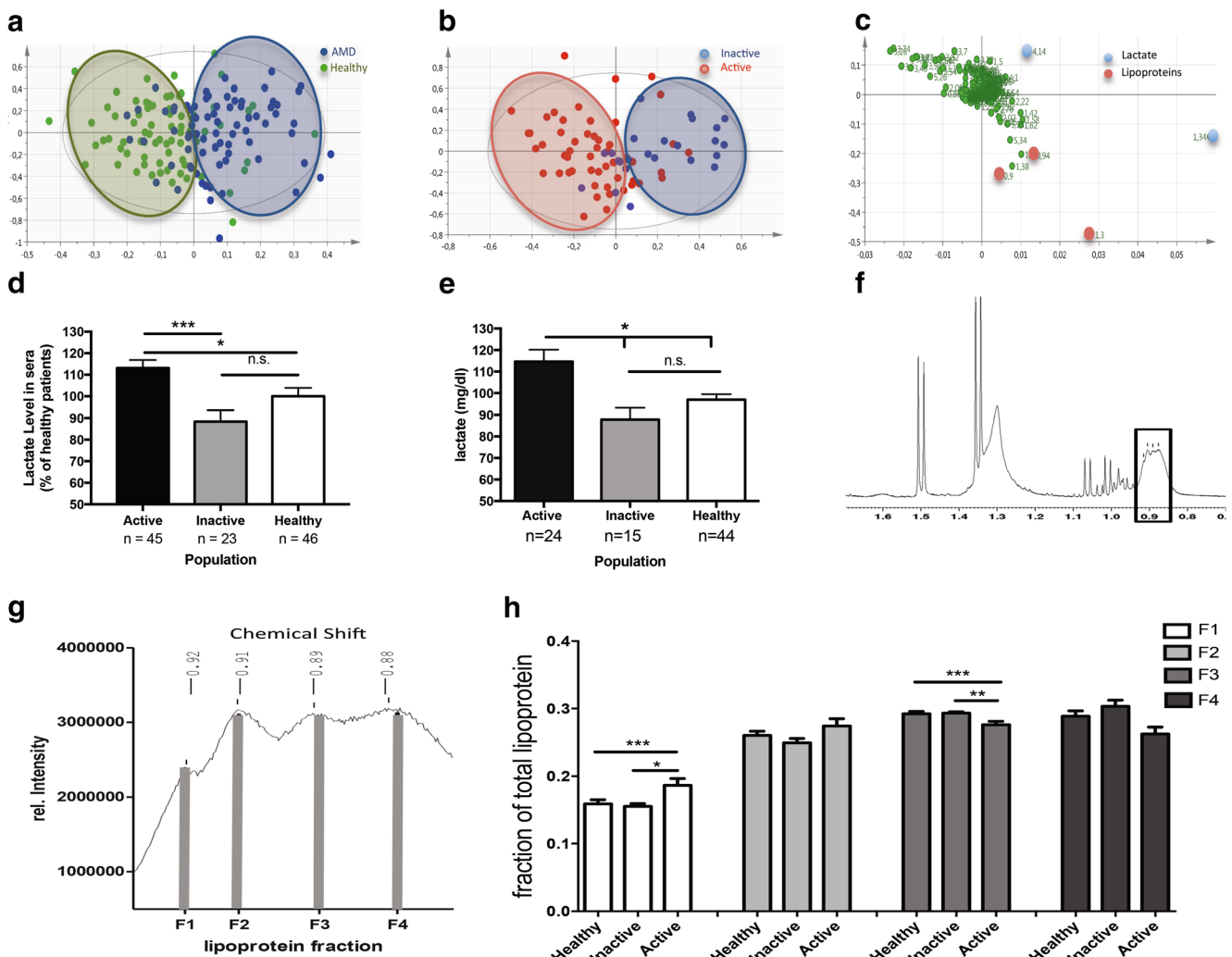
Fig. 1 Lactate and lipoproteins are the main increased discriminant metabolite in serum of patients with nAMD. **a** Score plot derived from an OPLS-DA of spectral data (3 components, $R^2 = 0.527$, $Q^2 = 0.151$) collected from patients with nAMD (blue dots, $n = 72$) and healthy volunteers (green dots, $n = 50$). Each data point represents an individual patient. **b** OPLS-DA score plot (2 components, $R^2 = 0.468$, $Q^2 = 0.21$) of spectral data collected from patients with active (red dots, $n = 49$) and inactive (blue dots, $n = 23$) nAMD. **c** Loading plot of spectral data collected from patients with active and inactive nAMD highlighting lactate and lipoproteins as biomarkers of active status. **d** NMR and **e** biochemical dosages of blood lactate in the serum of healthy volunteers and patients with active and inactive forms of AMD. Data are expressed as the percentage of healthy donors. **f** NMR spectrum of human serum highlighting the lipoprotein profiles (from 0.88 to 0.92 ppm). **g** Enlarged view of the lipoprotein NMR spectral zone showing the chemical shift corresponding to the maximum intensity of the signal of the 4 lipoprotein classes. **h** Modification of the lipoprotein profile during CNV development. Fraction 1 is mainly composed of VLDL, while fraction 4 is mainly composed of HDL. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. Error bars are SEM

(F2), IDL (F3), and HDL (F4) (Fig. 1g). The analyses of the lipoprotein spectral zone in the 3 patient groups again revealed a shift in the lipoprotein profile. Indeed, both VLDL and LDL (F1 and F2) proportions were higher, while IDL and HDL (F3 and F4) fractions were reduced in the active group compared to inactive and healthy individuals (Fig. 1h).

Metabolomics of mice subjected to CNV led to identification of CNV-linked metabolites

When mice were subjected to laser-induced CNV, neovascular lesions appeared at day 5 after induction and were the largest at day 7 (Fig. 2a–d), in agreement with previous reports [12, 27, 36]. Concomitant longitudinal metabolomics analyses performed on blood samples highlighted a good concordance between CNV occurrence and the evolution of the metabolome patterns observed during the first week following laser burn (Fig. 2e–g). The most significant discrimination between control (CTL) and induced mice occurred at days 5 and 7 post-laser burn when the CNV was detectable on flat-mounted choroids (Fig. 2b, c, f, g). In line with the clinical data generated in the human study, lactate and lipoprotein profiles were the main discriminating metabolites (Fig. 2h–j) between the two experimental groups. Consistently, lactate
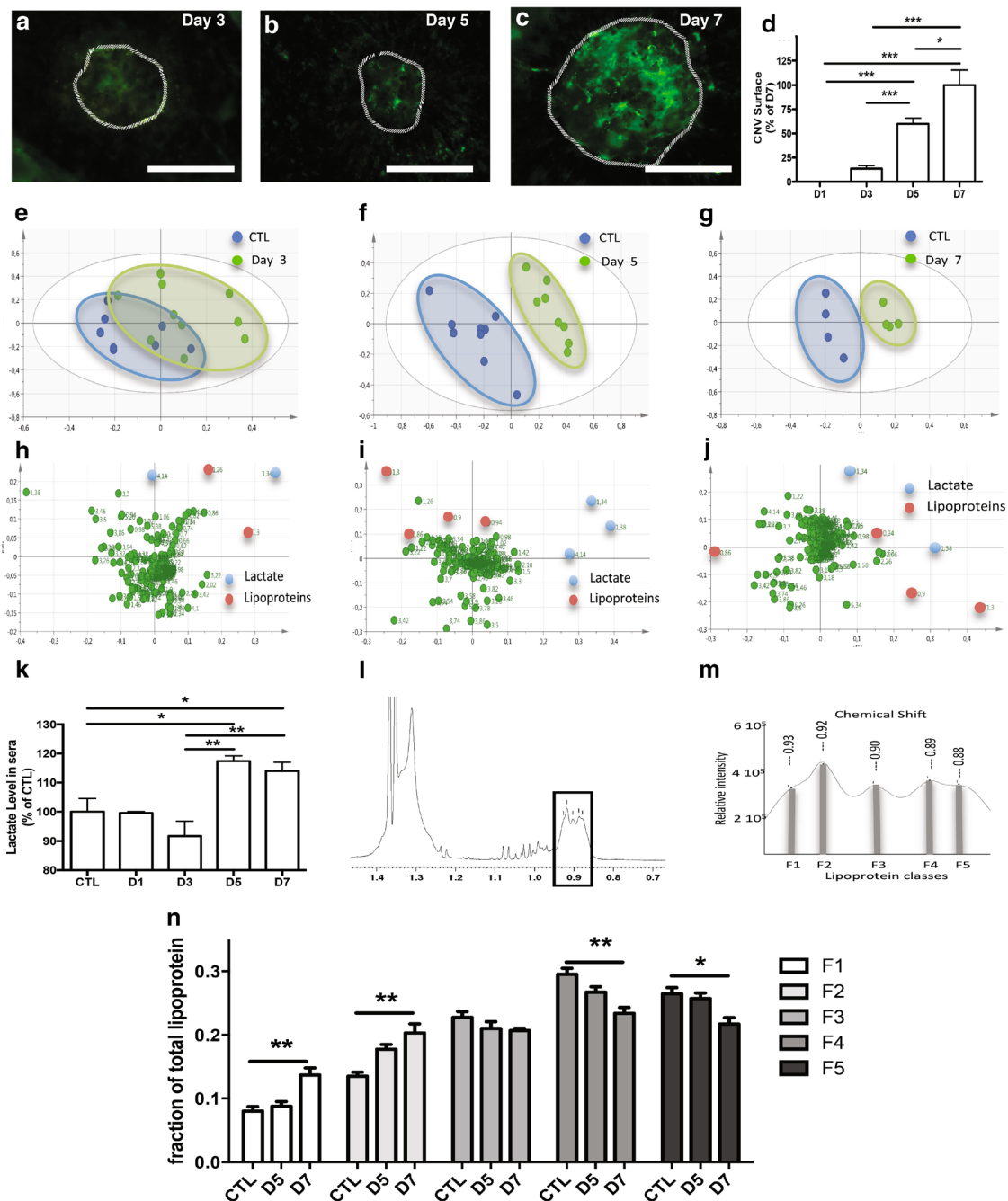
Fig. 2 Lactate and lipoproteins are the main increased discriminant metabolites in the serum of mice subjected to CNV. Mice were subjected (CNV) or not (CTL) to a laser burn. FITC-dextran-labeled flat-mounted choroid observed at day 3 (**a**), 5 (**b**) or 7 (**c**) after laser induction. Dashed lines delineate the lesion. Scale bars, 100 μm. Quantification of fluorescent neovessel area with ImageJ software ($n \geq$ 4 mice/group, $n \geq 12$ laser impacts/group) at days 1, 3, 5, and 7 (**d**). *$P < 0.05$; ***$P < 0.001$. Error bars indicate SEM. Score plot resulting from a PLS-DA analysis of spectral data performed at day 3 (**e**), day 5 (**f**), and day 7 (**g**) after laser induction (control vs J3: 2 components, $R^2 =$ 0.579 and $Q^2 = 0.119$; control vs J5: 3 components, $R^2 = 0.669$ and $Q^2 =$ 0.841; control vs J7: 3 components, $R^2 = 0.793$ and $Q^2 = 0.734$). Laser-induced mice (green dots) were distinguishable from non-induced mice (blue dots) at day 5 and day 7. Each data point represents an individual

mouse ($n \geq 4$ mice/group). Loading plot resulting from a PLS-DA analysis of spectral data performed at day 3 (**h**), day 5 (**i**), and day 7 (**j**) after laser induction. Lactate (blue dots) and lipoproteins (red dots) are the main discriminant metabolites. NMR dosage of serum lactate levels at day 7 ($n \geq 5$ mice/group) (**k**). The results are expressed as the percentage of the control. *$P < 0.05$; **$P < 0.01$. Error bars indicate SEM. NMR spectrum of mouse serum highlighting the lipoprotein profiles (from 0.88 to 0.93 ppm) (**l**). Enlarged view of the lipoprotein NMR spectral zone showing the chemical shift corresponding to the maximum intensity of the signal of the 5 lipoprotein classes (**m**). Modification of the lipoprotein profile during CNV development (**n**). Fraction 1 is mainly composed of VLDL while fraction 5 is mainly composed of HDL ($n \geq 5$ mice/group). *$P < 0.05$; **$P < 0.01$. Error bars are SEM

concentrations measured by NMR were increased by approximately 15% in the blood of laser-induced mice at day 5 (Fig. 2k) ($P < 0.05$). The NMR signal corresponding to the main different classes of lipoproteins (Fig. 2l) was separated into 5 fractions, namely F1 to F5, which progressively shifted from a fraction rich in VLDL (F1), to LDL (F2), to one richest in HDL (F5) (Fig. 2m). The kinetic analysis of lipoprotein profiles revealed that the proportion of LDL/VLDL-containing fractions (mainly F1 and F2) progressively increased, while HDL fractions (mainly F4 and F5) decreased during CNV development (Fig. 2n). Collectively, these data highlight two main discriminant metabolites between control and CNV-induced mice: lactate whose increased levels were detected at day 5, and a shift in the lipoprotein profile (VLDL-LDL/HDL ratio), which was the highest at day 7.

Lactate has a functional role in CNV development

Lactate is a targetable metabolite whose level can be modulated by interfering with pyruvate dehydrogenase kinase (PDK), which inactivates the pyruvate dehydrogenase (PDH) involved in pyruvate conversion into acetyl-CoA in the mitochondria (Fig. 3a). In the mouse model, PDK activity was blocked by treatment with dichloroacetic acid (DCA) [37]. In line with our expectation, DCA treatment decreased blood lactate levels at day 5 (Fig. 3b) and led to a significant reduction in CNV formation of 46%, at day 7 (Fig. 3c–e). These data demonstrate the functional role of lactate in CNV development. Notably, similar effects on blood lactate levels (Fig. 3g) were observed by using anti-VEGF antibodies (injected intravitreously at day 0), which are highly potent anti-angiogenic agents (Fig. 3f). In line with the clinical data, DCA treatment restored the lipoprotein profiles (Fig. 3h) with a distribution of the fractions similar to that seen under control conditions (lower F1/F2 and higher F4/F5 proportions) (Fig. 3h). Thus, DCA treatment induced a normalization of lipoprotein profiles. Interestingly, anti-VEGF antibodies also modified the lipoprotein profile with reduced VLDL/LDL (F1–F3 fractions) and increased HDL (F4 and F5 fractions) proportions as compared to control mice (Fig. 3i).

Lactate produced locally and systemically modulates inflammation and macrophage polarization

To determine the cellular source of lactate detected in the serum at day 5 (Fig. 2k), we conducted NMR dosages in different organs at days 3 and 5 post-laser induction. Increased lactate levels were detected locally in injured eyes as early as day 3 (Fig. 4a). At this time point, no significant modification of lactate levels was detected in the serum (Fig. 2k), bone marrow, spleen, or liver (Fig. 4b–d). Interestingly, at day 5, a substantial increase in lactate levels was noted in the bone marrow (Fig. 4b). Notably, DCA-treated mice displayed

normalized lactate levels in eyes at day 3 (Fig. 4a) and in the bone marrow at day 5 (Fig. 4b). A slight decrease of lactate level was seen in spleen after DCA treatment at day 3 (Fig. 4c). These data reveal two sites of lactate production: an early and local production of lactate in the injured eye, followed by a release of lactate in the blood circulation by bone marrow–derived inflammatory cells.

Ly6b2+ neutrophil infiltration occurred at day 1 post-laser burn and then rapidly decreased from days 3 to 7 (Fig. 4e–i), while the F4/80+ macrophage density progressively increased after CNV induction and peaked at day 5 (Fig. 4j–m). Upon DCA treatment, neutrophil recruitment was slightly decreased at day 1 (Fig. 4i). An earlier infiltration of macrophages (Fig. 4n) was observed upon DCA treatment and associated with a 1.85-fold reduction in the percentage of pro-angiogenic M2-like macrophages (CD45+, F4/80+, and CD206+ cells) (Fig. 4o). The impact of lactate on macrophage migration and polarization was next evaluated in vitro. Human peripheral monocytes were first polarized in vitro into M1 or M2 subtypes and then subjected to a lactate gradient in a Boyden chamber assay. Lactate was a potent chemoattractant for M1-like macrophages (at 10 mM and 20 mM) and M2-like macrophages (at 1 and 2 mM) (Fig. 4p, q). We next hypothesized that lactate can repolarize M1-like macrophages into M2 macrophages. To address this issue, macrophages were stimulated with LPS and IFN-γ to polarize them into an M1-like phenotype (Fig. 4r). Under those conditions, approximatively 30% of the macrophage population were M1-like and < 4% were M2-like macrophages. The subsequent addition of lactate for 2 days led to a shift in the macrophage population, which was then composed of approximatively 25% M2-like macrophages and < 2% M1-like macrophages. Collectively, these data reveal that lactate modulates macrophage recruitment and the M1/M2 balance in favor of M2-like macrophages by converting M1 macrophages to M2 macrophages.

DCA is more efficient during the angiogenesis phase

Our experimental data allowed us to correlate the systemic increase in lactate levels with macrophage recruitment and the angiogenic phase (CNV formation) (Fig. 5a). Interestingly, CNV formation and macrophage recruitment in the eye were concomitant with the increase in lactate levels observed in the bone marrow and in the serum. To define the best therapeutic window for DCA treatment, we designed several treatment schedules to target several steps of disease development: (i) the whole process (D0–D7); (ii) the neutrophilic response (D0–D2); (iii) the initial macrophage recruitment associated with increased local lactate levels (D0–D4); and (iv) the boost of systemic lactate production associated with early (D3–D5) and late (D4–D7) CNV formation (Fig. 5b). CNV inhibition was optimal when DCA treatments were applied during the whole process (D0–D7) (Fig. 5c). In
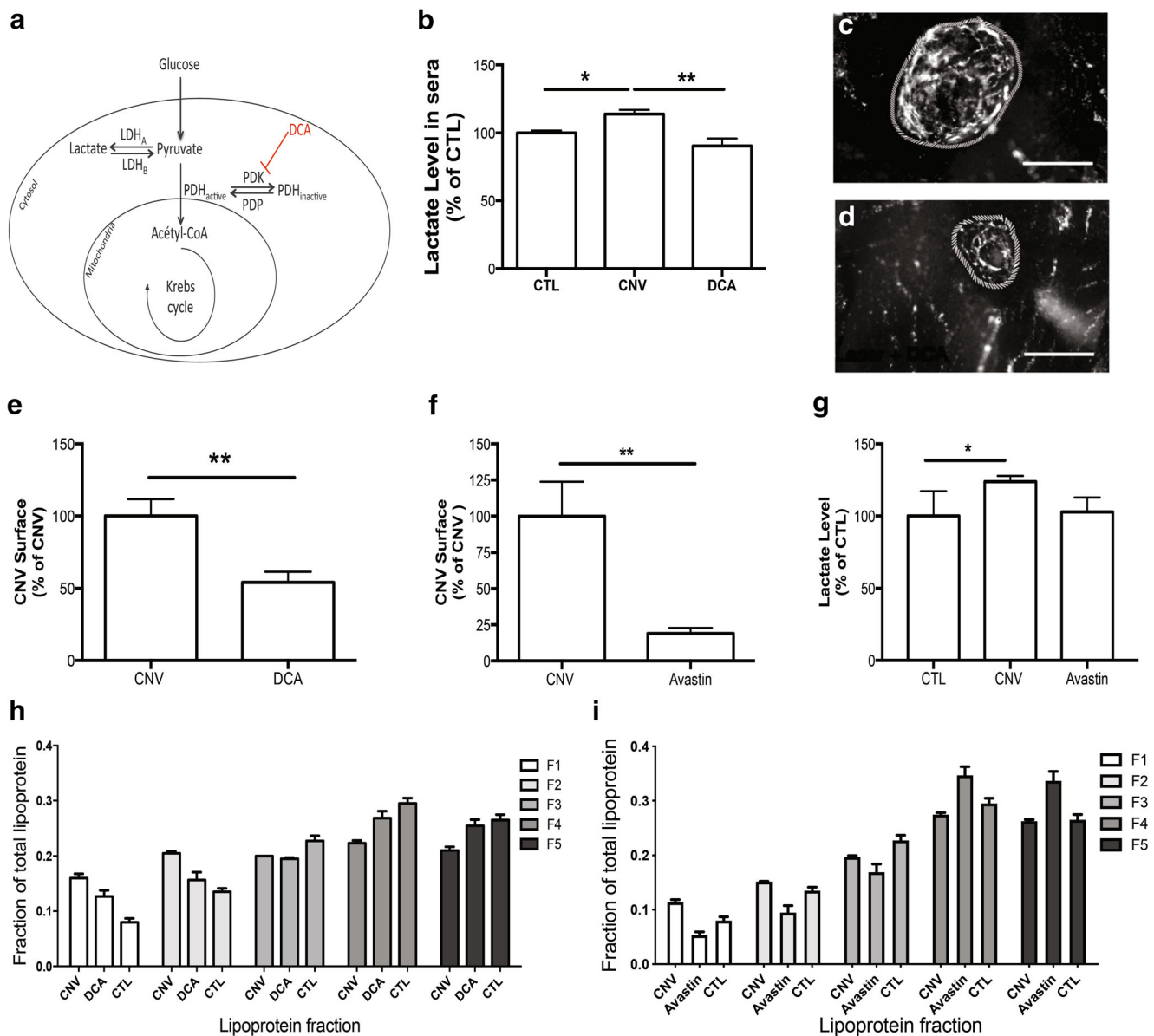
Fig. 3 DCA treatment normalizes lactate level and reduces CNV surface. **a** Schematic overview of how DCA can impact lactate levels. DCA inhibits mitochondrial PDK activity, thereby maintaining PDH in its (unphosphorylated) active form and facilitating the decarboxylation of pyruvate to acetyl-CoA. As the flux of pyruvate is accelerated, the equilibrium between lactate and pyruvate is unbalanced towards pyruvate. **b–e** Mice subjected to laser-induced CNV were treated or not with DCA (3 mg DCA/day/mouse) ($n \geq 5$ mice/group). Untreated mice (laser) were used as controls. **b** NMR dosage of serum lactate level at day 7 after DCA treatment ($n \geq 5$ mice/group). The results are expressed as the percentage of laser-induced mice without treatment. *$P < 0.05$; **$P < 0.01$. Error bars correspond to SEM. Flat-mounted choroid of **c** an untreated mouse and **d** a DCA-treated mouse at day 7: Dashed lines delineate the lesion.

Scale bars, 100 μm. **e** Quantification of CNV after DCA treatment at day 7 ($n \geq 6$ mice/group, $n = 26$–27 laser impacts/group). The results are expressed as the percentage of laser-induced mice without treatment. **$P < 0.01$. Error bars correspond to SEM. **f** Quantification of CNV after Avastin treatment at day 7 ($n \geq 6$ mice/group, $n = 26$–27 laser impacts/group). The results are expressed as the percentage of laser-induced mice without treatment. **g** NMR dosage of serum lactate level at day 7 after Avastin treatment; the results are expressed as the percentage of laser-induced mice without treatment ($n \geq 6$ mice/group). **$P < 0.01$. Error bars correspond to SEM. **h** Modification of the lipoprotein profile at day 7 after DCA treatment ($n \geq 6$ mice/group). **i** Modification of the lipoprotein profile at day 7 after Avastin treatment. Fraction 1 is mainly composed of VLDL while fraction 5 is mainly composed of HDL

contrast, only a partial CNV impairment was observed upon treatment during earlier periods (D0–D2, D0–D4, and D3–D5). Of note, when we reproduced the clinical situation by applying DCA treatment during the late-angiogenic phase

(D4–D7), CNV inhibition was optimal and reached that observed with a treatment applied during the whole process. These findings emphasize the interest of the PDK/lactate axis as a therapeutic target to treat patients with active AMD.
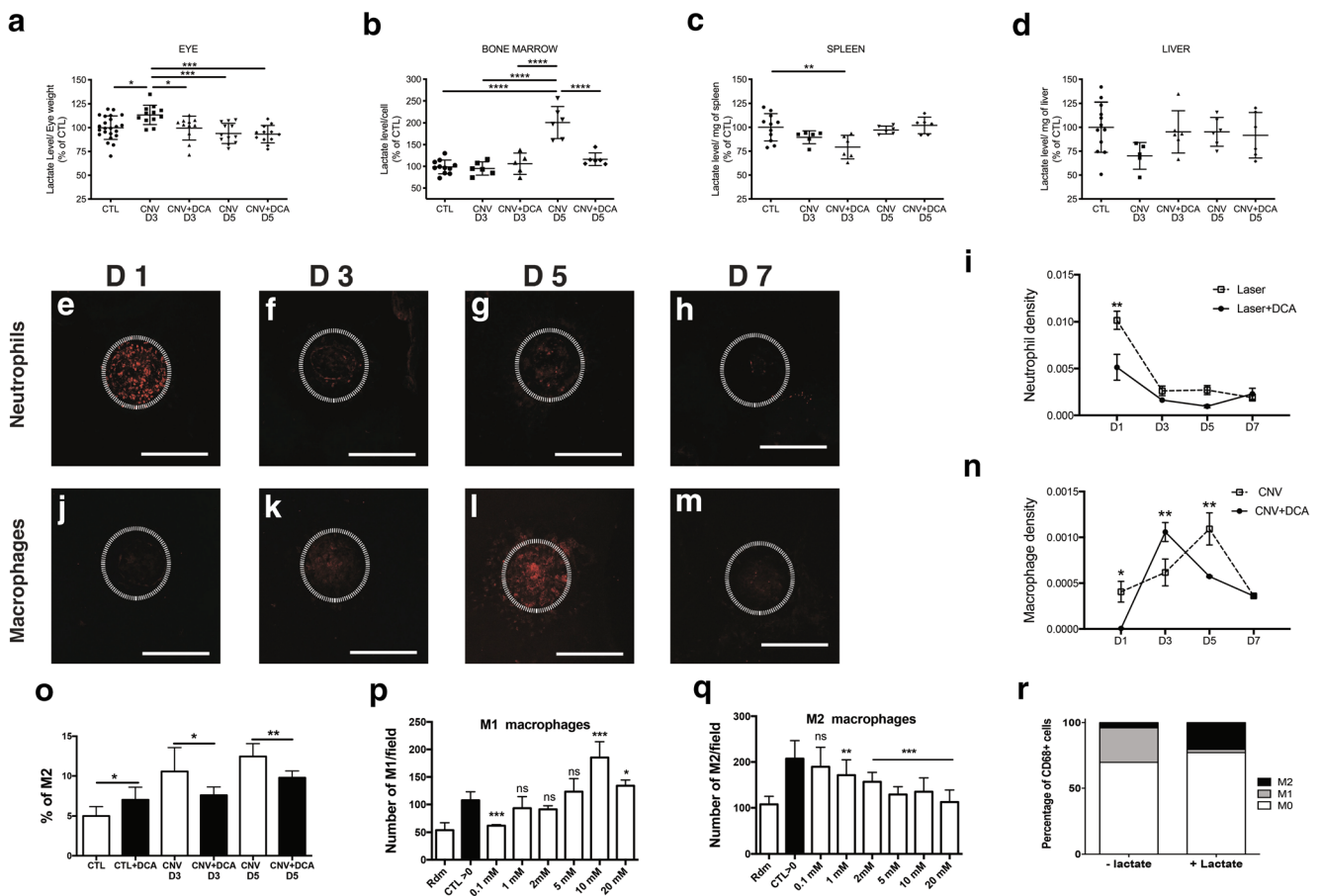
Fig. 4 Lactate mediates the recruitment of inflammatory cells during CNV. Mice subjected to laser-induced CNV were treated or not with DCA (3 mg DCA/day/mouse). NMR dosage of the lactate level at day 3 and day 5 in eye (**a**), bone marrow (**b**), spleen (**c**), and liver (**d**) ($n \geq 5$ mice/group). *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; data are expressed as the mean ± SD. Immunostaining of neutrophils (**e–i**) and macrophages (**j–n**) on flat-mounted choroids collected at day 1 (D1: **e**, **j**), day 3 (D3: **f**, **k**), day 5 (D5: **g**, **l**), and day 7 (D7: **h**, **m**) after laser induction. Scale bars, 100 μm. Dashed lines delineate the lesion. Macrophage/neutrophil density defined as the volume occupied by cells divided by the total laser impact volume with or without treatment with DCA (**i**, **n**). Percentage of M2 macrophages among CD45+ cells determined by flow cytometry ($n = 9$ eyes/group) (**o**). CTL corresponds to mice not subjected to laser burn. Boyden chamber assay showing the differential migration of human macrophages polarized in vitro into the M1 (**p**) or M2 (**q**) subtype ($n = 6$ wells/condition). M1 cells were attracted by high and M2 by low-lactate concentrations. Rdm corresponds to random cell migration in the absence of chemoattractant and CTL>O to cell migration in response to BSA-containing medium (positive control). Cultures of M1 macrophages were treated or not with 10 mM lactate for 48 h (**r**). The results obtained after phenotypic analysis showed that lactate induced the conversion of M1 into M2 macrophages ($n = 5$). *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ns, nonsignificant. Data are expressed as the mean ± SEM (**i**, **n**, **o**) or mean ± SD (**p**, **q**). See also Fig. S1 for the gating strategy used in FACS analyses to identify the M2 macrophage subset in posterior eye segments after laser-induced CNV

Discussion

The present holistic study highlights that metabolomics is a relevant tool to obtain new insights into nAMD. Here, we assign a functional role for the PDK/lactate axis in AMD and in CNV progression that holds promise for new treatment. Moreover, a shift in the lipoprotein profile towards higher VLDL-LDL proportions was associated with CNV development providing novel markers of pathology progression that could be suitable for patient treatment follow-up and personalized medicine. The clinical relevance of our innovative

findings is supported by similar metabolomics changes detected in samples of patients with nAMD and in murine samples in a pre-clinical model.

The most important finding of our human study is the detection of metabolomics differences between patients with active AMD (in an active exudative phase of the pathology) and healthy or non-active patients. This original observation underlines the underappreciated nonlinear features of this chronic disease and reveals that the exudative phase in patients is concomitant with serum changes detectable by metabolomics. Notably, increased serum lactate levels and a shift in the
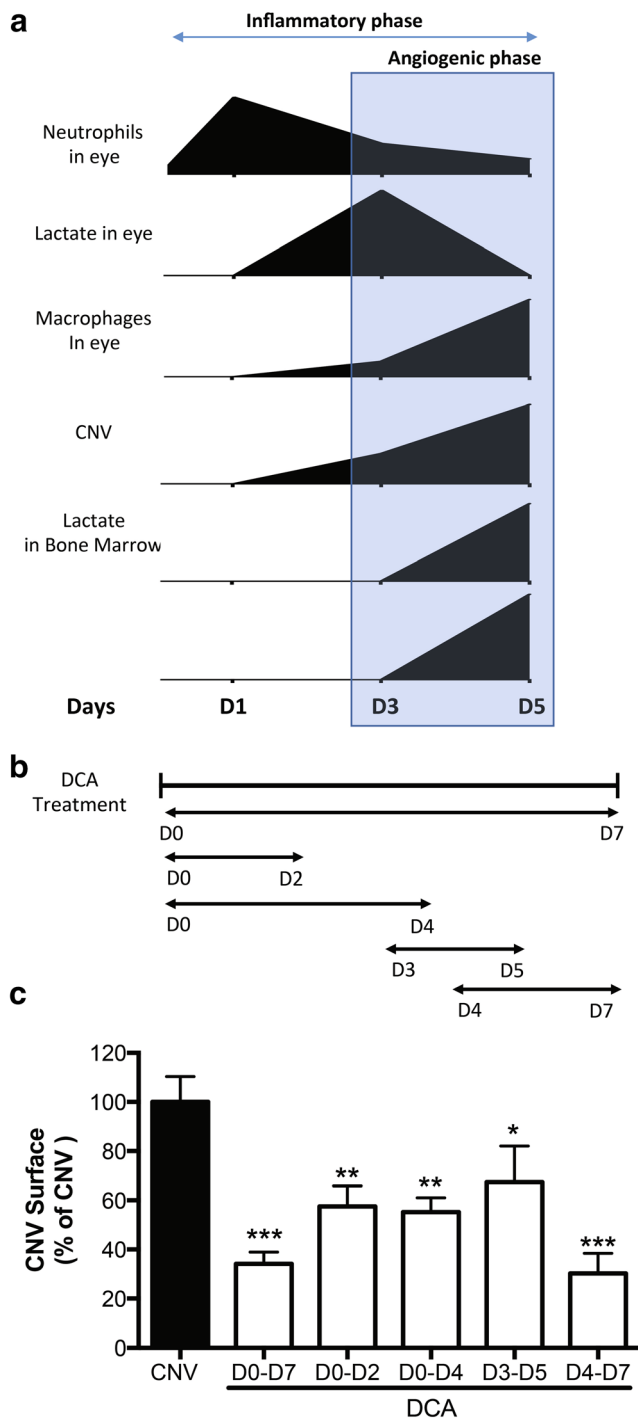
Fig. 5 DCA treatment during the angiogenic phase of CNV formation is as efficient as anti-VEGF treatment in the murine model. **a** Summary diagram of the kinetics of inflammation (neutrophils and macrophages), CNV formation, and lactate levels in eyes and sera. **b** The schedules of DCA treatment (3 mg DCA/day/mouse) are indicated below. **c** Quantification of fluorescent neovessel area ($n \geq 5$ mice/group) showing the maximal reduction of CNV development obtained with the DCA treatments applied from day 0 to day 7 or from day 4 to day 7. The results are expressed as the percentage of laser-induced mice. *$P < 0.05$; ***$P < 0.001$; ns, nonsignificant. Error bars are SEM

lipoproteins profile (increased VLDL-LDL/HDL ratio) were noted during the active phase of the pathology. Furthermore, the increase in the VLDL-LDL proportion is in line with reported data that associated AMD with a higher level of LDL [38]. Similar metabolic alterations as those seen in the human study (increase in lactate levels and changes in the VLDL profile) were identified in mice during CNV formation. The good correlation between the metabolomics data generated from mouse and human samples validates our concept and also indicates that the animal model correctly reflects the active phase of nAMD [27].

We next focused our investigation on the PDK/lactate axis as a putative therapeutic target for AMD treatment. The implication of lactate in angiogenesis and inflammation, the two underlying causes of AMD, is well documented [39, 40]. However, its role in ocular disease has not yet been reported. In the present study, we provide evidence that lactate controls the recruitment of macrophages, which are well recognized as key cellular regulators of CNV formation [41, 42]. Importantly, the normalization of lactate levels by pharmacological inhibition of PDK in mice modified the kinetics of macrophage recruitment and regulated M1/M2 balance by reducing the percentage of the M2 subtype. M2 macrophages are considered pro-angiogenic cells that contribute to different pathologies including AMD [43, 44]. Our data reveal that lactate can exert direct and differential effects on M1 and M2 macrophage properties. Lactate is a chemoattractant for M1- and M2-like macrophages with different optimal concentrations for each macrophage subtype. Of note, a key finding here is that lactate promotes the repolarization of M1-like to M2-like macrophages. Our observations are in line with a previous study implicating lactate in the M2 macrophage polarization in the context of cancer [45]. Therefore, increasing concentrations of lactate could attract M1 macrophages that can locally repolarize into pro-angiogenic M2-like macrophages. Altogether, our data support the interest of targeting lactate metabolism as a promising approach, not only for cancer [46] but also for AMD.

The source of lactate and its increased levels detected in the serum are intriguing issues. We are providing evidence that the increased lactate levels detected in eyes preceded those in sera. This result supports the local production of lactate at an early stage following eye injury, which can be caused by laser burn in mice and likely by lifelong exposure of the retina and the underlying retinal pigment epithelium (RPE) to different stimuli (light, oxidative stress) in aging patients. At the cellular level, lactate could be secreted by cells exhibiting aerobic glycolysis, such as RPE [47] and endothelial cells [48]. Hypoxia could also trigger glycolysis and lactate production by resident and inflammatory cells. A progressive enhancement of lactate production was observed in eyes until day 3, excluding neutrophil involvement in this secretion. A very important finding is the increased lactate production observed

at day 5, concomitantly in bone marrow extracts and sera. We provide evidence for lactate production in two sites: a local one caused by ocular tissue injury and a systemic one induced by circulating inflammatory cells.

This observed dual production of lactate explains the similarity in metabolomics fingerprints detected in patients with nAMD and in the experimental model, which could be considered intriguing. In the course of AMD, age-related chronic para-inflammation (an intermediate between basal and inflammatory states) occurs in the retinochoroidal tissues [7, 49]. Constant exposition to light stimulation and other oxidative stresses generates large amounts of oxidized materials and reactive oxygen species (ROS) that promote complement activation, cell damage, drusen formation, and irreversible photoreceptor degeneration. Prolonged ROS augmentation, hypoxia, and chronic inflammation can then trigger angiogenic signaling leading to CNV, a hallmark of advanced stage of the disease. In the experimental CNV mouse model, the laser injury induced Bruch's membrane rupture, cell damage, and hypoxia, finally causing CNV formation. Although the events leading to human CNV are clearly different from those in laser-induced trauma, the subsequent angiogenic response that occurs in human CNV due to the breakdown of Bruch's membrane is reproduced in the laser injury model [27]. Thus, the murine laser-induced CNV model reproduces the late events occurring in nAMD progression. This concept is further supported by our following findings: (1) in clinical samples, higher lactate levels and lipoprotein profile shifts are detected in patients with active AMD; (2) in the experimental model, metabolite changes are detected from day 5 in blood and are concomitant to late-inflammatory and angiogenic phases; (3) lactate level normalization through a pharmacological approach is more potent when applied at late (> day 4 post-cauterization) than at early (< day 4) time points; and finally (4) the inhibition of angiogenesis through anti-VEGF antibody administration normalized lactate levels. These observations further support the relevance of the laser-induced CNV model in reproducing the AMD pathology [27].

Of great interest is our finding that lactate levels and lipoprotein profiles are normalized and modified upon anti-VEGF and PDK inhibitor treatments. Our study opens a new horizon for personalized medicine and the follow-up of patients with nAMD during anti-angiogenic therapy. Indeed, to determine the number and frequency of anti-VEGF intravitreous injections, clinical practitioners are currently mostly following general guidelines in the absence of evidence-based methods to guide them. Personalized-based protocols of treatment are not yet available, and tools are lacking for providing a rationale for decision-making and guiding clinicians. In this context, our metabolomics results offer an opportunity to monitor some metabolite levels in biofluids such as lactate and the lipoprotein profiles. The lactate level is followed as a prognostic/predictive marker in cancers [50, 51], and combined with the lipoprotein profile, it could be used for the follow-up of patients with nAMD during anti-angiogenic treatment and/or in order to personalize therapeutic interventions for nAMD. Moreover, the normalization of lactate levels by oral PDK modulators clearly represents a new therapeutic option and/or a complementary treatment to reduce CNV progression.

In conclusion, our study provides new conceptual insights into the pathogenesis and evolution of poorly understood nAMD and supports the innovative concept of the PDK//lactate axis as a functional, traceable, and targetable mediator of CNV and nAMD progression. In addition to offering a novel therapeutic option for AMD treatments, our work suggests that metabolomics profiling and lactate/lipoprotein level monitoring during treatment are worth considering for the follow-up of patients with AMD in future clinical studies and offer new tools to help clinicians in personalized therapeutic design.

Compliance with ethical standards

Ethics Animal experiments were performed in compliance with the Animal Ethical Committee of the Liège University (Liège, Belgium) after the approval of the local Animal Ethical Committee. The human study was conducted under protocols approved by the Ethical Committee of the University Hospital of Liège, B7072006295 (Belgium). Informed consent was obtained from all study subjects before participation.

Code availability Metabolomics data were processed using Bruker Topspin 3.5 and AMIX 3.9 software. Metabolomics analyses were performed using SIMCA 14.1 software. Statistical analyses were performed using GraphPad Prism 7.0 software.

References

1. Colijn JM, Buitendijk GHS, Prokofyeva E, Alves D, Cachulo ML, Khawaja AP, Cougnard-Gregoire A, Merle BMJ, Korb C, Erke MG, Bron A, Anastosopoulos E, Meester-Smoor MA, Segato T, Piermarocchi S, de Jong PTVM, Vingerling JR, Topouzis F, Creuzot-Garcher C, Bertelsen G, Pfeiffer N, Fletcher AE, Foster PJ, Silva R, Korobelnik J-F, Delcourt C, Klaver CCW, EYE-RISK consortium, European Eye Epidemiology (E3) consortium (2017) Prevalence of age-related macular degeneration in Europe: the past and the future. Ophthalmology 124:1753–1763

2. Schmidt-Erfurth U, Klimscha S, Waldstein SM, Bogunović H (2017) A view of the current and future role of optical coherence tomography in the management of age-related macular degeneration. Eye (Lond) 31:26–44

3. Lai T-T, Hsieh Y-T, Yang C-M, Ho T-C, Yang C-H (2019) Biomarkers of optical coherence tomography in evaluating the treatment outcomes of neovascular age-related macular degeneration: a real-world study. Sci Rep 9. https://doi.org/10.1038/s41598-018-36704-6,

4. Cascella R, Strafella C, Caputo V, Errichiello V, Zampatti S, Milano F, Potenza S, Mauriello S, Novelli G, Ricci F, Cusumano A, Giardina E (2018) Towards the application of precision medicine in age-related macular degeneration. Prog Retin Eye Res 63:132–146

5. DeAngelis MM, Owen LA, Morrison MA, Morgan DJ, Li M, Shakoor A, Vitale A, Iyengar S, Stambolian D, Kim IK, Farrer LA (2017) Genetics of age-related macular degeneration (AMD). Hum Mol Genet 26:R45–R50

6. Liu K, Chen LJ, Lai TY, Tam PO, Ho M, Chiang SW, Liu DT, Young AL, Yang Z, Pang CP (2014) Genes in the high-density lipoprotein metabolic pathway in age-related macular degeneration and polypoidal choroidal vasculopathy. Ophthalmology 121:911–916

7. Nita M, Grzybowski A, Ascaso FJ, Huerva V (2014) Age-related macular degeneration in the aspect of chronic low-grade inflammation (pathophysiological parainflammation). Mediat Inflamm 2014:930671–930610

8. Lavalette S, Raoul W, Houssier M, Camelo S, Levy O, Calippe B, Jonet L, Behar-Cohen F, Chemtob S, Guillonneau X, Combadiere C, Sennlaub F (2011) Interleukin-1beta inhibition prevents choroidal neovascularization and does not exacerbate photoreceptor degeneration. Am J Pathol 178:2416–2423

9. Doyle SL, Ozaki E, Brennan K, Humphries MM, Mulfaul K, Keaney J, Kenna PF, Maminishkis A, Kiang AS, Saunders SP, Hams E, Lavelle EC, Gardiner C, Fallon PG, Adamson P, Humphries P, Campbell M (2014) IL-18 attenuates experimental choroidal neovascularization as a potential therapy for wet age-related macular degeneration. Sci Transl Med 6:230ra44

10. Xi H, Katschke KJ Jr, Li Y, Truong T, Lee WP, Diehl L, Rangell L, Tao J, Arceo R, Eastham-Anderson J, Hackney JA, Iglesias A, Cote-Sierra J, Elstrott J, Weimer RM, Campagne MV (2016) IL-33 amplifies an innate immune response in the degenerating retina. J Exp Med 213:189–207

11. Colijn JM, den Hollander AI, Demirkan A, Cougnard-Grégoire A, Verzijden T, Kersten E, Meester-Smoor MA, Merle BMJ, Papageorgiou G, Ahmad S, Mulder MT, Costa MA, Benlian P, Bertelsen G, Bron AM, Claes B, Creuzot-Garcher C, Erke MG, Fauser S, Foster PJ, Hammond CJ, Hense H-W, Hoyng CB, Khawaja AP, Korobelnik J-F, Piermarocchi S, Segato T, Silva R, Souied EH, Williams KM, van Duijn CM, Delcourt C, Klaver CCW, Acar N, Altay L, Anastosopoulos E, Azuara-Blanco A, Berendschot T, Berendschot T, Bergen A, Bertelsen G, Binquet C, Bird A, Bobak M, Larsen MB, Boon C, Bourne R, Brétillon L, Broe R, Bron A, Buitendijk G, Cachulo ML, Capuano V, Carrière I, Chakravarthy U, Chan M, Chang P, Colijn J, Cougnard-Grégoire A, Cree A, Creuzot-Garcher C, Cumberland P, Cunha-Vaz J, Daien V, De Jong E, Deak G, Delcourt C, Delyfer M-N, den Hollander A, Dietzel M, Erke MG, Faria P, Farinha C, Fauser S, Finger R, Fletcher A, Foster P, Founti P, Gorgels T, Grauslund J, Grus F, Hammond C, Heesterbeek T, Hense H-W, Hermann M, Hoehn R, Hogg R, Holz F, Hoyng C, Jansonius N, Janssen S, de Jong E, Khawaja A, Klaver C, Korobelnik J-F, Lamparter J, Le Goff M, Lehtimäki T, Leung I, Lotery A, Mauschitz M, Meester M, Merle B, Meyer zu Westrup V, Midena E, Miotto S, Mirshahi A, Mohan-Saïd S, Mueller M, Muldrew A, Murta J, Nickels S, Nunes S, Owen C, Peto T, Pfeiffer N, Piermarocchi S, Prokofyeva E, Rahi J, Raitakari O, Rauscher F, Ribeiro L, Rougier M-B, Rudnicka A, Sahel J, Salonikiou A, Sanchez C, Schick T, Schmitz-Valckenberg S, Schuster A, Schweitzer C, Segato T, Shehata J, Silva R, Silvestri G, Simader C, Souied E, Speckauskas M, Springelkamp H, Tapp R, Topouzis F, van Leeuwen E, Verhoeven V, Verzijden T, Vingerling H, Von Hanno T, Williams K, Wolfram C, Yip J, Zerbib J, Ajana S, Arango-Gonzalez B, Arndt V, Bhatia V, Bhattacharya SS, Biarnés M, Borrell A, Bühren S, Calado SM, Colijn JM, Cougnard-Grégoire A, Dammeier S, de Jong EK, De la Cerda B, Delcourt C, den Hollander AI, Diaz-Corrales FJ, Diether S, Emri E, Endermann T, Ferraro LL, Garcia M, Heesterbeek TJ, Honisch S, Hoyng CB, Kersten E, Kilger E, CCW K, Langen H, Lengyel I, Luthert P, Maugeais C, Meester-Smoor M, BMJ MI, Monés J, Nogoceke E, Peto T, Pool FM, Rodríguez E, Ueffing M, Ulrich Bartz-Schmidt KU, van Leeuwen EM, Verzijden T, Zumbansen M (2019) Increased high-density lipoprotein levels associated with age-related macular degeneration. Ophthalmology 126:393–406

12. Noel A, Jost M, Lambert V, Lecomte J, Rakic JM (2007) Anti-angiogenic therapy of exudative age-related macular degeneration: current progress and emerging concepts. Trends Mol Med 13:345–352

13. Cheung GCM, Lai TYY, Gomi F, Ruamviboonsuk P, Koh A, Lee WK (2017) Anti-VEGF therapy for neovascular AMD and polypoidal choroidal vasculopathy. Asia Pac J Ophthalmol (Phila) 6:527–534

14. Nagai N, Suzuki M, Uchida A, Kurihara T, Kamoshita M, Minami S, Shinoda H, Tsubota K, Ozawa Y (2016) Non-responsiveness to intravitreal aflibercept treatment in neovascular age-related macular degeneration: implications of serous pigment epithelial detachment. Sci Rep 6:29619

15. Sun X, Yang S, Zhao J (2016) Resistance to anti-VEGF therapy in neovascular age-related macular degeneration: a comprehensive review. DDDT 1857. https://doi.org/10.2147/DDDT.S97653

16. Schmidt-Erfurth U, Chong V, Loewenstein A, Larsen M, Souied E, Schlingemann R, Eldem B, Mones J, Richard G, Bandello F (2014) Guidelines for the management of neovascular age-related macular degeneration by the European Society of Retina Specialists (EURETINA). Br J Ophthalmol 98:1144–1167

17. Frédérich M, Pirotte B, Fillet M, de Tullio P (2016) Metabolomics as a challenging approach for medicinal chemistry and personalized medicine. J Med Chem 59:8649–8666

18. Beger RD, Dunn W, Schmidt MA, Gross SS, Kirwan JA, Cascante M, Brennan L, Wishart DS, Oresic M, Hankemeier T, Broadhurst DI, Lane AN, Suhre K, Kastenmüller G, Sumner SJ, Thiele I, Fiehn O, Kaddurah-Daouk R, for "Precision Medicine and Pharmacometabolomics Task Group"-Metabolomics Society Initiative (2016) Metabolomics enables precision medicine: "a white paper, community perspective". Metabolomics 12:149

19. Li B, He X, Jia W, Li H (2017) Novel applications of metabolomics in personalized medicine: a mini-review. Molecules 22. https://doi.org/10.3390/molecules22071173

20. Draoui N, de Zeeuw P, Carmeliet P (2017) Angiogenesis revisited from a metabolic perspective: role and therapeutic implications of endothelial cell metabolism. Open Biol 7:170219

21. Mills E, O'Neill LAJ (2014) Succinate: a metabolic signal in inflammation. Trends Cell Biol 24:313–320

22. Brown CN, Green BD, Thompson RB, den Hollander AI, Lengyel I (2019) Metabolomics and age-related macular degeneration. Metabolites 9. https://doi.org/10.3390/metabo9010004

23. Luo D, Deng T, Yuan W, Deng H, Jin M (2017) Plasma metabolomic study in Chinese patients with wet age-related macular degeneration. BMC Ophthalmol 17:165

24. Mitchell SL, Uppal K, Williamson SM, Liu K, Burgess LG, Tran V, Umfress AC, Jarrell KL, Cooke Bailey JN, Agarwal A, Pericak-Vance M, Haines JL, Scott WK, Jones DP, Brantley MA (2018) The carnitine shuttle pathway is altered in patients with neovascular age-related macular degeneration. Invest Ophthalmol Vis Sci 59:4978–4985

25. Osborn MP, Park Y, Parks MB, Burgess LG, Uppal K, Lee K, Jones DP, Brantley MA Jr (2013) Metabolome-wide association study of neovascular age-related macular degeneration. PLoS One 8:e72737

26. Laíns I, Duarte D, Barros AS, Martins AS, Gil J, Miller JB, Marques M, Mesquita T, Kim IK, Cachulo M d L, Vavvas D, Carreira IM, Murta JN, Silva R, Miller JW, Husain D, Gil AM (2017) Human plasma metabolomics in age-related macular degeneration (AMD) using nuclear magnetic resonance spectroscopy. PLoS One 12:e0177749

27. Lambert V, Lecomte J, Hansen S, Blacher S, Gonzalez ML, Struman I, Sounni NE, Rozet E, de Tullio P, Foidart JM, Rakic JM, Noel A (2013) Laser-induced choroidal neovascularization model to study age-related macular degeneration in mice. Nat Protoc 8:2197–2211

28. Adeva M, González-Lucán M, Seco M, Donapetry C (2013) Enzymes involved in l-lactate metabolism in humans. Mitochondrion 13:615–629

29. Roche TE, Hiromasa Y (2007) Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer. Cell Mol Life Sci 64:830–849

30. Bian L, Josefsson E, Jonsson IM, Verdrengh M, Ohlsson C, Bokarewa M, Tarkowski A, Magnusson M (2009) Dichloroacetate alleviates development of collagen II-induced arthritis in female DBA/1 mice. Arthritis Res Ther 11:R132

31. Matheus N, Hansen S, Rozet E, Peixoto P, Maquoi E, Lambert V, Noel A, Frederich M, Mottet D, de Tullio P (2014) An easy, convenient cell and tissue extraction protocol for nuclear magnetic resonance metabolomics. Phytochem Anal: PCA 25:342–349

32. Sounni NE, Cimino J, Blacher S, Primac I, Truong A, Mazzucchelli G, Paye A, Calligaris D, Debois D, De Tullio P, Mari B, De Pauw E, Noel A (2014) Blocking lipid synthesis overcomes tumor regrowth and metastasis after antiangiogenic therapy withdrawal. Cell Metab 20:280–294

33. Mia S, Warnecke A, Zhang X-M, Malmström V, Harris RA (2014) An optimized protocol for human M2 macrophages using M-CSF and IL-4/IL-10/TGF-β yields a dominant immunosuppressive phenotype. Scand J Immunol 79:305–314

34. Edin S, Wikberg ML, Rutegård J, Oldenborg P-A, Palmqvist R (2013) Phenotypic skewing of macrophages in vitro by secreted factors from colorectal cancer cells. PLoS One 8:e74982. https://doi.org/10.1371/journal.pone.0074982

35. Hubert P, van den Brûle F, Giannini SL, Franzen-Detrooz E, Boniver J, Delvenne P (1999) Colonization of in vitro-formed cervical human papillomavirus-associated (pre)neoplastic lesions with dendritic cells. Am J Pathol 154:775–784

36. Pennesi ME, Neuringer M, Courtney RJ (2012) Animal models of age related macular degeneration. Mol Asp Med 33:487–509

37. Michelakis ED, Webster L, Mackey JR (2008) Dichloroacetate (DCA) as a potential metabolic-targeting therapy for cancer. Br J Cancer 99:989–994

38. Kersten E, Paun CC, Schellevis RL, Hoyng CB, Delcourt C, Lengyel I, Peto T, Ueffing M, Klaver CCW, Dammeier S, den Hollander AI, de Jong EK (2018) Systemic and ocular fluid compounds as potential biomarkers in age-related macular degeneration. Surv Ophthalmol 63:9–39

39. Porporato PE, Payen VL, De Saedeleer CJ, Préat V, Thissen J-P, Feron O, Sonveaux P (2012) Lactate stimulates angiogenesis and accelerates the healing of superficial and ischemic wounds in mice. Angiogenesis 15:581–592

40. Beckert S, Farrahi F, Aslam RS, Scheuenstuhl H, Konigsrainer A, Hussain MZ, Hunt TK (2006) Lactate stimulates endothelial cell migration. Wound Repair Regen 14:321–324

41. Skeie JM, Mullins RF (2009) Macrophages in neovascular age-related macular degeneration: friends or foes? Eye (Lond) 23:747–755

42. Sakurai E, Anand A, Ambati BK, van Rooijen N, Ambati J (2003) Macrophage depletion inhibits experimental choroidal neovascularization. Invest Ophthalmol Vis Sci 44:3578–3585

43. Zandi S, Nakao S, Chun KH, Fiorina P, Sun D, Arita R, Zhao M, Kim E, Schueller O, Campbell S, Taher M, Melhorn MI, Schering A, Gatti F, Tezza S, Xie F, Vergani A, Yoshida S, Ishikawa K, Yamaguchi M, Sasaki F, Schmidt-Ullrich R, Hata Y, Enaida H, Yuzawa M, Yokomizo T, Kim YB, Sweetnam P, Ishibashi T, Hafezi-Moghadam A (2015) ROCK-isoform-specific polarization of macrophages associated with age-related macular degeneration. Cell Rep 10:1173–1186

44. Jetten N, Verbruggen S, Gijbels MJ, Post MJ, De Winther MPJ, Donners MMPC (2014) Anti-inflammatory M2, but not proinflammatory M1 macrophages promote angiogenesis in vivo. Angiogenesis 17:109–118

45. Colegio OR, Chu NQ, Szabo AL, Chu T, Rhebergen AM, Jairam V, Cyrus N, Brokowski CE, Eisenbarth SC, Phillips GM, Cline GW, Phillips AJ, Medzhitov R (2014) Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. Nature 513:559–563

46. Dhup S, Dadhich RK, Porporato PE, Sonveaux P (2012) Multiple biological activities of lactic acid in cancer: influences on tumor growth, angiogenesis and metastasis. Curr Pharm Des 18:1319–1330

47. Hurley JB, Lindsay KJ, Du J (2015) Glucose, lactate, and shuttling of metabolites in vertebrate retinas. J Neurosci Res 93:1079–1092

48. De Bock K, Georgiadou M, Schoors S, Kuchnio A, Wong BW, Cantelmo AR, Quaegebeur A, Ghesquiere B, Cauwenberghs S, Eelen G, Phng LK, Betz I, Tembuyser B, Brepoels K, Welti J, Geudens I, Segura I, Cruys B, Bifari F, Decimo I, Blanco R, Wyns S, Vangindertael J, Rocha S, Collins RT, Munck S, Daelemans D, Imamura H, Devlieger R, Rider M, Van Veldhoven PP, Schuit F, Bartrons R, Hofkens J, Fraisl P, Telang S, Deberardinis RJ, Schoonjans L, Vinckier S, Chesney J, Gerhardt H, Dewerchin M, Carmeliet P (2013) Role of PFKFB3-driven glycolysis in vessel sprouting. Cell 154:651–663

49. Parmeggiani F, Romano MR, Costagliola C, Semeraro F, Incorvaia C, D'Angelo S, Perri P, De Palma P, De Nadai K, Sebastiani A (2012) Mechanism of inflammation in age-related macular degeneration. Mediat Inflamm 2012:546786–546716

50. Brizel DM, Schroeder T, Scher RL, Walenta S, Clough RW, Dewhirst MW, Mueller-Klieser W (2001) Elevated tumor lactate concentrations predict for an increased risk of metastases in head-and-neck cancer. Int J Radiat Oncol Biol Phys 51:349–353

51. Quennet V, Yaromina A, Zips D, Rosner A, Walenta S, Baumann M, Mueller-Klieser W (2006) Tumor lactate content predicts for response to fractionated irradiation of human squamous cell carcinomas in nude mice. Radiother Oncol 81:130–135

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.